



DOCTORAL SCHOOL  
UNIVERSITY OF MILANO-BICOCCA

Department of Informatics, Systems and Communication  
Ph.D. Program in Computer Science – Cycle XXXI

# **Computer-Assisted Analysis of Biomedical Images**

Ph.D. Thesis of: Leonardo Rundo

Supervisor: Prof. Giancarlo Mauri

Tutor: Prof. Giuseppe Vizzari

Ph.D. Coordinator: Prof. Stefania Bandini

ACADEMIC YEAR 2017-2018

*To my parents and Miriana,  
for your loving support.*

## Acknowledgements

First of all, I want to thank my parents to instill in me the strong values about the scientific work and its social impact. I am essentially the result of your infinite love and perseverance. Obviously, the motivations for my Ph.D. have been strengthened by the presence of a unique person. Miriana, you always supported my choices, even the most difficult to accept, with love and passion. Many thanks also to every single component of your family, especially your parents Carmela and Francesco for treating me like a son during my short weekends. A loving thought goes to my little nieces (Maria Francesca, Noemi, Maria), nephews (Lorenzo, Angelo, Francesco), and cousins (Giuseppe and Antonino) who managed to gladden me even during the most difficult moments in these intense years. I am grateful to my entire family, from the South to the North of Italy, for cheering me throughout this fantastic journey. My uncles Benedetto e Marianna have been like parents and my cousins Benedetta and Adriano, Cristina and Salvatore like real brothers and sisters. In addition, my cousins Anna Maria and Salvatore hosted me when my Ph.D. admission seemed to be a crazy idea, while Mariella and Nino were fundamental in my settling-in period in Milan.

I would like to sincerely thank all the people who discovered potential in me, even when I was a shy Master's student. Prof. Roberto Pirrone, Prof. Orazio Gambino, and Dr. Vincenzo Cannella presented the fascinating medical imaging context for my Master's Thesis activities at the University of Palermo. I thank Prof. Maria Carla Gilardi, Dr. Carmelo Militello, Dr. Giorgio Russo, Dr. Alessandro Stefano, for providing valuable opportunities for my Research Fellowship at the Institute of Molecular Bioimaging and Physiology – Italian National Research Council. In particular, Carmelo has been my affectionate “elder brother” in the last five years for my scientific advances. I am also grateful to Prof. Salvatore Vitabile (University of Palermo) and Prof. Vincenzo Conti (University of Enna KORE) for trustfully involving me in challenging projects.

My Ph.D. program has been an amazing experience mainly thanks to Prof. Giancarlo Mauri. He is not just my Supervisor, but he has been a perfect mentor who always demonstrated to believe in me. In addition to a significant impact on my education, his kindness forged my approach to research and daily life. I wish to thank the people working at DISCo for their friendship and for welcoming me from the first moment. I absolutely learned a lot from all of you. Especially, the research group composed of Prof. Daniela Besozzi, Prof. Paolo Cazzaniga, Dr. Marco Nobile, Andrea Tangherloni, and Simone Spolaor. In addition to introducing me to the Computational Systems Biology field, I have been growing up as a scientist thanks to your methodological rigor and dedication. I am very lucky for really enjoying these years with you and your students (especially Simone and Riccardo). Andrea, recalling our memories at the Vanderbilt University and at the University of Cambridge, we are going to spend a lot of new memorable experiences together in Cambridge. A great thank to Prof. Giuseppe Vizzari for his tutoring and precious support in every moment, in particular at the very beginning of my Ph.D. activities.

Special thanks are directed to: Prof. Carlos Lopez, along with Prof. Vito Quaranta and Prof. Darren Tyson, for inviting me at the Vanderbilt University, Nashville, TN, USA; Prof. Hideki Nakayama (The University of Tokyo, Tokyo, Japan) together with Changhee Han (a real friend who literally hosted me in his own home, so spending wonderful Japanese moments with the other guys in the Lab) for our international collaboration project; last but not least, Prof. Pietro Liò (University of Cambridge, Cambridge, UK) who practically transmits the love for science in everyday life. The experience as a Visiting Scholar in your laboratories and the fruitful collaborations with your group members positively changed my vision about scientific research.

In this regard, I am grateful to Prof. Evis Sala who is investing in me for my next exciting adventure at the University of Cambridge. Tons of thanks go to my new colleagues at the Department of Radiology, since you have made me feel at ease from the first moment.



## Abstract

Nowadays, the amount of heterogeneous biomedical data is increasing more and more thanks to novel sensing techniques and high-throughput technologies. In reference to biomedical image analysis, the advances in image acquisition modalities and high-throughput imaging experiments are creating new challenges. This huge information ensemble could overwhelm the analytic capabilities needed by physicians in their daily decision-making tasks as well as by biologists investigating complex biochemical systems. In particular, quantitative imaging methods convey scientifically and clinically relevant information in prediction, prognosis or treatment response assessment, by also considering radiomics approaches. Therefore, the computational analysis of medical and biological images plays a key role in radiology and laboratory applications. In this regard, frameworks based on advanced Machine Learning and Computational Intelligence can significantly improve traditional Image Processing and Pattern Recognition approaches. However, conventional Artificial Intelligence techniques must be tailored to address the unique challenges concerning biomedical imaging data.

This thesis aims at proposing novel and advanced computer-assisted methods for biomedical image analysis, also as an instrument in the development of Clinical Decision Support Systems, by always keeping in mind the clinical feasibility of the developed solutions. The devised classical Image Processing algorithms, with particular interest to region-based and morphological approaches in biomedical image segmentation, are first described. Afterwards, Pattern Recognition techniques are introduced, applying unsupervised fuzzy clustering and graph-based models (i.e., Random Walker and Cellular Automata) to multispectral and multimodal medical imaging data processing. Taking into account Computational Intelligence, an evolutionary framework based on Genetic Algorithms for medical image enhancement and segmentation is presented. Moreover, multimodal image co-registration using Particle Swarm Optimization is discussed. Finally, Deep Neural Networks are investigated: (i) the generalization abilities of Convolutional Neural Networks in medical image segmentation for multi-institutional datasets are addressed by conceiving an architecture that integrates adaptive feature recalibration blocks, and (ii) the generation of realistic medical images based on Generative Adversarial Networks is applied to data augmentation purposes. In conclusion, the ultimate goal of these research studies is to gain clinically and biologically useful insights that can guide differential diagnosis and therapies, leading towards biomedical data integration for personalized medicine. As a matter of fact, the proposed computer-assisted bioimage analysis methods can be beneficial for the definition of imaging biomarkers, as well as for quantitative medicine and biology.



# Table of contents

<b>List of figures</b>	<b>ix</b>
<b>List of tables</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Aims and motivations . . . . .	1
1.2 Computer science contributions to biomedical imaging . . . . .	3
1.3 Scientific production and thesis overview . . . . .	3
1.4 Other publications . . . . .	8
<b>2 Biomedical image analysis and computing</b>	<b>11</b>
2.1 Background . . . . .	11
2.1.1 Medical imaging . . . . .	11
2.1.2 Bioimage informatics . . . . .	18
2.2 Investigated contexts and applications . . . . .	18
2.2.1 Medical image analysis . . . . .	18
2.2.2 Biological image analysis . . . . .	39
<b>3 Classical Image Processing algorithms</b>	<b>47</b>
3.1 Region-based approaches . . . . .	47
3.1.1 Image thresholding . . . . .	47
3.1.2 Region splitting-and-merging . . . . .	65
3.1.3 Region growing . . . . .	67
3.2 Morphological approaches . . . . .	74
3.2.1 Watershed transform . . . . .	74
<b>4 Pattern Recognition techniques</b>	<b>87</b>
4.1 Unsupervised fuzzy clustering . . . . .	87
4.1.1 Brain tumor segmentation in neuro-radiosurgery . . . . .	90

4.1.2	NeXt: necrosis extraction in neuro-radiosurgery . . . . .	99
4.1.3	Prostate gland segmentation on multispectral MRI . . . . .	109
4.2	Graph-based methods . . . . .	125
4.2.1	Shortest Path with Cellular Automata . . . . .	126
4.2.2	Random Walker . . . . .	138
<b>5</b>	<b>Computational Intelligence</b>	<b>163</b>
5.1	Genetic Algorithms . . . . .	163
5.1.1	Medical image enhancement . . . . .	166
5.1.2	Computational framework to improve MR image thresholding . . . . .	187
5.2	Particle Swarm Optimization . . . . .	197
5.2.1	Multimodal biomedical image registration using PSO . . . . .	201
<b>6</b>	<b>Deep Neural Networks</b>	<b>211</b>
6.1	Convolutional Neural Networks . . . . .	211
6.1.1	Cross-dataset generalization abilities of CNNs: Prostate zonal segmentation of multi-centric MRI datasets . . . . .	214
6.2	Generative Adversarial Networks . . . . .	231
6.2.1	GAN-based synthetic brain MR image generation . . . . .	232
6.2.2	PGGAN-based DA for MRI brain tumor detection . . . . .	239
<b>7</b>	<b>Conclusion</b>	<b>249</b>
7.1	Final remarks . . . . .	249
7.2	Future work . . . . .	251
	<b>Bibliography</b>	<b>253</b>
	<b>Appendix A Validation methodology</b>	<b>305</b>
A.1	Medical image segmentation evaluation metrics . . . . .	305
A.2	Medical image enhancement evaluation metrics . . . . .	308
A.3	Quantitative validation in clonogenic assays . . . . .	311
A.3.1	Conventional evaluation . . . . .	311
A.3.2	Area-based evaluation . . . . .	312
A.4	Clinical evaluation based on a Likert scale . . . . .	312
A.5	Visual Turing test . . . . .	313



# List of figures

1.1	Conceptual scheme of the main topics treated in the thesis . . . . .	4
3.1	Flow diagram of the IOTS-based fibroid segmentation method . . . . .	49
3.2	Uterine fibroid segmentation results achieved by the IOTS-based method . .	50
3.3	Overall flow diagram of the processing pipeline of the proposed fully automatic approach for ACC . . . . .	52
3.4	Examples of the control strategy used to enhance the automatic well detection performed using the CHT . . . . .	56
3.5	Correct detection of the wells in typical operating conditions . . . . .	57
3.6	Examples of colony extraction performed using the proposed approach based on local adaptive thresholding . . . . .	59
3.7	Examples of cell colony detection in different real scenarios . . . . .	60
3.8	Scatter diagrams, showing the correlation of the automated ACC-based SF evaluation ( $SF_{\%Area}$ ) versus mean manual colony counting measurements ( $SF_{Manual}$ ) . . . . .	61
3.9	Flow diagram of the SM & RG method . . . . .	66
3.10	Uterine fibroid segmentation examples achieved by the SM & RG incremental procedure . . . . .	70
3.11	Examples of input uterine fibroid T1w MR sagittal slices . . . . .	71
3.12	Workflow of uterine fibroid MRgFUS treatment response evaluation . . . .	72
3.13	Uterine fibroid ROT segmentation results: IOTS versus SM & RG . . . . .	73
3.14	Boxplots of the uterine fibroid ROT segmentation results . . . . .	73
3.15	Examples of the analyzed microscopy fluorescence images provided by the Department of Biochemistry of the Vanderbilt University . . . . .	76
3.16	Examples of the analyzed microscopy fluorescence from the 2018 Data Science Bowl dataset . . . . .	77
3.17	Flow diagram of the proposed ACDC pipeline for cell nuclei segmentation .	79
3.18	Pre-processing results in the ACDC pipeline . . . . .	80

3.19	Segmented images after the refinement steps . . . . .	82
3.20	Examples of cell nuclei segmented by ACDC . . . . .	84
3.21	Scatter plots depicting the ACDC results compared to the gold standard in terms of cell nuclei counting . . . . .	85
4.1	Flow diagram of the FCM-based brain lesion segmentation method . . . . .	91
4.2	Examples of brain MR axial slices . . . . .	92
4.3	ROI selection tools for brain tumor segmentation . . . . .	93
4.4	Necrotic area handling examples . . . . .	95
4.5	Results achieved by the FCM-based segmentation method . . . . .	95
4.6	Boxplots of the evaluation metrics results of the FCM-based brain tumor segmentation approach . . . . .	96
4.7	Volume rendering of GTVs . . . . .	97
4.8	Three instances of input T1w CE-MR axial slices with the rectangular bounding area, selected by the user . . . . .	100
4.9	Necroses extracted from the brain tumors using the proposed NeXt approach	103
4.10	Necrosis segmentation results on representative brain tumors representative necrotic regions . . . . .	103
4.11	Volumetric reconstruction of brain tumors with necrotic core . . . . .	104
4.12	Boxplots of the spatial overlap-based metrics values achieved by the implemented necrosis segmentation approaches . . . . .	105
4.13	Boxplots of the spatial distance-based metrics achieved by the implemented necrosis segmentation approaches . . . . .	106
4.14	Input prostate T1w and T2w MRI axial slice pairs . . . . .	110
4.15	Flow diagram of the proposed prostate gland MRI segmentation method . . . . .	112
4.16	Details of the prostate on a pair of fused T1w and T2w MR images . . . . .	113
4.17	Pre-Processing phase of the prostate gland MRI segmentation . . . . .	115
4.18	Results of the FCM clustering algorithm with $C = 3$ for prostate gland ROI segmentation . . . . .	117
4.19	Prostate gland ROI segmentation instances . . . . .	119
4.20	Examples of segmentation results obtained on apical, central, and basal slices	120
4.21	Boxplots of the spatial overlap-based metrics values achieved by the proposed prostate gland ROI segmentation approach based on the FCM algorithm . . . . .	121
4.22	Boxplots of the spatial distance-based metrics values achieved by the proposed prostate gland ROI segmentation approach based on the FCM algorithm	122
4.23	Tridimensional reconstruction of the segmented prostate gland ROI on the 3D orthogonal tri-planar view . . . . .	122

---

4.24	Classical neighborhoods in the 2-dimensional case: (a) von Neumann neighborhood; (b) Moore neighborhood . . . . .	127
4.25	Three instances of input T1w contrast-enhanced MR axial slices with the rectangular bounding area, selected by the user . . . . .	129
4.26	Flow diagram of the GTVCUT segmentation approach . . . . .	130
4.27	Adaptive seed selection according to the rectangular boundary regions (including the tumor) dragged by the user . . . . .	132
4.28	Brain tumor segmentation using GTVCUT in different anatomic areas of human brain . . . . .	134
4.29	Bland-Altman plot of brain tumor volume measurements . . . . .	135
4.30	Boxplots of the spatial overlap-based metrics values achieved by GTVCUT and the FCM-based method . . . . .	136
4.31	Boxplots of the spatial distance-based metrics values achieved by GTVCUT and the FCM-based method . . . . .	137
4.32	Examples of original input brain MR and PET image pairs concerning patients with brain tumors who underwent Gamma Knife . . . . .	143
4.33	Flow diagram of the proposed fully automatic multimodal PET/MRI segmentation method . . . . .	144
4.34	Instances of PET/MRI co-registration using SPM <i>via</i> maximization of NMI . . . . .	147
4.35	ROI bounding region generation on MR images . . . . .	149
4.36	Examples of brain GTV segmentation on MR images, starting from the ROI bounding region . . . . .	150
4.37	Results achieved by the proposed multimodal PET/MRI segmentation method . . . . .	152
4.38	Graphical examples of co-segmentation results achieved by the proposed multimodal PET/MRI approach . . . . .	153
4.39	Segmentation results obtained by the single modality PET method (BTV) versus those obtained by the multimodal PET/MRI approach after the PET data updating based on the results of the MRI brain lesion segmentation (BTV <sub>MRI</sub> ) . . . . .	154
4.40	Scatter diagrams of tumor volume measurements on multimodal PET/MRI data . . . . .	155
5.1	Examples of input MR images characterized by nearly bimodal histograms . . . . .	173
5.2	Plots of the global non-linear intensity transformations for image enhancement: (a) Gamma Transformation; (b) Sigmoid intensity Transformation . . . . .	175
5.3	Workflow of MedGA . . . . .	177
5.4	Graphical representation of the single-point crossover strategy . . . . .	180
5.5	Enhanced uterine fibroid MR image and histogram obtained by MedGA . . . . .	182

5.6	Enhanced brain tumor MR image and histogram obtained by MedGA . . .	183
5.7	Comparison of the ABF achieved by MedGA considering different parameterizations . . . . .	184
5.8	Flow diagram of the proposed computational framework that integrates MedGA as a pre-processing step for MR image segmentation based on the efficient IOTS algorithm . . . . .	190
5.9	Segmentation results achieved by the IOTS algorithm on the MR images pre-processed by MedGA . . . . .	192
5.10	Segmentation results achieved by MedGA and the competing methods on uterine fibroids . . . . .	193
5.11	Segmentation results achieved by MedGA and the competing methods on brain tumors . . . . .	193
5.12	Tridimensional reconstruction of the segmented ROIs . . . . .	194
5.13	Boxplots of the overlap-based and distance-based metrics obtained by MedGA and the competitor methods on the uterine fibroid MRI dataset . . . . .	195
5.14	Boxplots of the overlap-based and distance-based metrics obtained by MedGA and the competitor methods on the brain tumor MRI dataset . . . . .	196
5.15	Different medical image modalities concerning the same subject affected by brain tumor . . . . .	201
5.16	Flow diagram of the biomedical image registration process, by using an intensity-based technique . . . . .	203
6.1	Workflow of the overall CNN-based prostate zonal segmentation approach .	216
6.2	Examples of prostate zonal segmentation in pre-training/fine-tuning . . . .	221
6.3	Scheme of the proposed USE-Net architecture . . . . .	222
6.4	Kiviat diagrams showing the <i>DSC</i> values achieved by each method under different conditions . . . . .	228
6.5	Segmentation results obtained by the six investigated methods (under the three-dataset training condition) . . . . .	229
6.6	Critical Difference (CD) diagram comparing the <i>DSC</i> values achieved by all the investigated CNN-based architectures using the Bonferroni-Dunn's <i>post hoc</i> test . . . . .	230
6.7	Critical Difference (CD) diagram comparing the <i>DSC</i> values achieved by all the investigated CNN-based architectures using the Bonferroni-Dunn's <i>post hoc</i> test . . . . .	230
6.8	Conceptual scheme of GANs . . . . .	231

---

6.9	Potential applications of the proposed GAN-based synthetic brain MR image generation . . . . .	233
6.10	Examples of real MR images from the BRATS 2016 training dataset used for GAN training . . . . .	234
6.11	Examples of synthetic MR images yielded by DCGAN . . . . .	236
6.12	Examples of synthetic MR images yielded by WGAN . . . . .	237
6.13	PGGAN-based DA for improved tumor detection . . . . .	240
6.14	Examples of real $256 \times 256$ MR images used for PGGAN training . . . . .	242
6.15	PGGANs architecture for synthetic $256 \times 256$ MR image generation . . . . .	242
6.16	Examples of real MR image and its geometrically-transformed synthetic images . . . . .	243
6.17	Examples of synthetic MR images yielded by PGGANs . . . . .	245
6.18	t-SNE result for tumor/non-tumor images . . . . .	247



# List of tables

3.1	Characteristics of the analyzed multi-well plates . . . . .	54
3.2	Pearson’s linear correlation coefficient calculated between the automatic ACC-based SF values ( $SF_{\%Area}$ ) and the conventional colony counting measurements ( $SF_{Manual}$ ) . . . . .	61
3.3	Acquisition parameters of the MRI uterine fibroid dataset . . . . .	70
3.4	Overlap-based and distance-based metrics achieved by the IOTS and SM & RG methods on the complete uterine fibroid MRI dataset . . . . .	72
3.5	Evaluation metrics on cell counting and segmentation achieved by ACDC on the analyzed time-lapse microscopy datasets . . . . .	85
4.1	MRI acquisition parameters of the brain tumor dataset . . . . .	91
4.2	Comparison among the developed brain lesion segmentation methods on the same 15 patients treated with Gamma Knife neuro-radiosurgery . . . . .	98
4.3	Values of the spatial overlap-based metrics regarding the necrosis extraction results achieved by the developed segmentation methods . . . . .	105
4.4	Values of the spatial distance-based metrics regarding the necrosis extraction results achieved by the developed segmentation methods . . . . .	106
4.5	Comparison of the execution times between NeXt and the implemented approaches . . . . .	107
4.6	MRI acquisition parameters of the prostate dataset . . . . .	110
4.7	SNR calculated on the entire MRI prostate dataset . . . . .	116
4.8	Values of the spatial overlap-based metrics regarding the achieved prostate gland ROI segmentation results . . . . .	120
4.9	Values of the spatial distance-based metrics regarding the achieved prostate gland ROI segmentation results . . . . .	121
4.10	Values of the spatial overlap-based metrics regarding the achieved GTV segmentation results . . . . .	136

4.11	Values of the spatial distance-based metrics regarding the achieved GTV segmentation results . . . . .	137
4.12	Brain tumor volume measurements for each metastatic tumor using MR and PET imaging . . . . .	156
4.13	Achieved tumor spatial overlap-based metrics for GTV, BTV, and $BTV_{MRI}$ segmentations . . . . .	157
4.14	Achieved tumor spatial distance-based metrics for GTV, BTV, and $BTV_{MRI}$ segmentations . . . . .	158
4.15	Clinical value of BTV and $BTV_{MRI}$ , for each brain metastasis, in CTV definition . . . . .	159
5.1	MRI acquisition parameters of uterine fibroids and brain metastatic tumor dataset . . . . .	172
5.2	Image enhancement evaluation metrics achieved by MedGA and the classical image pre-processing approaches on the uterine fibroid MRI dataset . . . . .	185
5.3	Image enhancement evaluation metrics achieved by MedGA and the classical image pre-processing approaches on the brain tumor MRI dataset . . . . .	185
6.1	Prostate zonal segmentation results of the three CNN-based architectures in 4-fold cross-validation assessed by the <i>DSC</i> metrics . . . . .	220
6.2	Prostate zonal segmentation results of the CNN-based architectures and the unsupervised continuous max-flow model . . . . .	227
6.3	Visual Turing test results by a physician for classifying real vs synthetic images	238
6.4	Binary classification results for detecting brain tumors with/without DA . . . . .	245
6.5	Visual Turing test results by a physician for classifying real vs synthetic images and tumor vs non-tumor images . . . . .	246



# Chapter 1

## Introduction

### 1.1 Aims and motivations

Nowadays, the amount of heterogeneous biomedical data is increasing more and more thanks to novel sensing techniques (i.e., wearable devices and ubiquitous computing paradigms) [284] and high-throughput technologies (i.e., high-content imaging and multi-omics studies) [256, 362], generating data typically characterized by an underlying complex structure [217]. In addition, electronic health (e-health) [224] and mobile health (m-health) [13] can be properly integrated to develop automated mining engines [11], as well as support real-time personalized diagnosis [455] and continuous monitoring (also by means of wearable biophysical and electrochemical sensors) [492]. These emerging trends bring life to the so-called medical Internet of Things (mIoTs), which provides network-based platforms for the ubiquitous connectivity of patients and medical doctors in future medical applications [406]. Therefore, specific network security mechanisms are required by such sensor-based medical devices [24]. Cutting-edge Information and Communication Technology (ICT) can enable the shift from an organization-centric to a patient-centric model, leading to collaborative multi-centric healthcare service delivery processes [555].

With particular interest to the biomedical image analysis field, the advances in image acquisition modalities [461] and high-throughput imaging experiments [558] are creating new challenges. This huge and valuable information ensemble, deriving from heterogeneous large-scale datasets, could overwhelm the analytic capabilities needed by physicians in their decision-making tasks as well as by biologists investigating complex biological systems [403]. In this regard, medical imaging comprises minimally invasive techniques for acquiring images that provide detailed information about the anatomy and physiology of the imaged organs [159], while live cell microscopy imaging allows for the visualization and analysis of

the dynamic processes of living specimens [402] that underwent treatments (e.g., radiation therapy or chemotherapy [411]).

Quantitative imaging methods convey scientifically and clinically relevant information in prediction, prognosis or treatment response assessment [680], by also considering radiomics approaches [5]. In such a context, Computational Intelligence and Machine Learning can significantly improve traditional Image Processing and Pattern Recognition techniques [648], playing a key role in radiology [657] and laboratory applications [462]. However, conventional Machine Learning and Computational Intelligence techniques must be tailored to address the unique challenges pertaining to biomedical images [336, 560]. Machine Learning studies computer algorithms that can learn automatically complex relationships or patterns from empirical data (i.e., features) and are able to make accurate decisions [292]. In addition, Computational Intelligence comprises a set of nature-inspired computational techniques to address complex real-world problems that cannot be tackled by means of exact algorithmic approaches. As a matter of fact, these tasks—which could be difficult to model with constrained-based approaches relying on dynamic or multi-parametric programming [175, 176]—might be stochastic in nature and can be suitably solved by exploiting global search metaheuristics [563]. Relevant models, features, and characteristics can be indeed automatically learned from images by relying on these computational frameworks [513].

However, interpretable computational models [86], allowing for the understandability of the results, must be taken into account in life sciences. The issues related to the interpretability of Machine Learning and Computational Intelligence methods in medicine are compelling for the adoption and the feasibility of a novel Clinical Decision Support System (CDSS). As a matter of fact, over-reliance on CDSS platforms could also cause a subtle loss of self-confidence and affect the willingness of a physician to provide a definitive interpretation or diagnosis [74]. Therefore, considering the unique challenges encountered in clinical scenarios, interactive segmentation algorithms [63] often represent a more feasible and safe solution for physician in clinical practice with respect to fully automatic approaches [243].

It is worth noting that in this Ph.D. thesis, the design choice among the devised computational methods has been always motivated by the availability of biomedical data (considering also their labeling). In addition, my research has been always tightly linked to the feasibility of the proposed clinical or laboratory applications.

## 1.2 Computer science contributions to biomedical imaging

Considering these compelling issues related to biomedical big data [498], proposing novel automated computational approaches can involve very diverse computer science areas and domains:

- Information Retrieval [156] and Data Mining [410] to extract knowledge from Electronic Health Records (EHRs)—for epidemiological research studies [482]—relying on structured and unstructured textual data;
- Computational Intelligence [16] and Machine Learning [648, 657] to enhance conventional digital Image Processing techniques [37] on pictorial data.

The conceptual “neural-based scheme” of the main topics treated in this thesis is shown in Fig. 1.1, aiming at representing the interconnections among the devised computational methods and biomedical image analysis and computing. Therefore, the principal contribution of computer science research in biomedicine concerns the proper fusion of diverse datasets [344, 556] (i.e., imaging modalities, high-throughput technologies, EHRs) to provide a complete clinical knowledge for precision medicine [66].

## 1.3 Scientific production and thesis overview

This Ph.D. thesis aims at proposing novel computer-assisted methods for biomedical image analysis, also as an instrument in the development of CDSSs, by always keeping in mind the clinical feasibility of the developed solutions. Therefore, my Ph.D. program has been directed towards a two-fold objective: (i) computer-assisted segmentation methods in cancer imaging, and (ii) automatic cell image analysis.

This thesis is structured as follows. The basic concepts related to medical imaging and bioimage informatics, as well as the investigated clinical and laboratory scenarios are recalled in **Chapter 2**. The devised classical Image Processing algorithms, especially region-based and morphological approaches in biomedical image segmentation, are described in **Chapter 3**. **Chapter 4** introduces Pattern Recognition techniques, with particular interest to unsupervised fuzzy clustering and graph-based models (i.e., Cellular Automata and Random Walker) to multispectral and multimodal medical imaging data processing. Computational Intelligence methods are presented in **Chapter 5**, wherein a novel evolutionary framework based on Genetic Algorithms for medical image enhancement and segmentation is presented. Moreover, multimodal image co-registration using Particle Swarm Optimization is discussed. Finally, Deep Neural Networks are investigated in **Chapter 6**: (i) the generalization abilities

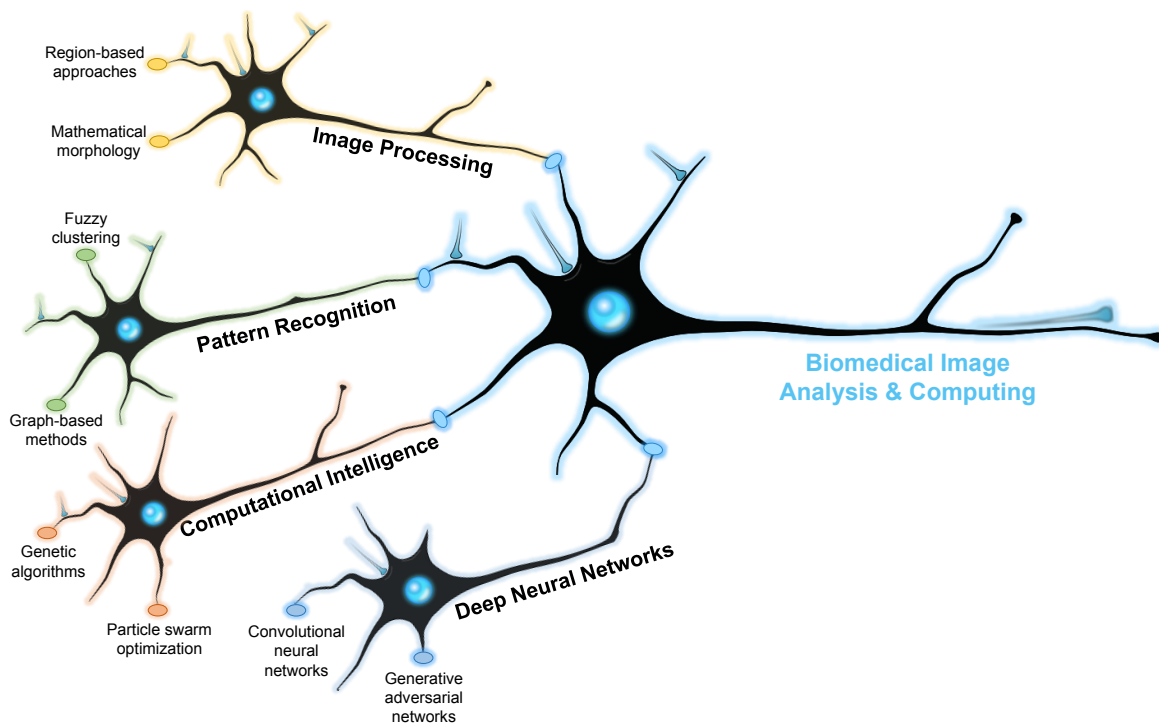


Fig. 1.1 Conceptual scheme of the main topics treated in this thesis. This “neural-based scheme” aims at showing the interconnections (represented by synapses) among the different computational methods and biomedical image analysis and computing. The terminal synapses report the approaches specifically encountered during my research activities, reflecting the structure of the thesis. The colored halos of the four left-most neurons denote the four computer science areas related to biomedical imaging: Image Processing (Chapter 3), Pattern Recognition (Chapter 4), Computational Intelligence (Chapter 5), Deep Neural Networks (Chapter 6). Under a biological inspiration, the axons of these four neurons send out their output signal to the dendrites of the central light-blue neuron, by activating its own synapses. The effective combination emerging from the different computational approaches involves the excitation of the neuron representing biomedical image analysis and computing, which suddenly generates an action potential originating at the soma and propagates along the axon.

of Convolutional Neural Networks in medical image segmentation of multi-institutional datasets are addressed by conceiving an architecture that integrates adaptive feature recalibration, and (ii) the generation of realistic medical images based on Generative Adversarial Networks is applied to data augmentation purposes. **Chapter 7** concludes with final remarks and future developments. **Appendix A** contains the evaluation metrics and procedures used for the validation of the proposed computational approaches.

**Journal articles**

- **Rundo L.\***, Tangherloni A.\*, Nobile M.S., Militello C., Mauri G., Besozzi D., Cazaniga P. (2019) MedGA: A novel evolutionary method for image enhancement in medical imaging systems. *Expert Systems with Applications*, 119, 387–399.  
DOI: 10.1016/j.eswa.2018.11.013.
- Gambino O.\*, **Rundo L.\***, Cannella V., Vitabile S., Pirrone R. (2018) A framework for data-driven adaptive GUI generation based on DICOM. *Journal of Biomedical Informatics*, 88, 37–52.  
DOI: 10.1016/j.jbi.2018.10.009
- **Rundo L.**, Militello C., Russo G., Vitabile S., Gilardi M.C., Mauri G. (2018) GTV-CUT for neuro-radiosurgery treatment planning: an MRI brain cancer seeded image segmentation method based on a Cellular Automata model. *Natural Computing*, 17(3), 521–536. Special Issue on “Cellular Automata for Research and Industry (ACRI 2016)”.  
DOI: 10.1007/s11047-017-9636-z
- **Rundo L.**, Militello C., Tangherloni A., Russo G., Vitabile S., Gilardi M.C., Mauri G. (2018) NeXt for neuro-radiosurgery: a fully automatic approach for necrosis extraction in brain tumor MRI using an unsupervised Machine Learning technique. *International Journal of Imaging Systems and Technology*, 28(1), 21–37.  
DOI: 10.1002/ima.22253
- Militello C., **Rundo L.**, Conti V., Minafra L., Cammarata F.P., Vitabile S., Mauri G., Gilardi M.C., Porcino N. (2017) Area-based cell colony surviving fraction evaluation: a novel fully automatic approach using general-purpose acquisition hardware. *Computers in Biology and Medicine*, 89, 454–465.  
DOI: 10.1016/j.combiomed.2017.08.005
- **Rundo L.**, Militello C., Russo G., Garufi A., Vitabile S., Mauri G., Gilardi M.C. (2017) Automated Prostate gland segmentation based on an unsupervised fuzzy C-means clustering technique using multispectral T1w and T2w MR imaging. *Information*, 8(2), 49, pp. 1–28, Special Issue on “Fuzzy Logic for Image Processing”.  
DOI: 10.3390/info8020049
- **Rundo L.**, Stefano A., Militello C., Russo, G. Sabini M.G., D’Arrigo C., Marletta F., Ippolito M., Mauri G., Vitabile S., Gilardi M.C. (2017) A fully automatic approach for

---

\*Equal contribution

multimodal PET and MR image segmentation in Gamma Knife treatment planning. *Computer Methods and Programs in Biomedicine*, 144, 77–96.

DOI: 10.1016/j.cmpb.2017.03.011

### Conference Proceedings and Book Chapters

- **Rundo L.**, Han C., Zhang J., Hataya R., Nagano Y., Militello C., Ferretti C., Nobile M.S., Tangherloni A., Gilardi M.C., Vitabile S., Nakayama H., Mauri G. (2018) CNN-based prostate zonal segmentation on T2-weighted MR images: a cross-dataset study. Accepted by the: *28th Italian Workshop on Neural Networks (WIRN 2018)*, Vietri sul Mare, Salerno, Italy, June 13–15, 2018.
- Han C., **Rundo L.**, Araki R., Furukawa Y., Mauri G., Nakayama H., Hayashi H. (2018) Infinite brain MR images: PGGAN-based data augmentation for tumor detection. Accepted by the: *28th Italian Workshop on Neural Networks (WIRN 2018)*, Vietri sul Mare, Salerno, Italy, June 13–15, 2018.
- Han C., Hayashi H., **Rundo L.**, Araki R., Shimoda W., Muramatsu S., Yujiro F., Mauri G., Nakayama H. (2018) GAN-Based synthetic brain MR image generation. Proceedings of the *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 734–738, Washington, DC, USA, April 4–7, 2018.  
DOI: 10.1109/ISBI.2018.8363678
- **Rundo L.**, Militello C., Tangherloni A., Russo G., Lagalla R., Mauri G., Gilardi M.C., Vitabile S. (2019) Computer-assisted approaches for uterine fibroid segmentation in MRgFUS treatments: quantitative evaluation and clinical feasibility analysis. In: Esposito, A., Faundez-Zanuy, M., Morabito, F.C., Pasero, E. (Editors), *Quantifying and Processing Biomedical and Behavioral Signals, Smart Innovation, Systems and Technologies*, Vol. 103, Chap. 22, pp. 229–241, Springer International Publishing, 2018. Proceedings of the *27th Italian Workshop on Neural Networks (WIRN 2017)*, Vietri sul Mare, Salerno, Italy, June 14–16, 2017.  
DOI: 10.1007/978-3-319-95095-2\_22
- **Rundo L.**, Tangherloni A., Militello C., Gilardi M.C., Mauri G. (2017) Multimodal medical image registration using particle swarm optimization: a review. Proceedings of the *2016 IEEE Symposium Series on Computational Intelligence (IEEE SSCI 2016) – Swarm Intelligence Symposium (SIS)*, pp. 1–8, Athens, Greece, December 6–9, 2016.  
DOI: 10.1109/SSCI.2016.7850261

- **Rundo L.**, Militello C., Russo G., D'Urso D., Valastro L.M., Garufi A., Mauri G., Vitabile S., Gilardi M.C. (2017) Fully automatic multispectral MR image segmentation of prostate gland based on the fuzzy C-means clustering algorithm. In: Esposito, A., Faundez-Zanuy, M., Morabito, F.C. (Editors), *Multidisciplinary Approaches to Neural Computing, Smart Innovation, Systems and Technologies*, Vol. 69, Chap. 3, pp. 23–37, Springer International Publishing, 2017. Proceedings of the 26th Italian Workshop on Neural Networks (WIRN 2016), Vietri sul Mare, Salerno, Italy, May 18-20, 2016.  
DOI: 10.1007/978-3-319-56904-8\_3

### Abstracts

- **Rundo L.**, Militello C., Vitabile S., Russo G., Pisciotta P., Sabini M.G., Marletta F., Ippolito M., D'Arrigo C., Midiri M., Gilardi M.C. (2016) A.197 – Clinical support in radiation therapy scenarios: MR brain tumor segmentation using an unsupervised fuzzy C-means clustering technique. In: 9th National Congress of AIFM 2016. *Physica Medica, European Journal of Medical Physics*, Vol. 32, Suppl. 1, Feb. 2016, p. 58.  
DOI: 10.1016/j.ejmp.2016.01.201
- Pisciotta P., Militello C., **Rundo L.**, Stefano A., Russo G., Vitabile S., Sabini M.G., D'Arrigo C., Marletta F., D'Urso D., Ippolito M., Midiri M., Gilardi M.C. (2016) A.179 – Using anatomic and metabolic imaging in stereotactic radio neuro-surgery treatments. In: 9th National Congress of AIFM 2016. *Physica Medica, European Journal of Medical Physics*, Vol. 32, Suppl. 1, Feb. 2016, p. 53.  
DOI: 10.1016/j.ejmp.2016.01.183

### Submitted papers and manuscripts in preparation

- **Rundo L.\***, Tangherloni A.\*, Nobile M.S., Cazzaniga P., Besozzi D., Russo G., Vitabile S., Gilardi M.C., Mauri G., Militello C. (2018) Genetic Algorithms improve thresholding-based segmentation of bimodal Magnetic Resonance images. Submitted to: *Computer Methods and Programs in Biomedicine* – Elsevier. ISSN: 0169-2607.
- **Rundo L.\***, Han C.\* , Nagano Y., Zhang J., Hataya R., Militello C., Tangherloni A., Nobile M.S., Ferretti C., Besozzi D., Gilardi M.C., Vitabile S., Mauri G., Nakayama H., Cazzaniga P. (2018) USE-Net: incorporating Squeeze-and-Excitation blocks into U-Net for prostate zonal segmentation of multi-centric MRI datasets. Submitted to: *Neurocomputing* – Elsevier. ISSN: 0925-2312.

- **Rundo L.\***, Tangherloni A.\*, Tyson D.R.\*, Betta R., Nobile M.S., Spolaor S., Militello C., Lubbock A.L.R., Mauri G., Quaranta V., Besozzi D., Lopez C.F., Cazzaniga P. (2019) ACDC: Automated Cell Detection and Counting for time-lapse fluorescence microscopy. *Medical & Biological Engineering & Computing* – Springer [Manuscript in preparation]
- Militello C., **Rundo L.**, Conti V., Toia P., La Grutta L., Midiri M., Vitabile S. (2019) A semi-automatic approach for epicardial adipose tissue segmentation and quantification on non-enhanced cardiac CT scans. *Symmetry* – MDPI [Manuscript in preparation].

## 1.4 Other publications

Along with my main research on biomedical image analysis, to achieve a broader perspective on personalized medicine, I have been involved in activities regarding Systems Biology and Genome Analysis. In particular, some challenging computational issues were addressed exploiting High-Performance Computing (HPC). These efforts resulted in some publications.

### Journal articles

- Tangherloni A., Spolaor S., **Rundo L.**, Nobile M.S., Cazzaniga P., Mauri G., Liò P., Merelli I., Besozzi D. (2018) GenHap: a novel computational method based on Genetic Algorithms for haplotype assembly. Accepted by: *BMC Bioinformatics* – BioMed Central; Special Issue of the 17th Workshop on Network Tools and Applications for Biology (NETTAB 2017). ISSN: 1471-2105. [In press]  
arXiv preprint [arXiv:1812.07689](https://arxiv.org/abs/1812.07689).
- Tangherloni A., Spolaor S., Cazzaniga P., Besozzi D., **Rundo L.**, Mauri G., Nobile M.S. (2018) Benchmark functions do not capture the complexity of biochemical parameter estimation. Submitted to: *Applied Soft Computing* (Virtual Special Issue on “Benchmarking of Computational Intelligence Algorithms”) – Elsevier. ISSN: 1568-4946.
- Tangherloni A., Nobile M.S., Cazzaniga P., Capitoli G., Spolaor S., **Rundo L.**, Mauri G., Besozzi D. (2018) FiCoS: a fine- and coarse-grained GPU-powered deterministic simulator for biochemical networks. *Scientific Reports* – Nature Publishing Group. [Manuscript in preparation]



**Conference proceedings**

- Spolaor S., Tangherloni A., **Rundo L.**, Nobile M.S., Cazzaniga P. (2019) Estimation of kinetic reaction constants: exploiting reboot strategies to improve PSO's performance. Proceedings of the *14th International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics (CIBB 2017)*, Cagliari, Italy, September 7–9, 2017.
- Tangherloni A., **Rundo L.**, Spolaor S., Nobile M.S., Merelli I., Besozzi D., Mauri G., Cazzaniga P., Liò P. (2018) High performance computing for haplotyping: models and platforms. In: Mencagli, G., *et al.* (Editors), Euro-Par 2018 Workshops, Lecture Notes in Computer Science, Vol. 11339, pp. 650–661, Springer International Publishing, 2019. *24th International European Conference on Parallel and Distributed Computing (Euro-Par 2018)*, Workshop on Advances in High-Performance Bioinformatics, Systems Biology (Med-HPC 2018), Turin, Italy, August 27–28, 2018.  
DOI: 10.1007/978-3-030-10549-5\_51
- Nobile M.S., Tangherloni A., **Rundo L.**, Spolaor S., Besozzi D., Mauri G., Cazzaniga P. (2018) Computational Intelligence for parameter estimation of biochemical systems. Proceedings of the *IEEE World Congress on Computational Intelligence (IEEE WCCI 2018) – 2018 IEEE Congress on Evolutionary Computation (IEEE CEC 2018)*, pp. 1–8, Rio de Janeiro, Brazil, July 8–13, 2018.  
DOI: 10.1109/CEC.2018.8477873
- Tangherloni A., **Rundo L.**, Spolaor S., Cazzaniga P., Nobile M.S. (2018) GPU-powered multi-swarm parameter estimation of biological systems: a master-slave approach. Proceedings of the *2018 Parallel, Distributed, and Network-Based Processing (PDP 2018)*, Special Session on: Parallel and distributed high-performance computing solutions in Systems Biology, pp. 698–705, Cambridge, UK, March 21–23, 2018.  
DOI: 10.1109/PDP2018.2018.00115
- Spolaor S., Tangherloni A., **Rundo L.**, Nobile M.S., Cazzaniga P. (2017) Reboot strategies in Particle Swarm Optimization and their Impact on parameter estimation of biochemical systems. Proceedings of the *IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology (IEEE CIBCB 2017) – Special Session on Parallel and Distributed High Performance Computing Solutions for Computational Intelligence Methods*, pp. 1–8, Manchester, United Kingdom, August 23–25, 2017.  
DOI: 10.1109/CIBCB.2017.8058550

- Tangherloni A., **Rundo L.**, Nobile M.S. (2017) Proactive particles in swarm optimization: a settings-free algorithm for real-parameter single objective optimization problems. Proceedings of the *2017 IEEE Congress on Evolutionary Computation (IEEE CEC 2017)* – Special Session & Competitions on Real-Parameter Single Objective Optimization, pp. 1940–1947, Donostia – San Sebastián, Spain, June 5–8, 2017. DOI: 10.1109/CEC.2017.7969538

In addition, the following works concern applications of Field Programmable Gate Arrays (FPGAs) in biometry and digital signal processing:

- Conti V., **Rundo L.**, Billeci G.D., Militello C., Vitabile S. (2018) Energy efficiency evaluation of dynamic partial reconfiguration in Field Programmable Gate Arrays: an experimental case study. *Energies*, 11(4), 739.  
DOI: 10.3390/en11040739
- Conti V., **Rundo L.**, Militello C., Mauri G., Vitabile S. (2017) Resource-efficient hardware implementation of a neural-based node for automatic fingerprint classification. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 8(4), 19–36.  
DOI: 10.22670/JOWUA.2017.12.31.019

# Chapter 2

## Biomedical image analysis and computing

### 2.1 Background

This section first introduces the basic concepts in medical imaging and bioimage informatics. Afterwards, the practical scenarios under investigation are described.

#### 2.1.1 Medical imaging

Nowadays, medical imaging plays a crucial role in the clinical workflow, thanks to its capability of representing anatomical and physiological features that are otherwise inaccessible for inspection, thus producing accurate imaging biomarkers and clinically useful information [345, 513]. Medical images are considerably different from the pictures usually analyzed in Pattern Recognition and Computer Vision, regarding the appearance of the depicted objects as well as the information conveyed by the pixels [214]. Indeed, medical imaging techniques exploit several different principles to measure spatial distributions of physical attributes of the human body, allowing us to better understand complex or rare diseases [612]. The effectiveness of such techniques can be reduced by a plethora of phenomena, such as noise and Partial Volume Effect (PVE) [578, 612], which might affect the measurement processes involved in imaging modalities and data acquisition devices. In addition, computer-aided medical image acquisition procedures generally include reconstruction methods (producing two, three, or even four dimensional imaging data), which could cause the appearance of artifacts. Image contrast and details might also be impaired by the procedures used in medical imaging, as well as by the physiological nature of the body part under investigation.

The most common imaging modalities as well as diagnostic workstation software issues are briefly outlined in what follows.

### **X-ray radiography and Computed Tomography**

X-ray (plain) radiography is a projectional 2D image acquisition technique that utilizes ionizing radiation. The X-rays are directed towards the body and then detected, in order to yield an image representing the amount of rays that reached the detector without being absorbed or scattered by the patient.

In the early 1970s, X-ray Computed Tomography (CT) gave rise to a significant impact in technological development in medicine as well as in industry [279]. CT is a diagnostic imaging modality that combines a finite number of conventional X-ray projections with computing algorithms. CT scans are acquired by rotating an X-ray source around the patient. The detectors are located on the opposite side with respect to the X-ray tube. In more recent CT scanners, the patient's bed is continuously moved to obtain a spiral acquisition mode [303]. X-ray beam intensity exponentially decreases according to the mass attenuation coefficient: the beam is more attenuated by tissues with a high atomic number. On the contrary, whether the beam crosses a low-density tissue, the resulting attenuation is lower. By so doing, high-density tissues appear as hyper-intense regions (i.e., higher attenuation) and, conversely, low-density tissue are hypo-intense (lower attenuation). After the tomographic reconstruction (usually by means of the filtered back projection algorithm [211]), the CT images measure the attenuation of each voxel according to the Hounsfield scale, where the air has a radio-density of  $-1000$  Hounsfield Units (HU) and distilled water has a radio-density of  $0$  HU. CT provides high resolution anatomical images of the body, with an excellent contrast between air, adipose tissue, soft-tissue, and bone. However, the soft-tissue contrast is poor compared to other imaging modalities, such as Magnetic Resonance Imaging (MRI). Contrast agents may additionally be used to highlight some regions, such as the gastrointestinal tract or the blood vessels.

### **Magnetic Resonance Imaging**

MRI relies on the different relaxation times of tissues after to a high and stimulated by means of a radiofrequency impulse [70, 398]. MRI basically measures a proton density map of the tissues [172]. This technique is a multiparametric modality that can acquire both anatomical and functional imaging data [69]. According to the different relaxation times, different sequences can be obtained:

- T1-weighted (T1w) images: the spin-lattice time (i.e., longitudinal) relaxation time is a measurement of the time needed by the protons to return to the initial equilibrium conditions, through the transfer of energy to the surrounding microenvironment (lattice), in order to obtain a T1w SE sequence, using a short relaxation time (TR) associated with a short echo time (TE). In T1w images, the cerebrospinal fluid is dark while the fat is bright;
- T2-weighted (T2w) images: the spin-spin (i.e., transverse) relaxation time is a measurement of the time taken by the spin of protons to get out of synchronization. This progressive desynchronization strongly reduces the transverse magnetization. A T2w sequence generally has a long TR associated with a long TE. Liquids or very hydrated tissues appear hyper-intense on T2w images.

The relaxation time of a tissue depends on its water content. MRI provides excellent soft-tissue contrast. For this reason, MRI is widely applied in the diagnosis and treatment of neurological, cardiovascular, musculoskeletal, liver, and gastrointestinal diseases. Moreover, image contrast can be further enhanced through the injection of a CE agent.

Advanced MRI techniques can be used to convey information about physiological tissue characteristics [398]:

- Dynamic Contrast Enhanced (DCE) imaging allows for inferring the perfusion and vascularity in the tumor micro-environment by means of the shortening of the T1 time induced by a Gadolinium-based contrast bolus [613];
- Diffusion Weighted Imaging (DWI) estimates the random Brownian motion of water molecules within tissue voxels [238]. The resistance of water molecule diffusion is quantitatively measured using the Apparent Diffusion Coefficient (ADC) value [348], calculating clinically useful ADC maps pertaining to the imaged body part;
- Magnetic Resonance Spectroscopic Imaging (MRSI) quantifies further information about metabolic activity, by providing spectroscopic information regarding metabolites in addition to anatomical MR images [480];
- functional MRI (fMRI) is valuable in neuroimaging [369] to measure the brain activity by detecting the changes associated with blood flow and oxygenation patterns, thus acquiring functional activation maps [188].

### **Medical ultrasonography**

Medical ultrasonography is well-established in clinical practice since it is relatively portable, cost-effective, safe (in terms of radiation risk), and allows for real-time capabilities in image-

guided interventional procedures [275]. In particular, this modality is useful for soft-tissue imaging. Ultrasounds (US), representing mechanical pressure waves with frequencies higher than 20 kHz, can be both emitted and received by a hand-held probe (i.e., a piezoelectric transducer). These transducers are composed of multiple piezoelectric crystals that can convert electrical signals to mechanical waves as well as mechanical pressure to electrical signals. In addition, US scans are tomographic, i.e., offering cross-sections in the imaging of anatomical structures.

US waves propagating through tissues are partially transmitted to deeper body parts, reflected back to the transducer as echoes, scattered, and also converted to heat. For imaging purposes, the echoes reflected back to the transducer are the most valuable to investigate the tissue properties [93]. The measurement of the echo resulting from the interaction with a tissue interface is performed by directly employing the acoustic impedance, which is an intrinsic physical property [71].

The echo principle defines the most commonly used diagnostic US-based techniques:

- Amplitude modulation (A-mode) refers to a one-dimensional signal acquisition technique based on the echoes received are displayed as vertical deflections;
- Brightness modulation (B-mode) is a technique wherein the echo amplitude is depicted on a gray-scale image. It is mostly used as a 2D B-scan by combining multiple ultrasound beams. This technique allows for real-time acquisitions thanks to its high frame rate;
- Motion modulation (M-mode) relies on a stationary ultrasound field. The images are continuously recorded over time. M-mode has its main application in echocardiography imaging;
- Doppler techniques exploit the Doppler effect as a further source of information. As a matter of fact, the US waves reflected by a moving object reveal a difference, in terms of frequency with respect to the emitted frequencies, which is proportional to the speed of the moving reflector. Color-Doppler techniques integrate Doppler information on the standard B-mode scans by means of pseudo-colored images [420].

### **Elastography and Tactile Imaging**

Mechanical Imaging (MI) is a relatively recent medical diagnostics modality based on the reconstruction of tissue structures and viscoelastic properties exploiting biomechanical sensors. This inverse problem relies on the acquisition of data regarding stress patterns on the surface of tissue compressed by a pressure sensor array. Thus, the imaged tissue structures

are presented in terms of their viscoelastic properties, so allowing for the characterization of the tissue, differentiating normal and diseased conditions and detecting tumors and other lesions [544]. Unlike the existing medical imaging methods that use sophisticated hardware (e.g., superconductive magnets, expensive CT equipment, and complex ultrasonic phased arrays), MI hardware consists of inexpensive mechanical sensors and a positioning system connected to a general-purpose computer. Therefore, the structures are visualized by sensing the pattern of mechanical stresses on the surface of an organ.

Elastography maps the elastic properties and stiffness of tissues, specifically enabling the visualization and assessment of mechanical properties of soft-tissues [545]. Mechanical properties of tissues—i.e., elastic modulus and viscosity—are highly sensitive to tissue structural changes related to several physiological and pathological processes.

Tactile Imaging (TI) yields a tissue elasticity map, by computationally emulating the manual palpation process by means of a pressure sensor array mounted onto a probe that compresses soft-tissues and detects the changes in the pressure patterns [165].

### **Nuclear Medicine**

Nuclear Medicine (NM) includes functional imaging techniques that require the delivery of a gamma-emitting radioisotope (i.e., a radionuclide) into the patient, generally by means of an injection [408].

Single-Photon Emission Computed Tomography (SPECT) is an NM tomographic imaging technique using gamma rays [458]. Extending the conventional NM planar imaging based on a gamma camera (i.e., scintigraphy), SPECT can acquire 3D information through cross-sectional slices.

Positron Emission Tomography (PET) is a non-invasive nuclear medical imaging technique for studying functional characteristics of the human body [656]. PET scans reveal functional processes that show complementary information with respect to anatomical imaging [171], by providing an *in vivo* measurement of the tumor biological processes [199].

Common clinical applications include brain imaging, neurology, oncology, and recently also in cardiology. The patient is injected with a dose of radioactive tracer isotope that selectively concentrates in tissues of interest in the body. Typically, the most active cells in the tissue reveal a higher metabolism, thus uptaking more tracer isotope than cells which are less active. The isotope undergoes radioactive decay involving the emission of a positron. As soon as the emitted positron interacts with an electron, this annihilation produces a pair of 511 keV gamma photons at almost 180 degrees to each other and is detected by the PET scanner. Gathering all these pairs allow for the PET measurements from which the distribution of the relevant radiotracer is reconstructed by means of direct or iterative

approaches [72]. Therefore, PET is a type of emissive imaging differently to CT, which represents a transmissive technique.

Among the different PET radiotracers labeled with radioactive or stable isotopes,  $^{18}\text{F}$ -fludeoxyglucose (FDG) is a glucose analog widely used in the evaluation of several neoplastic pathologies as well as in radiotherapy planning. The FDG uptake is increased in tissues with a high metabolic rate, such as tumor or inflammation regions. These areas appear as hyper-intense regions on PET images. Accordingly, PET studies are frequently used in oncological imaging, offering an adequate staging and patient follow-up tool to several cancer types. In addition, metabolic changes are often faster and more indicative of the effects of the therapy with respect to morphological changes [644]. Unfortunately, due to the poor resolution in PET images, an accurate quantification is strongly affected by the PVE [578]. Several literature methods dealt with the PVE correction [198, 583]. Modern scanners can estimate attenuation maps using integrated X-ray CT equipment [693], allowing for attenuation correction relying on CT scans.

Hybrid PET/MRI scanners can acquire both PET and MR images simultaneously, so conveying new insights in molecular imaging [301, 468]. These MRI-compatible PET systems were realized thanks to technological developments in the detectors of PET scanners, i.e., replacing scintillators coupled to photomultiplier tubes with magnetic field-insensitive avalanche photodiodes [87].

### **Medical photography and video acquisition**

High-resolution digital cameras have been revolutionizing medical applications concerning photography and video acquisition.

Considering m-health applications, modern smartphone cameras can be considered as medical devices for disease diagnosis or medical condition monitoring [8], with potential applications in telemedicine for chronic diseases [628]. In this context, smartphones equipped with electrochemical sensors can be suitably used for real-time health monitoring [232]

In the context of video acquisition, endoscopy has become fundamental in many surgical settings. The endoscopy procedure uses an endoscope to examine the interior of an organ or cavity inside the body. Unlike other medical imaging techniques, endoscopes are introduced directly into the patient's body. During surgery, the patient's anatomy can be displayed as a 2D video (captured by a camera mounted onto the endoscopy probe) and specialized surgical instruments are introduced into the body accordingly. More recently, Video Capsule Endoscopy (VCE) has revolutionized the diagnostic workflow in the field of small bowel diseases [285]. Although endoscopy is widely used to guide surgical treatments, it allows the surgeon to investigate only the anatomical surface of the surgical site with a limited



FOV. The 2D conventional endoscopy does not provide any depth information of the surgical scene. Therefore, Augmented Reality (AR) techniques have been employed to address these limitations [649]. Moreover, CT scans can be used to non-invasively reconstruct the corresponding virtual colonoscopy, such as in the case of polyp detection [593].

### **Diagnostic medical imaging software**

Computer applications for diagnostic medical imaging generally provide a wide range of tools to support physicians in their daily diagnosis activities. Unfortunately, some functionalities are specialized for specific diseases or imaging modalities, while other ones are useless for the images under investigation [77]. Nevertheless, the corresponding Graphical User Interface (GUI) widgets are still present on the screen reducing the image visualization area. As a consequence, the physician may be affected by cognitive overload and visual stress, causing a degradation of performances mainly due to unuseful widgets. In clinical environments, a GUI must represent a sequence of steps for image investigation following a well-defined workflow.

These issues were addressed in a software framework proposed in [200]. Specifically, we designed a Digital Imaging and COmmunications in Medicine (DICOM) standard [58, 432] based mechanism of data-driven GUI generation, referring to the examined body part and imaging modality as well as to the medical image analysis task to perform. In this way, the self-configuring GUI is generated on-the-fly, so that just specific functionalities are active according to the current clinical scenario. Such a solution provides also a tight integration with the DICOM standard, which considers various aspects of the technology in medicine but does not address GUI specification issues.

The proposed workflow is designed for diagnostic workstations with a local file system on an interchange media inside or outside the hospital ward, not for Picture Archiving and Communication System (PACS) platforms—connected by means of the Radiology Information System (RIS)—which have functionalities mainly oriented to fundamental imaging tools and network-based services. The framework is designed to manage high-level multiple functionalities, such as image segmentation, image fusion, and advanced measurements. As a matter of fact, a diagnostic workstation is not only installed in a hospital ward, but it may also belong to an external facility for a second opinion given by highly specialized physicians [695] or multi-disciplinary boards [204] who have to examine the data. This is a typical scenario of a second opinion, where an external physician who does not belong to the hospital institution is consulted to perform an alternative reporting/diagnosis.

We implemented a proof-of-concept as a suitable plug-in of the OsiriX imaging software (Pixmeo SARL, Bernex, Geneva, Switzerland) [472, 510], by exploiting the DICOM standard

as an enabling technology for an auto-consistent solution in medical diagnostic applications. Accordingly, the DICOMDIR conceptual data model, defined by the hierarchical structure of the Basic Directory Information Object Definition (IOD) Information Model [433], is exploited and extended to include the GUI information thanks to a new Information Object Module (IOM), which reuses the DICOM information model.

### **2.1.2 Bioimage informatics**

Modern biological research increasingly relies on bioimages as a primary source of information in unraveling the complex mechanisms and functioning of live cells [461]. The quantity and complexity of the data generated by the most recent microscopes preclude visual or manual analyses and require advanced computational methods to fully explore this wealth of information [403]. As a matter of fact, such a kind of peculiar challenges have given rise to a separate research area that focuses on the development of novel image processing, data mining, and visualization techniques to extract and integrate the biological knowledge in these data-intensive problems. Nowadays, this emerging scientific field involving bioinformatics and bioimage computing can be called “bioimage informatics” [462].

## **2.2 Investigated contexts and applications**

This section provides the background about the clinical and biological scenarios considered in this thesis.

### **2.2.1 Medical image analysis**

#### **MRgFUS treatments for uterine fibroids**

Magnetic Resonance guided Focused Ultrasound Surgery (MRgFUS) represents a non-invasive surgical technique that exploits thermal ablation principles to treat several oncological and neurological pathologies, carefully preserving neighboring healthy tissues (Organs at Risk, OARs) [299]. High Intensity Focused Ultrasound (HIFU) waves are used to rapidly increase the temperature (reaching temperatures above 50°C) inside the target solid tumors [288] leading the neoplastic tissue to either apoptosis or coagulative necrosis [665]. Theoretically, the HIFU technology can be applied in both small and large irregularly-shaped tumors in any anatomical district where the path to the focus is free of bones and air interfaces [119]. Image guidance is needed to treat these deep-lying tumors, because accurate HIFU beam positioning and reliable acoustic power delivery are mandatory. Especially, MRI allows for

excellent soft-tissue contrast as well as real-time visualization of the heat distribution for thermal dose calculation on target areas, thanks to the Proton Resonance Frequency (PRF) shift thermometry [7].

Uterine leiomyomas, more commonly called fibroids or myomas, are benign clonal tumors growing from the smooth-muscle tissue of the uterus. They are clinically apparent in about 25% of women during middle and late reproductive years [73, 532], and with recent *in vivo* acquisition modalities for soft-tissue imaging, especially US and MRI, the actual clinical prevalence may be higher, so representing a relevant public health problem [132]. Most fibroids are asymptomatic, but many women have significant symptoms that could negatively impact on their life quality, so requiring effective therapies. Generally, these symptoms include abnormal uterine bleeding, pelvic pressure or pain, and reproductive dysfunction [587]. Depending on the severity of the symptoms, different treatment options for uterine fibroids are available: traditional surgery (i.e., abdominal/laparoscopic hysterectomy or myomectomy), mini-invasive treatments (i.e., hysteroscopic myomectomy, uterine artery embolization), and non-invasive approaches (i.e., MRgFUS, pharmacological therapy) [379].

In MRgFUS uterine fibroid therapy [96, 501], MR image analysis is required in all clinical phases: (i) imaging-aided diagnosis; (ii) treatment planning and real-time temperature monitoring, by detecting the fibroids to be treated—i.e., Region of Treatment (ROT)—as well as the OARs near the uterus—i.e., Region of Interest (ROI)—to be preserved from the HIFU beam; (iii) patient's follow-up for measuring the Non-Perfused Volume (NPV), which is the fibroid area actually ablated with HIFU therapy [412].

Nowadays, the treatment outcome is evaluated on complete infarction of the ablated fibroids and normal perfusion of the surrounding myometrium [390]. This estimation is accomplished on post-treatment T1w contrast-enhanced (CE) MRI, wherein the ablated ROT does not uptake the administered Gadolinium-based contrast medium and are imaged as unenhanced hypo-intense with respect to the uterus. In clinical practice, the NPV is measured by means of manual contouring by an experienced radiologist. The current operator-dependent methodology could affect the subsequent follow-up phases, due to the lack of results repeatability. In addition, this fully manual procedure is definitely time-consuming, considerably increasing treatment times. These critical issues can be addressed only by accurate and efficient computer-assisted MR image segmentation approaches based on Pattern Recognition techniques.

**Related work** Guyon *et al.* [234] developed a tool called Volume Estimation and Tracking over Time (VETOT) to track the volume of fibroids in patient's follow-up. VETOT provides two different types of Level Set Functions (LSFs) for fibroid segmentation: (i) fast marching

level sets, and (ii) geodesic active contours. In [47], a two-step semi-automatic method was described: an initial automatic segmentation, based on LSFs, is followed by an interactive manual refinement (with user feedback). This method was applied to uterine fibroid MR scans acquired prior to MRgFUS treatment execution. The authors of [681] proposed a semi-automatic approach based on LSFs, by combining the fast marching level set and the Laplacian level set methods. However, this approach was not tailored to HIFU treatments. In [318], a two-step method for segmentation of uterine fibroids was developed: (i) a coarse semi-automatic segmentation is performed using the Chan-Vese level set method; (ii) segmentation refinements based on the prior-shape model are applied, by exploiting training data regarding an ellipse model based on Principal Component Analysis (PCA). Fallahi *et al.* [178] proposed an automated approach for fibroid volume evaluation on MR images. By exploiting the method described in [177] the uterus is firstly segmented using the Fuzzy C-Means (FCM) algorithm from the T1w CE-MRI series. Some refinement operations are applied, and redundant parts are removed by masking the co-registered T1w MRI series. The fibroids are then segmented by applying first the Modified Possibilistic Fuzzy C-Means (MPFCM) algorithm on registered T2w MR images and finally by performing some post-processing operations. Image registration steps are mandatory and are performed using the external software tool Medical Image Processing, Analysis, and Visualization (MIPAV) [397]. In [22], an automatic method for treated fibroid segmentation is proposed. Both the post-operative CE-MR image stack (interpolated to a 3D volume) and the treatment plan (i.e., file that stores the sonication cells used during the MRgFUS treatment) are required. In fact, the coordinates and the size of each treatment cell are exploited as *a priori* information for the fibroid localization. In the first step, a surface model is created on the sonication cell data and it is deformed with an Active Shape Model (ASM) [127] by moving the mesh nodes towards the strongest image edge gradients. Finally, the voxels are classified in the treated NPV and the untreated (Perfused Volume, PV) tissues, by an Expectation-Maximization (EM) [423] segmentation algorithm.

In recent years, our research group has already proposed computer-assisted segmentation approaches for uterine fibroids in MRgFUS treatments [415, 416, 524]. The clinical feasibility of these methodologies was evaluated and critically discussed in [522].

### **Neuro-radiosurgery for brain tumors**

Stereotactic radiosurgery allows for an accurate external irradiation (with a single, high dose and steep dose gradient) to minimize the doses delivered to adjacent critical brain structures. Stereotactic neuro-radiosurgery is a well-established mini-invasive surgical therapy for the treatment of intracranial diseases [377], especially brain metastases and highly invasive

cancers [437, 574]. Early clinical trials showed that tumor control rates are superior to Whole Brain Radiotherapy (WBRT) alone [19, 333]. As a result, WBRT combined with stereotactic radiosurgery has been widely adopted for patients with a limited number of brain metastases (i.e., 1-4 metastases). The main difference with conventional radiotherapy is that the treatment is executed in one session alone using a single and high dose by means of a conformal technique.

In particular, Leksell Gamma Knife<sup>®</sup> (Elekta, Stockholm, Sweden) is a stereotactic radiosurgical device for the treatment of different brain disorders that are often inaccessible for conventional surgery, such as benign or malignant tumors, arteriovenous malformations and trigeminal neuralgia [353, 377]. The gamma rays (generated by Cobalt-60 [<sup>60</sup>Co] radioactive sources) are focused on the target through a metal helmet. Multiple separate radiation rays converge onto the target tumor, delivering precise high dose radiation to the tumor, but a negligible dose to surrounding health tissues [210]. Unlike surgical tumor resection, which eliminates the tumor at the time of the operation, stereotactic radiosurgery causes the tumor to undergo necrosis or growth arrest over time [424, 465]. The Leksell stereotactic frame is applied onto the patient (under mild sedation and a local anesthesia) before MRI scanning and treatment, to accurately locate the target areas and to prevent head movement during imaging and treatment. A personalized treatment plan is implemented using the Leksell GammaPlan<sup>®</sup> Treatment Planning System (TPS). Currently, the Gamma Plan software does not support DICOM-RT objects (generated for the radiation therapy workflow [347]) and does not permit to import/export any external ROIs (neither manually nor automatically calculated). In addition, the Gamma Plan software does not provide automatic image processing methods.

CyberKnife<sup>®</sup> (Accuray Inc., Sunnyvale, CA, USA) is an alternative neuro-radiosurgery system that employs different technologies and treats several anatomical districts [4, 94]. As a matter of fact, the main difference between Gamma Knife and CyberKnife is represented by the design concept of the two devices. Gamma Knife uses [<sup>60</sup>Co] Cobalt-60 radioactive sources, placed inside the machine in well-defined positions, and the patient is moved during the treatment. In Leksell Gamma Knife<sup>®</sup> Model C, the patient is fixed onto the Automatic Positioning System (APS) that is part of the patient bed. In the most recent Gamma Knife<sup>®</sup> Icon<sup>™</sup>, an X-ray system is employed to acquire a cone beam CT. CyberKnife is a frameless system that utilizes a linear accelerator mounted onto a robotic arm, which allows for dose delivery with several degrees of freedom while the patient stays hold on the treatment table. The CyberKnife<sup>®</sup> uses just a thermoplastic mask, since the position of the skull is checked online by means of two X-ray tubes.

MRI provides high-quality detailed images, especially concerning the extent of the oncological disease, providing a support in the radiotherapy planning of solid brain tumors [43]. MRI is a prominent modality in radiation therapy planning and patient follow-up, complementing the use of CT in target delineation. MRI is considered to be superior to CT in determining the extent of tumor infiltration, although the literature shows that some histologic evidence of malignancy may extend beyond the margin of enhancement [160, 225, 297, 298]. In particular, MRI is currently the most used imaging modality in radiation therapy for soft-tissue anatomical districts, such as brain and prostate, by providing several advantages over CT. Furthermore, the excellent soft-tissue contrast of MRI has led to an increasing role of this technique in target volume delineation for therapy applications, such as MRI-guided surgery and radiotherapy treatment, and for patient follow-up (staging and assessing tumor response) [317]. Indeed, MRI is considered the best imaging modality, with significant improvement, in terms of contrast, among lesions and normal tissue [407]. In clinical practice, typical MRI protocols, used for stereotactic radiosurgery treatment planning, include pre- and post-contrast (Gadolinium-enhanced) T1w MR images of the head volume. Because of the infiltrative nature of glial and metastatic tumors, accurate delineation is difficult and the possibility for a precise determination of the appropriate target volume for treatment is limited. Moreover, in patients who underwent surgery, CT and MR imaging may not effectively define tumor recurrence. Furthermore, CT attenuation maps, used in radiotherapy for dose planning, are not available in MRI and geometric distortion could be manifested due to the static magnetic field non-uniformities [172]. Nevertheless, stereotactic dose calculation algorithms utilize a simplified isocentric technique with a known beam profile and a constant linear attenuation through tissue [425], not performing dose painting for non-uniform radiation dose distribution. The Gamma Plan system also assumes that the brain is composed entirely of unit density material [662] and dose delivery is based on the unit “shot” (i.e. a dose distribution approximately with spherical shape). Multiple shots are used in a single Gamma Knife treatment to deliver a conformal dose to an irregular radiosurgical target [664].

Gross Tumor Volume (GTV) segmentation is a great challenge in stereotactic neuro-radiosurgery therapy. In treatment planning phase, this delineation task is generally carried out by expert neurosurgeons or radiation oncologists through a fully manual procedure, since the current TPS platforms do not implement automatic or semi-automatic segmentation tools. This methodology is time-consuming and strongly dependent on the operator: target contouring is affected by high variability in radiation therapy treatment plans [17]. Only automatic or semi-automatic methods, which assist the clinicians in the segmentation task, can enable the repeatability of the delineated lesion boundaries [297], representing an

operator-independent solution for the assessment of the treatment response during patient's follow-up and cancer staging. Nowadays, the GTV is usually delineated on anatomical MRI alone, acquired a few hours before treatment, by means of a fully manual process. Manual segmentation has two main drawbacks: (i) time-expensiveness, because dozens of slices have to be manually segmented in a short time; (ii) operator-dependence in target volume definition. The repeatability of the tumor volume delineation can be ensured only by using effective computer-assisted methods, to support the treatment planning phase. Automated or semi-automated approaches may be of great help, providing higher intra- and inter-operator reliability compared to conventional manual tracing [401]. Consequently, stereotactic neuro-radiosurgery treatment effectiveness can be optimized using automatic methods to support clinicians in the planning phase and to improve treatment response assessment.

In addition, detecting and analyzing necrotic material within the whole tumor region is an important problem in cancer progression assessment [243]. These pathological necrotic regions in heterogeneous tumors are generally composed of cells adapted to survive within a hypoxic microenvironment. Hypoxia is implicated in several aspects of tumor development, angiogenesis, and growth in many different human cancers. The characteristics of highly aggressive brain tumors, particularly the glioblastoma multiforme (GBM) and metastases with necrotic tissue, are supposed to be affected similarly by hypoxia. This pathological condition is hypothesized to be the mechanism by which an aggressive tumor phenotype is developed through increased invasion, metastatic potential, loss of apoptosis, as well as chemoresistance, resistance to anti-angiogenic therapy, and radiation resistance. In this context, evaluating hypoxic regions could predict the likelihood of metastases and tumor recurrence [293]. Sub-volumes reflecting hypoxic cell populations can be identified by considering functional imaging techniques, such as CE-MRI [80]. Also in treatment response assessment, quantifying the extent of necrotic cores within a tumor could be relevant: delayed radiation necrosis, which typically occurs three months or more after the treatment, is indeed the primary risk associated with stereotactic radiosurgery [105, 557]. Necrotic material could develop after aggressive radiation therapy options as well as by the tumor progression itself, such as in high grade gliomas and metastases [349].

The integration of metabolic PET imaging may add another layer of sophistication to the use of radiosurgery in the treatment of gliomas and metastases. In addition, metabolic changes are often faster and more indicative of the effects caused by the therapy with respect to anatomical imaging [644]. Levivier *et al.* [356] used PET functional imaging in stereotactic conditions for the management of brain tumors, and found that it conveys independent metabolic information that is complementary to the anatomical information derived from CT or MR imaging.

In this regard, [ $^{11}\text{C}$ ]-Methionine (MET) seems to have a potential role in conveying additional information, although MRI remains the gold standard for diagnosis, treatment planning, and follow-up in radiation therapy [230, 422]. Methionine is a natural amino acid avidly taken up by brain cancer cells, whereas its uptake by normal cells is low. As a matter of fact, MET uptake is mainly from the activation of the L-mediated and A-mediated amino acid transport at the level of the Blood-Brain Barrier (BBB) [262]. In this way, MET-PET discerns malign and benign tissues in brain tumors with great sensitivity and specificity, by localizing selectively in cancer regions of the brain. Numerous studies have shown that the specificity of the MET-PET for marking tumor delineation and for the differentiation relapse versus radiation necrosis is higher compared with MRI. In [229], metabolic imaging was used for biological target delineation in 36 patients that showed a significantly longer median survival compared with the group of patients in which target volume was merely defined by MRI. In contrast, FDG-PET for metabolic target delineation, which was systematically analyzed by Gross *et al.* [227], showed limited clinical value when comparing brain tumor volume defined by PET with the corresponding volume defined by PET/MRI fusion images. Only in few patients, additional information was derived from FDG-PET for radiation treatment planning because of the low contrast between viable tumor and normal brain tissue, although FDG uptake is regionally related to anaplastic areas. The FDG distribution is thus not limited to malignant tissue because FDG is uptaken according to glucose transport mechanism [147].

The higher diagnostic accuracy of MET-PET is the rationale for using this diagnostic technique in target volume delineation of brain tumors: T1w MRI alone cannot differentiate between treatment-related changes and residual tumor after neurosurgery, chemotherapy or radiation therapy [228]. Although the high contrast between tumor and normal tissue on PET images can reduce the inter- and intra-observer variability in tumor localization, the variability in tumor delineation with the qualitative use of PET is still high and often inconsistent with anatomically defined GTV [228]. Due to the nature of PET images (low spatial resolution, high noise and weak boundary), the Biological Target Volume (BTV) varies substantially depending on the algorithm used to segment functional lesions: the choice of a standard method for PET volume contouring is a very challenging yet unresolved step [585]. Anatomical GTV often does not match with metabolic BTV at all. For this reason, MET-PET metabolic imaging is valuable and it could be used to provide additional information useful for treatment planning and enhanced tumor characterization [209]: the BTV may be used to modify the GTV in order to treat the actual cancer region more precisely. To improve the determination of the lesion margin, it is necessary to combine the complementary information of tissues from both anatomical and functional imaging. Therefore, a reproducible multimodal PET/MRI segmentation method, which contextually segments tumors in each image domain,



is mandatory. This task, named joint segmentation or co-segmentation, is a challenging problem due to: (i) unique demands and peculiarities brought by each imaging modality, and (ii) a lack of one-to-one region and boundary correspondences of lesions in different imaging modalities [675]. The implementation of molecular imaging, such as PET, into Gamma Knife treatments allows us to better understand the biological activity about the cancer imaged on MRI. Especially for the brain tumors gaining deep insights about the cancer region is not always easy investigating MRI alone and it is needed to compare and, often, combine the information acquired by different imaging modalities. Applying contextually PET and MRI in Gamma Knife clinical scenarios showed that sometimes the BTV may contain metabolically active regions outside the GTV. In this case, combining both PET and MR imaging modalities is very important to improve the clinical outcome. Moreover, PET imaging requires computer-assisted segmentation methods to obtain an accurate BTV, minimizing operator-dependence and increasing result repeatability. As a matter of fact, this data integration approach is prominent for effective brain lesion therapy.

### Related work

**Brain tumor segmentation** Computer-assisted brain tumor segmentation on structural medical images, i.e., CT and MRI, represents a widely-investigated research field [666].

Current literature does not present works closely related to the segmentation of target volumes for Gamma Knife treatment support. Consequently, a broader bibliographical research about the state-of-the-art in lesion segmentation on brain MRI was carried out. This section focuses particularly on studies dealing with tumor identification on brain MR images, and especially with target segmentation in radiotherapy.

Hamamci *et al.* [243, 244] presented a segmentation tool for radiosurgery clinical applications, especially in CyberKnife<sup>®</sup> treatments. The proposed semi-automatic method for brain tumors segmentation uses a seeded Cellular Automata (CA) [311] approach on input contrast enhanced T1w MR images. First, the iterative CA framework solves the shortest path problem. Then, an implicit Level Set surface is evolved on a tumor probability map constructed from CA states to handle the heterogeneous tumor regions. Because there are heterogeneous intensity distributions in many application fields using MR images, such as cancer imaging, Aslian *et al.* [27] argued that it is better to use local image statistics instead of a global one. A localized region-based Active Contour semi-automatic segmentation method (by means of LSFs) was applied on brain MR images of patients to obtain the target volumes. A similar method was reported in [671], in which Hybrid Level Sets (HLSs) are exploited. The segmentation is driven by region and boundary information simultaneously.

In [609], owing to the non-overlapping property of Zernike polynomials, MR images are analyzed and the different shapes of the tumors present in the brain were retrieved. Local statistics values obtained from the low- and high-order Zernike moments are thus used to calculate the appropriate threshold value for efficient tumor region extraction.

It was also found that segmentation approaches borrowed from Artificial Intelligence techniques are frequently used. Mazzara *et al.* [396] presented and evaluated two fully automated brain MRI tumor segmentation approaches: supervised  $k$ -Nearest Neighbors (kNN) and automatic Knowledge-Guided (KG) methods. Analogously, an approach based on Support Vector Machine (SVM) discriminative classification is proposed in [42]. However, for these supervised methods, a training step is always required. Hall *et al.* [242] extensively compared literal and approximate FCM (i.e., FCM [53] and AFCM [78], respectively) unsupervised clustering algorithms, and a supervised computational Neural Network, in brain MRI segmentation [54]. The comparison study suggested comparable experimental results (qualitative visual evaluation made by expert radiologists) between supervised and unsupervised learning. Another system for unsupervised feature extraction for brain MRIs was presented in [700]. In particular, T1w brain MR images are partitioned into three main tissue types: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). First, a multidimensional feature vector is constructed based on intensity, low-frequency sub-band of Dual-Tree Complex Wavelet Transform (DT-CWT), and spatial position image information. Then, a Spatial Constrained Self-Organizing Tree Map (SCSOTM) is used as the actual segmentation system. SCSOTM applies a dual-thresholding method for automatic growing of the tree map, which uses the information from the high-frequency sub-bands of DT-CWT. Several Soft Computing techniques have been thus proposed for brain tumor region identification, and a complete automatic segmentation method using optimization and clustering techniques was proposed in [220]. The resulting hybrid FCM algorithm, combined with Particle Swarm Optimization (PSO) [316] during the centroid vector initialization step, segments the tissues and identifies the edema and tumor affected regions in the brain. On the other hand, Verma *et al.* [633] presented an Improved Fuzzy Entropy Clustering (IFEC) algorithm to segment brain MR images, characterized by noisy data. A new fuzzy factor, which incorporates both local spatial and gray-level information, was introduced. The developed IFEC method handles noisy data, preserves image details during clustering, and is setting-free.

**Brain tumor tissue classification** Several approaches aimed to process multispectral MRI data, by generally employing supervised Machine Learning techniques even though large-scale labeled input datasets are required to learn highly discriminative features. How-

ever, brain tumor segmentation and tissue classification is still a difficult problem, due to the large number of cancer types that differ greatly in size, shape, location, tissue composition, and tissue homogeneity [652]. Bauer *et al.* [42] proposed a supervised fully automatic approach for brain tumor tissue classification using an SVM model for multispectral MR image intensities/textures classification. Since this SVM assumes voxels to be independent from their neighbors, a subsequent hierarchical regularization based on Conditional Random Fields (CRFs) was integrated to introduce spatial constraints and neighboring context information that encounter the dependence of voxel labels on their neighbors. The whole process results in a two-step classification: (i) healthy and tumor tissue distinction; (ii) sub-classification in CSF, WM, GM and necrotic, active, edema regions, respectively. The approach was tested on a dataset composed of 10 patients affected by gliomas, including T1w, T1w-CE, T2w, and Fluid Attenuation Inversion Recovery (FLAIR) MR images. Similarly, in Wu *et al.* [666] multispectral MRI data (T1w, T1w-CE, T2w, and FLAIR images) are first segmented into superpixels to alleviate the sampling issue and to improve the sample representativeness. Afterwards, features were extracted from the superpixels using multi-level Gabor wavelet filters. According to the features, an SVM classifier and an affinity metric model for tumors were trained to overcome the limitations of previous generative models. Based on the output of the SVM and spatial affinity models, CRFs were applied to segment the tumor using a maximum *a posteriori* (MAP) estimation, given the smoothness prior defined by the affinity model. The same classification procedure is repeated for edema segmentation. Finally, a structural denoising process on the labels is performed by relying on symmetrical and continuous characteristics of the tumor in spatial domain. Li *et al.* [360] presented a model that combines sparse representation and Markov Random Fields (MRFs) to solve the spatial and structural variability problem. The tumor segmentation problem was formulated as a multi-classification task wherein each voxel is labeled estimating the MAP probability, by introducing the sparse representation into a likelihood probability and an MRF into the prior probability. Since the MAP is an NP-hard problem, this estimation is converted in a minimum energy optimization problem and the Graph cuts algorithm is exploited to find the solution to the MAP estimation. The Brain Tumor Segmentation (BRATS) 2015 database was used for the evaluation [404]. Zikic *et al.* [706] presented a method for automatic segmentation of high-grade gliomas and the following differentiation between active cells, necrotic core, and edema. The authors integrated a discriminative model, based on decision Random Forests (RFs) using context-aware spatial features, with a generative model of tissue appearance probabilities, obtained by tissue-specific Gaussian Mixture Models (GMMs), as additional input for the RF. The approach was tested on a set of multi-channel 3D MRI data for 40 patients suffering from high-grade gliomas.

In the latest year, Deep Neural Networks (DNNs) have been used to learn a hierarchy of increasingly complex features from the processed data, enabling multiple levels of abstraction [350]. Havaei *et al.* [255] adapted Convolutional Neural Network (CNN) architectures specifically to the task of brain tumor segmentation on multispectral MRI data. The prediction of the label of each pixel is influenced by local features (i.e., details of the region around that pixel) as well as more global contextual features (i.e., considering a larger “context” about the position of the patch is in the brain), simultaneously. Various architectures are compared by concatenating feature maps from different layers. In such cascaded CNN architectures, a two-phase training procedure was employed to deal with imbalanced label distribution. An alternative deep architecture, consisting in an 11-layers three-dimensional CNN, is proposed in Kamnitsas *et al.* [305]. A dense training scheme was used, by processing adjacent 3D image patches (i.e., set of connected voxels) into one pass adapting to the inherent class imbalance. A dual pathway architecture for local and larger contextual information simultaneous processing. Finally, a CRF is used as a post-processing step for regularization, by removing false positives. The proposed pipeline was evaluated on three challenging tasks of lesion segmentation in multispectral MRI patient datasets concerning traumatic brain injuries, brain tumors, and ischemic strokes. Top ranking performance was achieved on the public benchmarks BRATS 2015 [404] and Ischemic Stroke Lesion Segmentation (ISLES) 2015 [384] challenges.

Most of the methods presented above used supervised Machine Learning techniques, thus requiring a training phase. Nowadays, the availability of medical imaging data for both training and validation phases represents still a serious issue, especially when a computational approach is tailored for a specific clinical application. Indeed, unsupervised techniques do not need input labeled images for the training and could be clinically feasible for supporting neurosurgeons and radiation oncologists in their decision-making tasks. Unsupervised Machine Learning techniques have gained ground in biomedical applications, particularly when dealing with an amount of input data that does not allow for a significant training sample. CA models have been also recently used for MRI brain tumor segmentation. Hamamci *et al.* [243] presented a semi-automatic segmentation tool, called “Tumor-Cut”, for brain tumors segmentation in Cyberknife<sup>®</sup> radiosurgery treatments. A method based on CA is used on input contrast enhanced T1w MR images. First, the iterative CA framework solves the shortest path problem. Secondly, an implicit Level Set surface is evolved on a tumor probability map, which combines the tumor and background strengths obtained separately by two independent CA algorithm executions. User interaction consists in drawing the maximum diameter of the tumor on the corresponding axial slice, according to the Response Evaluation Criteria In Solid Tumors (RECIST) rules [167]. Subsequently, a bounding volume

is built, by enlarging the traced line, and the seeds are computed. The method was validated on three different T1w CE-MRI datasets: 5 synthetically generated brain tumors from the University of Utah, 10 real brain tumors from the Harvard Brain Tumor Repository, 19 tumors from patients who underwent CyberKnife radiosurgery treatment. A further processing phase based on CA was proposed to distinguish necrotic and enhancing regions within the previously segmented tumor. The initial seed selection procedure exploits the threshold computed by the Otsu's adaptive thresholding method [448], by employing the 25% of the most hypo-intense necrotic and the 25% of the most hyper-intense enhanced volumes as necrosis (i.e., foreground) and enhancement (i.e., background) seeds, respectively. This approach for enhancement and necrotic region differentiation was evaluated on 6 tumors with necrosis taken from the clinical dataset of patients who underwent neuro-radiosurgery. Sompong and Wongthanavas [575] presented the Improved Tumor-Cut (ITC) to cope with the robustness of seed growing in the standard Tumor-Cut algorithm. The multimodal MRI brain tumor dataset BRATS 2013, including T1w, T2w, T1w-CE, and FLAIR images, was processed. The authors focused on the segmentation of edema tissues, which generally exhibit ambiguous tumor boundaries and are imaged with a brighter hyper-intense signal than normal brain on T2w MRI. Firstly, MR images are enhanced using a novel Gray Level Co-occurrence Matrix based CA (GLCM-based CA) [250, 251], by transforming the original MRI in a new GLCM feature space. Secondly, the proposed CA algorithm maps these features to the target image, by employing a patch-weighted feature in the similarity function and consequently defining a patch-weighted distance. Cluster analysis methods have been widely used also due to the computational efficiency since it does not require training data, so increasing operator-independency and responding to potential inconsistencies of supervised classification techniques [185]. A superpixel-based graph for GBM from multimodal MR images was used in [591]. First, a local  $K$ -means clustering with weighted distances is employed to segment the multispectral MR images, including T1w, T1w-CE, T2w and FLAIR, into superpixels. Then, the spectral clustering algorithm is utilized to extract the enhanced tumor, necrosis and non-enhancing T2w hyper-intense regions by considering the superpixel map as a graph. Another critical issue is represented by the border between pathological and normal tissues that is often not well imaged in heterogeneous tumors. Moreover, PVEs (i.e., one pixel may be actually composed of multiple tissue types) and the noise due to the MRI acquisition process can affect segmentation accuracy. Fuzzy Logic techniques can cope with these difficulties, by modeling multiple membership degrees for each voxel and fusing multispectral MRI information [154]. Therefore, introducing Fuzzy Logic in clustering-based classification has represented an active research area in recent years. Prakash and Kumari [400] proposed an automated segmentation method of MR brain

images into three tissue types: WM, GM, and CSF based on FCM clustering, by integrating spatial information and contrast enhancement. Similarly, Agnello *et al.* [6] proposed an unsupervised tissue classification method of brain for Voxel-Based Morphometry (VBM) analysis on T1w MRI data alone. The FCM algorithm for defining the training set, composed of tuple wherein each intensity value of the MRI data corresponds to a specific cluster, in order to train a Fully Connected Cascade Neural Network (FCCNN) to learn the brain tissue classification in WM, GM, and CSF. Govindaraj and Murugan [220] presented a hybrid FCM algorithm, combined with the PSO technique in the centroid vector initialization step, to segment brain tumor tissues and identify edema on multispectral MRI. Hemanth *et al.* [266] introduced an additional data reduction pre-processing step in the conventional FCM algorithm to minimize the computational complexity and the convergence rate. The resulting Fast and Accurate FCM (FAFCM) algorithm consists of two stages: (i) binary representation and pre-clustering for intensity value reduction; (ii) FCM clustering with the reduced MRI sequence. Kaya *et al.* [315] focused on reduction techniques, based on different PCA versions, as a data pre-processing step to allow for an appropriate low-dimensional MRI data representation for the subsequent clustering process using the  $K$ -means and FCM algorithms.

**Multimodal PET/MRI segmentation** A huge number of monomodal segmentation methods on single image modalities (i.e. CT, MRI, or PET) are present in the literature. Only few studies address the segmentation of multimodal imaging data. In particular, this section reviews the approaches that combine the complementary functional information conveyed by PET with the corresponding anatomical image scans (CT or MRI). In fact, in clinical use, it is highly desirable to acquire both functional and structural quantifiable information by means of a single scan so the disease can be both identified and localized, potentially resulting in an earlier diagnosis and more effective treatment plan. In parallel to the developments of multimodal scanners (PET/CT and PET/MRI), there have been recent literature works trying to bring the usefulness of integrating anatomical and functional information into a common place for separating tumor tissues from normal structures. A concise but complete review of the state-of-the-art regarding multimodal co-segmentation approaches is reported in the following, wherein each research paper is briefly explained.

Several studies deal with PET/CT tumor identification and characterization in radiation therapy scenarios. In [687], co-registered FDG-PET/CT images were used for the textural characterization of head and neck cancer (HNC) in radiotherapy treatment planning. After a manual segmentation on co-registered PET/CT images (performed by an experienced radiation oncologist), useful textural features were selected for distinguishing tumor from normal tissue in HNC subjects. Both  $k$ -NNs and decision tree (DT)-based  $k$ -NN classifiers

were employed to discriminate images of cancerous and healthy tissues. Han *et al.* [248] presented an MRF-based co-segmentation of the PET/CT image pair with a regularized term that penalizes the segmentation difference between PET and CT. This graph-based method utilizes the strength of PET and CT modalities for target delineation in a group of 16 patients with HNC. Background and foreground seed voxels must be always manually identified by the user. A similar approach is reported in [576], where the segmentation is seen as a minimization problem of an MRF model, which encodes the information from both modalities. This optimization is solved using a Graph cuts based method [64, 65], by constructing two sub-graphs for PET and CT segmentation, respectively, and adding inter-subgraph arcs. The overall approach is semi-automatic, because initial seed points are required for the graph-based co-segmentation: one center point and two radii are given by the user. The algorithm was validated in robust delineation of lung tumors on 23 PET/CT datasets and two HNC subjects. A further MRF-based systematic solution for the automated co-segmentation of brain PET/CT images into GM, WM, and CSF regions is exposed in [668]. A PET/CT image pair and its segmentation result are modeled as an MRF triplet, and segmentation is eventually achieved by solving a maximum *a posteriori* (MAP) problem using the EM [148] algorithm with simulated annealing. The overall MRF-MAP model was tested both on simulated and real patient PET/CT data. The same authors presented in [667] a more advanced brain PET/CT segmentation model. This dual-modality image segmentation approach converts PET/CT image segmentation into an optimization process controlled simultaneously by PET and CT voxel values and spatial constraints. A modality discriminatory power (MDP) coefficient is applied as a weighing scheme to adaptively combine functional and anatomical information to separate voxels from different tissue types. The authors of [481] proposed a method for automated delineation of tumor boundaries in whole-body PET/CT by jointly using information from both PET and diagnostic CT images. After an initial robust hot-spot detection and segmentation performed in PET, a model for tumor appearance and shape in corresponding CT structures is learned by the weighted non-parametric density estimate. This voxel-based CT classification is then probabilistically integrated with the PET classification using the joint likelihood ratio test technique to derive the final segmentation. The algorithm was tested on patient studies with lung and liver tumors identifiable in both the PET and CT images acquired by the same scanner. In this context, a recent study [624] aims at developing a radiomics model from joint FDG-PET and MRI texture analysis for the early evaluation of lung metastasis risk in soft-tissue sarcomas (STSs). The creation of new composite textures from the combination of FDG-PET and MR imaging information, to better identify aggressive tumors, was investigated. The results showed that FDG-PET and MRI texture features could act as strong prognostic factors of STSs.

Yezzi *et al.* [685] introduced a geometric variational framework that uses active contours to simultaneously segment and register features from multiple images. The key aspect of this approach is that multiple images may be segmented by evolving a single contour as well as the mappings of that contour into each image during feature-based realignment steps. The results of three experiments on MRI/CT images of the head and the spine were reported. In [659], a variational segmentation method, based on the minimization of the total variation semi-norm and a convex formulation, was used for segmenting lung tumors and lymph node disease on thoracic PET/CT image pairs, in the context of radiotherapy planning. Also the authors of [169] developed a method based on multi-valued level set deformable models for simultaneous 2D or 3D segmentation of multimodality images consisting of combinations of co-registered PET, CT, or MRI datasets. In particular, only three patients are considered: a non-small cell lung cancer case with PET/CT, a cervix cancer case with PET/CT, and a prostate patient case with CT/MRI. In addition, CT, PET, and MRI phantom data were used for quantitative validation of the proposed multimodality segmentation approach. However, here we focus on PET/MRI joint segmentation approaches reported in the literature. An automatic algorithm for the co-segmentation of HNC based on PET/MRI data was proposed in [352], in order to standardize tumor volume delineation. For both imaging modalities tumor probability maps were derived, assigning each voxel a probability of being cancerous according to its signal intensity. A combination of these maps was subsequently segmented using a threshold level set algorithm. The algorithm processes both the anatomical T2w MRI and FDG-PET data concerning 10 HNC patient datasets acquired by a combined PET/MRI system. The group led by Bagci and Mollura developed some co-segmentation approaches in multimodal medical imaging, which mostly used the Random Walker (RW) algorithm [221]. In [31] a computer-assisted volume quantification method for PET/MRI dual modality images using automated PET-guided RW MR image co-segmentation was presented. A small-animal breast cancer model was built by using 30 female nude mice. A more comprehensive and general approach was proposed in [32]. The RW algorithm is extended for jointly delineating multiple objects from multimodal images by unifying graph representation of each image modality in a single product lattice. The overall method results in an automatic and unified framework, providing an automated object detection *via* an Interesting Uptake Region (IUR) algorithm to avoid users having to provide foreground and background seeds. Afterwards, prior to the initiation of the segmentation process, these identified seeds are propagated to the corresponding anatomical images. Although no significant anatomical and functional changes between the scans have to be assumed, the study used PET, PET/CT, MRI/PET, and fused MRI/PET/CT scans (77 studies in total) from 56 patients who had various lesions in different body regions. Finally, a 3D computer-aided co-segmentation tool for lesion



detection and quantification from hybrid PET/MRI and PET/CT images was described in [675]. This method used a modality-specific visibility weighting scheme built upon a Fuzzy Connectedness (FC) image segmentation algorithm.

### **Prostate MRI analysis**

According to the American Cancer Society, in 2018 the Prostate Cancer (PCa) is the most common malignant tumor for American males [564]. The only well-established risk factors for prostate cancer are increasing age, a family history of the disease, and genetic conditions. Prostate image analysis is still a critical and challenging issue in diagnosis, therapy, and staging of PCa.

Currently, high-resolution multiparametric MRI has been shown to have higher accuracy than Transrectal Ultrasound (TRUS) when used to ascertain the presence of prostate cancer [367]. In [503], MRI-guided biopsy showed a high detection rate of clinically significant PCa in patients with persisting cancer suspicion, also after a previous negative systematic TRUS-guided biopsy, which represents the gold-standard for diagnosis of prostate cancer [259]. In addition, MRI scanning before a biopsy can also serve as a triage test in men with raised serum Prostate-Specific Antigen (PSA) levels, in order to select those for biopsy with significant cancer that requires treatment [9]. For instance, PCa treatment by radiotherapy requires an accurate localization of the prostate. In addition, neighboring tissues and organs at risk, i.e., rectum and bladder, must be carefully preserved from radiation, whereas the tumor should receive a prescribed dose [327]. CT has been traditionally used for radiotherapy treatment planning, but MRI is rapidly gaining relevance because of its superior soft-tissue contrast, especially in conformal and intensity-modulated radiation therapy [635]. Manual detection and segmentation of both prostate gland and prostate carcinoma on multispectral MRI is a time-consuming task, which has to be performed by experienced radiologists [245]. In fact, in addition to conventional structural T1w and T2w MRI protocols, complementary and powerful functional information about the tumor can be extracted [104, 213, 355] from: DCE [174], DWI [240], and MRSI [197, 393]. A standardized interpretation of multiparametric MRI is very difficult and significant inter-observer variability has been reported in the literature [507]. The increasing number and complexity of multispectral MRI sequences could overwhelm analytic capabilities of radiologists in their decision-making processes. Therefore, automated and computer-aided segmentation methods are needed to ensure results repeatability. Accurate prostate segmentation on different imaging modalities is a mandatory step in clinical activities. For instance, prostate boundaries are used in several treatments of prostate diseases: radiation therapy, brachytherapy, HIFU ablation, cryotherapy, and transurethral microwave therapy [213]. Prostate volume, which can be directly calculated

from the prostate ROI segmentation, aids in diagnosis of benign prostate hyperplasia and prostate bulging. Precise prostate volume estimation is useful for calculating PSA density and for evaluating post-treatment response. However, these tasks require several degrees of manual intervention, and may not always yield accurate estimates [615].

Given a clinical context, several imaging modalities can be used for PCa diagnosis, such as TRUS, CT, and MRI. For an in-depth investigation, structural T1w and T2w MRI sequences can be combined with the functional information from DCE-MRI, DWI, and MRSI [355]. Despite the recent technological advances in MRI scanners and coils, prostate images are prone to artifacts related to magnetic susceptibility. Although the shift from 1.5 T to 3 T magnetic field strength systems theoretically results in a doubled Signal-to-Noise-Ratio (SNR), it also involves different T1 (spin-lattice) and T2 (spin-spin) relaxation times and greater magnetic field inhomogeneity in the organs and tissues of interest [75, 512]. On the other hand, 3 T MRI scanners allow for using pelvic coils instead of endorectal coils, obtaining good quality images and avoiding invasiveness as well as prostate gland compression and deformation [260, 319]. So, challenges for automatic segmentation of the prostate in MR images include the presence of imaging artifacts due to air in the rectum (i.e., susceptibility artifacts) and large intensity inhomogeneities of the magnetic field (i.e., streaking or shadowing artifacts), the large anatomical variability between subjects, the differences in rectum and bladder filling, and the lack of a normalized “HU” for MRI, like in CT scans [327]. Therefore, MRI plays a decisive role in PCa diagnosis and disease monitoring (even in an advanced status [450]), revealing the internal prostatic anatomy, prostatic margins, and PCa extent [635]. According to the zonal compartment system proposed by McNeal, the prostate Whole Gland (WG) can be partitioned into the Central Gland (CG) and Peripheral Zone (PZ) [554]. In prostate imaging, T2w MRI serves as the principal sequence [547], thanks to its high resolution that enables to differentiate the hyper-intense PZ and hypo-intense CG in young male subjects [269].

Besides manual detection/delineation of the WG and PCa on MR images, distinguishing between the CG and PZ is clinically essential, since the frequency and severity of tumors differ in these regions [106, 435]. As a matter of fact, the PZ harbors 70-80% of PCa and represents a target for prostate biopsy [237]. Furthermore, the PZ volume ratio (i.e., the PZ volume divided by the WG volume) can be considered for PCa diagnostic refinement [94], while the CG volume ratio may help monitoring prostate hyperplasia [324]. Therefore, according to the Prostate Imaging-Reporting and Data System version 2 (PI-RADS<sup>TM</sup> v2) [654], radiologists must perform a zonal partitioning before assessing the suspicion of PCa on multi-parametric MRI.

## Related work

**Whole prostate gland segmentation** From a computer science perspective, especially in automatic MR image analysis for PCa segmentation and characterization, the computed prostate ROI is very useful for the subsequent more advanced processing steps [355]. Klein *et al.* [327] presented an automatic method, based on non-rigid registration of a set of pre-labeled atlas images, for delineating the prostate in 3D MR scans. Each atlas image is non-rigidly registered, employing Localized Mutual Information (LMI) as similarity measure [588], with the target patient image. The registration is performed in two steps: (i) coarse alignment of the two images by a rigid registration; (ii) fine non-rigid registration, using a coordinate transformation that is parameterized by cubic B-splines [514]. The registration stage yields a set of transformations that can be applied to the atlas label images, resulting in a set of deformed label images that must be combined into a single segmentation of the patient image, considering majority voting (VOTE) and Simultaneous Truth And Performance Level Estimation (STAPLE) [653]. The method was evaluated on 50 clinical scans (probably T2w images, but the acquisition protocol is not explicitly reported in the paper), which had been manually segmented by three experts. The Dice Similarity Coefficient (*DSC*) [152], to quantify the overlap between the automatic and manual segmented regions, and the shortest Euclidean distance, between the manual and automatic segmentation boundaries, were computed. The differences among the used four fusion methods are small and mostly not statistically significant. For most test images, the accuracy of the automatic segmentation method was, on a large part of the prostate surface, close to the level of inter-observer variability. With the best parameter setting, a median *DSC* of around 0.85 was achieved. The segmentation quality is especially good at the prostate-rectum interface, whereas most serious errors occurred around the tips of the seminal vesicles and at the anterior side of the prostate. Vincent *et al.* [636] proposed a fully automatic model-based system for segmenting the prostate in MR images. The segmentation method is based on Active Appearance Models (AAMs) [261] built from 50 manually segmented examples provided by the Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2012 Prostate MR Image Segmentation 2012 (PROMISE12) team [365]. High-quality correspondences for the model are generated using a variant of the Minimum Description Length approach to Groupwise Image Registration (MDL-GIR) [128], which finds the set of deformations for registering all the images together as efficiently as possible. A multi-start optimization scheme is used to robustly match the model to new images. To test the performance of the algorithm, *DSC*, mean absolute distance and the 95%-th Hausdorff distance (*HD*) [82] were calculated. The model was evaluated using a Leave-One-Out Cross Validation (LOOCV) on the training data

obtaining a good degree of accuracy and successfully segmented all the test data. The system was used also to segment 30 test cases (without reference segmentations), considering the results very similar to the LOOCV experiment by a visual assessment. In [203], a unified shape-based framework to extract the prostate from MR images was proposed. First, the authors address the registration problem by representing the shapes of a training set as point clouds. In this way, they are able to exploit the more global aspects of registration *via* a particle filtering based scheme [541]. In addition, once the shapes have been registered, a cost functional is designed to incorporate both the local image statistics and the learned shape prior (used as constrain to perform the segmentation task). Satisfying experimental results were obtained on several challenging clinical datasets, considering *DSC* and the 95%-th *HD*, also compared to [620], which employs the Chan-Vese Level Set model [92] assuming a bimodal image. Martin *et al.* [388] presented the preliminary results of a semi-automatic approach for prostate segmentation of MR images that aims to be integrated into a navigation system for prostate brachytherapy. The method is based on the registration of an anatomical atlas computed from a population of 18 MRI studies along with each input patient image. A hybrid registration framework, which combines an intensity-based registration [265] and a Robust Point-Matching (RPM) algorithm introducing a geometric constraint (i.e., the distance between the two point sets) [493], is used for both atlas building and atlas registration. This approach incorporates statistical shape information to improve and regularize the quality of the resulting segmentations, using Active Shape Models (ASMs) [126]. The method was validated on the same dataset used for atlas construction, using the LOOCV. Better segmentation accuracy, in terms of volume-based and distance-based metrics, was obtained in the apex zone and in the central zone with respect to the base of the prostate. The same authors proposed a fully automatic algorithm for the segmentation of the prostate in 3D MR images [389]. The approach requires the use of a probabilistic anatomical atlas that is built by computing transformation fields (mapping a set of manually segmented images to a common reference), which are then applied to the manually segmented structures of the training set in order to get a probabilistic map on the atlas. The segmentation procedure is composed of two phases: (i) the processed image is registered to the probabilistic atlas and then a probabilistic segmentation (allowing the introduction of a spatial constraint) is performed by aligning the probabilistic map of the atlas against the patient's anatomy; (ii) a deformable surface evolves towards the prostate boundaries by merging information coming from the probabilistic segmentation, an image feature model, and a statistical shape model. During the evolution of the surface, the probabilistic segmentation allows for the introduction of a spatial constraint that prevents the deformable surface from leaking in an unlikely configuration. This method was evaluated on 36 MRI exams. The results

showed that introducing a spatial constraint increases the segmentation robustness of the deformable model compared to another one that is only driven by an image appearance model. Toth *et al.* [615] used a Multifeature Active Shape (MFA) model [127], which extends ASMs by training the multiple features for each voxel using a GMM based on the log-likelihood maximization scheme, for estimating prostate volume from T2w MRI. Using a set of training images, the MFA learns the most discriminating statistical texture descriptors of the prostate boundary *via* a forward feature selection algorithm. After the identification of the optimal image features, the MFA is deformed to accurately fit the prostate border. The MFA-based volume estimation scheme was evaluated on a total 45 *in vivo* T2w MRI studies, corresponding to both 1.5 T and 3.0 T field strengths. A correlation with the ground truth volume estimations showed that the MFA obtained higher Pearson's correlation coefficient values with respect to the current clinical volume estimation schemes, such as ellipsoid, Myschetzky, and prolate spheroid methodologies. In [615], the same research team proposed a scheme to automatically initialize an ASM for prostate segmentation on endorectal *in vivo* multiprotocol MRI, exploiting MRSI data and identifying the MRI spectra that lie within the prostate. As a matter of fact, MRSI is able to measure relative metabolic concentrations and the metavoxels near the prostate have distinct spectral signals from metavoxels outside the prostate. A replicated clustering scheme is employed to distinguish prostatic from extra-prostatic MR spectra in the midgland. The detected spatial locations of the prostate are then used to initialize a multi-feature ASM, which employs also statistical texture features in addition to intensity information. This scheme was quantitatively compared with another ASM initialization method for TRUS imaging by Cosìo [129], showing superior average segmentation performance on a total of 388 2D MRI sections obtained from 32 3D endorectal *in vivo* patient studies. Other segmentation approaches are also able to segment the zonal anatomy of the prostate [106], differentiating the CG from the PZ [278]. The method in [385] is based on a modified version of the Evidential C-Means (ECM) clustering algorithm [391], introducing spatial neighborhood information. *A priori* knowledge of the prostate zonal morphology was modeled as a geometric criterion and used as an additional data source to enhance the differentiation of the two zones. 31 clinical MRI series were used to validate the accuracy of the method, taking into account inter-observer variability. The mean *DSC* was 89% for the CG and 80% for the PZ, calculated against an expert radiologist segmentation.

**Prostate zonal segmentation** Due to the crucial role of MR image analysis [355] in PCa diagnosis and staging, researchers have paid specific attention to automatic WG detection/segmentation. Classic methods mainly leveraged atlases [327, 388] or statistical shape priors [389]: atlas-based approaches realized accurate segmentation when new

prostate instances resemble the atlas, relying on a non-rigid registration algorithm [389, 616]. Unsupervised clustering techniques allowed for segmentation without manual labeling of large-scale MRI datasets [518, 519]. In the latest years, Deep Learning techniques [364] have achieved accurate prostate segmentation results by using deep feature learning combined with shape models [233] or location-prior maps [595]. Moreover, CNNs were used with patch-based ensemble learning [294] or dense prediction schemes [418].

Differently to WG segmentation, less attention has been paid to CG and PZ segmentation despite its clinical importance in PCa diagnosis [435]. In this context, classic Computer Vision techniques have been mainly exploited on T2w MRI. For instance, early studies combined classifiers with statistical shape models [15] or deformable models [686]; Toth *et al.* [616] employed active appearance models with multiple level sets for simultaneous zonal segmentation; Qiu *et al.* [487] used a continuous max-flow model—the dual formulation of convex relaxed optimization with region consistency constraints [688]; in contrast, Makni *et al.* [385] fused and processed 3D T2w, DWI, and T1w CE-MR images by means of an ECM algorithm [391]. Litjens *et al.* [363] proposed a zonal prostate segmentation approach based on Pattern Recognition techniques. First of all, they use the multi-atlas segmentation approach in [327] for prostate segmentation, by extending it for using multi-parametric data. T2w MR images and the quantitative ADC maps were registered simultaneously. As a matter of fact, the ADC map conveys additional information on the zonal distribution within the prostate in order to differentiate between the CG and PZ. Then, for the voxel classification, the authors determined a set of features that represent the difference between the two zones, by considering three categories: anatomy (positional), intensity and texture. This voxel classification approach was validated on 48 multiparametric MRI studies with manual segmentations of the whole prostate. *DSC* and Jaccard index (*JI*) [290] showed good results in zonal segmentation of the prostate and outperformed the multiparametric multi-atlas based method in [385] in segmenting both CG and PZ.

As the first CNN-based method, Clark *et al.* [115] detected DWI MR images with prostate relying on Visual Geometry Group (VGG) net [567], and then sequentially segmented WG and CG using U-Net [506]. However, no literature method so far coped with the generalization ability among multi-institutional MRI datasets, making their clinical applicability difficult [12]. In a preliminary study [517], we compared existing CNN-based architectures—namely, SegNet [30], U-Net [506], and pix2pix [289]—on two multi-institutional MRI datasets. According to our results, U-Net generally achieves the most accurate performance.

## 2.2.2 Biological image analysis

### Clonogenic assays

Clonogenic assay (or colony formation assay) is an *in vitro* cell survival technique based on the ability of a single cell to grow into a colony. It is widely used to study the number and the size of cell colonies after irradiation or cytotoxic agent administration, serving as a measure for the anti-proliferative effect of these treatments. After plating, incubation and treatment of cells, the standard procedure consists in colony counting under a stereomicroscope [486]. Traditionally, clonogenic assay evaluation has been performed by counting the colonies composed of at least 50 densely-packed cells [486]: to evaluate the effect of the treatment on cell survival, the Plating Efficiency (PE), which is the fraction of colonies obtained from untreated cells, and the Surviving Fraction (SF) of cells after any treatment, are measured [191]. It is well known that cells after any type of treatment have a reduction in growth rate, and the colonies are clearly smaller than the ones of the untreated control. The conventional counting process can underestimate the actual effect of the therapy because, as explained in [191], the treatment could not decrease the colony number while reducing the colony size. This counting procedure is definitely time-consuming and tedious. Especially, accurate counting with high numbers of Colony Forming Units (CFUs) on a plate is difficult and error-prone, since it has to be carefully carried out by human operators. Manual counting is strongly subjective and it was shown that the results may also be affected by intra-operator variability [376]. From the variation observed in the measurements, the automated counts are generally more consistent with respect to the manual ones, representing the only approach able to ensure result repeatability and operator independence. The concept of computer-aided colony counting has been already investigated in literature. As a matter of fact, several methods to segment images of cell colonies were proposed [38, 49, 103, 138, 206, 235], but none of them has been widely used in practice. Their apparent failure to be adopted could be explained by: (i) the inability to split adjacent merged colonies, representing a common problem with mammalian cell cultures, and (ii) long processing times or, more generally, the need for interactive inputs since fully automatic approaches have not been proposed yet. Moreover, no related works address all the problems simultaneously. Some approaches analyze only single-well images, while the ones that use multi-well plates do not perform a fully automatic well detection, by strongly relying on plate measurements and characteristics during well extraction phases [138, 235]. In addition, some methods use different image processing techniques depending on the type of the analyzed cells, leaving the choice of the most suitable method case-by-case to the user [38, 138].

Considering the above mentioned open issues and that the well Area Covered by Colonies (ACC) is directly correlated with the number and size of colonies, we decided to tackle these challenges by measuring the percentage of ACC. In this way, we overcome the necessity to split the possible merged colonies, by taking into account just the colony size, which has not currently been considered, since it is correlated with the effect of the treatment delivered to the cells [249, 395].

**Related work** In the literature, numerous computer-aided approaches were proposed for colony counting in clonogenic assays, even though they strongly depend on the analyzed cell type. Especially, when cell colonies are small, characterized by a circular shape or with few adjacent merging colonies, such as in isolated and diluted bacteria cultures, the problem is easier to solve and estimating the number of colonies is not particularly difficult [49, 103, 206, 116]. On the contrary, colonies with irregular shapes and greater sizes, as well as colonies deriving from plating a conspicuous number of cells, often cause colony merging and overlapping, which could represent a critical problem when the aim is determining the number of grown colonies. Moreover, the conventional counting process can underestimate the effect of therapy, since it is possible that the treatment does not reduce the colony number but shrinks the colony size, as reported in [376]. In [138], the authors designed a cheap and efficient device, which employs a flatbed scanner to acquire 12 Petri dishes at the same time, for automated counting of mammalian cell colonies. A sequence of traditional image processing algorithms, including low-pass and high-pass filtering, was applied to enhance the image for shape analysis of colonies. Colony detection was performed by placing the ROIs automatically onto a fixed location of the Petri dishes. Finally, gray level global thresholding was performed by means of user interaction on the histogram of each ROI. Guzmán *et al.* [235] developed ColonyArea, a plug-in for the open-source image analysis software ImageJ (National Institutes of Health, Bethesda, MD, USA) [552], optimized for an immediate analysis of focus formation assays. ColonyArea processes image data of multi-well dishes for colony formation quantification, after separating, concentrically cropping and background correcting individual well images. Instead of counting the number of colonies, ColonyArea determines the percentage of area covered by crystal violet stained cell colonies. In the well separation and cropping steps, this approach uses a mask obtained considering the size and shape of the used plate. The characteristics and typical dimensions of 6- to 24-well plates, as reported in the datasheets from a specific manufacturer, are stored in the plug-in. In [103], an automated counting system was proposed, by using a hardware apparatus (designed and built ad hoc by the authors themselves) for Petri dish image acquisition. The processing pipeline is based on several steps: (i) the image is converted from color to gray levels using PCA; (ii)



the gradient magnitude image is calculated and thresholded to detect the imaged dish; (iii) the bacterial colonies are extracted using the Otsu's method [448]. Colonies in the central area and in the rim area are analyzed differently, both in the extraction and counting phases. In the case of overlapping colonies, the distance transform [321, 394] and the watershed transform [51, 638] are used to divide the overlapping colonies. Clarke *et al.* [116] described a colony counting tool, named NIST's Integrated Colony Enumerator (NICE), exploiting either an ordinary camera or a flatbed scanner. The developed software combines extended minima, employed for colony center localization, and automatic or manual thresholding approaches [448]. Image denoising is performed by means of Gaussian smoothing and then extended minima are computed. The author of [206] presented OpenCFU: a method for circular object segmentation on images and video sequences. The input color image, acquired by means of either a high-definition camera or a webcam (relying on a white trans-illuminator to optimize image contrast in both cases), is pre-processed separately channel-by-channel to correct background intensity inhomogeneities and expand the contrast. The first processing stage, which is applied on the image resulting from the normalization and fusion of the multiple image channels, yields a score-map based on valid region annotation, by taking advantage of thresholding with multiple values and then evaluating the connected-components by means of particle filtering. The second stage consists in labeling the connected-components in the thresholded score-map and combining the distance transform and the watershed algorithm for the segmentation [387, 431]. Sometimes, cell counting approaches require to use specific hardware (designed *ad hoc* or commercially available) for the acquisition of the images to be processed. Barber *et al.* [38] developed an automated colony counter composed of a hardware device and a software for image processing. The system was tested with four different cell types and the achieved results were compared against the corresponding manual counting measurements from a staff composed of four experts. The specific hardware enabled high-quality image acquisition of the culture flasks. As a matter of fact, a monochrome Charge-Coupled Device (CCD) camera with a wide-angle lens ( $f = 2.8$  mm) was used, defining a field of view of  $75 \times 56$  mm at the flask bottom. In the acquisition phase, the average of 10 images was utilized to deal with random noise. However, a geometric deformation was employed for coping with the barrel distortion due to the wide-angle lens. Two different approaches based on classic image processing techniques were used to identify cell colonies: the first method is tailored to highly-contrasted cell images; the second one deals with irregular-shaped colonies with fuzzy boundaries. However, the most suitable between these two image processing methods is selected by the user according to the cell type. In [49], a model-based image segmentation method is introduced, by relying on prior knowledge concerning the colony. This strategy addressed the issue related to recognizing

isolated, touching and overlapping cell colonies. As first step, a statistical method proposes a set of candidate Gaussian model instances, by estimating the model parameters and the discrepancy between an instance and the current image. Afterwards, the MDL principle selects the model instances that better match the distinct colonies in the cell image. This method was tested on images of Chinese hamster lung fibroblast cell lines. The hardware apparatus included a customized homogeneous diffuse lighting module that provides a Petri dish holder. A CCD monochrome camera with C-mount lens ( $f = 50$  mm) was placed over the dish and connected to a Personal Computer. The acquired 8-bit images had a size of  $512 \times 512$  pixels. However, the computational method could not consider colonies lying on the Petri dish rim, since this border region is excluded *a priori*. In this case, a masking operation for each image is sufficient, by exploiting the fixed position and dimension of the imaged Petri dishes.

### Fluorescence microscopy

Thanks to the advances in microscopy imaging technologies, it is possible to visualize living specimens' dynamic processes by time-lapse microscopy images [306]. Typical application examples, such as high-throughput/high-content phenotyping and atlas building for model organisms, reveal the importance of bioimage informatics [461, 462]. However, these compelling challenges require specialized image analysis approaches. In addition, the imaging data captured during even a single experiment may consist of hundreds of objects over thousands of images, making manual inspection a tedious and rather unfeasible task [306]. In practice, usually a single biological image stack has a large size (hundreds of MB or even GB) over multiple channels [208]. The objects of interest in cellular images might be characterized by high variations of morphology and intensity from image to image [462]. The high dimensionality and heterogeneity of the cell data acquired by means of the modern imaging technologies play a key role in the understanding of complex biological processes. With more detail, cellular morphology is an important phenotypic feature that is indicative of the physiological state of a cell. Therefore, the problem of cell segmentation has gained increasing attention during the past years since accurate cell boundaries are often required for subsequent analysis of intra-cellular processes and cell interactions. These aspects give rise to new challenges with respect to existing image processing techniques, due to the high complexity and information content in bioimages.

Experiments where live cell specimens are imaged over prolonged time periods can potentially help us to understand how cells behave and respond to changes in the local environment. High-content time-lapse imaging monitors the responses in living cells over a specific period of time exploiting the activity of fluorescently labeled and unlabeled cells.

These technologies have enabled large-scale imaging experiments for studying cell biology and for drug screening: indeed, the modern systems can acquire hundreds of thousands of microscopy images per-day, however their utility is hindered by inefficient image analysis methods. Modern imaging experiments, with ever increasing scale, point out the limits of manual/visual image processing and analysis [535], requiring new effective automated methods. One of the main problems in many automated analysis approaches, especially those following individual cells over time, is the image segmentation. Unfortunately, automated segmentation of unstained cells imaged by bright-field and fluorescence microscopy is typically very challenging [402]. In practice, automated analysis is highly desirable due to manual analysis being subjective, biased and extremely time-consuming for large-scale datasets [222].

**Related work** This section briefly outlines and reviews the state-of-the-art of cell fluorescence microscopy image segmentation. As a matter of fact, several tools exist for automated bioimage analysis and processing [170].

The most common free and open-source software tools for microscopy applications in laboratory routine are: (i) ImageJ [552] or Fiji (Fiji Is Just ImageJ) [550] (National Institutes of Health, Bethesda, MD, USA), and (ii) CellProfiler [83] (Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA, USA). Although these tools offer customization capabilities, they do not provide suitable functionalities for fast and efficient high-throughput cell image analysis. Moreover, CellAnimation [208] is similar to CellProfiler but provides a modular architecture as well as high-throughput options. CellProfiler Analyst [140] was designed for the exploration and visualization of image-based data, along with the classification of complex biological phenotypes, by means of an interactive graphical user interface. This tool implements classic supervised Machine Learning (e.g., RFs, SVMs).

Mathematical morphology has been extensively used in cell imaging [402]. Wählby *et al.* [645] presented a region-based segmentation approach in which both foreground and background seeds are selected by means of morphological filtering—based on  $h$ -maxima and  $h$ -minima, respectively—applied on the original image and the gradient magnitude of the image, respectively. These seeds are then used as starting points for the watershed segmentation of the gradient magnitude image. Since more than one seed could be assigned to a single object, initial over-segmentation might occur, i.e., a boundary is created also in the case of weak edges. Finally, a merging procedure is performed on neighboring objects according to gradient magnitude. In [304], the authors developed an automated approach based on the Voronoi tessellation [271] built from the centers of mass of the cell nuclei, to estimate morphological features of epithelial cells. The full procedure is optimized for the

imaging of cell mono-layers, where nuclei of cells do not overlap. The authors leveraged the watershed algorithm in order to correctly segment cell nuclei. In this case, the Voronoi tessellation is well-suited for measuring morphological properties of cells (such as area, perimeter, elongation) in a tissue composed of adjacent cells.

Machine Learning techniques have been applied to bioimage analysis [21]. CellCognition aims at annotating complex cellular dynamics in live-cell microscopic movies [264]. This computational framework combines Machine Learning methods for supervised classification and Hidden Markov Models (HMMs)—exploiting the Viterbi algorithm [641] to incorporate the time information—in order to evaluate the progression by means of morphologically distinct biological states.

More recently, it has been shown that methods based on CNNs [110] can successfully address segmentation problems that are difficult to solve exploiting traditional image processing methods [535]. For instance, Deep Learning has been applied in biological image analysis, with particular interest to yeast images [336] and also digital histopathology [254]. Recently, an ensemble of multiple CNNs was used for segmentation of cell images [268]. A gating network automatically divides the input image into several sub-problems and assigns them to specialized networks, allowing for a more efficient learning with respect to a single CNN. This semantic segmentation approach was assessed on the segmentation problem of cell membrane and nucleus. Abousamra *et al.* [2] proposed a CNN-based architecture for localization, classification, and tracking in 4D fluorescence microscopy imaging. In particular, this CNN-based approach was tested in the case of microtubule fibers' bridge formation during the cell division of zebrafish embryos.

Traditional image segmentation approaches often require experiment-specific parameter tuning, whilst CNNs require a huge amount of training data. Large amounts of high-quality annotated samples or ground truth is typically required for the CNN training. The ground truth, representing the extent to which an object is actually present, is usually delineated by a domain expert. Obviously, visual and manual annotation is tedious, cumbersome and time-consuming. The annotated samples for one dataset may not be useful for another dataset, so new ground truth generation may be needed for every new dataset, limiting the effectiveness of CNNs. The authors of [535] proposed an approach for automatically creating high-quality experiment-specific ground truth for segmentation of bright-field images of cultured cells based on end-point fluorescent staining (for nucleic and cytoplasmic regions). This gold standard was obtained by means of standard automated methods available in CellProfiler [83], exploiting the information conveyed by the fluorescent channels. Finally, these automatically created segmentation results were exploited as labels along with the corresponding bright-field image data to train a CNN [335]. In general, applying CNNs to

microscopy images remains still challenging due to the lack of large datasets labeled at the single cell level [335]. Thus, unsupervised image analysis techniques that do not require a training phase represent valuable solutions in this context [264].



# Chapter 3

## Classical Image Processing algorithms

### 3.1 Region-based approaches

#### 3.1.1 Image thresholding

##### Global image thresholding

The simplest Image Processing technique for automatic image segmentation is global thresholding, which generally consists in classifying the pixels according to fixed criteria, usually specified as ranges of intensities [504]. In particular, binarization is a segmentation approach that partitions the input image into two classes by considering a certain intensity threshold value  $\theta$ . Despite its simplicity, this strategy provides an efficient and effective segmentation technique, according to the different intensities in the foreground and background regions of an image. The threshold value  $\theta$  must be carefully chosen, considering the features of the image underlying the pixel intensity values. Consequently, given an image  $\mathcal{I}$  consisting of  $M \times N$  pixels, this intensity threshold defines two different classes, by dividing the histogram of the gray levels  $\mathcal{H}$  into two parts, namely  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , according to the threshold intensity value  $\theta$ . The pixels in the image  $\mathcal{I}$  are partitioned into the two sub-regions  $\mathcal{R}_1 = \{\mathcal{I}(x,y) : \mathcal{I}(x,y) > \theta\}$  and  $\mathcal{R}_2 = \{\mathcal{I}(x,y) : \mathcal{I}(x,y) \leq \theta\}$ , for every  $x = 1, \dots, M$  and  $y = 1, \dots, N$ .

Several literature methods have been proposed to implement adaptive thresholding methods, able to automatically select a proper value for each analyzed image. The most widespread algorithms for dynamic thresholding are: the Iterative Optimal Threshold Selection (IOTS) [500], the method proposed by Otsu [448], and the Minimum-Error Thresholding (MET) method proposed by Kittler and Illingworth [325] later extended by Ye and Danielsson [684]. All these approaches are closely related and strongly rely on images characterized by bimodal

histograms (see [677]). In the case of two-class image segmentation, the two sub-histograms (i.e., associated to the foreground and background pixels, respectively), assumed to be nearly Gaussian distributions, should be characterized by approximately equal size and variance [340]. When these assumptions are not satisfied—i.e., the pixel intensity distribution is not approximately bimodal—the aforementioned algorithms show some limitations. As a matter of fact, the optimal threshold  $\theta_{\text{opt}}$ —especially in the case of the Otsu’s method—either over- or under-estimates the ROI, since the computed threshold tends to split the class with larger size and to bias towards the class with larger variance. Under these conditions, the IOTS method [500, 619] could provide better results than the Otsu’s method [448] when the sizes of the two classes are highly different [674]. In addition, Medina-Carnicer *et al.* in [399] showed that the above mentioned algorithms often perform poorly with unimodal distributions of gray levels. Moreover, in the case of images affected by intensity overlap, the IOTS algorithm is less likely to either over- or under-estimate the threshold, when compared to other techniques selecting a threshold in the valley between two peaks of the histogram, even if the histogram is not strongly bimodal [577], in particular when applied to medical images [342].

Algorithm 1 shows the pseudo-code of the IOTS algorithm [500], by using an efficient implementation that computes the image histogram only once [619]. Therefore, a single histogram (i.e., the initial image histogram) is processed during the iterations until convergence. This iterative approach aims at dividing the whole histogram  $\mathcal{H} = h(r)$  (with  $r \in [0, 1, \dots, L - 1]$ , where  $L$  is the maximum gray level) into two classes  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , separated by an adaptive global threshold  $\theta_{\text{opt}}$ .

**Uterine fibroid segmentation on MR images** In [415], uterine fibroids were segmented using the efficient version of the IOTS algorithm [500, 619] described in Algorithm 1. After the uterus segmentation based on the FCM clustering algorithm presented in [415], the particular characteristics of the region were identified using an adaptive thresholding applied on gray levels, because fibroids have intensity values much lower than the untreated uterus pixels. Although IOTS works generally well on images characterized by a basically bimodal histogram, some post-processing and mathematical morphological refinements are applied on the detected fibroid ROT to improve segmentation reliability. The overall processing pipeline is presented in Fig. 3.1. Fig. 3.2 shows three segmentation examples of the uterus ROI (blue contour) and of the fibroid ROT (red contour).

The experimental fibroid MRI dataset as well as the achieved segmentation evaluation metrics are reported in Section 3.1.3, wherein a quantitative comparison against the method in [524] is shown (according to the evaluation metrics provided in Appendix A.1).



---

**Algorithm 1** Pseudo-code of the iterative optimal threshold selection algorithm.

---

**Input:** Image  $\mathcal{I}$

**Output:** Optimal threshold  $\theta_{\text{opt}}$

- 1: Calculation of the histogram  $\mathcal{H} = h(r)$ , with  $r \in [0, 1, \dots, L-1]$ 
    - ▷ Initial estimated value for the threshold  $\theta^{(t)}$  (for example, the average intensity of  $\mathcal{I}$ )
  - 2:  $\theta^{(0)} \leftarrow \mu_{\text{global}} = \frac{\sum_{r=0}^{L-1} h(r) \cdot r}{\sum_{r=0}^{L-1} h(r)}$
  - 3: **while**  $|\theta^{(t+1)} - \theta^{(t)}| \leq \varepsilon_{\text{tol}}$  **do** ▷ Convergence condition achievement
  - 4:     The histogram  $\mathcal{H}$  is divided into two classes  $\mathcal{H}_1$  and  $\mathcal{H}_2$  separated by  $\theta^{(t)}$ 
    - ▷ Average intensity values  $\mu_1$  and  $\mu_2$  for the sub-histograms  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively
  - 5:      $\mu_1^{(t)} = \frac{\sum_{r=0}^{\theta^{(t)}} h(r) \cdot r}{\sum_{r=0}^{\theta^{(t)}} h(r)}$ ;  $\mu_2^{(t)} = \frac{\sum_{r=\theta^{(t)}+1}^{L-1} h(r) \cdot r}{\sum_{r=\theta^{(t)}+1}^{L-1} h(r)}$
  - 6:      $\theta^{(t+1)} \leftarrow \frac{\mu_1^{(t)} + \mu_2^{(t)}}{2}$  ▷ Update the current threshold value
  - 7: **end while**
  - 8:  $\theta_{\text{opt}} \leftarrow \theta^{(t+1)}$
- 

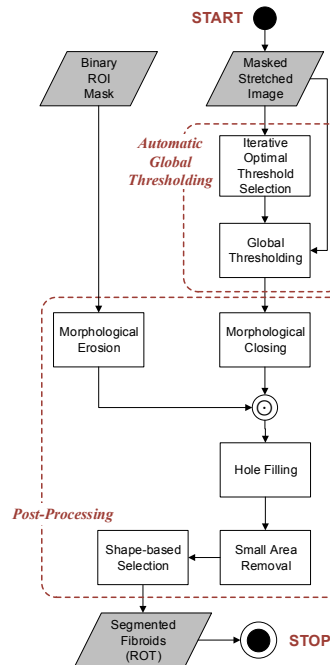


Fig. 3.1 Flow diagram representing the steps of the fibroid segmentation method based on the IOTS algorithm.

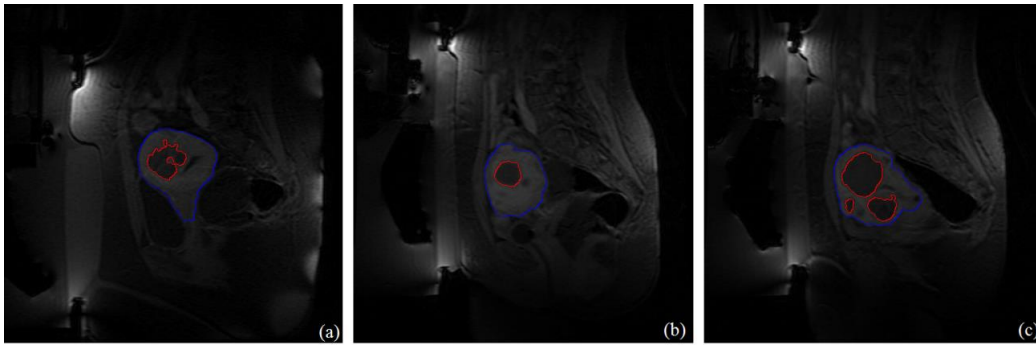


Fig. 3.2 Uterine fibroid segmentation results achieved by the IOTS-based method. The blue contour represents the uterus, while the red contour represents the ROT fibroids.

### Local image thresholding

Local adaptive thresholding techniques estimate the threshold locally over sub-regions of the entire image, by considering only a neighborhood (i.e., sliding window) with a specified size and exploiting local image properties to calculate a variable threshold [218].

These local approaches can be classified into two strategies. The first one refers to the approach in [107], in which the entire image is divided into an array of overlapping sub-images: the optimal threshold is searched for each sub-image by investigating its histogram. The threshold for each single pixel is found by interpolating the results of the sub-images. The main drawback of this method is the computational cost, making it often not suitable for real-time applications. The second class, also called local adaptive thresholding, finds the threshold by locally examining the intensity values of the neighborhood of each pixel according to image intensity statistics. The most appropriate first-order statistics depends strongly on the analyzed input image [546]. Mean and median of the local intensity distribution are the most used measures of central tendency. The size of the examined neighborhood has to be large enough to cover sufficient foreground and background pixels, so defining a significant statistical sample. On the other hand, choosing too large regions could violate the assumption of approximately uniform illumination. This method is less computationally expensive than the approach in [107] and yields good segmentation results in several practical applications, such as in document binarization [41, 546], cell image analysis [282, 703], and medical image analysis [331]. The assumption underlying these methods is that smaller image regions are more likely to show approximately uniform illumination, thus being more suitable for thresholding.

In our implementation, to calculate the local threshold  $\theta_i$ , we used local adaptive thresholding considering the mean value in the neighborhood as first-order statistics. In particular, let  $\mathcal{N}_i$  be the  $i$ -th neighborhood in which the input  $\mathcal{I}$  (with size  $M \times N$ ) was divided, the

corresponding local threshold  $\theta_i$  is computed for each pixel according to Eq. (3.1), where the index  $i = 1, 2, \dots, M \times N$  scans the vectorization of the image matrix  $\mathcal{I}$ :

$$\theta_i = \text{mean} \left\{ \sum_{(x,y) \in \mathcal{N}_i} \mathcal{I}(x,y) \right\}, \text{ with } i = 1, 2, \dots, M \times N. \quad (3.1)$$

The size of each neighborhood is  $\text{size}_{\mathcal{N}}$ , depending only on the size of the well sub-image  $\mathcal{I}$ , is a pair of positive odd integers according to Eq. (3.2):

$$\text{size}_{\mathcal{N}} = \left( 2 \times \left\lfloor \frac{M}{16} \right\rfloor + 1, 2 \times \left\lfloor \frac{N}{16} \right\rfloor + 1 \right). \quad (3.2)$$

**Area-based cell colony surviving fraction evaluation** This section describes a novel computational method for automated SF evaluation, based on ACC alone. Fig. 3.3 shows the overall flow diagram with the processing steps performed in the pipeline that, starting from the acquired plate image, automatically detects the wells, extracts the colonies, calculates the ACC, making it possible to measure the SF. The proposed method was entirely developed using the MatLab<sup>®</sup> environment (The MathWorks, Natick, MA, USA).

The proposed approach exploits the Circle Hough Transform (CHT) and local adaptive thresholding algorithms: these two effective and efficient techniques are combined in an original pipeline, not yet considered in the literature, enabling to address several open issues in clonogenic assay SF evaluation.

The proposed fully automatic approach presents several key novelties and contributions with respect to related literature works:

- independence on the colony number, which in some cases may lead to inconsistencies between the computed SF values and the real cell growth. As a matter of fact, plates with the same number of colonies but different ACCs, due to different growth rate, are indicative of different treatment responses. The traditional approach, based on the number of colonies alone, is not able to differentiate among these cases;
- evaluation of the SF based on the percentage of the ACC, which allows us to establish a correlation between the growth rate and the dose administered during the treatment;
- unlike some of the literature works that use specific cameras or equipment (sometimes expensive), our approach uses general-purpose acquisition hardware (a conventional flatbed scanner);
- no calibration and parameter-settings are required, differently to [206];

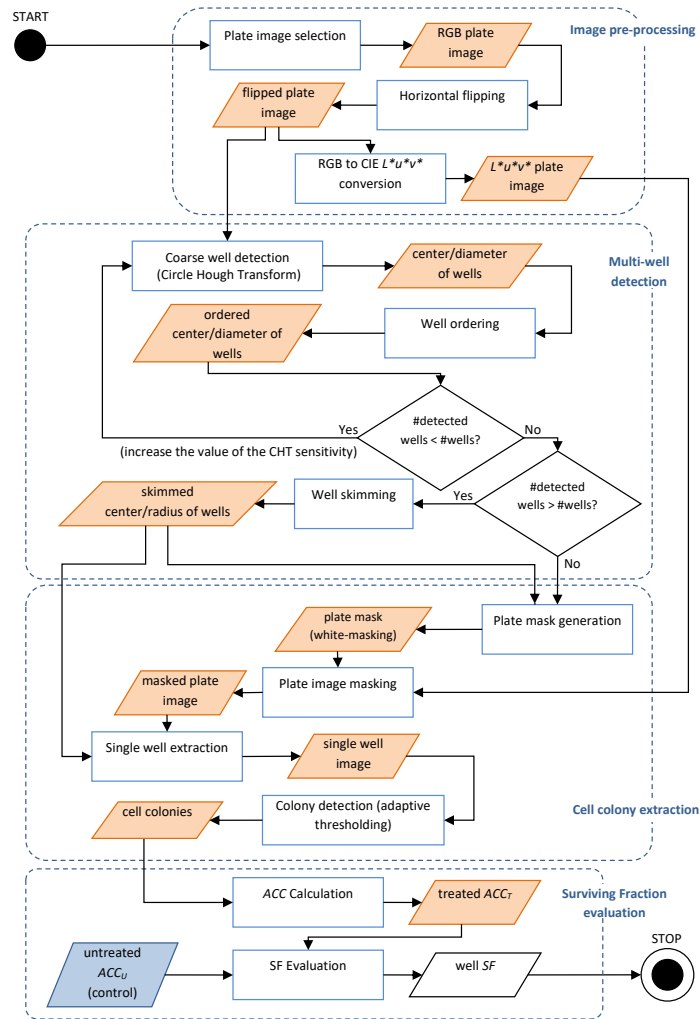


Fig. 3.3 Overall flow diagram of the processing pipeline of the proposed fully automatic approach for ACC. Four main stages can be identified: (i) image-preprocessing, basically dealing with the RGB to CIE  $L^*u^*v^*$  conversion; (ii) multi-well detection, which identifies and selects the wells inside the plate image; (iii) colony extraction, which segments cell colonies from the well background; (iv) SF evaluation, which estimates the SF of the current well according to the treated  $ACC_T$  against the reference untreated  $ACC_U$  (control).

- only a pre-processing step, performing a Red Green Blue (RGB) to International Commission on Illumination (CIE)  $L^*u^*v^*$  color space conversion, is applied to reduce luminance inhomogeneities (decoupling the intensity and chromaticity components), thus not affecting the original pictorial data [542];
- automated wells detection inside the plate image by means of the CHT [158, 690], unlike several literature approaches [49, 138, 235] that require the plate dimensions and the spatial arrangement of the wells;
- colony segmentation exploiting just a local adaptive thresholding algorithm [218, 291], which deals with both small and large colonies;
- no “aggressive” post-processing operations are needed (e.g., morphological refinement could have stronger effects when performed on small colonies and, consequently, can lead to significant errors in the ACC calculation).

**Materials** In the following, the characteristics of the analyzed cell cultures as well as the procedure for multi-well plate image acquisition are reported.

This study analyzed MCF7 and MCF10A cell lines: MCF7 is a human breast epithelial carcinoma estrogen receptor-positive cell line, while MCF10A is a human non-tumorigenic breast epithelial cell line. At the end of a 12-15 day incubation period, the cultures were fixed and stained with crystal violet according to the protocol described in [191] and were analyzed manually by the experienced staff at the Cellular and Genomic Methodologies Laboratory of the Institute of Molecular Bioimaging and Physiology (IBFM-CNR) in Cefalù (PA), Italy. The obtained SF values, averaged over three counting measurements for each well, performed by a staff composed of three different expert biologists, represent our ground-truth used as a reference for the fully automated approach.

MCF7 and MCF10A cellular lines were placed on 6-well plates (Corning<sup>®</sup> Inc., Corning, NY, USA). Several culture plates were prepared, considering different configurations of either irradiation or cytotoxic agent. The proposed approach is not just limited to 6-well plates, since it is able to automatically detect the wells, regardless of the well spatial arrangement and location in the plate (only the well number and well diameter are needed). Table 3.1 depicts the main characteristics of the different plate types employed in this study. The plate images were acquired using the Epson Expression 10000XL (Seiko Epson Corporation, Nagano, Japan), which is a conventional flatbed scanner, with a spatial resolution of 800 dots-per-inch (dpi) and 24 bits (8 bits per-color channel) as bit-depth. The scanned images were saved in the uncompressed TIFF format. The used resolution represents a good compromise between

Table 3.1 Some characteristics of the multi-well plates employed in this study. Differently to literature approaches that use static techniques (i.e., plate measurements and well spatial arrangement) for well detection, the proposed approach, based on the CHT, requires only the well number and diameter.

Plate type	Number of wells	Well diameter [ms]	Spatial arrangement of wells [pixels]	Plate size [mm <sup>2</sup> ]
Corning <sup>®</sup> 6 wells	6	34.80	2 × 3	127.76 × 85.47
Corning <sup>®</sup> 12 wells	12	22.11	3 × 4	127.89 × 85.60
Corning <sup>®</sup> 24 wells	24	15.62	4 × 6	127.89 × 85.60
Corning <sup>®</sup> 48 wells	48	11.05	6 × 8	127.89 × 85.60

image quality (lower resolutions could lead to the elimination of small colonies) and the required processing-time.

**Image pre-processing** After the multi-well plate image acquisition phase, some image artifacts and inhomogeneities could affect the accuracy achieved in the subsequent processing steps. Thus, some pre-processing steps are performed.

Firstly, the image is horizontally flipped to correct the specular reflection that occurs in the acquisition phase, allowing for the correct correspondence between the well number and its real position. However, due to small parallax errors, a circular dark shadow on the rim of each well was observed. In order to reduce these shadowing artifacts and make the illumination more uniform, a conversion of the color space from RGB to CIE  $L^*u^*v^*$  is performed [542]. As a matter of fact, the  $L^*u^*v^*$  is an excellent intensity (represented by lightness  $L^*$ ) and chromaticity (denoted by  $u^*$  and  $v^*$  components) decoupler [218]. The lightness  $L^*$  component, which conveys greater information related to intensity, is not considered in the subsequent processing steps for colony cell segmentation. By exploiting this decoupling between intensity and chromatic components, we are able to deal better with possible brightness non-uniformities inside the wells.

**Multi-well detection based on circle Hough transform** The wells inside the plate are detected automatically applying the HT on the horizontally flipped RGB plate image. The HT is an efficient feature extraction technique for detecting objects with a specific shape [276], which brings the problem of curve detection in a simpler search procedure for maxima in an accumulator array [76]. Especially, the CHT was used since it is specialized to find circle-shaped objects [158, 683, 690]. A plane curve  $\mathcal{F}(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is defined analytically using a set of parameters. Eq. (3.3) maps the parameters to Cartesian coordinates in implicit form:

$$\mathcal{F}((x, y); (a_1, a_2, \dots, a_n)) = 0, \quad (3.3)$$

where  $(x, y)$  is a point of the two-dimensional curve, and  $(a_1, a_2, \dots, a_n)$  is an  $n$ -ple of values that defines a point in the parameter space (i.e., a point in the space of parameters uniquely identifies an analytical curve). Each point in the image space corresponds to a hypersurface in the parameter space and  $n$  points in the image space, belonging to the same curve, generate  $n$  surfaces that intersect at the same point in the parameter space. An intersection of many surfaces in the parameter space is the clue to the presence of a particular instance of the analytical curve to be found. In our case, the purpose is to detect circles, which represent the wells in the plate image. The CHT is able to find circles of known radius as well as circles having any radius. Since the well diameter  $r_w$  is known, we are interested in the first case, in which the parameter space is two-dimensional and the curve generated from each point  $(x, y)$  in the image space is a circle itself, according to Eq. (3.4). The parameter space is generated, considering each pixel  $(x, y)$ , by the center coordinates  $\mathbf{c} \equiv (x_c, y_c)$ :

$$(x - x_c)^2 + (y - y_c)^2 = r_w^2. \quad (3.4)$$

Generally, a number of points at least equal to the number of parameters has to be considered to define a curve (i.e., its features). Therefore, the CHT allows to convert a searching problem for curves in a simpler surface intersection search by means of a so-called accumulator matrix [287]. The circle candidates are produced by voting in the parameter space center points  $\mathbf{c}$ , and then the local maxima are selected in the accumulator matrix. Formally, each point  $(x_c, y_c)$  in the parameter space, representing a circle center, must meet the following conditions in polar representation:

$$\begin{cases} x_c = x - r_w \cdot \cos \theta \\ y_c = y - r_w \cdot \sin \theta \end{cases}, \quad (3.5)$$

where  $r_w$  is the radius of the well and  $\theta \in [0, 2\pi]$  is the angular coordinate with respect to the positive  $x$ -axis, by using a complex phase coding scheme [28]. The pseudo-code of the implemented robust well detection strategy based on the CHT technique is outlined in Algorithm 2.

In all the analyzed plates, the well detection using the CHT is performed robustly without errors directly in the original RGB color space. However, in order to make the procedure more robust, in case of incorrect well identification, a more selective control strategy was implemented. In particular, if the initial CHT identified less than the expected number of wells (i.e., 6, 12, 24 or 48), then the CHT is repeated increasing the sensitivity factor for the accumulator matrix (from 0.98 to 0.99), including also weak and partially occluded or overlapping circles (e.g., possible shadowing artifacts in well edges or rims). Otherwise, if the

---

**Algorithm 2** Pseudo-code of the well detection strategy based on the CHT technique.

---

**Input:** plate image  $\mathcal{I}_{\text{plate}}$ , radius of each well  $r_w$ , total number of wells  $N_{\text{wells}}$

**Output:**  $N_{\text{wells}}$  circles identified by the center coordinates  $\mathbf{c} \equiv (x_c, y_c)$

```

1:  $\mathbf{A}(\theta, r_w) \leftarrow \mathbf{0}$  ▷ Define and initialize the accumulator matrix
2:  $N_{\text{DetWells}} \leftarrow 0$  ▷ Number of detected wells
3: for each center-candidate point  $\mathbf{c} \equiv (x_c, y_c) \in \mathcal{I}_{\text{plate}}$  do
4:   Increase all the corresponding points in  $\mathbf{A}$  ▷ Local maxima in the accumulator
   correspond to circular objects
5:   if  $N_{\text{DetWells}} < N_{\text{wells}}$  then
6:     Increase the CHT sensitivity value
7:     continue
8:   end if
9:   if  $N_{\text{DetWells}} > N_{\text{wells}}$  then ▷ Well-skimming procedure (according to CHT metrics)
10:    Select the first  $N_{\text{wells}}$  detected circles sorted by circle strength value
11:   end if
12: end for

```

---

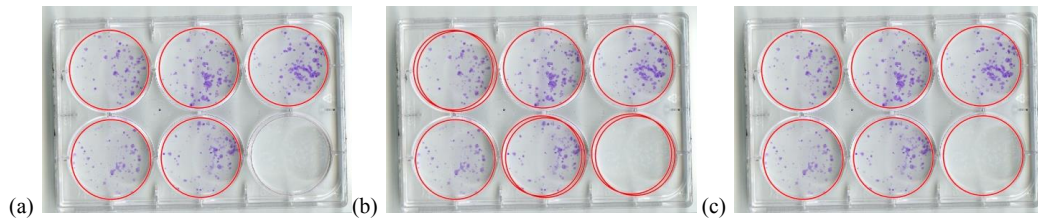


Fig. 3.4 Some examples of the control strategy used to enhance the automatic well detection performed using the CHT in the case of 6-well plates: (a) the initial CHT detects only 5 wells; (b) after the sensitivity increasing, the CHT detects 9 wells; (c) in order to skim the correct wells, only the first 6 detected wells are selected after well ordering according to the CHT metrics (i.e., circle strength value).

identified wells are less than the expected number, then a skimming procedure is performed considering the value of the metrics that the CHT computes for each found circular object. We noticed remarkable differences between the metrics values of the correct wells and those of incorrectly identified wells. For this reason, we decided to select exactly the number of expected wells reporting the highest CHT metrics (i.e., the first 6, 12, 24 or 48 wells in the candidate list sorted by the circle strength value). Fig. 3.4 shows a particular case, in which the initial CHT detects only 5 wells (Fig. 3.4a). The following CHT iteration with increased sensitivity detects 9 wells (Fig. 3.4b). The final skimming phase extracts the correct wells (Fig. 3.4c). Overall, the CHT enhanced with this control strategy was shown to be robust and reliable, allowing for an accurate well localization in 100% of the examined plates.



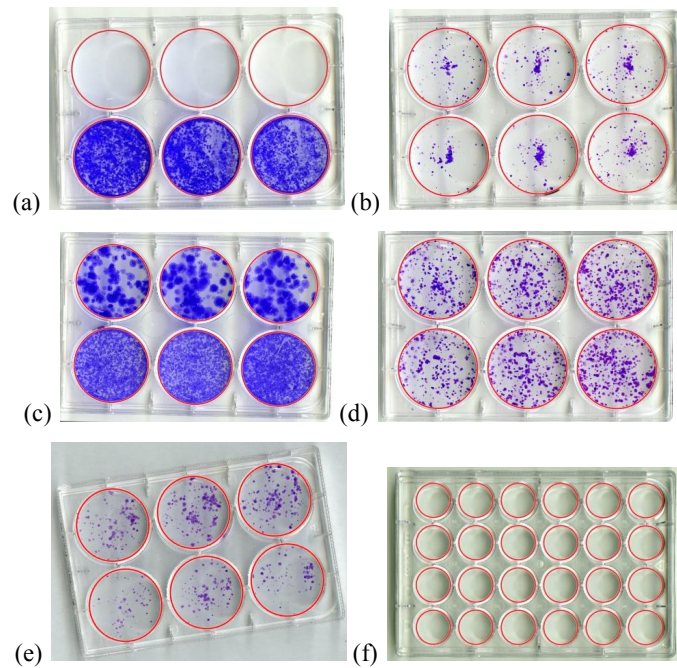


Fig. 3.5 The proposed approach correctly detects wells in typical operating conditions (a-d), as well as when the acquired plate image is not aligned (e), and (f) with different plate types (e.g., 24-well plate).

Unlike some literature works [138, 235] that identify the wells statically, by considering either the plate characteristics (e.g., well radius, horizontal and vertical distance between wells, thickness of well wall, number of wells per-row, number of well per column) or by using a fixed position mask [49], our method is fully automatic. An approach based on specific sizes can lead to errors in case of misalignment between the plate and the reference position. Differently, the well detection implemented in our method requires only the number of wells in the plate and the well radius regardless the type and position of the plate. The examples reported in Fig. 3.5 show the robustness of the well detection phase to rotation and translation. In Figs. 3.5(a, b, c, d), some well detection results in four different typical real conditions are depicted. Moreover, as depicted in Fig. 3.5e, it is possible to identify automatically the wells even if the image is not perfectly aligned horizontally (due to arbitrary misalignments in the acquisition phase) and regardless the well spatial arrangement (2 rows and 3 columns in 6-well plates). Fig. 3.5f shows an example of detection on a 24-well plate. After the well detection, an ordering step is performed, to establish a unique correspondence between the center coordinates of the found circles and their own position in the plate.

**Cell colony extraction using local adaptive thresholding** Cell colony segmentation is performed using a local adaptive thresholding algorithm. Both global and local thresholding techniques aim to separate desirable foreground objects from the background in an image, considering differences in pixel intensities. Global thresholding determines a single threshold for all pixels and works only if the histogram of the input image contains well-separated peaks corresponding to the desired subject and background. Consequently, such an approach cannot deal with images containing a strong illumination gradient [291]. As a matter of fact, robust and accurate image segmentation methods represent an active research topic in image processing and graphics [442].

Nevertheless, the RGB to CIE  $L^*u^*v^*$  conversion allows us to reduce most illumination issues, in some areas remain some non-uniformities that adaptive thresholding can easily handle, making it possible to handle changes in lighting conditions in the image.

Several literature approaches [103, 115, 138, 206, 235] use global thresholding for colony extraction. Although global adaptive thresholding algorithms, such as [325, 448, 500], generally yield excellent segmentation results, they could fail under certain conditions. For instance, when the colonies cover the whole neighborhood, the adaptive thresholding could not work properly. As a matter of fact, if the analyzed local neighborhood contains only cells, the local adaptive thresholding algorithm is not able to estimate a valid local threshold  $\theta_i$  and, consequently, this area will be discarded. These conditions occur if a huge number of cells are plated, or if cells have a high growth rate to cover the whole well (or at least the currently examined neighborhood). To overcome this issue, which may occur mainly in areas near the edge of the plate, instead of a typical black-mask we adopted a white-mask, obtained automatically by considering centers and diameters of the circles identified in the previous well detection phase, to mask the wells in the plate image. Using this simple solution, the adaptive thresholding works correctly, so considerably enhancing colony extraction.

Fig. 3.6 shows two series of results: in the first two rows a well with 500 plated MCF7 cells, and in the last two rows a well with 150 plated MCF10 cells. It is appreciable how the used white-masking operation (Figs. 3.6(c1, c2, f1, f2)) enhances the output with respect to adaptive thresholding after typical black-masking (Figs. 3.6(b1, b2, e1, e2)). After colony detection, the only post-processing step is represented by a hole filling algorithm [571], aimed to include the central region of the colonies that, appearing generally darker, can be excluded from the detected regions.

**Results** The SF values obtained by considering the percentage of ACC were compared against the conventional measurement based on colony counting alone, performed manually by three different expert biologist operators (see Appendix A.3 for a detailed explanation).

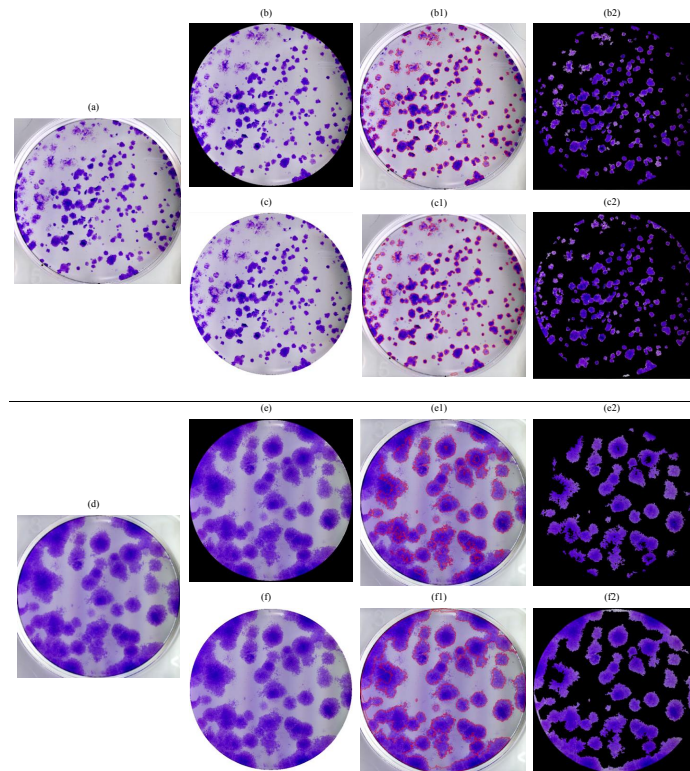


Fig. 3.6 Examples of colony extraction performed using the proposed approach based on local adaptive thresholding using white-masking compared against the same local adaptive thresholding with previous black-masking. Two wells with 500 MCF7 and 150 MCF10 plated cells are shown in the first two rows and in the last two rows, respectively: (a, d) input wells before colony detection; (b1, e1) and (b2, e2) colonies detected using local adaptive thresholding and black-masking (image with superimposed red contour and masked with the achieved segmentation, respectively); (c1, f1) and (c2, f2) colonies detected using local adaptive thresholding and white-masking (image with superimposed red contour and masked with the achieved segmentation, respectively). The segmentation combining adaptive thresholding and white-masking enhances colony extraction close to the well rim, where cell growth created a uniform area, which classic adaptive thresholding with black-masking is not able to segment properly.

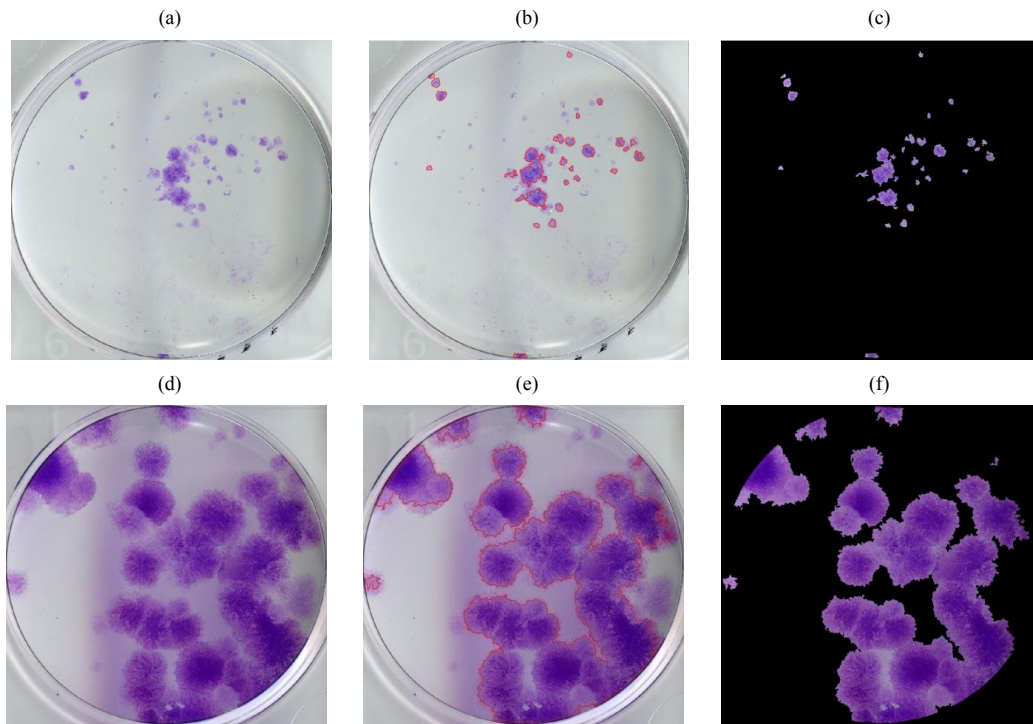


Fig. 3.7 Examples of cell colony detection in different real scenarios. The first row concerns MCF7 cell colonies with very small size: (a) original well; (b) and (c) show the detected colonies with red contour and background masking, respectively. The second row concerns MCF10A cell colonies characterized by a high growth rate: (d) original well; (e) and (f) show the detected colonies with red contour and background masking, respectively.

The performed experimental tests aim to demonstrate the validity of the proposed approach and the accuracy of the obtained results, in terms of correlation between the SF calculated considering manually counted cell colonies ( $SF_{\text{Manual}}$ ) and the SF based on ACC computed using our automated method ( $SF_{\%Area}$ ). As shown in Fig. 3.7, the proposed approach allows for the correct detection of the colonies in a wide variety of scenarios, including cells characterized by a high growth rate, which develop in very large colonies with significantly variable shape (and that merging together they would create problems to conventional approaches based on counting the number of colonies), as well as with cells that form very small colonies.

Fig. 3.8 shows the scatter diagrams concerning the automated SF evaluation based on the percentage of ACC ( $SF_{\%Area}$ ) versus the manual colony counts ( $SF_{\text{Manual}}$ ) averaged over three measurements by three different biologists. The achieved results ( $\rho = 0.9791$  on MCF7 cells and  $\rho = 0.9682$  on MCF10A cells obtained on 79 and 40 samples (i.e., different wells), respectively) showed a high level of correlation between our fully automatic method and

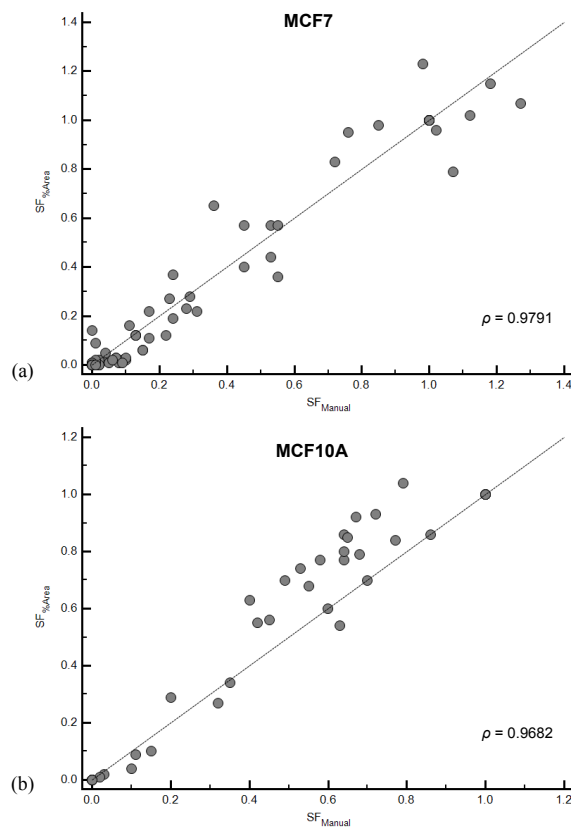


Fig. 3.8 Scatter diagrams, showing the correlation of the automated ACC-based SF evaluation ( $SF_{\%Area}$ ) versus mean manual colony counting measurements ( $SF_{Manual}$ ), for: (a) MCF7 and (b) MCF10A cell lines. The values were obtained on MCF7 and MCF10 cell cultures (79 and 40 samples, respectively) treated with different configurations of radiation or cytotoxic agent. The equality line through the origin is drawn as a dashed line and the correlation coefficients are shown in the bottom right-hand corner of each diagram.

Table 3.2 Pearson's linear correlation coefficient calculated between the automatic ACC-based SF values ( $SF_{\%Area}$ ) and the conventional colony counting measurements ( $SF_{Manual}$ ) on the analyzed MCF7 and MCF10A cell lines (79 and 40 samples, respectively). The corresponding significance level ( $p$ -value) and 95% confidence interval are also reported.

Cell culture	Number of samples	Pearson's correlation coefficient	Significance level ( $p$ -value)	95% confidence interval
MCF7	79	$\rho = 0.9791$	$p < 0.0001$	0.9674 to 0.9866
MCF10A	40	$\rho = 0.9682$	$p < 0.0001$	0.9403 to 0.9832

those conventionally obtained with manual cell colony counting. The achieved high Pearson's coefficient  $\rho$  values prove the accuracy of the proposed approach, so carrying out a validation against the conventional counting procedure by three experienced biologists. More details about the Pearson's linear correlation coefficient calculation are depicted in Table 3.2.

A strong positive correlation between the SF computed automatically by the proposed method and the conventional manual SF measurements was observed for the two cell lines analyzed in this study. Automated SF measurements are highly consistent with the corresponding manual evaluations for the MCF7 cell samples, by denoting also a unity slope, as shown in Fig. 3.8a. In the case of MCF10A cell line (Fig. 3.8b), the actual regression straight-line reveals a positive offset in the automated SF measurement based on ACC alone. As a matter of fact, the SFs calculated by the proposed method slightly over-estimated the manual SF measurements in approximately 45% of cases. This experimental finding could depend on the colony morphology and size: the conventional cell counting procedure for the SF evaluation is based on the colony number alone and does not consider the colony area. Especially, correct SF measurements could be affected in cell lines with high growth rate, such as MCF10A. As an example, two different wells with the same cell colony number but with different area coverage will have the same  $SF_{\text{Manual}}$  value. On the contrary, the proposed SF evaluation methodology based on the percentage of ACC can appropriately deal with these situations, by discriminating both small and large cell colonies using the local adaptive thresholding algorithm, and yields coherent  $SF_{\%Area}$  measurements. In conclusion, according to the findings in [193], the conventional counting procedure underestimated the effect of the treatment in highly proliferative cell lines, because the therapy could not reduce the colony number by influencing the colony size alone.

**Discussion** As reported in the literature, cells after any type of treatment have a reduction in growth rate, and the colonies are generally smaller than the colonies of the untreated control [249, 395]. The conventional counting procedure, based on colony number alone, can underestimate the effect of treatment, because sometimes it is possible that the therapy does not reduce the colony number but decreases just the colony size [193]. Several authors highlighted the problems associated with detecting colonies near the rim of a culture well/flask. Possible solutions can be to exclude the rim area from processing or to use different segmentation methods for central and rim zones of the well [103]. In addition, when cell colonies present adherence in areas near the rim, discarding the rim area could strongly affect the SF evaluation, so worsening result accuracy. Many literature approaches [103, 138, 206, 235] use algorithms based on global thresholding techniques. Otsu's method [448] alone is not able to process the well/flask edges appropriately nor is suitable for less-discrete colonies. The approaches proposed in [103, 116, 138] were tested on small cell colonies, such as bacterial ones that are generally isolated with a circular shape and adjacent colonies rarely merging together. Conversely, human cells colonies often present an irregular shape and a greater size giving rise to large overlapping colonies. These issues have also proved to be difficult

to address [611]. More sophisticated approaches (e.g., watershed and distance transforms [502]) were also employed to overcome the problems regarding less-discrete or merging colonies with fuzzy boundaries by combining multiple counting measured, thus resulting in total colony area based statistical counts [427] with higher computational cost. Overlapping and heterogeneous colonies should be processed properly and effectively, without requiring statistical correction or estimation.

For these reasons, the proposed approach takes into account the actual size of each colony, by using a method that measures the ACC. In this way, the effect of the treatment on cell growth rate is also considered. Although estimating manually the ACC is unfeasible, this measurement conveys valuable information on the effects of the treatment. By considering the percentage of ACC, the number of colonies is not mandatory to evaluate the SF. The correct identification of the area wherein the cells were plated (e.g., well, Petri dish, flask) is crucial, because it can affect all the subsequent processing steps. Some literature approaches use binary masks with fixed position [49, 138, 235] or require manual alignment performed by the operator [235]. Often user intervention is needed also in the following steps. In [138] the operator interacts for threshold selection, while in [38] two different methods were implemented and the choice is made by the operator according to the cell type (colonies presented a low or a high contrast with respect to the background). To reduce operator dependence and possible errors in the processing pipeline (e.g., due to a misalignment between the well mask and the plate when fixed-positions/fixed-masks are used), we propose a fully automated approach. First of all, we employed a robust strategy, based on the CHT, for well detection. A control system was implemented *ad hoc* to enhance the CHT technique when errors occur, making it possible to detect all the wells properly in 100% of the cases. Successively, to detect cell colonies, the extracted wells are processed by using the local adaptive thresholding algorithm.

Finally, the need for dedicated hardware for the images acquisition is not to be underestimated. In [38, 49, 103], specific hardware is required (i.e., CCD cameras, lenses, light diffuser, or positioning devices) that can be also expensive. We exploited a conventional flatbed scanner to acquire multi-well plate images.

**Conclusion** Nowadays, clonogenic cell survival assays are evaluated by means of manual procedures, which could lead to inconsistencies related to operator dependence, so affecting the result repeatability. Unlike other literature approaches, the proposed method aims to address the issues that are yet open in the evaluation of the clonogenic assay (i.e., subjectivity in cell colony counting, need for dedicated hardware for acquire the analyzed

plate images, manual well delineation in multi-well plates), by proposing a fully automatic approach that attempts to overcome the limitations of the related works.

The proposed approach represents an alternative to the problem of clonogenic assays, presenting some novelties not yet considered in literature:

- fully automated well detection based on the CHT (properly enhanced with a control system to handle possible wrong detections) to recognize the wells inside the plate image (unlike several literature approaches that require plate dimensions and the spatial arrangement of the wells);
- differently to some related works, which use special cameras or hardware, our approach does not require any specific acquisition device;
- only a pre-processing step is applied, performing a conversion in the CIE  $L^*u^*v^*$  color model to reduce luminance inhomogeneities with respect to classical grayscale or RGB color space;
- colony segmentation exploiting adaptive thresholding to extract cell colonies from the background, which makes it effective for both small and large colonies;
- no post-processing is needed (for instance, morphological refinement [68, 571] can heavily affect ACC calculation, especially when performed on small cell colonies);
- SF evaluation based on the percentage of ACC instead that on the colony number (not applicable when the high growth rate leads to adjacent colony merging), which allows to establish a correlation between the growth rate and the dose administered during the treatment.

Comparisons on the SF calculated using the percentage of ACC against the conventional counting, averaged over three measurements by three expert operators, were performed for each experiment. The achieved results and the correlation between automatic and manual SF, demonstrate the reliability (in terms of robust identification of the wells in different plate types as well as correct colony extraction in various operative conditions) and accuracy (in terms of high correlation degree between the SFs measured by the human operators and those calculated automatically) of our approach. The proposed methodology to calculate the SF enables to overcome some of the most challenging open issues in clonogenic assay evaluation. In our opinion, this approach represents a valuable and effective solution to be integrated in an expert system to support biologists in their own decision making tasks. Future work is aimed at extending and improving the proposed approach on other cell lines, which specifically show a lower violet color uptake during colony fixation/staining phase



(e.g., invasive breast cancer, such as MDA cells). A more sensitive classification approach could be realized by extracting appropriate textural feature descriptors, such as the Haralick's features [250, 251].

### 3.1.2 Region splitting-and-merging

The split-and-merge algorithm [273] is a direct region detection approach that works in a hybrid fashion: the initial image can be alternatively subdivided into a set of disjoint regions by splitting (top-down procedure) and/or merging (bottom-up procedure) them by attempting to match the segmentation conditions in terms of the homogeneity criteria expressed by a Boolean predicate  $P$  [179]. Therefore, the split-and-merge algorithm starts with an arbitrary partition  $\mathcal{I}$  (i.e., entire image region) and yields an output composed of uniform sub-regions:  $\mathcal{R}_i$ , with  $i = 1, 2, \dots, R$ . At the  $t$ -th step, a generic region  $\mathcal{R}_i$  is split if uniformity measures are not matched ( $P(\mathcal{R}_i) = \text{false}$ ). Therefore, the initial square image region  $\mathcal{I}$  (whose dimensions are a power of 2) can be recursively divided into four smaller quadrants (quad-regions).

In this context, the most critical choice is the logical predicate  $P$ , as it is fundamental to get satisfactory outputs. For instance, homogeneity criteria can be defined in terms of statistical measures computed in the ROI (e.g., mean value or standard deviation). In order to control the fibroid segmentation accuracy, the parameter  $\rho_{\min}$  sets the minimum quad-region size beyond which no further splitting is carried out. This parameter must be a trade-off between precision and computational effort. In our case, the best results are obtained with  $\rho_{\min} = 4 \times 4$  pixels, because even small regions can be detected.

#### Uterine fibroid detection on MR images

The aim of the overall approach is to automatically segment the ablated fibroid ROT that are within the uterus, in order to optimize the current methodology for MRgFUS treatments.

Our approach combines region splitting-and-merging and multi-seeded region growing that relies on two phases: (i) automatic seed-region selections and (ii) uterine fibroid segmentation (Fig. 3.9). By taking inspiration from [533, 534], where a method for brain lesion segmentation in DWI is described, ROT detection is accomplished through a split-and-merge algorithm [273] in order to determine meaningful seeds. Moreover, the split-and-merge output obtained with our approach, after some refinement operations, is directly exploited as a partial result for the subsequent region growing step by implementing an incremental procedure. Region growing can segment a single area at a time, since a seed must be chosen for each fibroid, in order to start the inclusion step based on homogeneity measures. For these reasons, a seed selection method based purely on intensity levels (i.e., histogram

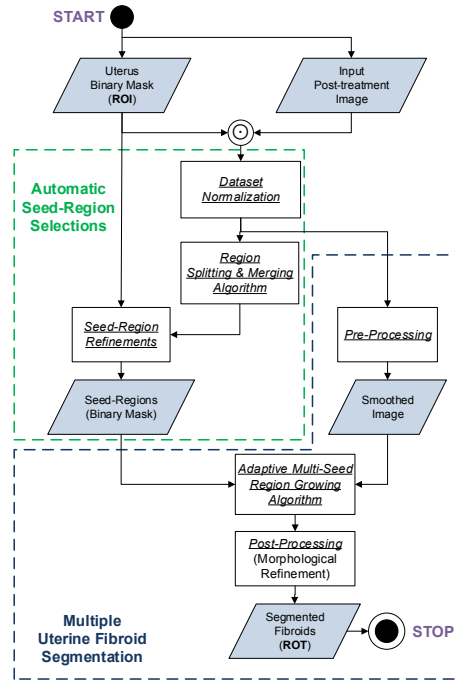


Fig. 3.9 Flow diagram of the SM & RG method, divided into two phases: (i) automatic seed-region selections; (ii) multiple uterine fibroid segmentation.

inspection) may not be completely reliable and sufficient to detect ROT in a pathological scenario with multiple fibroids, where the number of lesions is not known *a priori*. Therefore, a region analysis method is required to correctly detect the fibroids within the uterus ROI. The split-and-merge algorithm represents a well suited solution [386], because it is able to find homogeneous regions in terms of uniformity criteria. Seeds are then exactly calculated by excluding connected-components located near the uterus ROI boundaries. Then, starting from the seed-regions selected by the split-and-merge approach, segmentation is achieved through a region growing procedure guided by appropriate similarity properties that describe ROT intensity features. Finally, some post-processing operations are used for refining the segmented fibroids.

In order to identify the ROT, homogeneity criteria regarding the growing-region  $\mathcal{R}_G$  are defined in terms of its mean value  $\mu_{\mathcal{R}_G}$ . By tuning different parameters, we found the following values:

$$P := \begin{cases} \text{true,} & \text{if } (\mu_{\mathcal{R}_G} > 0.0) \wedge (\mu_{\mathcal{R}_G} < 0.58) \\ \text{false,} & \text{otherwise} \end{cases}. \quad (3.6)$$

However, many splits are inevitably performed at the uterus boundaries due to the strong intensity difference between ROI edge pixels and the background, which involves

non-compliance with the homogeneity criteria defined by the predicate  $P$ . Thus, these connected-components are removed before seed-region selection. First, loose connections between the splitting-and-merging output and ROI edges are removed by morphological opening (more conservative than erosion) with a diamond-shaped structuring element, to easily disconnect also diagonally connected pixels among 8-neighborhood, which are very frequent in ROI boundaries. Diamond-shaped structuring elements yield satisfying results even with vertical/horizontal edges and are preferred to squared or circular filters when object boundaries are roughly in the diagonal direction [572, 594].

The remaining parts at ROI edges must be cut out before seed selections through a connected-component labeling procedure using a flood-fill algorithm [151]. In fact, we calculate the pixel intersection (logical product) between each connected-component and the dilated ROI edge image, using a squared structuring element. Constraints on the cardinality of this intersection are defined according to the different granularity of the various connected-region areas, in order to consider the occurrence of large fibroids situated at the ROI boundary. Hence, the current analyzed connected-component is included into the output binary mask only if the control is matched, otherwise it will be removed. The splitting-and-merging output, besides being used for ROT detection, is thus exploited as an initial result for the subsequent fibroid segmentation through the region growing algorithm, making it possible to considerably optimize execution time.

### 3.1.3 Region growing

Region growing is a bottom-up segmentation method that produces a homogeneous region by successively merging primitive regions (sub-regions or also single pixels) [218]. Therefore, the region growing algorithm starts from seed values and attempts to find a local connected-region depending on given homogeneity criteria [534]. Since the traditional region growing approach is able to perform segmentation of a single area only, when the currently examined patient can be affected by more than one fibroid, it is necessary to implement a multi-seeded region growing segmentation. As mentioned before, in a basic seeded region growing algorithm, each region begins its own growth from a seed. Therefore, the initial set of seeds  $\mathcal{S} = \{s_1, s_2, \dots, s_n\}$  must be a representative and expressive sample of the region to segment. Since seed choice affects overall segmentation results, this task is fundamental and very critical [3]. Another issue is the stopping rule formalization. In fact, the region growing procedure must be interrupted if no more pixels match the membership criterion [218]. The inclusion in a region can be stated in terms of a segmentation threshold  $\Theta$ , which specifies the maximum allowed deviation of pixel features within that region. Therefore, segmentation threshold  $\Theta$  is a region-dependent parameter that guides the homogeneity decision test [95].

These criteria can be local (i.e., pixel intensity [416]) or region descriptor-based (region shape or size, likeness between a candidate pixel and the current grown region in terms of image features [201]) that are more powerful because they take into account the “history” of region growth when some *a priori* knowledge about the region is available. As explained in [3], to prevent a poor starting estimate, it is recommended that seed-regions be used instead of single pixels (seed-points). In fact, when dealing with noisy image segmentation, each seed-area should be sufficiently large to ensure a stable estimate of its region’s mean. Moreover, these seed-regions are used as input for the region growing incremental process, exploiting the previous split-and-merge segmentation results.

### Uterine fibroid segmentation on MR images

The algorithm proceeds by considering one seed-region at a time, which identifies a single region growth. The current region  $\mathcal{R}_G$  is iteratively grown by analyzing all unallocated pixels belonging to the boundary list  $\mathcal{B}$ , obtained by considering 8-neighborhoods as candidate pixels. As stated in [416], region growing is ruled by three conditions concerning the examined pixel  $\mathbf{p}_B \equiv (\bar{x}, \bar{y})$ :

- $\mathbf{p}_B$  is inside the image  $\mathcal{I}$  (with size  $M \times N$  pixels);
- $\mathbf{p}_B$  has not already been included in  $\mathcal{R}_G$ ;
- $\mathbf{p}_B$  value satisfies similarity criteria based on the segmentation threshold  $\Theta$  in terms of the region mean intensity, i.e.,  $\mu_{\mathcal{R}_G} \leftarrow \frac{\sum_{(x,y) \in \mathcal{R}_G} f(x,y)}{|\mathcal{R}_G|}$ .

These constraints can be more formally expressed by the following membership predicate:

$$P := \begin{cases} \text{true,} & \text{if } (0 \leq \bar{x} \leq M - 1) \wedge (0 \leq \bar{y} \leq N - 1) \wedge (\mathbf{p}_B \notin \mathcal{R}_G) \wedge (|f(\bar{x}, \bar{y}) - \mu_{\mathcal{R}_G}| < \theta) \\ \text{false,} & \text{otherwise} \end{cases} \quad (3.7)$$

This similarity criterion is thus defined in terms of absolute distance from the region mean intensity  $\mu_{\mathcal{R}_G}$ , which is the most representative property in images with speckle noise, such as MRI [183, 473]. Moreover, the seed-regions  $\mathcal{S}$  accurately estimate mean intensity for each growth region  $\mathcal{R}_G$ . The value of  $\mu_{\mathcal{R}_G}$ , the region mean intensity, is also exploited for the automatic and unsupervised estimation of the segmentation threshold  $\Theta$ , because the manual selection of a single threshold, suitable for all datasets, could be time-consuming and error-prone. In fact, since normalized images are characterized by a bimodal histogram, an optimal threshold  $\theta_{\text{opt}}$  can be selected through the Otsu’s method [448] that attempts to

maximize the inter-class variance between two regions with different intensity values. The adaptive threshold  $\Theta_{\text{adapt}}^{(t)}$  is defined as:

$$\Theta_{\text{adapt}}^{(t)} = \theta_{\text{opt}} - \mu_{\mathcal{R}_G}^{(t)}, \quad (3.8)$$

and it is updated during each iteration  $t$ , by estimating the corresponding average value  $\mu_{\mathcal{R}_G}^{(t)}$  while the region grows. In this way, a different stopping rule is defined for each growth region  $\mathcal{R}_G$  by calculating an adaptive segmentation threshold  $\Theta_{\text{adapt}}$  for each fibroid in multi-fibroid pathological contexts.

In the initialization step, morphological erosion on the initial seed binary masks  $\mathcal{B}$  is necessary to avoid that boundary points of seed-regions are eventually outside the ROT, resulting in leaking during the region growing process. Furthermore, seed-points  $\delta_j$  are determined by sampling 1 every 5 edge pixels from the array  $\mathcal{D}$  to prevent the overlap between the 8-neighborhoods of adjacent edge pixels. Fig. 3.10 shows some examples of the incremental process results. It is visible as most of the ROT has already been detected and successfully segmented by the split-and-merge algorithm. The use of seeded region growing, described in [95], requires seed-points for both fibroids and the rest of the uterus. In fact, the entire image is divided into a set of disjointed homogeneous regions, achieving a tessellation similar to the watershed algorithm [565]. This mechanism is more computationally expensive in terms of homogeneity tests (uterus region growth in addition to fibroid segmentation) and the need of a selection policy when a given pixel would be assigned simultaneously to several regions during the growth process. Moreover, small dark areas (due to acquisition artifacts) could be present in post-treatment images, invalidating the overall segmentation process.

**Results** To quantify the efficiency of the proposed method in terms of execution time, the (sequential) implementation of the proposed incremental region growing approach is compared with a classic multi-seed region growing implementation (where a single seed-point is defined for each fibroid by calculating centroids of detected regions in Phase  $(i)$ ). Our method was developed and tested under MatLab, on a PC with a 2.8 GHz Intel® Core™ i5 processor and 8 GB RAM. In particular, this incremental procedure reduces the segmentation time with a  $1.5\times$  speed-up factor, when compared to the traditional multi-seeded region growing approach (developed with a single seed-point selection for each fibroid).

**MRI dataset composition** The study was carried out on an MRI dataset composed of 18 patients (average age:  $42.22 \pm 6.29$  years) affected by symptomatic uterine fibroids underwent MRgFUS therapy. The total number of the examined fibroids was 27, since some patients presented a pathological scenario with multiple fibroids. The analyzed images were acquired

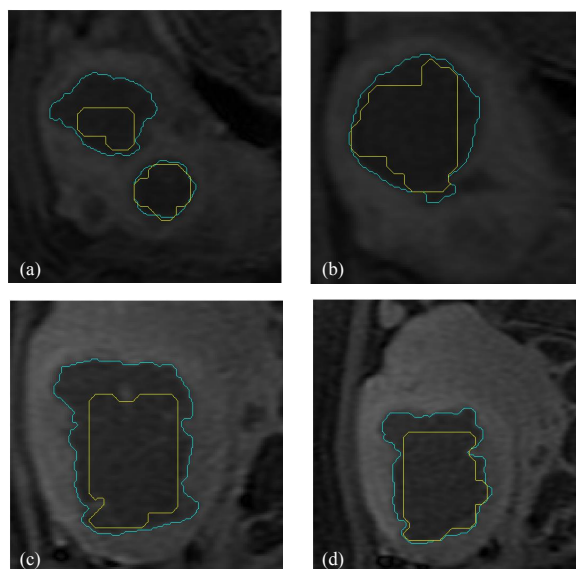


Fig. 3.10 Uterine fibroid segmentation examples achieved by the SM & RG incremental procedure: (a) patient with multiple fibroid; (b, c, d) patients with large fibroids. Yellow contours represent Phase (i) outputs, while cyan contours are the final region growing results (Phase (ii)).

Table 3.3 Acquisition parameters of the MRI uterine fibroid dataset.

MRI sequence	TR [ms]	TE [ms]	Matrix size [pixels]	Slice spacing [mm]	Slice thickness [mm]	Pixel spacing [mm]	Slices' number
T1w FSPGR+FS+C	150-260	1.392-1.544	512 × 512	0.6641-0.7031	5.0	6.0	16-29

using a Signa HDxt MRI 1.5T scanner (General Electric Medical Systems, Milwaukee, WI, USA) at two different institutions. Segmentation tests were performed on MR images acquired after the MRgFUS treatment, executed with the ExAblate 2100 (Insightec Ltd., Carmel, Israel) HIFU equipment. The considered MR slices were scanned using the T1w “Fast Spoiled Gradient Echo + Fat Suppression + Contrast mean” (“FSPGR+FS+C”) protocol. This MRI protocol is usually employed for NPV assessment [390], since ablated fibroids appear as hypo-intense areas due to low perfusion of the contrast mean. Sagittal MRI sections were processed, in compliance with current clinical practice for therapy response assessment. MRI data are encoded in the 16-bit DICOM format. Table 3.3 depicts the relevant MRI acquisition parameters, by reporting the range when different values were used. Two instances of input uterine MR images are shown in Fig. 3.11, where the white contour and the white arrows indicate the uterus ROI and the fibroid ROTs, respectively.

**Quantitative evaluation and clinical feasibility analysis** We evaluated and compared two validated computer-assisted segmentation methods, which we have already presented in [411] and [524], for uterine fibroid segmentation in MRgFUS treatments. A quantitative

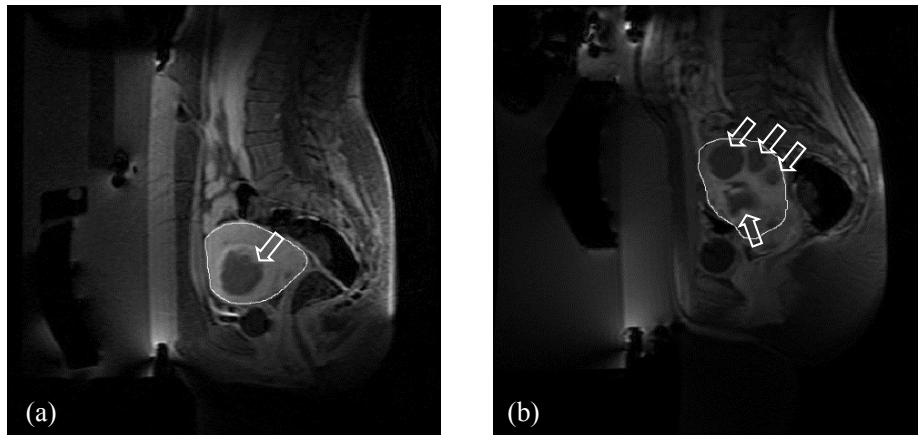


Fig. 3.11 Examples of input uterine fibroid T1w “FSPGR+FS+C” MR sagittal slices: (a) large fibroid; (b) multi-fibroid pathological scenario. The uterus ROI is delineated (white contour) and uterine fibroid ROTs are indicated with white arrows.

comparison on segmentation accuracy and reliability, in terms of spatial overlap-based and distance-based metrics, was performed. The clinical feasibility of these approaches was assessed from physicians’ perspective, proposing an integrated solution that considers also the workflow in real environments. Fig. 3.12 sketches the overall workflow for uterine fibroid MRgFUS treatment response assessment on post-treatments MRI series.

The two methods presented in [415] and [526] may be interchangeably used in the whole workflow defined in Fig. 3.12, representing two different options to effectively support radiologists in accurate NPV evaluation. In both cases, the overall medical image analysis pipeline is unsupervised in all its components, enabling its integration in a comprehensive decision support system that provides clinically useful and quantifiable information (i.e., NPV measurement, ROT shape and location) for personalized medicine. Fig. 3.13 shows the uterine fibroid ROT delineations, achieved by the IOTS and SM & RG segmentation methods (over-imposed gray and white contours, respectively).

The values of both overlap-based and distance-based metrics, achieved by the IOTS [415] and SM & RG [526] methods on the experimental dataset analyzed in this study, are shown in Table 3.4. To provide a comprehensive graphical statistical summary of the different evaluation metrics, boxplots are illustrated in Fig. 3.14.

*DSC* and *JI* mean values are approximately equal, proving high segmentation accuracy in both cases. Although SM & RG results are characterized by slightly higher standard deviation, the interquartile range in boxplots is more concentrated and two outliers are observed because leaking (i.e., over-estimation of fibroid size) in the case of poor contrast images could occur. *SEN* values show that the IOTS-based global thresholding solution is

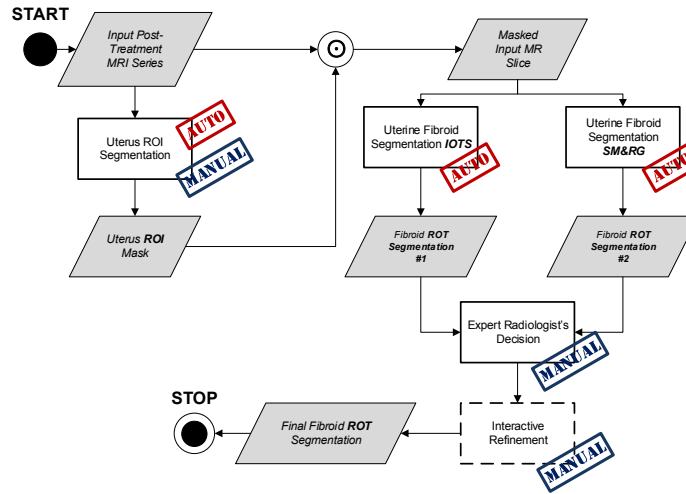


Fig. 3.12 Workflow of uterine fibroid MRgFUS treatment response evaluation. The whole process can be divided into two main stages: (i) uterus ROI segmentation, performed either by an automated approach (for instance, the FCM-based approach in [415]) or manually by an expert physician; (ii) uterine fibroid ROT segmentation, using the proposed approaches based either on the Iterative Optimal Threshold Selection (IOTS) algorithm [415] or on combined Split-and-Merge and Region Growing (SM & RG) [524]. All the steps involving the radiologist's intervention are also reported, by labeling automatic or manual operations. Solid line boxes are mandatory processes, while dashed line blocks represent interactive steps that are not always required.

Table 3.4 Values of overlap-based and distance-based metrics achieved by the IOTS and SM & RG methods on the complete MRI dataset composed of 18 patients undergone MRgFUS therapy for uterine fibroid ROT ablation. The results are expressed as average value  $\pm$  standard deviation.

Method	DSC	JI	SEN	SPC	AvgD	MaxD	HD
IOTS	87.25 $\pm$ 5.86	78.75 $\pm$ 8.74	88.16 $\pm$ 8.48	87.88 $\pm$ 4.77	3.308 $\pm$ 4.863	7.997 $\pm$ 6.952	3.100 $\pm$ 0.502
SM & RG	87.47 $\pm$ 6.29	78.57 $\pm$ 9.30	86.07 $\pm$ 8.45	90.86 $\pm$ 8.23	2.498 $\pm$ 2.702	7.890 $\pm$ 6.988	3.123 $\pm$ 0.494

often more sensitive than the SM & RG approach, especially in little fibroid detected in initial or final ROT slices. Indeed, the region splitting-and-merging procedure could not detect small regions with fuzzy boundaries. However, as visible from the obtained *SPC* values and distance-based metrics, the region growing algorithm generally delineates the ROT contours more precisely than the IOTS algorithm.

Our approaches have been already compared with similar literature works in the original publications [415, 526]. In both cases, the proposed methods remarkably outperformed the other implemented methods. In addition, the clinical feasibility of many literature approaches is limited. For instance, LSF-based methods are highly sensitive to initial user input and require a labor-intensive parameter setting procedure by clinicians to obtain satisfactory



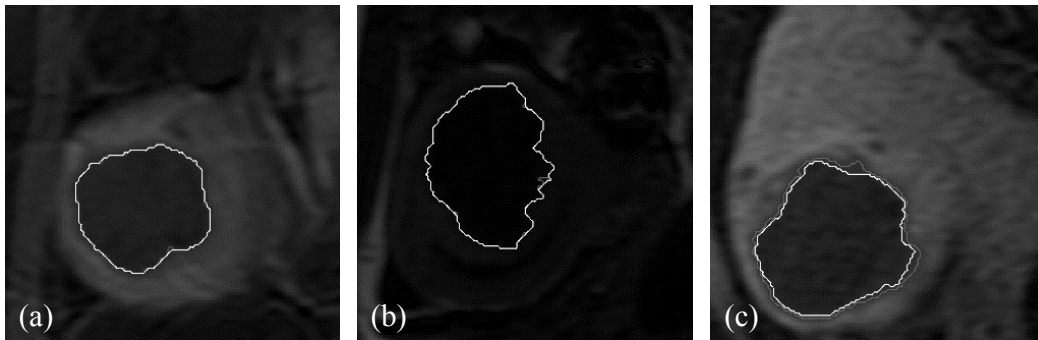


Fig. 3.13 Uterine fibroid ROT segmentation results: segmentation based on IOTS (gray contour) and segmentation using SM & RG (white contour). ( $8\times$  zoom factor).

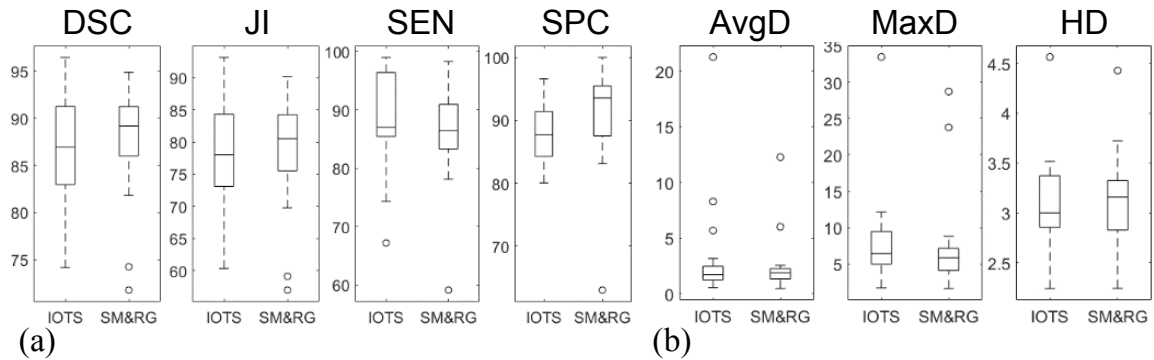


Fig. 3.14 Boxplots of the uterine fibroid ROT segmentation results: (a) overlap-based metrics; (b) distance-based metrics.

segmentations. Differently to the method proposed in [177, 178], where multispectral MRI sequences and external co-registration operations are mandatory, our methods require T1w contrast-enhanced MRI sequence alone and can be easily integrated in a single software tool.

**Discussion and conclusions** The evaluation measures achieved by these two methods, namely IOTS [415] and SM & RG [524] segmentation pipelines, were comparable in terms of delineation accuracy: the former showed higher sensitivity and the latter was characterized by higher specificity. Hence, we can argue that the IOTS and SM & RG uterine fibroid segmentation approaches have almost complementary properties, allowing for an integrated image processing pipeline by adapting on the different analyzed images. According to Fig. 3.12, a possible clinically feasible solution could be described by the following use case that is easily implementable in the real workflow for uterine fibroid MRgFUS treatment response evaluation:

1. the segmentation results obtained by the two different automatic approaches based on Pattern Recognition algorithms are proposed, by avoiding cognitive overload;
2. the radiologist chooses the best result between the two proposed segmentations, also combining them appropriately in a late integration phase;
3. the user can interactively refine the achieved segmentation by selecting only the correct ROTs (i.e., removing false positives and adding false negatives), as well as by moving the control points that define the computed ROT boundaries.

A volumetric reconstruction of uterus ROI and ablated fibroid ROT could be also useful for radiologists during their own decision-making tasks. Especially, in pathological scenarios with multiple fibroids. However, tridimensional uterine fibroid MR image segmentation approaches are not easy to realize, since fibroids, particularly if they are pedunculated, can be displaced by presenting gaps also in adjacent slices [681]. In such cases, leaking is likely to occur, especially in 3D region growing implementations [414].

In conclusion, the resulting segmentation pipeline is an integrated solution that attempts to solve the clinical problems regarding MRgFUS treatment response evaluation for uterine fibroid, by improving significantly the current operative methodology in terms of segmentation accuracy, result repeatability, and execution time. Generally, reliable segmentation results were observed, even when dealing with fibroid ROTs characterized by irregular or inhomogeneous necrotic material, so that operator de-pendency is reduced. Future developments aimed at the realization of advanced image enhancement methods to improve the segmentation process for medical images characterized by bimodal histograms, such as the approach presented in Section 5.1.1.

## 3.2 Morphological approaches

### 3.2.1 Watershed transform

The watershed transform [573] is one of the most used approaches in cell image segmentation [402], even though it was originally proposed in the field of mathematical morphology [571].

The intuitive description of this transform is straightforward: assuming an image as a topographic relief, where the height of each point is directly related to its gray level, and considering rain gradually falling on the terrain, then the watersheds are the lines that separate the resulting catchment basins [638]. This technique is valuable because the watershed lines generally correspond to the most significant edges among the markers [51] and is useful to separate overlapping objects. Even when no strong edges between the markers exist, the

watershed method is able to detect a contour in the area. This contour is detected on the pixels with higher contrast [573]. As a matter of fact, edge-based segmentation techniques—which strongly rely on local discontinuities in gray levels—often do not yield unbroken edge curves, thus heavily limiting their performance in cell segmentation [645]. Unfortunately, it is also well-known that the watershed transform may be affected by over-segmentation issues, so requiring further processing [223].

From a computational perspective, the watershed algorithm analyzes a gray-scale image by means of a flooding-based procedure. Since the flooding process is performed either on a gradient image or edge map, the basins should emerge along the edges. As a matter of fact, during the watershed process, the edge information allows for a better discrimination of the boundary pixels with respect to the original image. Finally, only the markers of the resulting foreground cells are selected.

Efficient versions of the watershed algorithm exploit a priority queue to store the pixels according to the pixel value (i.e., the height in the gray-scale image landscape) and the entry order into the queue (giving precedence to the closest marker). More specifically, during the flooding procedure, this process sorts the pixels in increasing order of their intensity value by relying on a breadth-first scan of the plateaus based on a first-in-first-out data structure [638].

### **Automatic Cell Detection and Counting**

In [530], we presented Automated Cell Detection and Counting (ACDC), a novel pipeline designed for time-lapse microscopy activity detection of fluorescently labeled cell nuclei that is capable of overcoming the practical limitations of the literature methods, mainly related to the training phases or time constraints, to be considered as a feasible solution in laboratory practice. The proposed fully automatic method is based on the watershed transform [573, 638] and morphological filtering [51, 571]. Bilateral filtering [614] on the original image is formerly applied to smooth the input cell images while preserving edge sharpness. ACDC efficiently yields reliable results just using few parameters, which may conveniently be altered by the biologist. Thus, the resulting interpretable segmentation model makes the user sufficiently aware of the underlying automated analysis, unlike sophisticated solutions that do not provide any interactive control suitable for the end users. The proposed computational approach was validated on two different cell imaging datasets in order to show its reliability in different experimental conditions. ACDC can also leverage multi-core architectures and computer clusters to reduce the running time required to analyze a large single image stack.

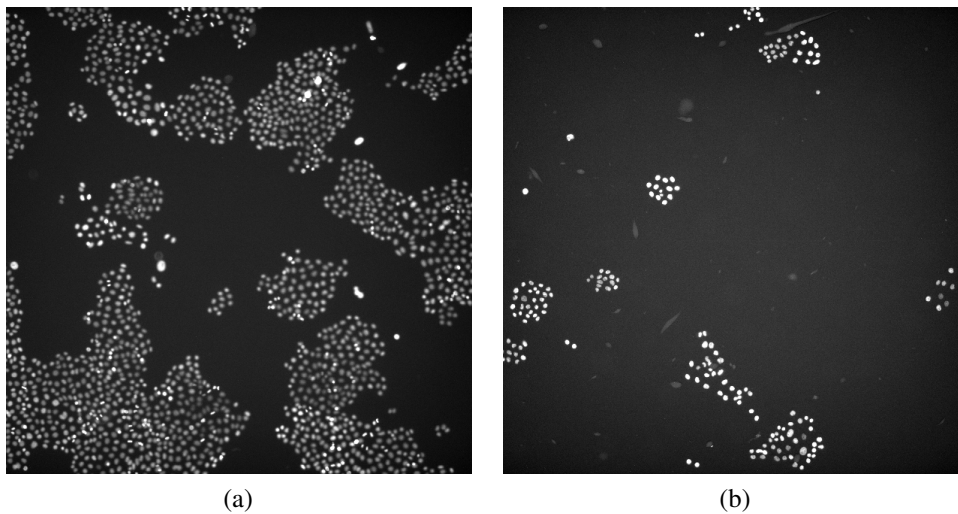


Fig. 3.15 Examples of the analyzed microscopy fluorescence images provided by the Department of Biochemistry of the VU. The images were displayed by automatically adjusting the brightness and contrast according to an histogram-based procedure.

### Fluorescence microscopy imaging data

**Vanderbilt University dataset** This dataset collects time-lapse microscopy images from two experiments performed at the Department of Biochemistry of the Vanderbilt University (VU) School of Medicine, Nashville, TN, USA. All the images were acquired using a Molecular Devices (San Jose, CA, USA) from ImageXpress Micro XL with a  $10\times$  objective in the red channel (Cy3); some images were acquired in the green channel (FITC) for the same well and location (overlapping information), using a PCO.SDK Camera (Kelheim, Germany). The imaged region is  $2160 \times 2160$  pixels and the pixel-depth is 16 bits. In this study, we consider the PC-9 human lung adenocarcinoma cell line [300, 622]. The total number of the analyzed images with the corresponding gold-standard was 46, related to two different experiments. Two examples of input images are shown in Fig. 3.15.

**2018 Data Science Bowl** In order to validate ACDC on imaging data coming from different sources, we considered a selection of the training set from the 2018 Data Science Bowl (DSB) [302], which is a competition organized by Kaggle (San Francisco, CA, USA). We had to use the training set alone, since the labels for the test set are not publicly provided by the organizers. The goal of the 2018 DSB regards the detection of cells' nuclei to identify each individual cell in a sample. This operation is mandatory to understand the underlying biological processes.

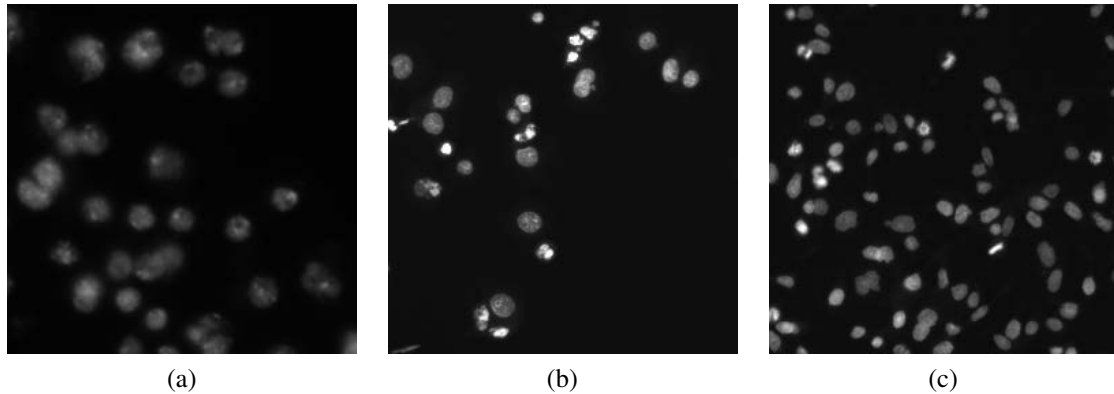


Fig. 3.16 Examples of the analyzed microscopy fluorescence from the 2018 DSB dataset.

This dataset includes a large number of labelled nuclei images acquired under a variety of conditions, magnification, and imaging modality (i.e., bright-field and fluorescence microscopy). Image size varies in  $256 \times 256$ ,  $256 \times 320$ ,  $260 \times 347$ ,  $360 \times 360$ ,  $512 \times 640$ ,  $520 \times 696$ ,  $603 \times 1272$ , and  $1040 \times 1388$  pixels. The dataset aims at evaluating the generalization abilities of computational methods when dealing with significantly different data. We first extracted only the fluorescence microscopy images from the training set. According to the type of images used during the design phase of ACDC, the selection criterion of the fluorescence microscopy images was based on the maximum size  $\max\{\text{size}(\mathcal{R}_{\text{cell}})\}$  of the region cells  $\mathcal{R}_{\text{cell}}$  segmented in the ground truth. Therefore, we selected only the images where  $\max\{\text{size}(\mathcal{R}_{\text{cell}})\} \leq 1000$  pixels. The number of the resulting images is 303. Three examples of input images characterized by highly different characteristics are shown in Fig. 3.16.

**ACDC: the proposed efficient automatic cell nuclei segmentation method** Time-lapse microscopy imaging allows for the visualization and analysis of living specimens' dynamic processes. A single experiment may consist of hundreds of objects depicted on thousands of images, leading to the need for computational methods that can enable the quantitative analysis of system-level biology automatically.

Advances in optics and imaging systems have enabled biologists to visualize live cell dynamic processes by time-lapse microscopy images. However, the imaging data recorded during even a single experiment may consist of hundreds of objects over thousands of images, which makes manual inspection a tedious, time-consuming and inaccurate option. Traditional segmentation techniques proposed in the literature generally exhibit low performance on live unstained cell images. These limitations are mainly due to low contrast, intensity-variant,

and non-uniform illuminated images [306]. Therefore, novel automated computational tools for quantitative system-level biology are required.

The proposed method is designed for time-lapse microscopy and aims at efficiently overcoming the limitations of the literature methods [83, 208, 264], especially in terms of efficiency and execution time. As a matter of fact, the resulting processing pipeline exploits classical image processing techniques in a smart fashion, allowing for a feasible solution in real laboratory environments. Fig. 3.17 outlines the overall flow diagram of the ACDC segmentation pipeline. Each single step is described hereafter. Our approach is fully automatic, so allowing for reproducible measurements [543]. It is worth noting that the number of fixed parameters that need to be set in the underlying image processing operations involves only the kernel size of spatial filters and the structuring element size in morphological operations, allowing for a robust yet simple solution. Therefore, these few parameters allow the user to have control over the achieved cell segmentation results. Unlike DCNN-based black or opaque boxes, ACDC offers an interpretable model for the biologists that may adjust conveniently adjust some parameters according to the cell lines currently analyzed. Differently to supervised Machine Learning approaches [140, 535], ACDC does not require any training phase, thus representing a reliable and practical solution even without labeled datasets.

**Pre-processing** The input microscopy image is pre-processed to yield a more convenient input to the downstream watershed-based segmentation relying on the following steps:

1. Bilateral filtering that allows for denoising the image  $\mathcal{I}$  while preserving the edges by means of a non-linear combination of nearby image values [614]. This noise-reducing smoothing filter combines gray levels (colors) according to both a geometric closeness function  $c$  and a radiometric (photometric) similarity function  $s$ . This combination is used to strengthen near values with respect to distant values in both spatial and intensity domains. More specifically, the intensity of each pixel is replaced with a weighted average of intensity values from neighboring pixels. The weights depend not only on the Euclidean distance between pixels, but also on the radiometric differences (i.e., range differences). This dependence on pixel intensity helps to preserve the edge sharpness. The image  $\mathcal{I}_{\text{bf}}$  smoothed through the bilateral filter can be denoted by the following Equation:

$$\mathcal{I}_{\text{bf}}(\mathbf{p}) = \frac{1}{\eta} \sum_{\mathbf{q} \in \Omega} \mathcal{I}(\mathbf{q}) \cdot c(\mathbf{p}, \mathbf{q}) \cdot s(\mathcal{I}(\mathbf{p}), \mathcal{I}(\mathbf{q})), \quad (3.9)$$

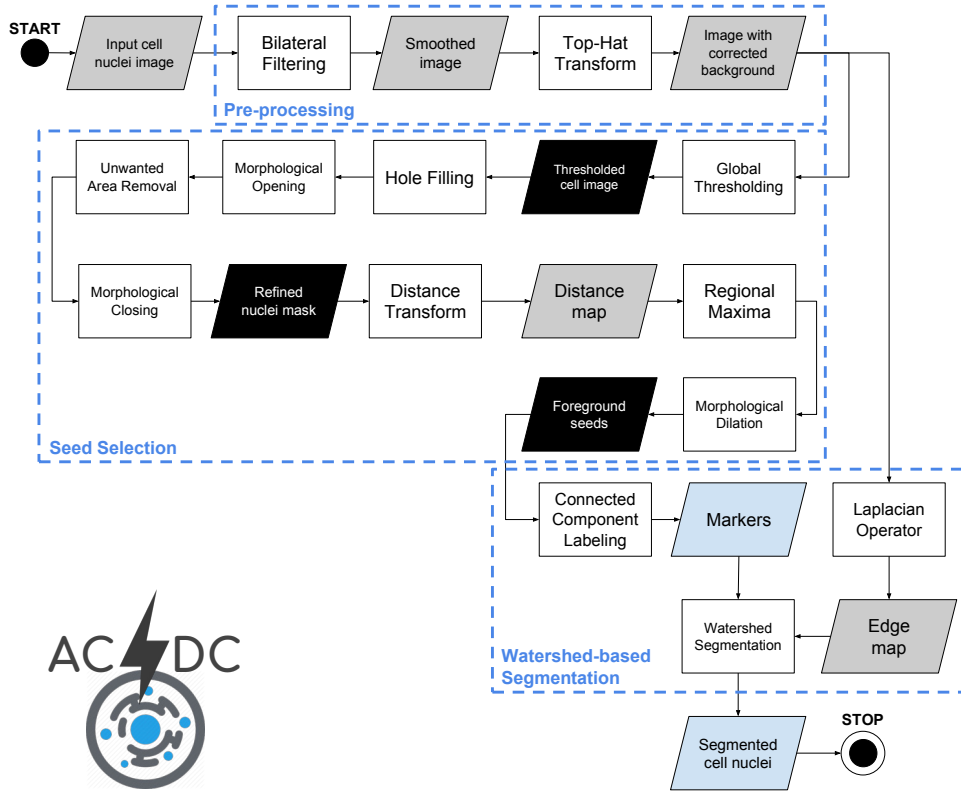


Fig. 3.17 Flow diagram of the proposed ACDC pipeline for cell nuclei segmentation. The gray, black and light-blue data blocks denote gray-scale images, binary masks and information extracted from the images, respectively.

where the normalization term  $\eta = \sum_{\mathbf{q} \in \Omega} c(\mathbf{p}, \mathbf{q}) \cdot s(\mathcal{I}(\mathbf{p}), \mathcal{I}(\mathbf{q}))$  constrains the sum of weights to one and preserves the local mean [97]. The window  $\Omega$ , centered on the current pixel  $\mathbf{p}$  for filtering, is a square whose side is calculated as:  $\text{size}(\Omega) = \max\{5, 2 \cdot \lceil 3 \cdot \sigma_s \rceil + 1\} = 7$  pixels. This simple yet effective strategy allows for contrast enhancement [548]. Bilateral filter has been shown to work properly in fluorescence imaging even preserving the directional information, such as in the case of the F-actin filaments [631]. This denoising technique was effectively applied to biological electron microscopy [295] as well as to cell detection [359], revealing better performance, with respect to low-pass filtering, in noise reduction without removing the structural features conveyed by strong edges. The most commonly used version of bilateral filtering is the shift-invariant Gaussian filtering, in which both the closeness function  $c$  and the similarity function  $s$  are Gaussian functions of the Euclidean distance between their arguments [614]. More specifically,  $c$  is radially symmetric:  $c(\mathbf{p}, \mathbf{q}) = e^{-\frac{1}{2} \left( \frac{\|\mathbf{p}-\mathbf{q}\|}{\sigma_s} \right)^2}$ . Analogously, the similarity function  $s$  can be defined as:  $s(\mathbf{p}, \mathbf{q}) = e^{-\frac{1}{2} \left( \frac{\|\mathcal{I}(\mathbf{p})-\mathcal{I}(\mathbf{q})\|}{\sigma_r} \right)^2}$ .

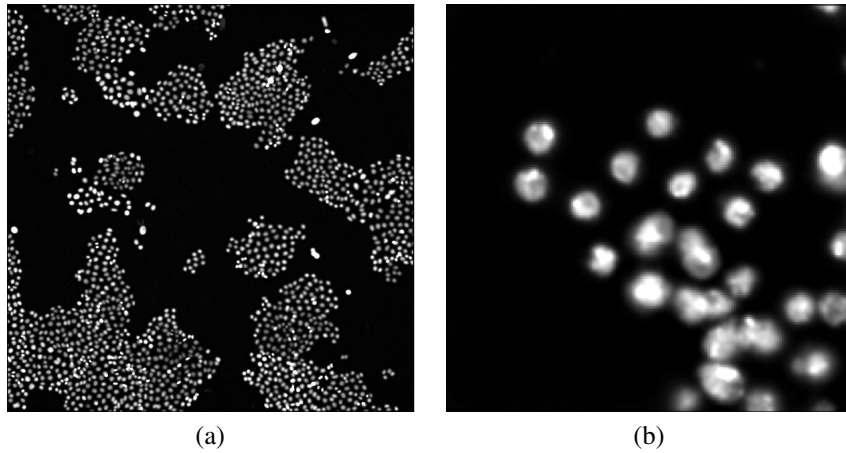


Fig. 3.18 Pre-processed images in the ACDC pipeline: (a) result on the image shown in Fig. 3.15a; (b) result on the image shown in Fig. 3.16a.

We set  $\sigma_c = 1$  and  $\sigma_s = \sigma_{\text{global}}$  (where  $\sigma_{\text{global}}$  is the standard deviation of the input image  $I$ ) for the standard deviation of the Gaussian functions  $c$  and  $s$ , respectively. This smart denoising approach allows us to keep the edge sharpness while reducing the noise of the processed image, so avoiding cell region under-estimation.

2. Top-hat transform for background correction with a binary circular structuring element (radius: 21 pixels) is applied on the smoothed image. This operation accounts for non-uniform illumination artifacts, by extracting the nuclei from the background. The white top-hat transform is the difference between the input image  $\mathcal{I}$  and the opening of  $I$  with a gray-scale structuring element  $b$ :  $\mathcal{T}_w = \mathcal{I} - \mathcal{I} \circ b$  [218].

The results of the pre-processing images applied to Figs. 3.15a and 3.16a are shown in Figs. 3.18a and 3.18b, respectively.

**Nucleus seed selection** The cell nuclei have to be accurately extracted from the pre-processed image. The following steps are applied to obtain a reliable seed selection:

1. a thresholding technique has to be first applied to detect cell regions. Both global and local thresholding techniques aim at separating desirable foreground objects from the background in an image, considering differences in pixel intensities [291]. Global thresholding determines a single threshold for all pixels and works well if the histogram of the input image contains well-separated peaks corresponding to the desired foreground objects and background [448]. Local adaptive thresholding techniques estimate the threshold locally over sub-regions of the entire image, by



considering only a neighborhood (i.e., a user-defined window) with a specified size and exploiting local image properties to calculate a variable threshold [218, 291]. These algorithms find the threshold by locally examining the intensity values of the neighborhood of each pixel according to image intensity statistics. In order to avoid unwanted pixels in the thresholded image, mainly due to small noisy hyper-intense regions caused by non-uniform illumination, we applied the Otsu's global thresholding method [448] instead of local adaptive thresholding based on the mean value in a neighborhood [411]. Moreover, global adaptive threshold techniques are significantly faster than local adaptive strategies;

2. hole filling is applied to remove possible holes in the detected nuclei due to small hypo-intense regions included in the nuclei regions;
3. morphological opening (using a 1-pixel disk as structuring element) is used to remove loosely connected components, such as in the case of almost overlapping cells;
4. unwanted area are removed based on the connected-components size. Especially, the detected candidate regions with areas smaller than 40 pixels are removed to refine the achieved segmentation results by robustly avoiding false positives;
5. morphological closing (using a 2-pixel radius circular structuring element) is applied to smooth the boundaries of the detected nuclei and avoid the under-estimation of the detected nuclei regions;
6. the approximate Euclidean distance transform from the binary mask, achieved by applying the Otsu's algorithm and refined by using the previous four steps, is used to obtain the matrix of distances of each pixel to the background by exploiting the  $\ell_2$  Euclidean distance [60] (with a  $5 \times 5$  pixel mask for a more accurate distance estimation). This algorithm calculates the distance to the closest background pixel for each pixel of the source image. Let  $\mathcal{G}$  be a regular grid and  $f : \mathcal{G} \rightarrow \mathbb{R}$  an arbitrary function on the grid, called a sampled function [181]. We define the distance transform  $\mathcal{D}_f : \mathcal{G} \rightarrow \mathbb{R}$  of  $f$  as:

$$\mathcal{D}_f(\mathbf{p}) = \min_{\mathbf{q} \in \mathcal{G}} (d(\mathbf{p}, \mathbf{q}) + f(\mathbf{q})), \quad (3.10)$$

where  $d(\mathbf{p}, \mathbf{q})$  is a measure of the distance between the pixels  $\mathbf{p}$  and  $\mathbf{q}$ . Owing to the fact that cells have a pseudo-circular shape, we used the Euclidean distance, so achieving the Euclidean distance transform (EDT) of  $f$ . In the case of binary images, with a set of points  $\mathcal{P} \subseteq \mathcal{G}$ , the distance transform  $\mathcal{D}_{\mathcal{P}}$  is a real-valued image of the

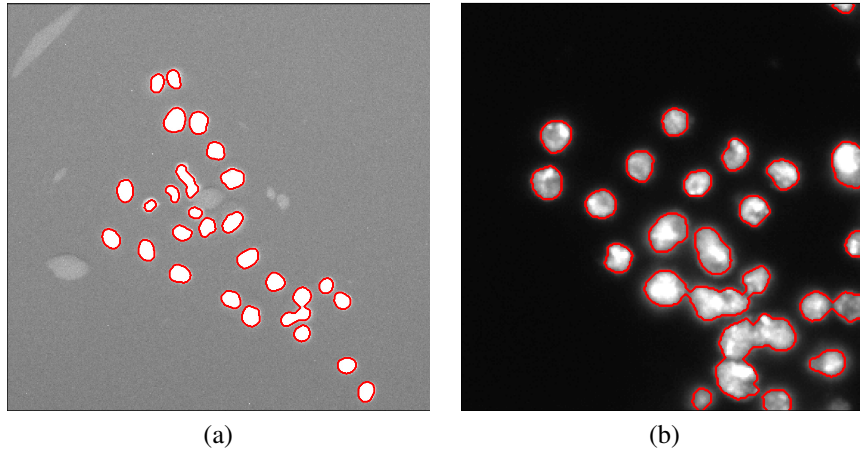


Fig. 3.19 Segmented images after the refinement steps, detail of: the (a) result on the image shown in Fig. 3.18a; the (b) result on the image shown in Fig. 3.18b.

same size:

$$\mathcal{D}_{\mathcal{P}} = \min_{\mathbf{q} \in \mathcal{P}} (d(\mathbf{p}, \mathbf{q}) + 1(\mathbf{q})), \quad (3.11)$$

where:

$$1(\mathbf{q}) = \begin{cases} 0, & \text{if } \mathbf{q} \in \mathcal{P} \\ \infty, & \text{otherwise} \end{cases}$$

is an indicator function for the membership in  $\mathcal{P}$  [181]. The computed distance map is normalized by applying contrast linear stretching to the full 8-bit dynamic range;

7. regional maxima computation allows for estimating foreground peaks on the normalized distance map. Regional maxima are connected-components of pixels with a constant intensity value, and whose external boundary pixels have all a lower value [571]. The resulting binary mask contains pixels that are set to 1 for identifying regional maxima; all other pixels are set to 0. A  $5 \times 5$  pixel square was employed as structuring element;
8. morphological dilation (using a 3-pixel radius circular structuring element) is applied to the foreground peaks previously detected for better defining the foreground regions and merging neighboring local minima into a single seed point.

The segmentation results on the Figs. 3.18a and 3.18b are shown in Figs. 3.19a and 3.19b, respectively. The detail in Fig. 3.18a shows that ACDC is highly specific to cell nuclei detection, so discarding non-cell regions related to acquisition artifacts.

**Cell nuclei segmentation using the watershed transform** Although the watershed transform can detect also weak edges, it often may not accurately detect the edge of interest in the case of blurred boundaries [223]. This sensitivity to noise could be worsened by the use of high pass filters to estimate the gradient and the edges, which amplify the noise. We address this issue by formerly applying the bilateral filter that reduces the halo effects [614]. Accordingly, we developed the following steps:

1. Connected-component labeling [598] of the foreground region binary mask for encoding the markers employed in the following watershed algorithm;
2. Laplacian operator to produce the edge image [218] to feed the edge map as input to the watershed transform;
3. Watershed segmentation on the edge image according to the previously defined markers [409, 431].

**Implementation details** The sequential version of ACDC has been entirely developed using the Python programming language (version 2.7.12), exploiting the following libraries and packages: NumPy, SciPy, OpenCV, scikit-image [626], and Mahotas [120]. The resulting processing pipeline exploits classical image processing techniques in a smart fashion [524], so enabling an efficient and feasible solution in time-lapse microscopy environments.

For laboratory feasibility purposes, an asynchronous job queue, based on a distributed message passing paradigm—namely Advanced Message Queuing Protocol (AMQP)—was developed using Celery [90] (implementing workers that execute tasks in parallel) and RabbitMQ [471] (exploited as a message broker to handle communications among workers) for leveraging modern multi-core processors and computer clusters.

**Results** Fig. 3.20 shows two pairs of cell detection results on the VU dataset and the 2018 DSB training dataset, respectively. The detected cells are displayed with different colors to highlight the separation among overlapping and merging cells. The analyzed images are characterized by a considerable variability in terms of low or high cell density.

The assessment on cell count is quantified by means of the Pearson's coefficient to measure the linear correlation between the automated and manual cell counts. Table 3.5 reports the results for both the analyzed datasets. These results are supported by the scatter plots in Fig. 3.21. The accuracy of the achieved segmentation results were quantitatively evaluated with respect to the real measurement by using the Intersection over Union (*IoU*) metrics, also known as *JJ* (see Appendix A.1).

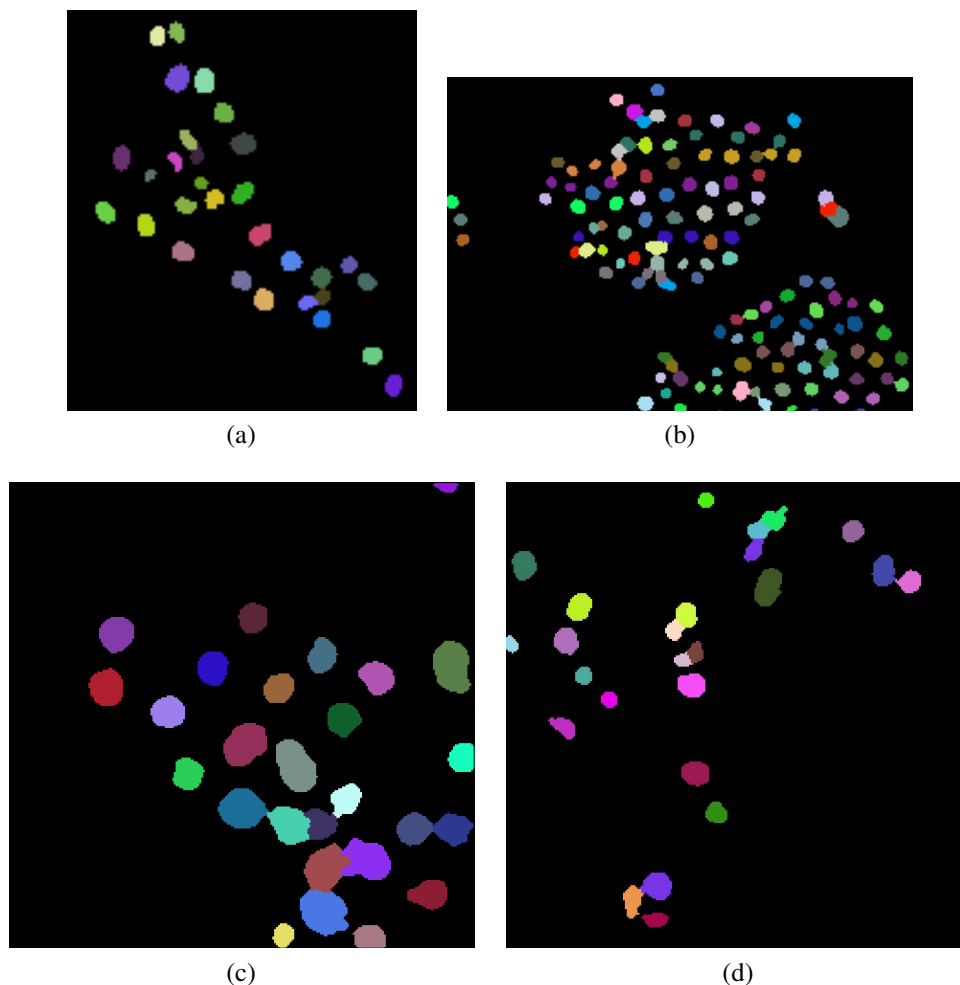


Fig. 3.20 Examples of cell nuclei segmented by ACDC: (a) and (b) show a detail of the results achieved on the images in Figs. 3.15a and 3.15b, respectively; (c) and (d) illustrate the results achieved on the images in 3.16a and 3.16b, respectively.

The achieved high  $\rho$  coefficient values confirm the effectiveness of the proposed approach, according to a validation performed against the expert's gold standard. A strong positive correlation between the cell counts computed automatically by ACDC and the manual measurements was observed for both datasets analyzed in this study. In particular, in the case of the VU dataset, the automated cell counting is strongly consistent with the corresponding manual measurements, by denoting also a unity slope, as shown in Fig. 3.21a. In the case of the 2018 DSB dataset (Fig. 3.21b), the regression straight-line reveals a negative offset in the ACDC measurements. This finding means that the cell counts achieved by ACDC slightly under-estimated the gold standard in approximately 55% of cases. In addition to the high variability of the fluorescence images included in the 2018 DSB dataset, this systematic error often depends on the gold standard binary masks that consider also partial

Table 3.5 Evaluation metrics on cell counting and segmentation achieved by ACDC on the analyzed time-lapse microscopy datasets. The results for the *IoU* metrics and the execution time measurements are expressed as mean value  $\pm$  standard deviation.

Dataset	Pearson's coeff.	IoU (%)	Exec. time [s]
Vanderbilt University	$\rho = 0.99974$	$62.84 \pm 8.58$	$7.49 \pm 0.30$
2018 Data Science Bowl	$\rho = 0.96121$	$80.37 \pm 10.58$	$0.12 \pm 0.09$

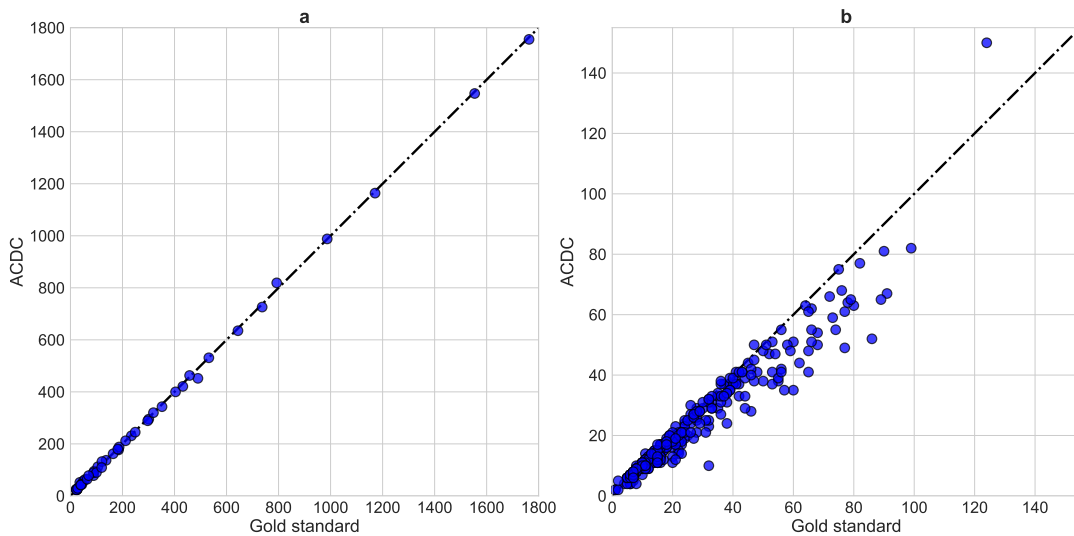


Fig. 3.21 Scatter plots depicting the ACDC results compared to the gold standard in terms of cell nuclei counting in the case of: (a) time-lapse fluorescence images from the VU; (b) selection of fluorescence images from 2018 DSB. The equality line through the origin is drawn as a dashed line.

connected-components of cell nuclei smaller than 40 pixels located at the image borders. Notice that ACDC's settings were not completely modified, so some very small partial cell components were removed according to the morphological refinements based on the connected-component size.

The *IoU* mean values calculated on the VU images reveals a good agreement between the segmentation results achieved by ACDC and the gold standard. The standard deviation values confirm the high variability encountered in the input datasets. In particular, in the case of the VU images, this evidence strongly depends on the density of the cells represented in the input image. As a matter of fact, the *IoU* metrics is highly affected by the size of the foreground regions with respect to the background. This behavior is confirmed by the 2018 DSB results, where the *IoU* values are considerably higher than those achieved on the VU images, even though the Pearson's coefficient is slightly lower in this case. Accordingly, the

high standard deviation for the DSB dataset is due to the intrinsic variety of the images—in terms of image size, zoom factor, and specimen type—included in this dataset (see Section 3.2.1).

The mean execution times concerning the segmentation tests are shown in Table 3.5. These experiments were run on a personal computer equipped with a quad-core Intel® Core™ 7700HQ (clock frequency 3.80 GHz), 16 GB RAM, and Ubuntu 16.04LTS operating system. The computational efficiency of ACDC is confirmed in both cases, completing the cell detection and counting tasks by coping with the time constraints imposed by high-throughput laboratory routine. As expected, the execution times are dependent on the image size. This trend is mainly due to the bilateral filtering operation.

To conclude, considering the two experimental datasets, ACDC achieved accurate results in terms of cell counting as well as segmentation accuracy. Therefore, ACDC showed to be a reliable and efficient solution even when dealing with multiple datasets characterized by significant variations, in terms of acquisition devices, specimens, experimental conditions, and imaging characteristics.

**Discussion and conclusion** A novel method for cell detection and counting, called ACDC, was proposed, by exploiting a fully automatic pipeline based on the watershed transform [573, 638] and morphological filtering operations [51, 571]. We benefited also from the edge-preserving smoothing achieved by the bilateral filtering [614]. We tested our approach on two different cell imaging datasets characterized by significantly different acquisition and experimental conditions. ACDC was shown to be accurate and reliable, thus representing a laboratory feasible solution also thanks to its computational efficiency. The current ACDC implementation can also distribute the computation on multi-core architectures and computer clusters to further reduce the running time required to analyze large single image stacks. To this end, an asynchronous job queue, based on a distributed message passing paradigm was developed exploiting Celery and RabbitMQ.

As future developments, we aim at exploiting the most recent Machine Learning techniques [336] to accurately refine the segmentation results. The improvements may be achieved by classifying the geometrical and textural features extracted from the detected cells [658]. This advanced computational analysis can allow us to gain biological insights in complex cellular processes [21]. As a biological application in the near future, we plan to focus on the Fluorescent, Ubiquitination-based Cell-Cycle Indicator (FUCCI) reporter system for investigating cell-cycle states [538], by combining accurate nuclei segmentation results with live cell imaging of cell cycle and division.

# Chapter 4

## Pattern Recognition techniques

### 4.1 Unsupervised fuzzy clustering

The FCM algorithm [52, 53] is a partitional clustering technique, i.e., an input data set is optimally classified in sub-groups [292]. Intuitively, a cluster collects a group of data points whose pairwise distances are lower with respect to the distances from points outside the cluster.

Thus, the FCM algorithm operates by minimizing the intra-cluster variance as well as by maximizing the inter-cluster variance, in terms of a distance metrics between the feature vectors [361]. These characteristics make this unsupervised clustering method suitable for pattern recognition and image segmentation, by representing each image region with a cluster. In addition, FCM is highly robust to initial conditions when well-separable data are analyzed [142]. The FCM technique search for the optimal partition of a data set  $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  composed of  $N$  feature vectors, which denote data samples  $\mathbf{x}_k \in \mathbb{R}^D$  belonging to a  $D$ -dimensional Euclidean space, into exactly  $C$  clusters (i.e., non-empty partitions of the input dataset), as introduced in [52, 53]. In Pattern Recognition, features are defined as measurable quantities that could be used to classify different regions [373]. An input dataset is partitioned into groups (regions), and each of them is identified by a centroid. Distance metrics are used to group data into clusters of similar types and the number of clusters is assumed to be known *a priori* [213, 361]. These feature vectors can be conveniently represented by a data matrix  $\mathbf{X} \in \mathbb{R}^{N \times D}$  with  $N$  rows containing the input  $D$ -dimensional feature vectors. Formally, a partition  $\mathcal{P}$  is defined as a set family  $\mathcal{P} = \{\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_C\}$ . The resulting clusters are always represented by connected-regions with unbroken edges. The goal of unsupervised clustering methods is to determine an intrinsic partitioning in a set of unlabeled data, which are associated with a feature vector [292].

The crisp  $K$ -means algorithm [368] imposes that the clusters must be a proper subset of  $\mathcal{X}$  ( $\emptyset \subset \mathcal{Y}_i \subseteq \mathcal{X}, \forall i$ ) and their set union must reconstruct the whole dataset ( $\bigcup_{i=1}^C \mathcal{Y}_i = \mathcal{X}$ ). Moreover, the various clusters are mutually exclusive ( $\mathcal{Y}_i \cap \mathcal{Y}_j = \emptyset, \forall i \neq j$ ), i.e., each feature vector may belong to only one group. Differently from the  $K$ -means algorithm [309] that yields a hard clustering output, in which each feature vector is strictly assigned to one group, the FCM algorithm defines a fuzzy partition  $\mathcal{P} = \{\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_C\}$  using a fuzzy set family [691]. Compared to the iterative EM algorithm [148, 423], which estimates the parameters of a Gaussian mixture model relying on a probabilistic framework [434], the  $K$ -Means splits the data into Voronoi cells of the cluster centroids.

The FCM scheme, which employs Fuzzy Logic [692], results in a soft computing solution allowing for partial membership to multiple classes [623] with varying degree [698]. Accordingly, the matrix  $\mathbf{U} \in \mathbb{R}^{C \times N}$  denotes a fuzzy  $C$ -partition of the set  $\mathcal{X}$  using  $C$  membership functions  $\mu_i : \mathcal{X} \rightarrow [0, 1]$ , whose values  $u_{ik} := \mu_i(\mathbf{x}_k) \in [0, 1]$  represent the similarity membership of each element  $\mathbf{x}_k$  to the  $i$ -th fuzzy set (i.e., the cluster  $\mathcal{Y}_i$ ) and have to meet the following conditions:

$$\begin{cases} 0 \leq u_{ik} \leq 1 \\ \sum_{i=1}^C u_{ik} = 1, \forall k \in \{1, \dots, N\} \\ 0 < \sum_{k=1}^N u_{ik} < N, \forall i \in \{1, \dots, C\} \end{cases} . \quad (4.1)$$

To clarify, the sets of all hard and fuzzy  $C$ -partitions of the input dataset  $\mathcal{X}$  are defined by 4.2 and 4.3, respectively:

$$\mathcal{M}_{\text{hard}} = \left\{ \mathbf{U} \in \mathbb{R}^{C \times N} : u_{ik} \in \{0, 1\} \right\}, \quad (4.2)$$

$$\mathcal{M}_{\text{fuzzy}} = \left\{ \mathbf{U} \in \mathbb{R}^{C \times N} : u_{ik} \in [0, 1] \right\}, \quad (4.3)$$

where  $\mathcal{M}_{\text{fuzzy}}$  improves the information conveyed by  $\mathcal{M}_{\text{hard}}$ , allowing for more flexibility than the crisp  $K$ -Means approach [281]. As a matter of fact, introducing the fuzzy sets for the partial membership of feature vectors keeps more information from the original image compared to the hard segmentation methods [467].

The key concept underlying the FCM clustering can be formalized by introducing a set of  $D$ -dimensional prototype vectors  $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_C\}$ , called centroids, which are associated with the  $C$  clusters. The membership function values are assigned to each feature vector in the data matrix  $\mathcal{X}$  according to the relative distance of each feature vector  $\mathbf{x}_k$  from the



prototype points in  $\mathcal{V}$ . The centroids are not forced to coincide with input data points. So, this clustering algorithm aims to optimize the objective function (also called distortion measure) in Eq. (4.4), representing a generalized least-squares error function:

$$\mathcal{J}_m(\mathbf{U}, \mathcal{V}; \mathbf{X}) = \sum_{i=1}^C \sum_{k=1}^N (u_{ik})^m \cdot d_{ik}, \quad (4.4)$$

where:

- $m \in [1, \infty)$  is the fuzzification constant that weighs the fuzziness of the classification process. If  $m = 1$ , the FCM algorithm approximates the crisp  $K$ -means version. The role of the weighting exponent  $m$  in the FCM model was analyzed in [451] by evaluating the quality of inferences about the validity of FCM  $(\hat{\mathbf{U}}, \hat{\mathcal{V}})$  pairs. The authors suggested that the best choice for  $m$  is in the interval  $[1.5, 2.5]$ , whose mean and midpoint  $m = 2$  is the most used value;
- $d_{ik} = \|\mathbf{x}_k - \mathbf{v}_i\|^2$  is the squared distance between the elements  $\mathbf{x}_k$  and  $\mathbf{v}_i$ , computed by means of an induced norm  $\|\cdot\|$  on  $\mathbb{R}^D$  (usually the Euclidean  $\ell_2$  norm).

Therefore, the data matrix  $\mathbf{X}$  is partitioned by iteratively searching for the optimal fuzzy partition  $\mathcal{P}$ , denoted by the pair  $(\hat{\mathbf{U}}, \hat{\mathcal{V}})$ , which minimizes the objective function  $\mathcal{J}_m$  by means of a local optimization technique. During the  $t$ -th iteration, the prototype set  $\hat{\mathcal{V}}^{(t)}$ , storing the centroid values  $\hat{\mathbf{v}}_i^{(t)}$ , as well as the elements of the matrix  $\hat{\mathbf{U}}^{(t)}$  are estimated and updated according to Eqs. (4.5) and (4.6), respectively:

$$\hat{\mathbf{v}}_i^{(t)} = \frac{\sum_{j=1}^N (\hat{u}_{ij}^{(t)})^m \mathbf{x}_j}{\sum_{j=1}^N (\hat{u}_{ij}^{(t)})^m} \quad (4.5)$$

$$\hat{u}_{ik}^{(t)} = \left[ \sum_{j=1}^C \left( \frac{\|\mathbf{x}_k - \hat{\mathbf{v}}_i^{(t)}\|}{\|\mathbf{x}_k - \hat{\mathbf{v}}_j^{(t)}\|} \right)^{\frac{2}{m-1}} \right]^{-1}, \text{ with } m > 1 \text{ and } \mathbf{x}_k \neq \hat{\mathbf{v}}_j^{(t)} \forall j, k. \quad (4.6)$$

Afterwards, each object  $\mathbf{x}_k$  is compared with the elements of the centroid vector and is then mapped to the nearest cluster. The iterative procedure ends when the convergence condition (i.e., the minimum improvement in objective function between two consecutive iterations  $\hat{\mathbf{U}}^{(t+1)}$  and  $\hat{\mathbf{U}}^{(t)}$  is less than a fixed tolerance value  $\varepsilon_{\text{tol}}$ ) or the maximum number of allowed iterations  $T_{\text{max}}$  is achieved.

---

**Algorithm 3** Pseudo-code of the FCM clustering algorithm.

---

**Input:** Data matrix  $\mathbf{X} \in \mathbb{R}^{N \times D}$  that contains the  $N$  feature vectors

**Output:** Fuzzy partition  $\mathcal{P}$  into  $C$  clusters  $\mathcal{M}_{\text{fuzzy}} = \{\mathbf{U} \in \mathbb{R}^{C \times N} : u_{ik} \in [0, 1]\}$ , identified by the prototype set  $\mathcal{V}$

```

    ▷ Initialization using a uniform pseudorandom number generator Uniform(0, 1)
1:  $u_{ij} \sim \text{Uniform}(0, 1), \forall u_{ij} \in \mathbf{U}^{(0)}$ 
2:  $\mathbf{v}_{i,k} \sim \text{Uniform}(0, 1), \forall \mathbf{v}_i \in \mathcal{V}^{(0)}, k \in \{1, 2, \dots, C\}$ 
    ▷ Iterate until convergence is not achieved or  $T_{\text{max}}$  iterations are executed
3: while  $\left( \left| \mathcal{J}_m^{(t)} - \mathcal{J}_m^{(t-1)} \right| > \epsilon_{\text{tol}} \text{ or } t \leq T_{\text{max}} \right)$  do
4:   for each  $\mathbf{x}_k \in \mathcal{X}$  do
5:     UPDATECENTROIDS( $\hat{\mathbf{V}}^{(t)}$ )           ▷ Update the centroids  $\hat{\mathbf{V}}^{(t)}$  Eq. (4.5)
6:     UPDATEMEMBERSHIPS( $\hat{\mathbf{U}}^{(t)}$ )       ▷ Update the membership matrix  $\hat{\mathbf{U}}^{(t)}$  Eq. (4.6)
7:   end for
8:   UPDATEOBJFUNC( $\mathcal{J}_m^{(t)}$ )           ▷ Update the objective function  $\mathcal{J}_m^{(t)}$  Eq. (4.4)
9:    $\hat{\mathbf{V}}^{(t)} \leftarrow \hat{\mathbf{V}}^{(t+1)}$ 
10:   $\hat{\mathbf{U}}^{(t)} \leftarrow \hat{\mathbf{U}}^{(t+1)}$ 
11:   $\mathcal{J}_m^{(t)} \leftarrow \mathcal{J}_m^{(t+1)}$ 
12:   $t \leftarrow t + 1$ 
13: end while

```

---

### 4.1.1 Brain tumor segmentation in neuro-radiosurgery

Hereby, to provide the proper prerequisites on brain lesion segmentation in neuro-radiosurgery, the semi-automatic GTV delineation method published in [413, 525] is recalled. The operator-dependence is minimized, because user intervention is just limited to select an area containing the tumor and no parameter setting is required. As illustrated in [297], semi-automated approaches provide more reproducible measurements compared to conventional manual tracing. This was proved by quantitatively comparing intra- and inter-operator reliability of the two methods. The flow diagram of the overall segmentation method is shown in Fig. 4.1.

#### Dataset description

The study was performed on a dataset made up of 15 patients with brain tumors to be treated with Leksell Gamma Knife<sup>®</sup> model C at Cannizzaro Hospital (Catania, Italy). All the MRI series were acquired on a Gyroscan Intera 1.5 T MRI scanner (Philips Medical System Eindhoven, the Netherlands), before treatment, for the planning phase. The total number of processed lesions was 20 and the average age of the examined patients was  $57.80 \pm 10.75$  years. MR images are T1w Fast Field Echo (“T1w FFE”) CE sequences. Imaging data were encoded in the 16-bit DICOM format. Thanks to the Gadolinium-based contrast agent,

Table 4.1 MRI acquisition parameters of the brain tumor dataset.

MRI sequence	TR [ms]	TE [ms]	Matrix size [pixels]	Slice spacing [mm]	Slice thickness [mm]	Pixel spacing [mm]
T1w FFE	25	1.808-3.688	256 × 256	1.5	1.5	0.977-1.0

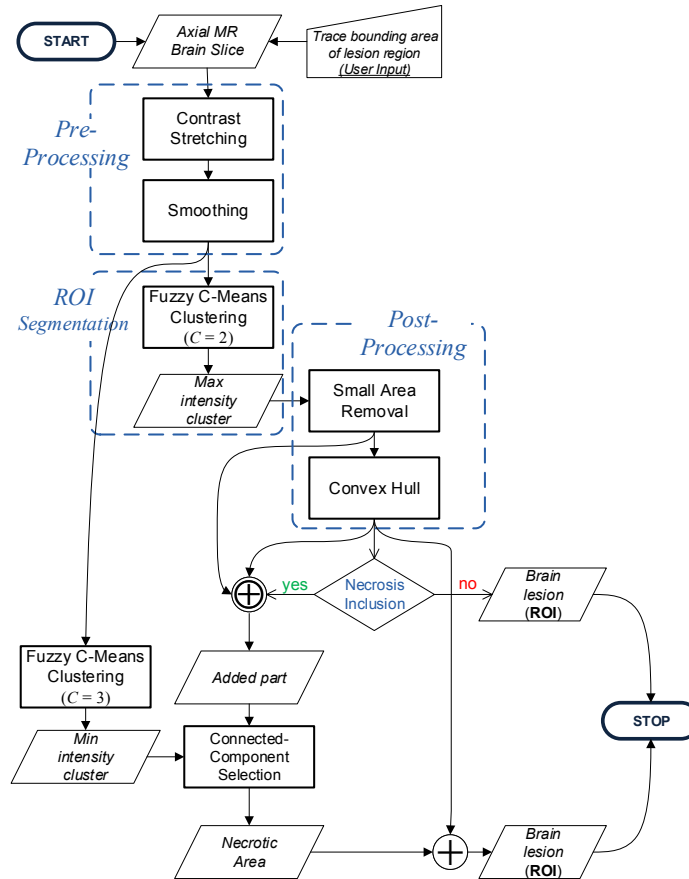


Fig. 4.1 Flow diagram of the FCM-based brain lesion segmentation method.

brain lesions appeared as hyper-intense zones. Sometimes a dark area might be present due to necrotic tissue. The main characteristics of the input data are reported in Table 4.1. This retrospective research study had no implication on the Leksell Gamma Knife<sup>®</sup> treatments involving the enrolled patients. No information concerning the subjects was accessed; consequently, institutional review board approval was not sought: the proposed image analysis is performed off-line and thus did not change the current treatment protocol.

These MRI series contain varying levels of segmentation difficulty, with the worst scans consisting of low contrast and large intensity inhomogeneities. Lesions with a non-homogeneous enhancement region, tumors with irregular shape or with inner necrosis are also present. Therefore, these case studies represent a meaningful sample of the possible real imaging data. Representative samples of input brain MR images are shown in Fig. 4.2.

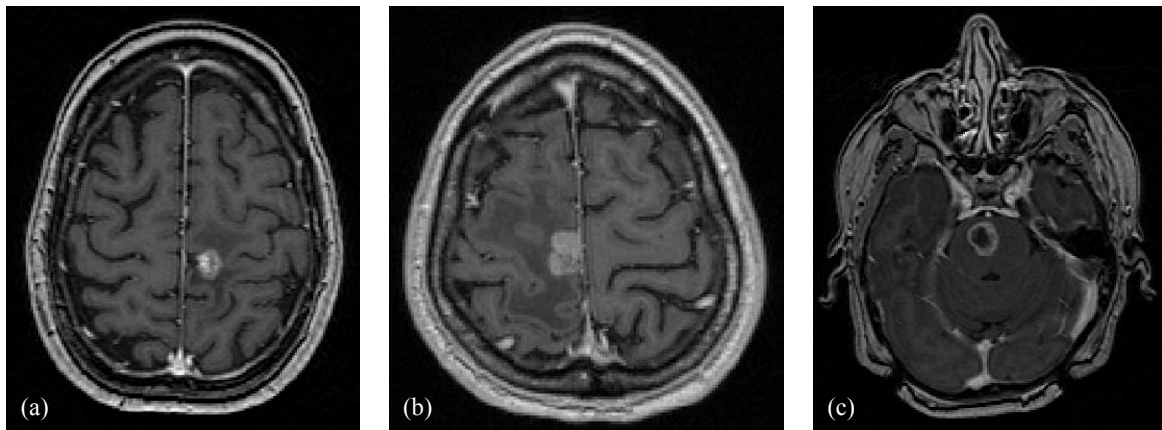


Fig. 4.2 Examples of brain MR axial slices: (a) lesion with non-homogeneous enhancement region; (b) irregular tumor shape; (c) tumor with inner necrosis (dark area).

### FCM-based brain tumor segmentation

**ROI selection tool** The area containing the lesion must be localized within the brain by the operator. Indeed, the users of medical imaging applications select in the image a specific area, i.e., drawing a boundary around the ROI.

Despite the fact that several selection tools are available in current clinical software, often they implement unnecessary or inefficient functions. On the basis of the evidence reported in [98], rectangular and free-hand ROI selection tools (bounding region of the lesion) were developed and tested. When the lesion is located near the cerebral cortex, a rectangular shaped selection tool could include some anatomic parts characterized by similar gray levels (i.e., skull bones or cerebral cortex), as shown in Fig. 4.3a. Thereby, the clustering algorithm may include, erroneously, these areas into the segmented ROI. Problems with an ellipsoidal selection tool also occur, if the lesion is near the superior sagittal sinus (both anterior and posterior extremities), as shown in Fig. 4.3b. Moreover, rectangular or circular settings [525] (i.e., resizing and moving) are more complex than free-hand selection during lesion enclosing under critical conditions. In order to avoid this issue, a free-draw selection enables a more precise contour containing the tumor [413]. In this way, result repeatability is ensured regardless of the form of free-shape bounding region. As a result, the operator-dependence of the proposed segmentation approach is reduced. After the freehand drawing, this selected part is then enclosed and zero-padded into the minimum box in order to deal with a rectangular image during the following processing phases.

**Pre-processing** In order to improve segmentation results, some pre-processing operations are applied on masked input images after the ROI selection by the user. First of all, the range

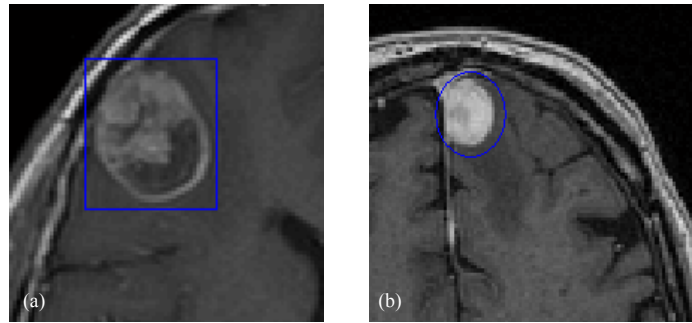


Fig. 4.3 ROI selection tools for brain tumor segmentation: (a) rectangular versus (b) ellipsoidal regions.

of intensity values of the selected part is expanded to extract the ROI more easily. Indeed, this operations is performed by means of unsupervised clustering. Thereby, a linear contrast stretching operation is applied.

**FCM-based ROI segmentation** The FCM algorithm is applied just on pixels included in the drawn bounding region and based upon image intensity levels. Therefore, the zero-valued background pixels, added with zero-padding, are never considered. Two clusters are employed ( $C = 2$ ) during this step in order to suitably distinguish hyper-intense lesion from the healthy part of the brain. As a matter of fact, in contrast-enhanced MR brain scans, brain metastases have a brighter core than periphery and distinct borders from the surrounding normal brain tissue [18]. Finally, during the defuzzification process, the pixels that have the maximum membership with the brightest cluster are selected.

**Post-processing** In some cases, central areas of the lesions could comprise necrotic tissue. During the treatment planning phase, also these necrotic areas (hypo-intense regions) must be included in the target volume. In order to delete any unwanted connected-components included by the highest intensity cluster, a small area removal operation is employed. These small regions may be due to anatomic ambiguities when the lesion is situated near high valued pixels, i.e. cranial bones or the corpus callosum. Brain lesions have a nearly spherical or pseudospherical appearance [18]. Owing to this fact, a convex hull algorithm [707] is used to envelope the segmented lesion into the smallest convex polygon that contains the automatically segmented region. Furthermore, this operation comprises also hole filling, in order to include any necrosis inside the segmented ROI. Even when the FCM clustering output is an arch-shaped ROI, perhaps due to poor contrast images, the internal necrosis is properly included by the convex hull execution.

**Necrotic area inclusion** Sometimes, because of images with necrotic tumors and characterized by low contrast, the FCM clustering process might not extract a region such that the following post-processing is not able to include internal or adjacent necrosis. A particular sequence of steps must be thus defined:

1. to detect the necrosis, imaged as a hypo-intense area, the FCM clustering is again applied on the pre-processed image, but this time with  $C = 3$  clusters. Thereby, the pixels that have the highest membership grade with the minimum intensity cluster are selected;
2. logical exclusive disjunction (XOR) is calculated on the ROI masks before and after convex hull algorithm execution. The added part by the convex hull operation is thus identified;
3. the only connected-component, whose intersection with the added part identified by Step 2 has a non-zero value, is selected and represents the desired necrotic area;
4. lastly, the necrotic area is added to the previous post-processing output (by means of logical OR between the two binary masks), and this represents the resulting binary mask.

A very critical case is reported in Fig. 4.4c, where the necrotic hypo-intense area is adjacent to the hyper-intense region of the tumor. The necrosis is thus correctly included in the resulting output. As shown in the flow diagram concerning the proposed method (Fig. 4.1), the physician can select, through a check button, the execution of the necrotic area inclusion steps mentioned above when considered as appropriate.

## Experimental results

To assess the effectiveness of the proposed approach, several comparative measures between the automatically segmented volume, obtained by the proposed method, and those manually drawn by a neurosurgeon (considered as our gold standard) are provided (for more details see Appendix A.1). Fig. 4.5 shows the segmentation results of the three MR slices reported in Fig. 4.2.

The achieved mean values of evaluation metrics (calculated on each dataset) are reported in the boxplots (Fig. 4.6). The short length of the boxplot in Fig. 4.6a means that values are very concentrated. All index values do not present outliers, thus demonstrating extremely low variability (whisker value is 1 in all cases). Fig. 4.6b depicts the boxplots of distance-based evaluation metrics. Even though there are outliers in boxplots (the whisker is set to 1 pixel,

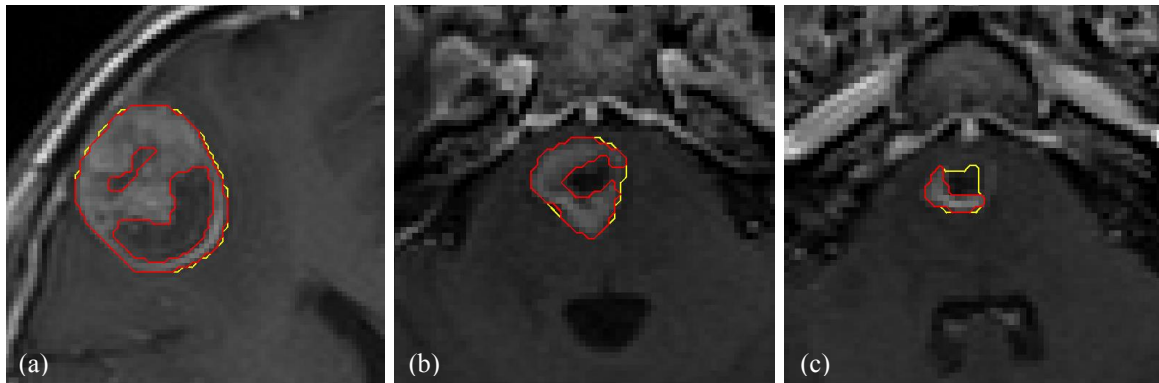


Fig. 4.4 Necrotic area handling examples: (a) ROI segmentation is refined using post-processing (internal necrotic “holes” are filled); (b) arch-shaped lesion is correctly segmented with post-processing steps; (c) necrosis is included using the post-processing steps. Both the boundaries without (red contour) and with the refinements (yellow contour) are superimposed (all images are zoomed by a  $4\times$  factor).

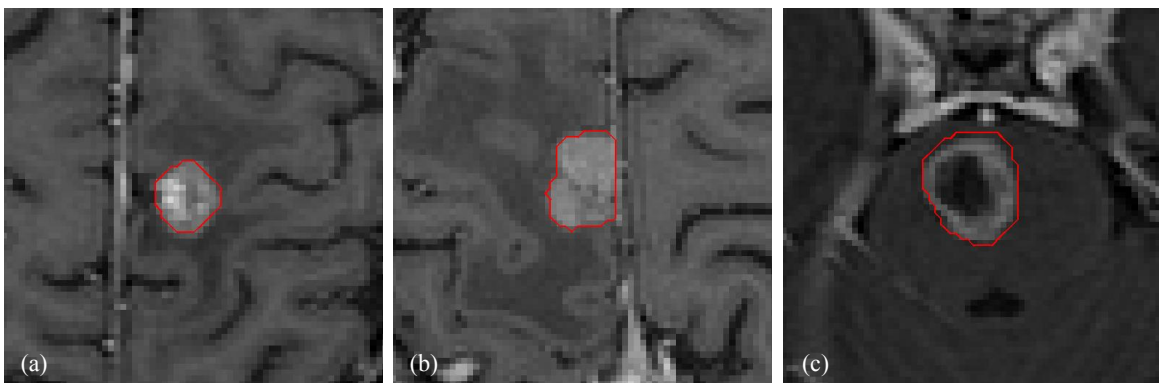


Fig. 4.5 Results achieved by the FCM-based segmentation method on the input images shown in Fig. 4.5(a-c), respectively.

which is the smallest imaging discrete element), they still indicate only a small deviation between the segmentations of the proposed method and those of the radiologist. Considering the surgery purpose of the application, referring to MRI spatial resolution (see characteristics in Table 4.1), all distance-based metrics mean values are in the range of 1 mm. This confirms the great validity and accuracy of the proposed segmentation method.

Fig. 4.7 displays two instances of tridimensional reconstruction of the tumors. In both cases, the patient’s head (extracted using a skull stripping algorithm) is represented by means of a transparent surface (with  $\alpha = 0.45$ ), and internal solid tumors (red volumetric models) are visible in their own actual position with respect to the tridimensional reference system.

For comparison, we developed four different semi-automatic segmentation approaches:

- edge detection [79], after bounding region selection;

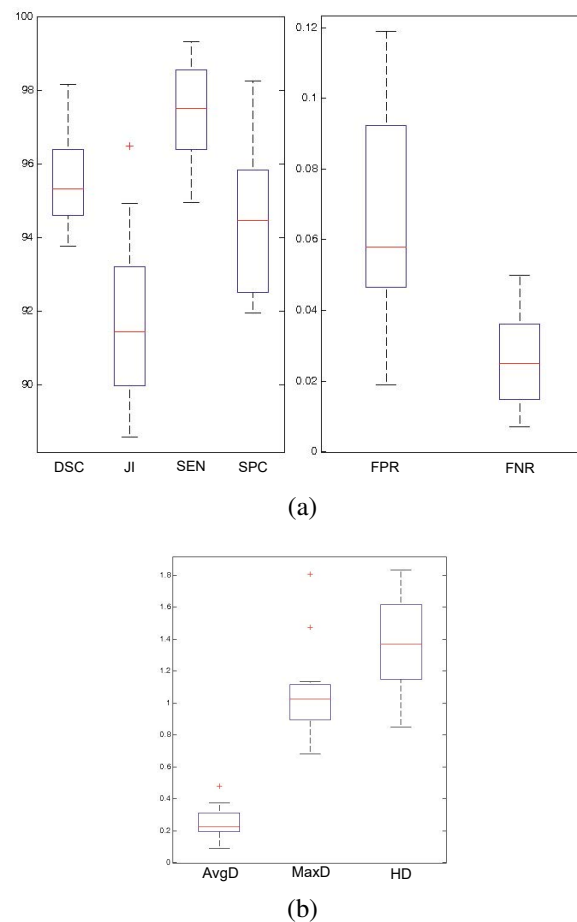


Fig. 4.6 Boxplots of the evaluation metrics results of the FCM-based brain tumor segmentation approach: (a) overlap-based metrics; (b) distance-based metrics.

- adaptive thresholding [448], after bounding region selection;
- region growing [416], after seed-point selection;
- interactive level set implementing the Distance Regularized Level Set Evolution (DRLSE) method [357], after initial region selection.

All the implementations were tested on our MRI datasets. The Canny method fails to successfully segment brain tumors because false edges are wrongly selected when non-homogeneous lesions are processed. Moreover, sometimes a totally unbroken edge is not returned, due to not connected edge curves. Lesion segmentation problems might occur with arch-shaped lesions, where an internal necrosis is present. Necrotic area inclusion is more difficult with respect to region-based segmentation. Even edge-linking [212] does not correctly detect the lesion. Since the FCM algorithm classifies input data, the achieved clusters are represented by connected-regions and the corresponding boundaries are closed



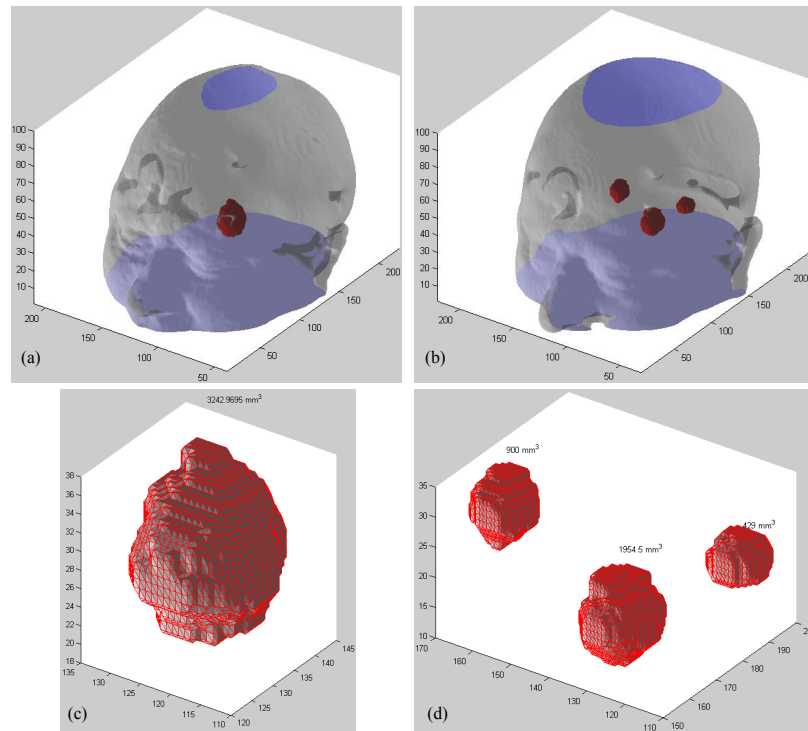


Fig. 4.7 Volume rendering of GTVs: (a) treated patient with single cerebral lesion, shown in detail in (c); (b) multi-metastasis patient, shown in detail in (d). In both cases, the patient's head is represented by means of a transparent surface (with  $\alpha = 0.45$ ), and internal solid tumors (red volumetric models) are visible in their own actual position with respect to the tridimensional reference system. Sub-figures (c) and (d) report also the volume (in  $\text{mm}^3$ ) of each segmented lesion.

even when holes are present. This fact is very important in radiotherapy or radiosurgery treatment planning, because lesion segmentation is a critical step that must be performed accurately and effectively.

For these reasons, we quantitatively compared the segmentation results obtained by our method to the aforementioned segmentation algorithms able to detect a totally unbroken edge (i.e., adaptive thresholding, region growing, interactive level set). Considering this, the Canny method was not considered in the comparison. All the developed approaches implement the same pre-processing and post-processing steps provided by the proposed segmentation method. Values of the main overlap-based and distance-based metrics, calculated for each segmentation method, are reported in Table 4.2.

Several difficulties arose during the trials of the developed segmentation approaches. Using adaptive thresholding, it is not possible to handle necrotic area inclusion not even with post-processing operations. In fact, the hypo-intense area cannot be successfully extracted neither through another thresholding operation. On the other hand, leaking occurs with

Table 4.2 Comparison among the developed brain lesion segmentation methods on the same 15 patients treated with Gamma Knife neuro-radiosurgery. The results are expressed as average value  $\pm$  standard deviation.

Method	DSC	JI	SEN	SPC	AvgD	MaxD	HD
Otsu's method [448]	79.59 $\pm$ 6.62	68.78 $\pm$ 8.36	90.49 $\pm$ 8.62	73.33 $\pm$ 7.65	1.179 $\pm$ 0.77	2.534 $\pm$ 1.17	1.910 $\pm$ 0.37
Region growing [416]	79.78 $\pm$ 10.93	70.72 $\pm$ 12.67	77.64 $\pm$ 11.74	87.61 $\pm$ 8.54	1.079 $\pm$ 1.02	2.425 $\pm$ 2.04	1.696 $\pm$ 0.28
Interactive Level Set [357]	65.16 $\pm$ 18.06	53.47 $\pm$ 19.05	76.06 $\pm$ 11.88	67.97 $\pm$ 23.96	2.291 $\pm$ 1.91	4.720 $\pm$ 3.39	2.021 $\pm$ 0.40
FCM [413]	95.59 $\pm$ 1.28	91.86 $\pm$ 2.28	97.39 $\pm$ 1.42	94.30 $\pm$ 2.04	0.246 $\pm$ 0.10	1.050 $\pm$ 0.28	1.365 $\pm$ 0.31

LSFs in low contrast datasets as well as in initial/final slices of the tumor to be treated. In such cases the actual ROI is over-estimated. The interactive level set method is highly affected by operator dependence due to the initial region selection. In addition, a careful parameter tuning must be performed depending on the either image contrast or the tumor region being processed (in particular, possible necroses or arch-shaped lesions). Weak ROI boundaries may sometimes cause leakage regardless of the parameter settings. For instance, if a tumor is adjacent to skull bones, the ROI could pass through the actual boundaries. This is evidenced by the false positive presence, as demonstrated by the higher sensitivity (*SEN*) than specificity (*SPC*). Region growing also requires a threshold selection according to image features. This task could be error-prone and tedious. However segmentation issues occur when there are non-homogeneties or loosely connected enhanced tumor regions (due to internal necroses). Ambiguities are encountered when the lesion is adjacent to a hyper-intense imaged anatomic area (e.g., sagittal sinus extremities) because the ROI grows towards it. The achieved distance-based metrics are consistent with overlap-based measures, thus confirming the above discussions Overall, the proposed segmentation method obtains the best performance with respect to the other implemented approaches.

## Discussion

A GTV delineation method based on unsupervised FCM clustering was presented. The developed approach was applied to the planning phase of Gamma Knife treatments for brain lesions. To evaluate the performance of the proposed approach, segmentation tests on 15 MRI sequences were performed. The obtained results, in terms of spatial overlap-based and distance-based metrics, demonstrated the great robustness of the proposed approach even when the MRI series were affected by acquisition noise. The good segmentation performance was also confirmed with poor contrast datasets or lesions characterized by irregular shapes. For further discussion, refer to Section 4.2.1.

### 4.1.2 NeXt: necrosis extraction in neuro-radiosurgery

This approach distinguishes necrotic and enhancement regions within the complete tumor region, which includes all tumor structures. Our method is tailored for radiosurgery applications and designed according to the available MRI data used in clinical practice. As a matter of fact, stereotactic neuro-radiosurgery treatment planning is generally performed on T1w CE-MR images alone by experienced neurosurgeons. So, the proposed necrosis extraction approach could represent a feasible solution in clinical environments, wherein multispectral MRI data are generally not available. From a computational perspective, the resulting automated segmentation approach represents an efficient Machine Learning solution.

NeXt should be considered as a second step after the GTV segmentation for a more precise brain cancer characterization. We exploit an accurate brain tumor segmentation, obtained by our validated GTV segmentation methods for supporting neuro-radiosurgery treatment planning in [413, 521]. These computer-assisted segmentation approaches—the former based on the FCM clustering algorithm and the latter on a CA model—have shown to be effective and operator independent solutions. Especially, both GTV segmentation methods were designed *ad hoc* to robustly deal with edemas and necroses, which are imaged differently to healthy brain tissues in MRI. As a matter of fact, these hypo-intense areas must be included during the treatment planning phase in the target volume that is going to receive the radiation dose delivered by the neuro-radiosurgery system. These interactive segmentation algorithms represent a more feasible and safe solution for physicians in clinical practice with respect to fully automatic approaches, as reported in [243]. The overall processing pipeline, built by cascading brain tumor GTV segmentation and necrosis/enhancement region distinction within the yielded GTV, results in an integrated two-stage segmentation scheme to support neuro-radiosurgery therapy.

In practice, necrotic and enhancement region distinction in brain tumors on MR images is clinically valuable in radiation therapy treatment planning as well as in patient's follow-up. Brain tumor necrosis extraction, yielded by NeXt, provides definitely more insights into the cancer nature and characteristics, gaining clinical significance in both radiation treatment planning and therapy outcome assessment to support personalized medicine [513]. As a matter of fact, necrotic areas are generally characterized by hypoxia, which is strongly involved in several aspects of cancer development.

Since tumors are characterized by heterogeneous and biological properties—such as hypoxia, cell density, proliferation, and irregular vascularization—some regions are more treatment-resistant than others. Thus, the potential of dose redistribution can gain relevant information by incorporating estimations of oxygen heterogeneity within imaging voxels for optimal dose determination [466]. In this context, NeXt could be actually exploited for dose

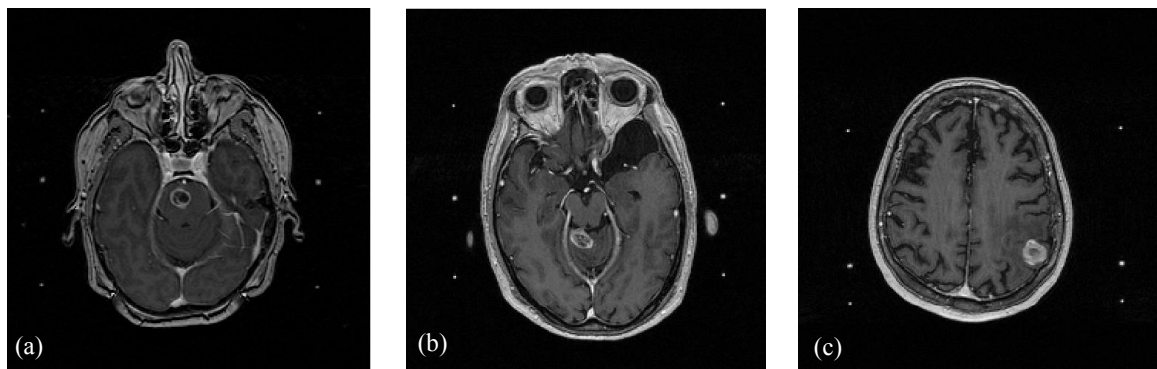


Fig. 4.8 Three instances of input T1w CE-MR axial slices with the rectangular bounding area, selected by the user, containing the tumor: (a) moderately hyper-intense enhancing region including a necrotic core; (b) highly hyper-intense enhancing region with inner necrotic material; (c) heterogenous cancer.

escalation, allowing for a more sophisticated strategy to selectively increase radiotherapy dose (i.e., boosting) in hypoxic radioresistant areas with respect to uniformly treat the whole tumor target volume [173]. NeXt is designed to be directly usable in radiation therapy scenarios, especially for neuro-radiosurgery treatments. In clinical routine, multispectral MRI data are usually not available during neuro-radiosurgery planning phases [243]: conventional MRI protocols for radiation treatment planning include T1w and T1w CE-MR imaging of the head volume [488]. Accordingly, our approach analyzes CE T1w MR images alone.

### Patient dataset description

The study was performed on a dataset composed of 32 brain metastatic cancers that underwent Gamma Knife neuro-radiosurgery. Among them, 20 brain tumors presented necrotic areas. These regions have to be carefully considered during treatment planning, since they include hypoxic areas that could lead to recurrent cancers and resistance to oncological therapies [293]. The characteristics of the analyzed T1w FFE CE-MR images are reported in Section 4.1.1. MRI acquisition parameters are reported in Table 4.1. Three representative examples of input MR images are shown in Fig. 4.8.

Typical MRI protocols, employed in stereotactic radiosurgery treatment planning, include T1w CE-MRI of the head volume [488]. After the administration of the Gadolinium-based contrast medium to the patient, the enhancing hyper-intense zones in the imaged tumors represent the increased and irregular neovascularity. On the other hand, hypo-intense areas depict either cystic or necrotic tissues that do not absorb the contrast medium.

**The proposed automatic necrosis extraction approach**

**Pre-processing** Starting from the preliminary complete tumor region segmentation results that include all tumor structures, obtained by using our previous GTV segmentation methods [413, 520, 521, 525], necrotic parts are distinguished from the tumor enhancing region. The segmented GTV is firstly eroded by means of a selective strategy to choose a suitable structuring element [68]:

- disk with 2-pixel radius, if the GTV area is greater than 80 pixels;
- disk with 1-pixel radius, otherwise.

Using this robust strategy any hypo-intense region in the segmented GTV boundaries is removed, even when the GTV has been overestimated during the preliminary automatic brain tumor segmentation. The input MR brain axial slice is then masked with the corresponding eroded GTV binary mask. This masked image is then linearly stretched in the full dynamic range, exploiting the double precision values in  $[0, 1]$  and enhancing detail discrimination. No smoothing operation is required, so avoiding the loss of imaging information about brain necroses.

**Necrosis and enhancement region distinction** In NeXt, the FCM algorithm is applied just on intensity values concerning the pixels included in previously segmented GTV tumor region (i.e.,  $\mathcal{I}_{\text{GTV}}$ ). Two clusters ( $C = 2$ ) are employed for this segmentation task to distinguish hypo-intense necrotic parts from the enhancement region of the brain tumor. Finally, to yield a quantifiable and clinically valuable measurement, the pixels that have the highest membership with the hypo-intense cluster are selected during the defuzzification process, so implementing a maximum membership segmentation scheme.

**Post-processing** Some morphological refinements [571] are also employed to enhance necrosis segmentation results and to prevent small hypo-intense region detection. The post-processing pipeline, which exploits binary mathematical morphology techniques, is the following:

1. a small area removal operation to remove small isolated hypo-intense regions and loosely connected pixels using 4-connectivity (i.e., connected-components with area less than 5 pixels). Especially, a 4-connected neighborhood was adopted to consider diagonally connected pixels as different connected-components, so allowing for a more selective labeling strategy;

2. a hole filling algorithm, based on morphological reconstruction [637], to remove possible hyper-intense regions in heterogeneous necrotic cores. Holes are determined as disjoint sets of background pixels that cannot be reached by a flood-filling algorithm starting from the background pixels in image borders.

## Results

The accuracy of the necrosis extraction results yielded by NeXt was quantitatively evaluated by means of the metrics defined in Appendix A.1. Both spatial overlap-based and distance-based metrics were calculated for all tumor slices of each MR image series. The proposed NeXt method was compared against the only other alternative necrosis/enhancement region segmentation method in literature, presented in Tumor-Cut [243]. For a more complete assessment, a further comparison with the most commonly used computer-assisted image segmentation techniques in medical imaging was carried out. Therefore, the following image processing methods were implemented and tested:

- adaptive thresholding (Otsu's method [448]), after the GTV segmentation;
- seeded region growing [524], after interactive seed-point selection;
- interactive LSFs [357], after seed-point selection for defining the initial region to segment;
- Tumor-Cut based on a CA model [243], after background and foreground seed initialization using Otsu's method [448] according to the previous GTV segmentation.

The necrosis/enhancement region segmentation method presented in Tumor-Cut was implemented by strictly following the description in the original paper [243]. For a fair comparison, the other developed approaches implement exactly the same pre-processing and post-processing pipelines used in NeXt.

Fig. 4.9 shows the necroses extracted by NeXt from the MRI brain tumors in Fig. 4.8. Our approach, based on the FCM clustering algorithm, suitably segments also inhomogeneous necrotic regions characterized by irregular or non-convex shapes. Additional necrosis segmentation results on representative brain tumors are displayed in Fig. 4.10. NeXt appropriately detects homogeneous dark regions enclosed by distinct hyper-intense enhancing rims (Fig. 4.10a), as well as disjoint necrotic regions when necrotic material is diffused within the tumor (Fig. 4.10b). In addition, NeXt correctly delineates challenging necrotic areas, characterized by very irregular shape or heterogeneous tissue (Fig. 4.10(c, d)).

For a more immediate graphical and clinically useful representation [646], Fig. 4.11 shows three instances of tridimensional reconstructions of brain tumors with a necrotic

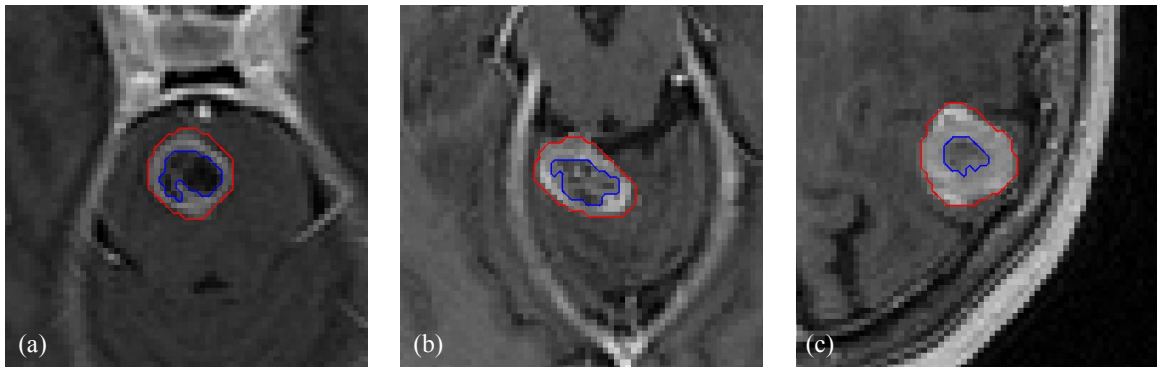


Fig. 4.9 Necroses (blue contour) extracted from the brain tumors (red contour) using the proposed NeXt approach, based on the FCM algorithm, on the images shown in Fig. 4.8. ( $4\times$  zoom factor).

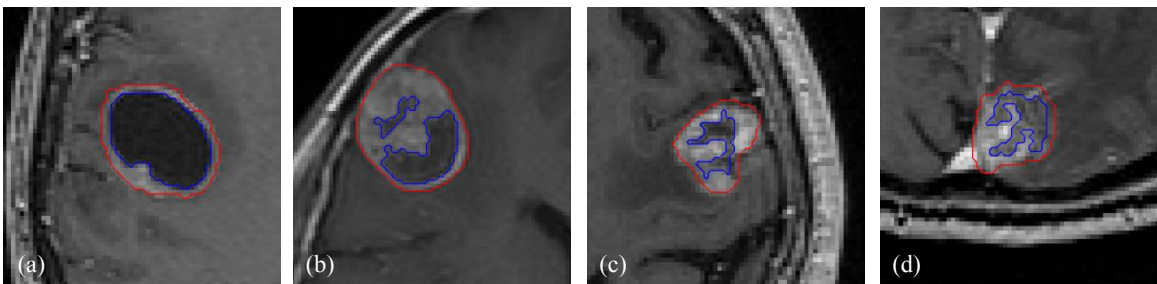


Fig. 4.10 Necrosis segmentation results on representative brain tumors representative necrotic regions, characterized by very irregular shape or heterogeneous tissue, extracted from the brain tumors using NeXt: (a) homogeneous dark region enclosed by distinct hyper-intense enhancing rims; (b) disjoint necrotic area with diffused necrotic material in the tumor; (c, d) challenging necroses characterized by very irregular shape or heterogeneous tissue. ( $4\times$  zoom factor).

core. The GTV is represented using a transparent red surface, allowing us to visualize the internal necroses (blue volumetric models) in their own actual position with respect to the tridimensional reference system.

### Experimental results

Table 4.3 depicts mean and standard deviation values of spatial overlap-based metrics achieved in the experimental tests by the developed necrosis extraction methods. Fig. 4.12 shows the boxplots of spatial overlap-based evaluation metrics to give a convenient representation of the statistical distribution of the achieved measurements. Overall, the achieved segmentation performance shows the great accuracy and reliability of the proposed approach. This evidence is also confirmed by the boxplots, where distributions for NeXt

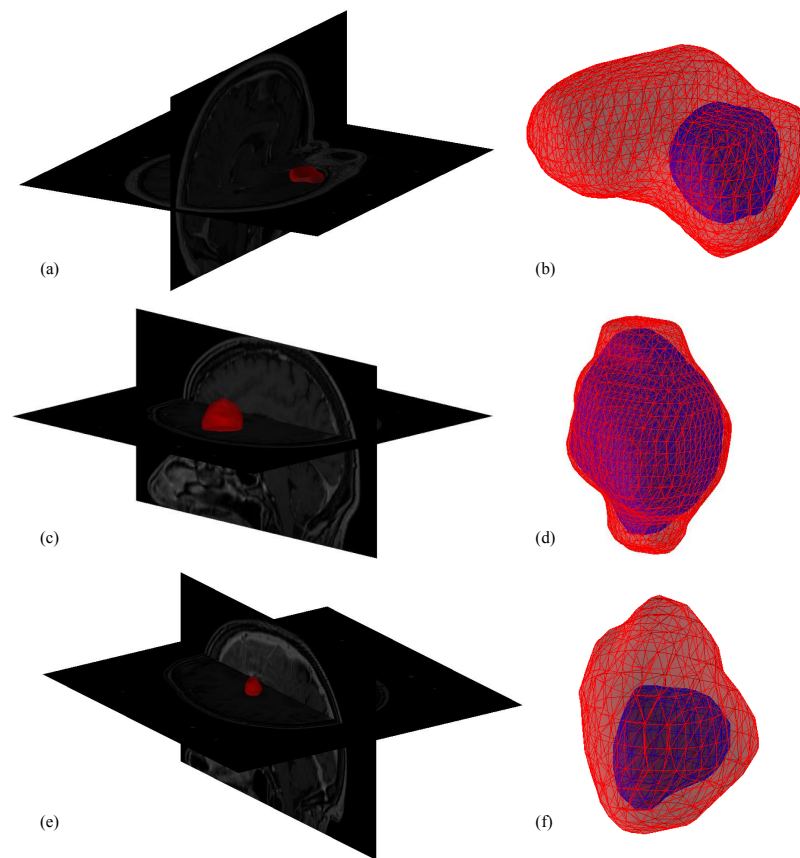


Fig. 4.11 Volumetric reconstruction of brain tumors with necrotic core: (a, c, e) brain tumors (red volumes) represented in a tridimensional space using orthogonal axial and sagittal planes to show their different anatomic positions in the corresponding whole brain MRI series; (b, d, f) the GTV is rendered using a transparent red surface and the inner necrosis (blue volumetric model) is visible in its own actual location. Transparent surfaces are rendered with Alpha blending ( $\alpha = 0.40$ ).

present approximately only 10% outliers in all the overlap-based metrics, thus evidencing extremely low statistical dispersion.

Since it is important to combine the distances from the misclassified voxels with respect to the gold standard for improving the information given by the overlap-based measures, average and standard deviation values of distance-based metrics, achieved on the whole MRI dataset, are reported in Table 4.4. Fig. 4.13 illustrates the boxplots of the achieved spatial distance-based evaluation metrics. The achieved spatial distance-based indices are consistent with overlap-based metrics, also observing the corresponding boxplots. Hence, good performances were obtained also in terms of difference between the automated and the manual boundaries.



Table 4.3 Values of the spatial overlap-based metrics regarding the necrosis extraction results achieved by the developed segmentation methods. The experimental results are expressed as average value  $\pm$  standard deviation.

Method	DSC	JI	SEN	SPC
Adaptive thresholding [448]	93.20 $\pm$ 6.69	88.65 $\pm$ 9.22	95.83 $\pm$ 3.38	93.20 $\pm$ 8.95
Seeded Region Growing [524]	58.98 $\pm$ 10.87	44.13 $\pm$ 10.94	51.97 $\pm$ 12.44	95.53 $\pm$ 5.55
Interactive LSFs [357]	54.81 $\pm$ 9.30	39.89 $\pm$ 9.84	84.44 $\pm$ 19.67	55.40 $\pm$ 23.17
Tumor-Cut based on CA [243]	78.98 $\pm$ 10.75	67.97 $\pm$ 13.34	86.19 $\pm$ 10.55	85.09 $\pm$ 10.97
NeXt based on FCM	95.93 $\pm$ 4.23	92.81 $\pm$ 6.56	98.09 $\pm$ 1.75	95.31 $\pm$ 6.56

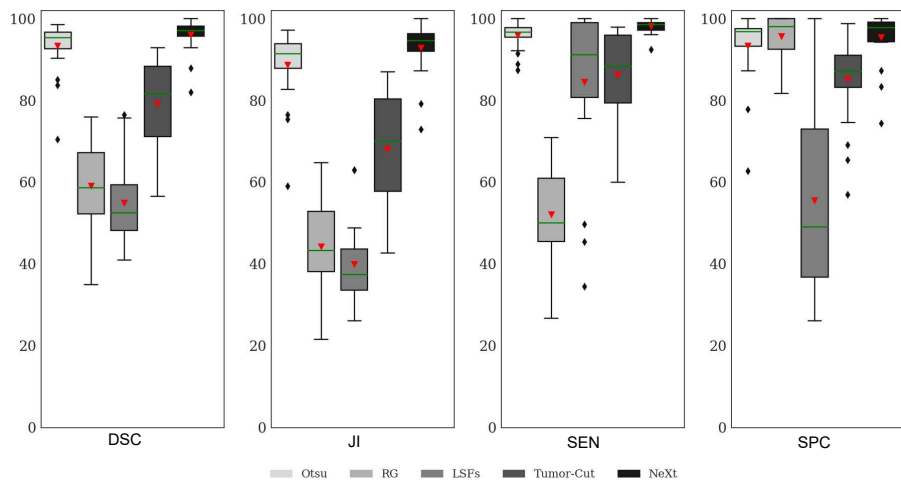


Fig. 4.12 Boxplots of the spatial overlap-based metrics values achieved by the implemented necrosis segmentation approaches. The lower and the upper bounds of each box represent the first and third quartiles of the statistical distribution, respectively. The median (i.e., the second quartile) and the mean values are represented by a green line and a red triangle, respectively. Whisker value is 1.5 in all cases and outliers are displayed as black diamonds.

**Comparison with other segmentation methods and implementation results** The proposed method NeXt, based on the FCM clustering algorithm, outperforms all the other implemented necrosis extraction approaches in terms of both spatial overlap-based and distance-based metrics. Adaptive thresholding using the Otsu's method [448] achieves good segmentation results, by exploiting the bimodal histogram distribution of the pixels included in the GTV region. However, the FCM clustering algorithm, thanks to the introduction of Fuzzy Logic enabling partial membership in classes, enables a more flexible classification process with respect to adaptive thresholding [448].

The CA-based method in Tumor-Cut [243] is characterized by lower sensitivity to small necrotic islands, since it strongly relies on the initial enhancement and necrosis seed selection, based on the Otsu's method [448], which considers the 25% of the necrotic and enhanced volumes as necrosis and enhancement seeds, respectively. So, the authors themselves stated

Table 4.4 Values of the spatial distance-based metrics regarding the necrosis extraction results achieved by the developed segmentation methods. The experimental results are expressed as average value  $\pm$  standard deviation.

Method	AvgD	MaxD	HD
Adaptive thresholding [448]	0.481 $\pm$ 0.524	1.411 $\pm$ 0.919	1.248 $\pm$ 0.405
Seeded Region Growing [524]	1.977 $\pm$ 3.097	3.700 $\pm$ 4.546	1.865 $\pm$ 0.334
Interactive LSFs [357]	2.589 $\pm$ 3.211	4.648 $\pm$ 4.225	1.864 $\pm$ 0.286
Tumor-Cut based on CA [243]	1.563 $\pm$ 2.266	3.093 $\pm$ 3.253	1.745 $\pm$ 0.221
NeXt based on FCM	0.225 $\pm$ 0.229	0.920 $\pm$ 0.581	1.060 $\pm$ 0.524

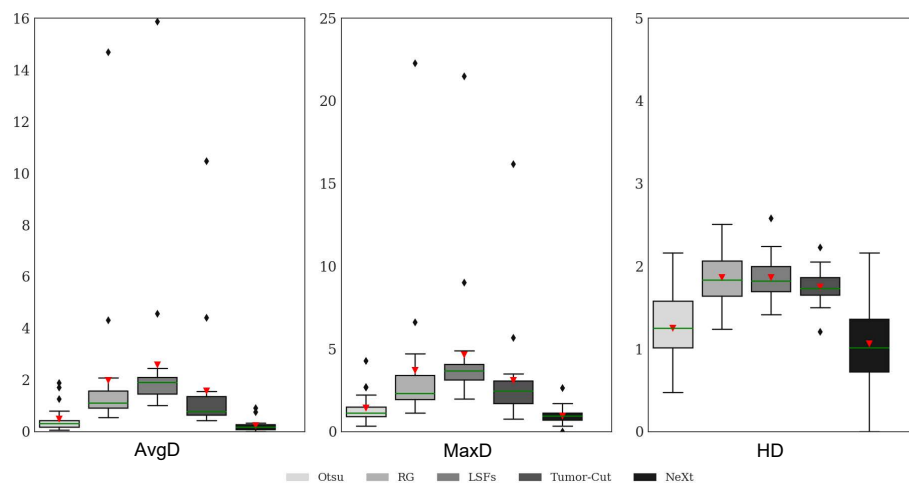


Fig. 4.13 Boxplots of the spatial distance-based metrics achieved by the implemented necrosis segmentation approaches. The lower and the upper bounds of each box represent the first and third quartiles of the statistical distribution, respectively. The median (i.e., the second quartile) and the mean values are represented by a green line and a red triangle, respectively. Whisker value is 1.5 in all cases and outliers are displayed as black diamonds.

that small islands of necrotic regions with ambiguous low contrast can be missed by the seed selection strategy. Moreover, this approach could be affected by the introduction of false positive seeds at the brain tumor boundaries, because no refinement steps are described in the original paper by Hamamci *et al.* [243]. The interactive LSF-based approach employed the DRLSE formulation [357]. Although this method attempts to address irregularities during the conventional Level Set evolution by means of a distance regularization term that maintains a desired shape of the LSF, leaking (i.e., over-estimation of the segmented necrosis ROI) could occur in the case of non-convex or weak ROI boundaries, see Fig. 4.10(c, d). This evidence is confirmed by FPs, as shown by the higher average values of SE with respect to SP, involving that the actual necrotic region is generally over-estimated. In addition, the LSF framework is highly sensitive to parameter setting, by considering image contrast and ROI characteristics (i.e., shape and boundaries). Region growing also requires a threshold

Table 4.5 Comparison of the execution times between NeXt and the implemented approaches.

Method	Average execution time [s]	Standard deviation
Adaptive thresholding [448]	$2.200 \times 10^{-3}$	$2.662 \times 10^{-4}$
Seeded Region Growing [524]	$7.900 \times 10^{-3}$	$6.608 \times 10^{-4}$
Interactive LSFs [357]	$1.683 \times 10^0$	$5.190 \times 10^{-1}$
Tumor-Cut based on CA [243]	$4.350 \times 10^{-2}$	$3.200 \times 10^{-3}$
NeXt based on FCM	$1.650 \times 10^{-2}$	$1.300 \times 10^{-3}$

selection according to image features, giving rise to an error-prone procedure. Segmentation errors could occur when inhomogeneities or loosely connected hypo-intense areas are present: the seeded region growing algorithm can suffer from highly heterogeneous areas, such as in the case of necrotic material. Especially, increasing the segmentation threshold in the homogeneity criteria could cause leaking, also expanding the segmented ROI to the whole brain region.

To quantify the efficiency of the proposed method in terms of execution times, the (sequential) version of the proposed necrosis extraction approach was compared against the other implemented image segmentation methods. All the approaches were developed and tested using MatLab<sup>®</sup> R2017a 64-bit, on a computer equipped with an Intel<sup>®</sup> Core<sup>™</sup> i7-7700HQ quad-core processor (clock frequency 2.80 GHz), 16 GB RAM and Windows 10 operating system.

Table 4.5 reports the performances, measured by the MatLab Profiler tool, in terms of average and standard deviation execution times calculated over the complete MRI dataset, by running the algorithms once on each single image. As expected, adaptive thresholding using Otsu's method [448], applied only on the MRI data included in the previously segmented GTV, was the fastest on average and with a low standard deviation. Since the necrosis ROI area to be segmented is small, seeded region growing achieved segmentation results in the order of milliseconds. Conversely, LSFs take more time to accomplish the segmentation based on an energy function minimization implementing a variational framework. However, these two methods require user interaction to select the initial seed-point. The approach presented in Tumor-Cut [243] for the segmentation of necrotic and enhancement regions is fully automatic since foreground and background seeds are selected according to the Otsu's method. Accordingly, threshold calculation has to be considered in the total execution time. Our method NeXt, based on FCM clustering, obtains on average a  $2.64\times$  speed-up factor with respect to the CA-based approach described in Tumor-Cut [243], showing also a lower standard deviation.

## Conclusion

We presented a fully automatic approach for necrosis extraction, called NeXt, using an unsupervised FCM clustering technique. Our method detects and delineates the necrotic regions within the planned GTV to be treated with neuro-radiosurgery. In compliance with the typical treatment planning practice, in which neurosurgeons generally use neuro-radiosurgery contrast-enhanced T1w MR images alone for GTV segmentation [488], NeXt does not need multispectral MRI data. So, NeXt represents a feasible solution in clinical routine, since the direct transfer of the proposed methodology in clinical environments is strongly reduced. An MRI dataset including 20 brain metastases with necrotic material, concerning patients who underwent stereotactic neuro-radiosurgery, was analyzed. The accuracy of the proposed approach was experimentally evaluated in terms of spatial overlap-based and distance-based metrics. A quantitative comparison with the most common literature methods for enhancement and necrosis region distinction was also performed. NeXt achieved the best segmentation accuracy, outperforming all the implemented alternative approaches.

Necrotic and enhancement region distinction in MRI brain tumors by using NeXt has certainly clinical value, in both radiation therapy treatment planning and patient's follow-up, since necrotic regions present hypoxic areas that could involve recurrent cancers and resistance to cancer therapy damages. However, conventional MRI could be also not appropriate for assessing the biological or metabolic activity of the tumor, since it does not provide a specific measurement of cancer activity [150]. To overcome these limitations, advanced imaging techniques enable the visualization of the actual biological changes of the cancer, by detecting proliferative activity, hypoxia, apoptosis, necrosis, and tumor vasculature. Advanced MRI techniques allow for the detection of various metabolic and physiological biomarkers. As a matter of fact, changes in metabolic tissue profile, tissue blood perfusion, microvessel permeability, and water diffusion can be estimated [80, 150].

In a complementary way, PET image analysis quantitatively assesses the metabolic state of cancer. As shown in [526], it is not always possible to identify the actual tumor extent using the MRI modality alone. As a matter of fact, MRI and MET-PET convey different but highly correlated imaging information. Consequently, the BTV, delineated on MET-PET images, is valuable and could be integrated with GTV in neuro-radiosurgery treatment planning. This applies also in dose painting, where quantitative analysis of MET-PET can provide aid in the differentiation of tumor recurrence from radiation necrosis [608]. In particular, dose painting and dose escalation have been already studied in some anatomic districts, such as non-small cell lung cancer [173] or prostate and cervix cancer [274]. Therefore, a comprehensive multimodal PET/MRI segmentation approach, which considers also metabolic and hypoxic sub-volumes, should represent a new way for a more selective and reliable identification of

the active cancer regions to be treated with radiosurgery systems that allow for dose painting [374] and personalized radiation therapy plans [497].

### 4.1.3 Prostate gland segmentation on multispectral MRI

In this section, an automated segmentation approach of the whole prostate gland from axial MRI slices is presented [518, 519]. The method is based on an unsupervised Machine Learning technique, resulting in an advanced application of the FCM clustering algorithm on multispectral MR anatomic images.

Nowadays, T2w FSE imaging is the standard protocol for depicting the anatomy of the prostate and for identifying prostate cancer considering reduced T2 signal intensity on MRI [339]. These sequences are sensitive to susceptibility artifacts (i.e., local magnetic field inhomogeneities), for instance due to the presence of air in the rectum, which may affect the correct detection of tissue boundaries [327]. On the other hand, because the prostate has uniform intermediate signal intensity at T1w MRI, the zonal anatomy cannot be clearly identified on T1w MRI series [106]. T1w MR images are used mainly to determine the presence of hemorrhage within the prostate and seminal vesicles and to delineate the gland boundaries [40]. Nevertheless, T1w MRI is not highly sensitive to prostate cancers, as well as to inhomogeneous signal in the peripheral zone or adenomatous tissue in the central gland with possible nodules. We exploit this uniform gray appearance, by combining prostate T2w and T1w MRI, to enhance the clustering-based classification.

Differently to the state-of-the-art approaches, this unsupervised Machine Learning technique does not require training phases, statistical shape priors or atlas pre-labeling. Therefore, our method could be easily integrated in clinical practice to support radiologists in their daily decision-making tasks. Another key novelty is the introduction of T1w MRI in the feature vector, composed of the co-registered T1w and T2w MR image series, fed to the proposed processing pipeline. Although T1w MR images were used in [456] for prostate radiotherapy planning, the authors did not combine both multispectral T2w and T1w MR images using a fusion approach. Their study just highlighted that the Clinical Target Volume (CTV) segmentation results on the T1w and T2w acquisition sequences did not differ significantly in terms of manual CTV. Our work aims to show that the early integration of T2w and T1w MR image structural information significantly enhances prostate gland segmentation by exploiting the uniform gray appearance of the prostate on T1w MRI [106].

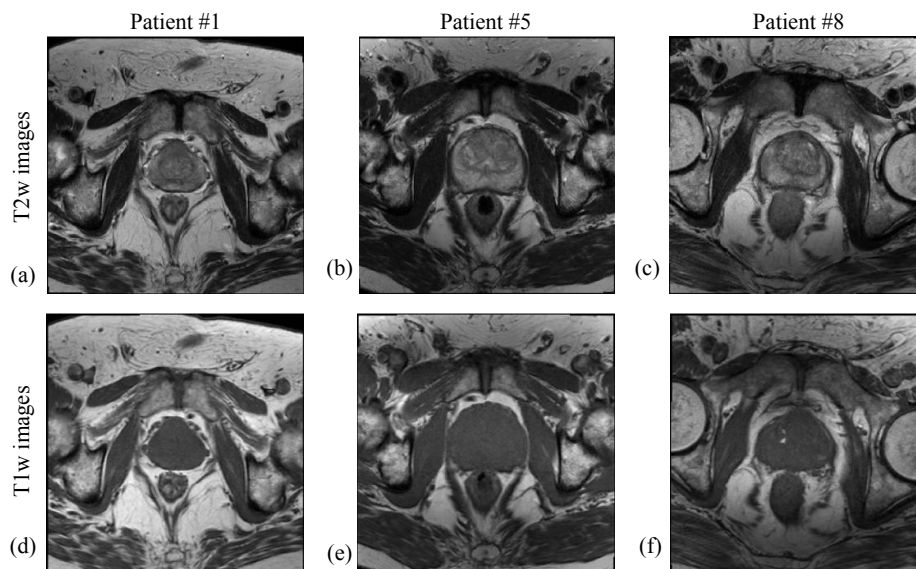


Fig. 4.14 Input prostate MRI axial slice pairs for the Patients #1, #5 and #8: (a, b, c) T2w MR images; (d, e, f) corresponding T1w MR image.

Table 4.6 MRI acquisition parameters of the prostate dataset.

MRI sequence	TR [ms]	TE [ms]	Matrix size [pixels]	Slice spacing [mm]	Slice thickness [mm]	Pixel spacing [mm]	Slices' number
T1w	515.3	10	256 × 256	4	3	0.703	18
T2w	3035.6	80	288 × 288	4	3	0.625	18

### Prostate MRI dataset

The study was performed on a clinical dataset composed of 21 patients with suspicious PCa. All the analyzed MR images were T1w and T2w Fast Spin Echo (FSE) sequences, acquired with an Achieva 3.0 T MRI Scanner (Philips Medical Systems, Eindhoven, the Netherlands) using a SENSE XL Torso coil (16 elements phased-array pelvic coil), at the Cannizzaro Hospital (Catania, Italy). The total number of the processed 2D MRI slices, in which prostate gland was imaged, was 185. Three pairs of corresponding T1w and T2w MR images are shown in Fig. 4.14. Imaging data were encoded in the 16-bit DICOM format. MRI acquisition parameters are reported in Table 4.6.

The average age of the examined patients is  $65.57 \pm 6.42$  (53-77) years. All the enrolled subjects have clinical suspicion about prostate cancer, due to high PSA levels or digital rectal examination (DRE) for prostate screening. Moreover, the MRI study can be also performed contextually with a previous biopsy, considered as gold standard, or combined with a further targeted biopsy according to the imaging evidences. In the positive cases, the possible clinical choices are three: prostatectomy, radiotherapy, or chemotherapy.

### The proposed whole prostate gland segmentation method

Since the organ to be imaged is always positioned approximately near the isocenter of the principal magnetic field to minimize MRI distortions, the whole prostate gland is represented in the imaged Field of View (FOV) center [343]. In a preliminary work in [518], after dividing the entire image in 9 equal-sized and fixed tiles, we considered the cropped image represented by the central one. Therefore, only a cropped image, whose size is 1/9 of the initial input image size, is processed for ROI image segmentation. Observing the prostate ROIs imaged on morphologic MRI, in terms of the segmented area in each slice, we observed that the prostate gland can be suitably contained in a patch with size 1/9 of the initial input image size. Using this strategy, user input is not required and processing time is certainly optimized.

Although the prostate gland is represented in the image FOV center, since the organ to be imaged is always positioned approximately near the isocenter of the principal magnetic field to minimize MRI distortions [343], sometimes it may appear shifted from the center or the patient could be not correctly positioned. This fact could affect segmentation accuracy especially in apical and basal prostate MRI slices, when the MR image center is considered for prostate gland segmentation. To address this issue, a rectangular ROI selection tool is dragged around the prostate by the user [511] for more reliable and precise delineation results [98]. Interactive segmentation is becoming more and more popular to alleviate the problems regarding fully automatic segmentation approaches, which are not always able to yield accurate results by adapting to all possible clinical scenarios [64, 65]. Thereby, operator-dependency is reduced but, at the same time, the physician is able to monitor and control the segmentation process. Accordingly, computerized medical image analysis has given rise to many promising solutions, but, instead of focusing on fully automatic computerized systems, researchers are aimed to propose computational techniques to aid radiologists in their clinical decision-making tasks [355].

The overall flow diagram of the proposed prostate segmentation method is represented in Fig. 4.15, and a detailed description of each processing phase is provided in the following subsections.

**MR image co-registration** Even though T1w and T2w MRI series are included in the same study, they are not acquired contextually because of the different employed extrinsic parameters that determine T1 and T2 relaxation times. Thus, an image co-registration step on multispectral MR images is required. Moreover, T1w and T2w images have different FOVs, in terms of pixel spacing and matrix size (see Table 4.6). However, in our case a 2D image

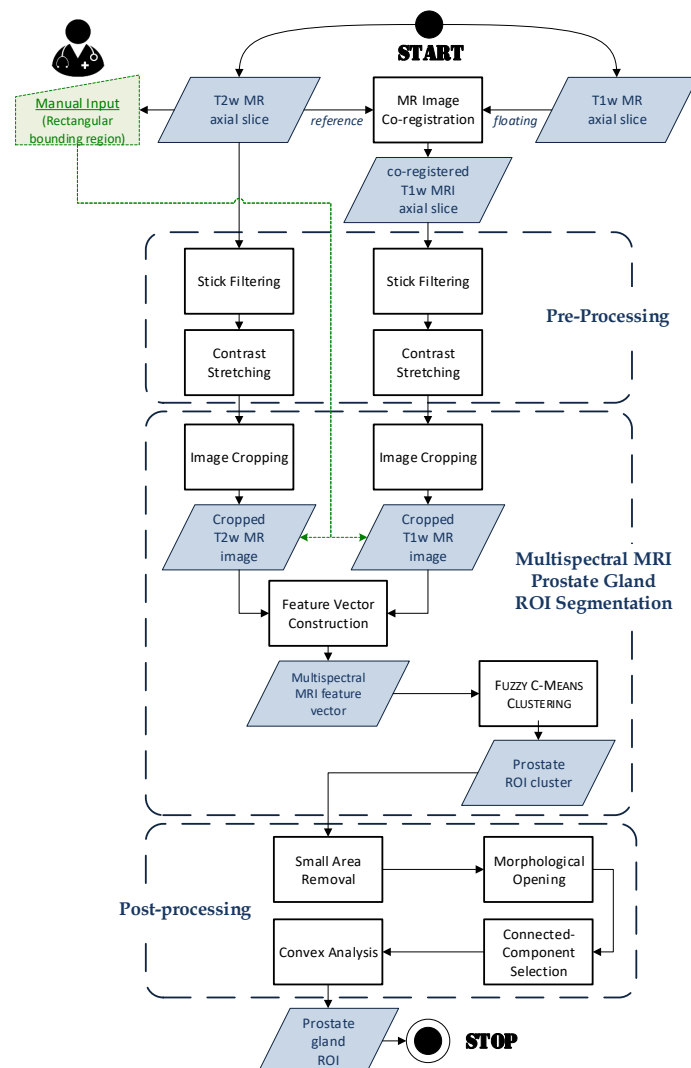


Fig. 4.15 Flow diagram of the proposed prostate gland MRI segmentation method. The pipeline can be subdivided into three main stages: (i) pre-processing, required to remove speckle noise and enhance the FCM clustering process; (ii) FCM clustering on multispectral MR images, to extract the prostate gland ROI; (iii) post-processing, well-suited to refine the obtained segmentation results.

registration method is effective because T1w and T2w MRI sequences are composed of the same number of slices and have the same slice thickness and interslice spacing values.

Image co-registration is accomplished by using an iterative process to align a moving image (T1w) along with a reference image (T2w). In this way, the two images will be represented in the same reference system, so enabling quantitative analysis and processing on fused imaging data. T2w MRI was considered as reference image, to use its own reference system in the following advanced image analysis phases for differentiating prostate anatomy



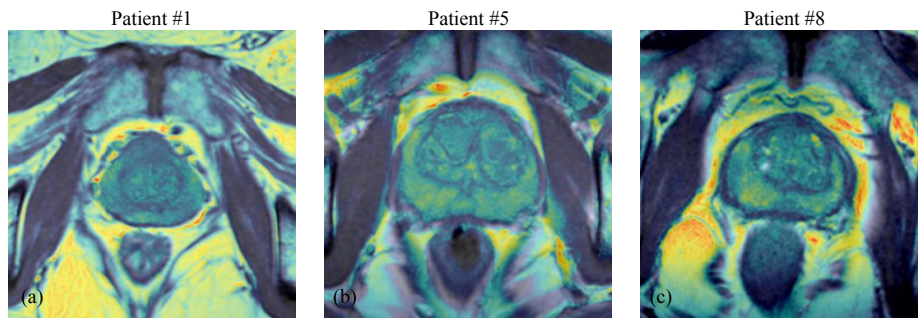


Fig. 4.16 Details of the prostate on the fused T1w and T2w MR image pairs in Fig. 4.14, using the gray-scale and the Jet colormap, respectively.

and for PCa detection and characterization, by processing also functional data (i.e., DCE, DWI, and MRSI).

In intensity-based registration techniques, the accuracy of the registration is evaluated iteratively according to an image similarity metrics (refer to Section 5.2.1 for further information). We used Mutual Information (MI) that is an information theoretic measure of the correlation degree between two random variables [474]. High mutual information means a large reduction in the uncertainty (entropy) between the two probability distributions, revealing that the images are likely better aligned [392]. An Evolution Strategy (ES), using a one-plus-one optimization configuration, is exploited to find a set of transformation that yields the optimal registration result. In this optimization technique, both the number of parents and the population size (i.e., number of offspring) are set to one [590]. Affine transformations, which combine rigid-body transformations (i.e., translations and rotations) with scaling and shearing, are applied to the moving T1w image to be aligned against the reference T2w image [528]. A detail of the prostate on the fused T1w and T2w images after the co-registration step is shown in Fig. 4.16.

**Pre-processing** Pre-processing steps are needed for MR image denoising, while preserving relevant feature details such as organ boundaries, and enhance the FCM clustering process. These operations are applied on T2w and the corresponding co-registered T1w input MR series in the same study.

Firstly, stick filtering is applied to efficiently remove the speckle noise that affects MR images [135, 136]. The main source of noise in MRI is the thermal noise in the patient [164], influencing both real and imaginary channels with additive white Gaussian noise [380]. The MR image is then reconstructed by computing the inverse discrete Fourier transform of the raw data, resulting in complex white Gaussian noise [398, 144]. The magnitude of the reconstructed MR image is commonly used for visual inspection and for automatic computer

analysis [473]. The magnitude of the MRI signal follows a Rician distribution, since it is calculated as the square root of the sum of the squares of two independent Gaussian variables. In low-intensity (dark) regions of the magnitude image, the Rician distribution tends to a Rayleigh distribution and in high intensity (bright) regions it tends to a Gaussian distribution [453]. This implies a reduced image contrast, since noise increases the mean value of pixel intensities in dark image regions [473]. By considering a local neighborhood around each pixel, the intensity value of the current pixel is replaced by the maximum of all the stick projections, which are calculated by the convolution of the input image with a long narrow average filter at various orientations [135, 136]. Using a stick filter with odd length  $\lambda$  and thickness  $\tau < \lambda$ , the square  $\lambda \times \lambda$  neighborhood can be decomposed into  $2\lambda - 2$  possible stick projections. Thereby, smoothing is performed on homogeneous regions, while preserving resolvable features and enhancing weak edges and linear structures, without producing too many false edges [670]. Stick filtering parameters are a good balance: the sticks have to be short enough that they can approximate the edges in images, but long enough that the speckle along the sticks is uncorrelated. Therefore, the projections along the sticks can smooth out speckle but not damage real edges in the image. Experimentally, considering the obtained filtering results to be the most suitable input for the FCM clustering algorithm and enhance the achieved classification, the selected stick filter parameters are: stick length  $\lambda = 5$  pixels; stick thickness  $\tau = 1$  pixel. All types of low-pass filtering reduce image noise, by removing high-frequency components of the MR signal. This could cause difficulties in tumor detection. Even though this smoothing operation could cause difficulties in cancer detection, our method is tailored for prostate gland segmentation alone and other pre-processing strategies could be employed for a better prostate cancer delineation. Afterwards, a contrast stretching operation is applied by means of a linear intensity transformation that converts the input intensities into the full dynamic range, in order to improve the subsequent segmentation process. The range of intensity values of the selected part is expanded to enhance the prostate ROI extraction. Thereby, a contrast stretching operation is applied through a global intensity transformation. This linear normalization converts input intensity values into the output full dynamics in  $[0, 1]$ . This operation implies an expansion of the values, improving detail discrimination. Since the actual minimum and the maximum intensity values of the MR input were considered, no region can have intensity outside of the cut-off values. So, no image content or details (e.g., tumors or nodules) might be missed during the following detection phases. The application of the pre-processing steps is shown in Fig. 4.17. The image noise is visibly suppressed in order to have a more suitable input for the subsequent clustering process, thus extracting the prostate ROI more easily.

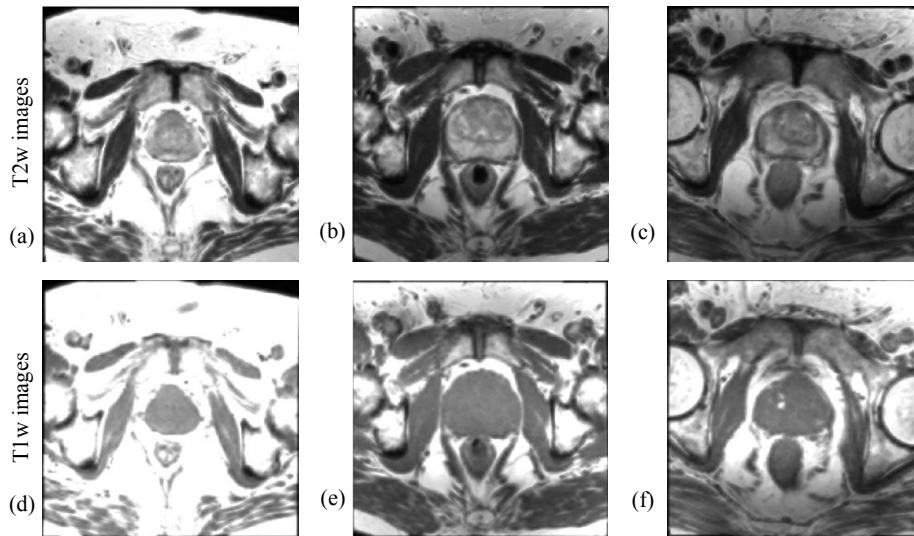


Fig. 4.17 Pre-Processing phase of the prostate gland MRI segmentation: (a, b, c) T2w images and corresponding (d, e, f) T1w images after pre-processing steps, consisting in stick filtering and linear contrast stretching, on the images in Fig. 4.14.

In order to quantitatively justify the selected smoothing operation, we calculated the SNR values on the prostate gland ROI segmentation results for the entire MRI dataset. Since we masked the different MR images with the corresponding prostate gland ROIs, forcing the background to uniform black, the SNR was computed, in compliance with routine MRI quality assurance testing [184], as:  $SNR = \mu_{ROI} / \sigma_{ROI}$ , where  $\mu_{ROI}$  and  $\sigma_{ROI}$  are the average value and the standard deviation of the signal inside the ROI, respectively. After the MR image co-registration step, we evaluated and compared SNR by multiplying pixel-by-pixel each ROI binary mask with the corresponding:

- original T1w and T2w MR images;
- T1w and T2w MR images pre-processed with stick filtering (length  $\lambda = 5$ , thickness  $\tau = 5$ );
- T1w and T2w MR images filtered with  $3 \times 3$  and  $5 \times 5$  Gaussian kernels;
- T1w and T2w MR images filtered with  $3 \times 3$  and  $5 \times 5$  average kernels.

For a fair comparison, in all the cases we normalized pixel intensities in  $[0, 1]$  using a linear stretching transformation, also to be consistent with the proposed pipeline. The SNR values achieved on the T2w as well as on the co-registered T1w images are reported in (Table 4.7). As it can be seen, the images pre-processed with stick filtering achieved the highest SNR in both T1w and T2w MRI series.

Table 4.7 SNR calculated on the entire prostate MRI dataset composed of 21 co-registered T1w and T2w MRI sequences. The used stick filtering was compared against the most common smoothing strategies.

MRI sequence	Original	Stick filtering	3 × 3 Gaussian filter	5 × 5 Gaussian filter	3 × 3 Average filter	5 × 5 Average filter
T1w	4.2133	5.2598	4.4069	4.4081	4.7523	4.9061
T2w	3.7856	5.0183	4.0032	4.0044	4.3479	4.8739

**Prostate gland segmentation using FCM algorithm on multispectral MRI** The feature vector, which can include more than one feature, is constructed by concatenating corresponding pixels on T2w and T1w MR images, after the image co-registration step, in an early integration scheme [556]. The clustering procedure is performed on the feature vector composed of corresponding T2w and T1w pixel values after the co-registration step. The cluster number is always set to  $C = 3$ , by considering both structural and morphologic properties of the processed prostate MR images. As a matter of fact, three different classes can essentially be seen by visual inspection according to gray levels. This choice was also justified and endorsed by experimental trials. Fig. 4.18 shows the clustering results achieved by the FCM algorithm with three classes during the prostate gland ROI segmentation stage. It is appreciable how the FCM clustering output on multispectral MRI data considerably enhances the results achieved by processing either T1w or T2w images alone. Especially, the introduction of T1w imaging, in the feature vector construction step, yields a more uniform overall clustering output in the prostate ROI cluster. The resulting feature vector helps decisively to separate the ROI from the other surrounding tissues and organs (i.e., bladder, rectum, seminal vesicles, levator ani muscle).

As visible in Fig. 4.18(b, e, h), FCM clustering on T2w images alone often does not detect the whole prostate gland because of the irregularities due to the prostate zonal anatomy or cancer. However, even though the FCM algorithm on T1w images alone obtains more regular clusters, it can include also external regions that can affect the final ROI segmentation results (see Fig. 4.18(f, i)). Therefore, performing the segmentation on the co-registered T1w and T2w MR image is the best solution for prostate gland segmentation. After the defuzzification step, which consists in assigning the pixels to the cluster with the greatest membership value, the prostate ROI cluster is identified.

**Post-processing** A sequence of morphological operations [68, 571] is applied to the obtained ROI cluster, in order to resolve possible ambiguities resulting from the clustering process and to improve the quality of the segmented prostate gland ROI. According to Fig. 4.15, the post-processing steps are the following:

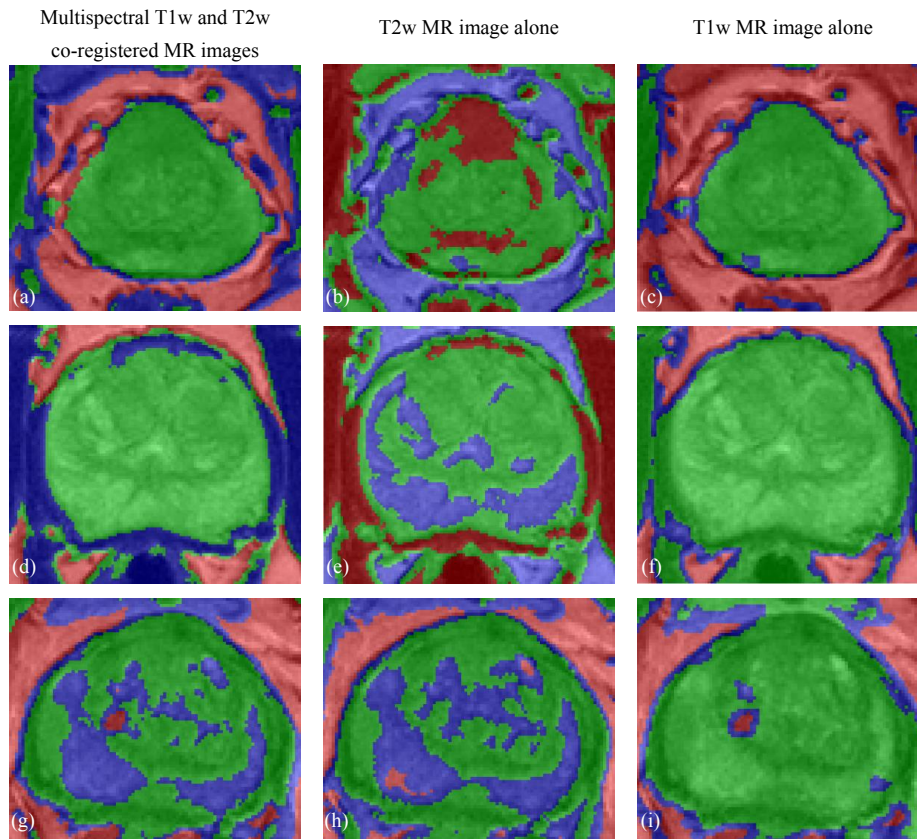


Fig. 4.18 Results of the FCM clustering algorithm with  $C = 3$  for prostate gland ROI segmentation on the input images in Fig. 4.14, by processing different feature vectors: in the first column (a, d, g) results on the multispectral T1w and T2w co-registered MR images; in the second column (b, e, h) results on T2w MR image alone; in the third column (c, f, i) results on T1w MR image alone. It is possible to appreciate the multispectral MRI contribution with respect to the single series clustering outputs. In all cases, the prostate gland ROI cluster is represented by the green regions. The clusters are superimposed on the original T2w MR image using a transparency with Alpha blending ( $\alpha = 0.40$ ).

1. a small area removal operation is employed to delete any unwanted connected-components, whose area is less than 500 pixels, which are included into the ROI cluster. These small regions can have similar gray levels to the prostate ROI and were classified into the same cluster by the FCM clustering algorithm. The areas less than 500 pixels certainly do not represent the prostate ROI, since the prostate gland in MRI slices is always characterized by a greater area, regardless of the used acquisition protocol. Thus, these small areas can be removed from the prostate ROI cluster to avoid a wrong connected-component selection in the next processing steps;

2. a morphological opening, using a  $5 \times 5$  pixel square structuring element, is a good compromise between precision in the actual detected contours and capability for unlinking poorly connected regions to the prostate ROI;
3. connected-component selection, using a flood-fill algorithm [663], in order to determine the connected-component that is the nearest to the cropped image center, since the prostate is always located at the central zone of the cropped image;
4. convex analysis of the ROI shape since the prostate gland appearance is always convex. A convex hull algorithm [707] is suitable to envelope the segmented ROI into the smallest convex polygon that contains it, so considering possible adjacent regions excluded by the FCM clustering output. Finally, a morphological opening operation with a circular structuring element, is performed for smoothing prostate ROI boundaries. The used circular structuring element with 3-pixel radius allows for smoother and more realistic ROI boundaries without deteriorating significantly the output yielded by the FCM clustering. Accordingly, the value of the radius does not affect the segmentation results and it is not dependent on image resolution.

### Experimental results

The results achieved by the proposed segmentation approach on multispectral T1w and T2w co-registered MR images were also compared by applying the same processing pipeline on the corresponding monoparametric T2w and T1w MR images alone. Fig. 4.19 shows three examples of prostate gland ROI obtained by the proposed automated segmentation approach. It is appreciable how correct segmentation results are achieved even with inhomogeneous and noisy input MRI data.

According to Fig. 4.19, it is possible to appreciate the differences by processing multispectral anatomic co-registered images, T2w MR image alone, and T1w image alone. Especially, as shown in Fig. 4.19(b, e, h), FCM clustering on T2w images alone often is not able to detect the whole prostate gland because of the irregularities due to the prostate zonal anatomy or cancer. On the other hand, even though FCM clustering on T1w images alone obtains an accurate result in Fig. 4.19c, it can segment also wrong anatomic parts, such as bladder or rectum, because they are imaged with similar gray levels to prostatic gland (Fig. 4.19(f, i)). The proposed method is able to correctly segment the midgland and detects reasonably well the apical to the basal prostate regions. An example of the achieved segmentations at the base, center (midgland) and apex of the prostate for the Patients #7 and #12 are shown in Fig. 4.20. The method is able to separate the prostate gland from seminal vesicles and bladder in basal slices as well as from rectum in central slices.

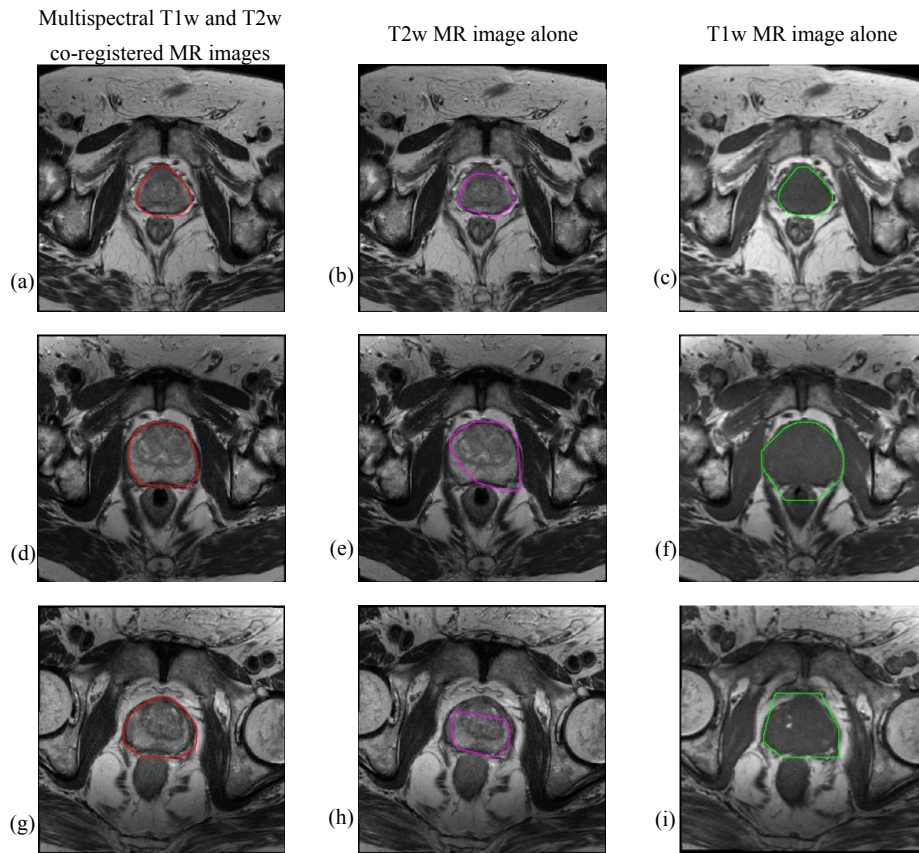


Fig. 4.19 Prostate gland ROI segmentation instances on the axial MRI slices shown in Fig. 4.14 by processing different input MRI data (i.e., feature vectors) of three different patients (one per-row): first column (a, d, g), results obtained on multispectral T1w and T2w co-registered MR image (red contour superimposed on T2w image); second column (b, e, h), results obtained on T2w MR image alone (magenta contour superimposed on T2w image); third column (c, f, i), results obtained on T1w MR image alone (green contour superimposed on T1w image).

Table 4.8 reports the mean and standard deviation values of the overlap-based metrics obtained in the experimental segmentation tests on 21 MRI series representing prostate gland executing the proposed segmentation method, based on the FCM clustering algorithm, on multispectral MRI anatomic images, T2w images alone, and T1w images alone, respectively. High *DSC* and *JI* mean values with very low standard deviation show the segmentation accuracy and reliability. In addition, *SEN* and *SPC* average values involve the correct detection of the “true” areas and the ability of not detecting “wrong” parts in the segmented prostate gland ROIs, respectively. This is also shown by the obtained *FPR* and *FNR* values. The multispectral approach significantly outperforms the monoparametric ones. The results associated to T1w processing are slightly more sensitive with respect to

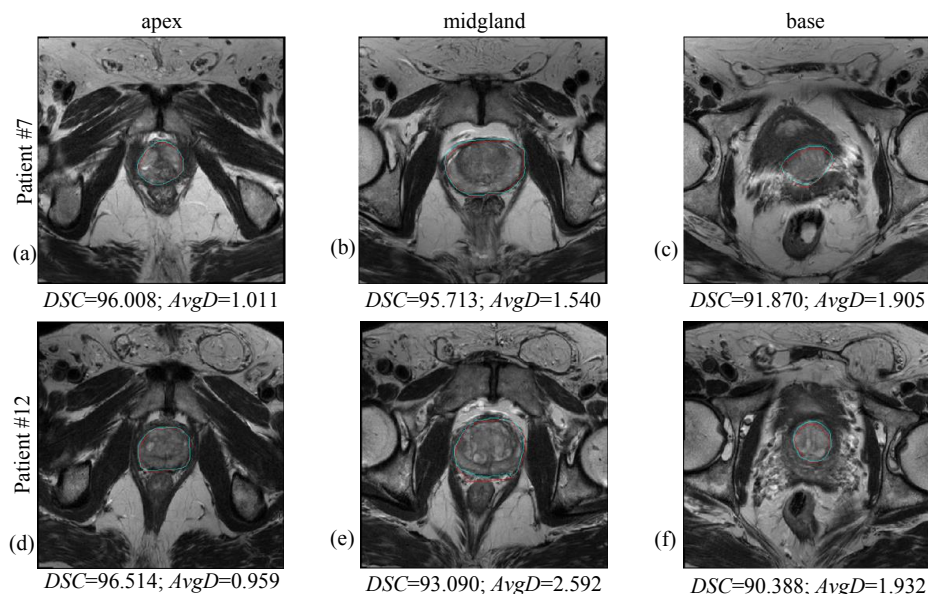


Fig. 4.20 Examples of segmentation results (red and cyan contours representing the proposed and the manual delineations, respectively, are superimposed on T2w axial prostate MR images) obtained on apical (a, d), central (b, e), and basal (c, f) slices concerning Patients #7 and #12, respectively. At the bottom of each sub-figure, the values for Dice similarity coefficient ( $DSC$ ) and mean average distance ( $AvgD$ ) are reported.

Table 4.8 Values of the spatial overlap-based metrics regarding the achieved prostate gland ROI segmentation results. The results are expressed as average value  $\pm$  standard deviation.

MRI data	DSC	JI	SEN	SPC	FPR	FNR
Multispectral	90.77 $\pm$ 1.75	83.63 $\pm$ 2.65	89.56 $\pm$ 3.02	99.85 $\pm$ 0.11	0.15 $\pm$ 0.11	6.89 $\pm$ 3.02
T2w alone	81.90 $\pm$ 6.49	71.39 $\pm$ 7.56	82.21 $\pm$ 8.28	99.63 $\pm$ 0.25	0.37 $\pm$ 0.25	12.58 $\pm$ 8.49
T1w alone	82.55 $\pm$ 4.93	71.78 $\pm$ 6.15	93.27 $\pm$ 4.87	98.85 $\pm$ 0.58	1.15 $\pm$ 0.58	3.58 $\pm$ 4.95

the proposed multispectral approach, even if they can include also wrong parts as shown by the higher  $FPR$  values. However, in all the cases, the variance of the results achieved by the monoparametric approaches is higher compared to the proposed multiparametric method. To provide a graphical representation of the statistical distribution of the results, the corresponding boxplots of spatial overlap-based evaluation metrics are also reported in Fig. 4.21. The short width of the interquartile range (i.e., difference between the third and first quartiles) represented in boxplots implies that the values are considerably concentrated. All index distributions for the multispectral approach do not present outliers, thus demonstrating extremely low statistical dispersion.

Table 4.9 shows distance-based metrics mean and standard deviation values obtained using the proposed segmentation approach on the dataset composed of 21 patients, by processing multispectral MRI anatomic images, T2w images alone, and T1w images alone,



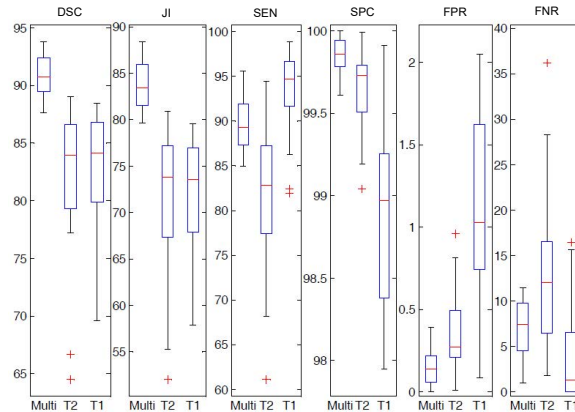


Fig. 4.21 Boxplots of the spatial overlap-based metrics values achieved by the proposed prostate gland ROI segmentation approach based on the FCM algorithm. The lower and the upper bounds of each box represent the first and third quartiles of the statistical distribution, respectively. The median value (i.e., the second quartile) is represented by a red line dividing the box. Whisker value is 1.5 in all cases.

Table 4.9 Values of the spatial distance-based metrics regarding the achieved prostate gland ROI segmentation results. The results are expressed as average value  $\pm$  standard deviation.

MRI data	AvgD [pixels]	MaxD [pixels]	HD [pixels]
Multispectral	$2.676 \pm 0.616$	$8.485 \pm 2.091$	$4.259 \pm 0.548$
T2w alone	$4.941 \pm 1.780$	$13.779 \pm 3.430$	$4.535 \pm 0.335$
T1w alone	$5.566 \pm 1.618$	$14.358 \pm 3.722$	$5.177 \pm 0.857$

respectively. The lower the distance values, the better the segmentation results. The achieved spatial distance-based indices are consistent with overlap-based metrics. Hence, good performances were obtained also with heterogeneous prostates. As for the spatial overlap-based metrics, also the results obtained on multiparametric MRI anatomic data outperform the monoparametric ones. Fig. 4.21 depicts the boxplots of the achieved spatial distance-based evaluation metrics.

Observing the boxplots, just a small deviation between the segmentations of the proposed method and those of the experienced radiologist can be denoted. Considering the practical purposes of the proposed prostate gland delineation method for volume calculation in clinical applications, referring to MRI spatial resolution (i.e., isotropic pixel resolution of 0.703 mm in Table 4.6), the proposed multispectral approach achieved on average  $5.965 \pm 1.470$  mm in the worst case (i.e., *MaxD*), so ensuring effective performance. This confirms the accuracy and reliability of the proposed segmentation method in real clinical contexts.

For a more complete and intuitive graphical representation, four different instances of volumetric reconstructions of segmented prostate glands are shown in Fig. 4.23. It is worth to note the variety of the prostate gland dimensions and shapes.

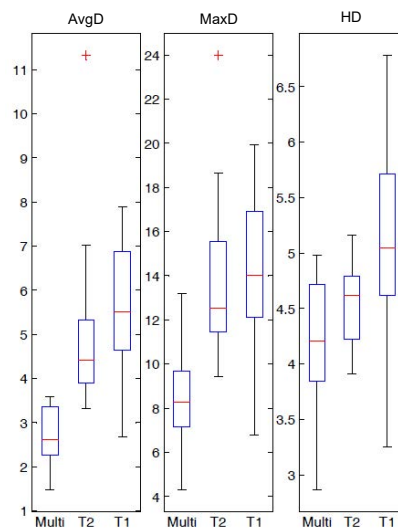


Fig. 4.22 Boxplots of the spatial distance-based metrics values achieved by the proposed prostate gland ROI segmentation approach based on the FCM algorithm. The lower and the upper bounds of each box represent the first and third quartiles of the statistical distribution, respectively. The median value (i.e., the second quartile) is represented by a red line dividing the box. Whisker value is 1.5 in all cases.

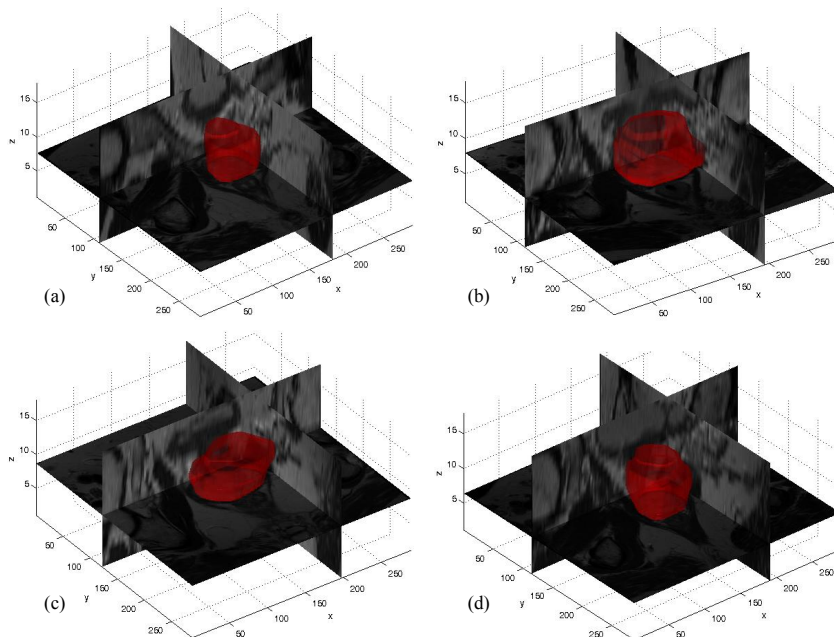


Fig. 4.23 Tridimensional reconstruction of the segmented prostate gland ROI on the 3D orthogonal tri-planar view for the Patients: (a) #1; (b) #5; (c) #7; (d) #12. Transparent surfaces of the prostate gland are rendered with Alpha blending ( $\alpha = 0.40$ ).

## Discussion

The proposed approach provided good segmentation results compared to the gold standard boundaries delineated by an experienced radiologist. Although volume-based metrics do not take into account the intersection, they represent a first and immediate measure of segmentation effectiveness. Just small differences between the automatic and the manual volume measurements were observed, showing the reliability of the proposed segmentation approach for automatic volume calculation in clinical applications. The achieved overlap-based indices, which are characterized by high average values and very low standard deviation, reveal the segmentation accuracy, involving the correct detection of the “true” pathological areas as well as the capability of not detecting wrong parts within the segmented prostate. The achieved spatial distance-based metrics agree with the overlap-based ones, corroborating the aforementioned experimental evidences.

The proposed method is based on an unsupervised Machine Learning technique, without requiring any training phase. On the contrary, the other literature works use or combine atlases [327, 363, 388, 389], AAMs [636], or statistical shape priors and ASMs [203, 388, 389, 616, 617], which require manual labeling of a significant image sample set performed by expert physicians. In addition, as stated in [389], atlas-based approaches may be affected by serious errors when the processed prostate instances are dissimilar to the atlas, despite the non-rigid registration. Literature approaches used T2w MR anatomic images, sometimes combined with ADC maps [363] or MRSI data [617]. Our method integrated T1w and T2w MRI anatomic data to enhance clustering segmentation results. To the best of our knowledge, we combined T1w and T2w MR image series for the first time in prostate gland segmentation. As it is stated in [355], the available public prostate MRI datasets for prostate gland segmentation and prostate cancer detection and delineation, such as PROMISE12 [365] and the benchmark proposed in [355], do not provide T1w MR images. Therefore, our method cannot be applied on these public datasets and compared with the prostate MR image segmentation approaches presented at the MICCAI 2012 PROMISE12 challenge, since our aim is to show that concatenating T2w and T1w pixels during the construction of the feature vector in an early integration phase. Although a comparison with state-of-the-art methods could be certainly interesting, it is also unfeasible because atlases should be built and supervised methods with *a priori* knowledge (such as Active Appearance Models or statistical shape priors) should be trained on our prostate MRI dataset composed of 21 patients. The number of samples is not sufficiently significant to apply suitably supervised Machine Learning techniques or shape-based models without using LOOCV, such as in [636] and [388]. As shown in Fig. 4.20, our approach is also able to correctly segment apical and basal prostate MRI slices, by differentiating also seminal vesicles (see Fig. 4.20f). The

segmentation quality was also good at the prostate-rectum and prostate-bladder interfaces. Whereas a segmentation error of a few millimeters is clinically acceptable at boundaries with muscular tissue, the interfaces with rectum and bladder need to be detected and distinguished very accurately [327]. Moreover, good performances were obtained also with prostate gland instances imaged as heterogeneous regions (i.e., inhomogeneous signal in the peripheral zone or adenomatous central gland). The principal experimental finding is that the FCM clustering on multispectral MRI anatomic data considerably enhances the achieved prostate ROI segmentations, by taking advantage of the prostate uniform intermediate signal intensity at T1w imaging [106]. The FCM clustering algorithm on multispectral MR anatomic images allows for more accurate prostate gland ROI segmentation with respect to the clustering results on either T2w or T1w images alone. However, T1w images are often not able to distinguish among different soft-tissues. For instance, FCM clustering on T1w images alone is not able to differentiate the prostate-rectum (Fig. 4.20f) or the prostate-muscle (Figs. 4.20f and 4.20i) interfaces. In conclusion, combining and fusing T2w and T1w MRI data, in the feature vector construction step, allow us to achieve better clustering outputs. The method was insensitive to variations in patient age, prostate volume, and the presence of tumors (i.e., suspicion of cancer in different prostate regions, inhomogeneous signal in the PZ or adenomatous CG with possible nodules), also considering radiotherapy or chemotherapy treatments, thus increasing its feasibility in clinical practice.

## Conclusions

NeXt, which is based on an unsupervised Machine Learning technique, was used to effectively support prostate gland delineation, such as in radiation therapy. The developed approach was tested on a dataset composed of 21 patients, considering T2w and T1w MRI series.

Spatial overlap-based and distance-based metrics were calculated to evaluate the performance, proving the accuracy of the proposed segmentation approach. The achieved experimental results demonstrated the great robustness of the proposed approach even when MRI series were affected by acquisition noise or artifacts. The good segmentation performance was also confirmed with prostate glands characterized by inhomogeneous signal in the peripheral zone or adenomatous central gland with possible nodules. The developed method could be clinically feasible, by addressing and overcoming typical issues affecting manual segmentation procedures, i.e., time-consuming and operator-dependency.

Future work could be directed towards a more selective segmentation technique, distinguishing between CG and PZ of the prostate anatomy [106, 278]. The advanced segmentation method can be definitely employed in a two-step prostate cancer delineation approach, in order to focus on pathological regions in the central gland and in the peripheral zone. The

overall prostate cancer segmentation method could be integrated into a Computer-Aided Diagnosis (CAD) systems, which include both Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx), enhancing the diagnosis performance of radiologists [355]. In this way, a more accurate PCa staging can be performed, and determining the disease burden may also lead to greater benefit for the patients who underwent treatment [9].

## 4.2 Graph-based methods

Sinop and Grady [569] presented a unifying graph theoretical framework for seeded image segmentation depending on the used norm in the objective function optimization. A weighted graph consists of a pair  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with vertices  $v \in \mathcal{V}$  and edges  $e \in \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . Each edge  $e_{ij}$ , connecting the vertices  $v_i$  and  $v_j$ , is associated to a weight  $w_{ij} \in \mathbb{R}^+$ . The graph is assumed to be connected and undirected (i.e.,  $w_{ij} = w_{ji}$ ). In seeded image segmentation, vertices correspond to image pixels while edge weights  $w_{ij}$  represent similarity measures between neighboring pixels according to image features. Moreover, each vertex  $v_i$  has an attribute  $x_i \in [0, 1]$ , which indicates the probability of a label.

Without loss of generality, a two-class image segmentation problem is considered. Starting from initial foreground  $\mathcal{F}$  and background  $\mathcal{B}$  seeds, the labeling process is solved by the general seeded segmentation optimization in Eq. (4.7):

$$x^{\text{opt}} = \arg \min_x \left[ \sum_{e_{ij} \in \mathcal{E}} (w_{ij} |x_i - x_j|)^q \right]^{\frac{1}{q}}, \quad (4.7)$$

with  $x(\mathcal{F}) = 1$ ,  $x(\mathcal{B}) = 0$ , and  $q \in \mathbb{R}^+$ . The final segmentation solution  $\mathcal{S}$  is defined by:

$$S_i = \begin{cases} 1, & \text{if } x_i \geq 0.5 \\ 0, & \text{if } x_i < 0.5 \end{cases}. \quad (4.8)$$

As described by Sinop and Grady [569], the solution has been shown to converge to various algorithms by considering different  $q$  values and the resulting  $\ell_q$  norm:

- Graph cuts for  $q = 1$  [65, 226];
- Random walker for  $q = 2$  [221];
- Shortest paths for  $q = \infty$  [45, 186].

The main advantage of using CA algorithms against graph-theoretic formulations is the ability to obtain a multi-label solution in a simultaneous iteration. For instance, the ‘‘Graph

cuts” algorithm finds a global optimum only for binary labeling; the extensions proposed in literature to achieve approximate solutions on multi-label problems exploit “Graph cuts” iteratively to a sequence of binary problems [64, 65]. In addition, the local transition rules are simple to interpret, and it is possible to impose prior knowledge, specific to the problem, into the image segmentation approach [243]. Thus, the state of the entire lattice advances in discrete time steps [314].

### 4.2.1 Shortest Path with Cellular Automata

Cellular Automata (CA) are one of the oldest models of bio-inspired computing [312] and were introduced by John von Neumann and Stanislaw Ulam in their seminal papers [642]. The goal was to design self-replicating artificial systems that make up a discrete universe consisting of a bi-dimensional mesh of finite state machines (cells), interconnected locally to each other, which will produce a complex global emergent behavior [660]. Each cell changes its own state synchronously depending on both: (i) its current state, and (ii) the states of the neighbor cells at the previous discrete time step, as determined by a local update rule. All cells use the same update rule so that the system is homogeneous like many physical and biological systems [311]. The most common CA are discrete, in both space and time, and process a lattice of  $d$ -dimensional sites  $\mathbf{p} \in \mathbb{Z}^d$ . The 1-D, 2-D, and 3-D cases are the most used topologies in practical problems [56].

CA have been also effectively applied in image processing [479, 509] and Pattern Recognition and classification tasks [490, 661], also with applications in cancer imaging [421]. In image processing,  $\mathbf{p}$  represents a pixel ( $d = 2$ ) or a voxel ( $d = 3$ ) of the digital image  $\mathcal{I}$ . The image segmentation problem can be solved using a labeling process, where  $K$  different types of self-replicating cells compete to occupy the grid space defined by the image  $\mathcal{I}$ , according to a CA model. Formally, a deterministic Cellular Automaton is defined by a triplet  $\mathcal{A} = (\mathcal{S}, \mathcal{N}, f)$ , where:  $\mathcal{S}$  is the non-empty state set;  $\mathcal{N}$  is the set of the  $v$  neighbors of a particular cell;  $f : \mathcal{S}^v \rightarrow \mathcal{S}$  is the local transition rule, which, given the states of the neighborhood cells at the current time step  $t$ , maps the cell state at the next  $t + 1$  time step. Let  $\mathbf{p} \equiv (p_1, p_2, \dots, p_d) \in \mathbb{Z}^d$  be a cell location, the most used neighborhood definitions are:

- von Neumann neighborhood  $\mathcal{N}(\mathbf{p}) = \{\mathbf{q} \in \mathbb{Z}^d : \|\mathbf{p} - \mathbf{q}\|_1 \leq 1\}$ , where  $\|\mathbf{p} - \mathbf{q}\|_1 = \sum_{i=1}^d |p_i - q_i|$  is the  $\ell_1$  Manhattan norm;
- Moore neighborhood  $\mathcal{N}(\mathbf{p}) = \{\mathbf{q} \in \mathbb{Z}^d : \|\mathbf{p} - \mathbf{q}\|_\infty \leq 1\}$ , where  $\|\mathbf{p} - \mathbf{q}\|_\infty = \max_{i=1, \dots, d} \{|p_i - q_i|\}$  is the  $\ell_\infty$  max-norm.

For greater clarity, the von Neumann and Moore neighborhoods for the 2-dimensional case are illustrated in Figs. 4.24a and 4.24b, respectively. See [694] for more details.

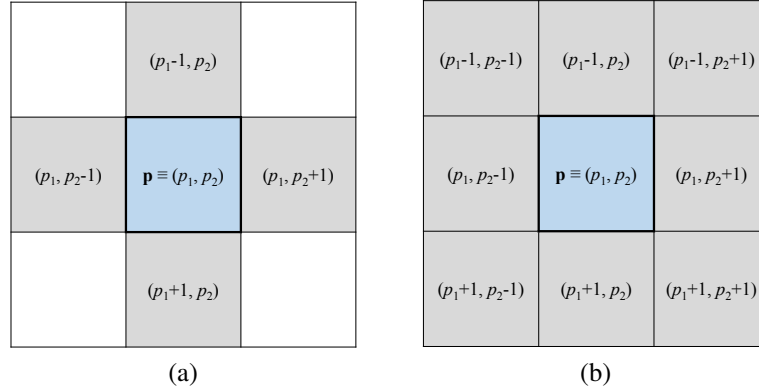


Fig. 4.24 Classical neighborhoods in the 2-dimensional case: (a) von Neumann neighborhood; (b) Moore neighborhood. Pixel indices are compliant with the conventional image coordinate system for representing digital images as arrays.

In the “GrowCut” CA formulation [634], the cell state is a triple  $S_{\mathbf{p}} := (l_{\mathbf{p}}, \theta_{\mathbf{p}}, \mathbf{c}_{\mathbf{p}})$ , where:  $l_{\mathbf{p}} \in \{1, 2, \dots, K\}$  is the label of the current cell in the  $K$ -labeled segmentation process;  $\theta_{\mathbf{p}} \in [0, 1]$  is the strength value of the current cell;  $\mathbf{c}_{\mathbf{p}} \in \mathbb{R}^n$  is the feature vector associated to the cell  $\mathbf{c}$  according to the image content. The automaton is initialized with the previously calculated seeds, by labeling foreground and background seeds. The maximum strength value is assigned to the labeled cells, while unlabeled cells are set  $\theta_{\mathbf{p}}^{(0)} = 0$ . Cell states are then updated synchronously at each iteration  $t$ , until the convergence condition is reached (i.e., any cell does not change its own state). State transitions occur if the attacker force of a neighbor cell  $\mathbf{q} \in \mathcal{N}(\mathbf{p})$  (Moore 8-neighborhood was used), defined by its strength  $\theta_{\mathbf{q}}^{(t)}$  and the distance between  $\mathbf{c}_{\mathbf{p}}$  and  $\mathbf{c}_{\mathbf{q}}$  feature vectors, is greater than the current cell strength  $\theta_{\mathbf{p}}^{(t)}$ :

$$g(\mathbf{c}_{\mathbf{p}}, \mathbf{c}_{\mathbf{q}}) \cdot \theta_{\mathbf{q}}^{(t)} > \theta_{\mathbf{p}}^{(t)}, \quad (4.9)$$

where  $g(\mathbf{c}_{\mathbf{p}}, \mathbf{c}_{\mathbf{q}})$  is a pixel similarity function, which is monotonically decreasing and bounded. This property ensures the convergence of the algorithm.

The utilized function in “GrowCut” [634] is based on the absolute image feature difference:

$$g_{\text{IFD}}(\mathbf{c}_{\mathbf{p}}, \mathbf{c}_{\mathbf{q}}) = 1 - \frac{\|\mathbf{c}_{\mathbf{p}} - \mathbf{c}_{\mathbf{q}}\|_2}{\max_{\forall \mathbf{c}_{\mathbf{i}} \in \mathcal{I}} \|\mathbf{c}_{\mathbf{i}}\|_2}. \quad (4.10)$$

The authors of [314] demonstrated that “GrowCut” is equivalent to the Bellman–Ford algorithm.

In [243, 244], the “Tumor-Cut” CA model implements the gradient magnitude:

$$g_{\text{GM}}(\mathbf{c}_{\mathbf{p}}, \mathbf{c}_{\mathbf{q}}) = e^{-\|\mathbf{c}_{\mathbf{p}} - \mathbf{c}_{\mathbf{q}}\|_2}. \quad (4.11)$$

The resulting algorithm is equivalent to the shortest path algorithm, as shown in [569].

Recalling the general graph-based theoretical framework:  $x_i$  can be seen as the cell strength  $\theta_i$ , while the edge weights  $w_{ij}$  are related to the values taken by the pixel similarity function  $g(\mathbf{c}_i, \mathbf{c}_j)$ .

### CA in brain tumor modeling

CA have been widely employed in dynamic complex system modeling and simulation [36]. Thanks to the capability of representing the individual state of each particle, CA are well suited for microscopic modeling and simulation approaches. Even if each cell takes a local decision, aggregate measures and complex emergent behaviors derive from these decisions. Thereby, CA models can be applied to diverse real phenomena and research fields.

In particular, CA were also adopted to simulate early tumor growth, by examining the roles of host tissue vascular density and tumor metabolism in the ability of a small number of monoclonal transformed cells to develop into an invasive tumor [457]. Computationally, the mathematical description of tumor growth can be formulated at different spatial scales: (i) simulating the growth at the cellular level (using CA), (ii) defining at macroscopic scale how the tumor density will evolve (with Partial Differential Equations (PDEs)). However, the models attempt to predict the mathematical law of the tumor growth in both cases [20]. Kansal *et al.* [308] developed a three-dimensional CA model of brain tumor growth, showing that macroscopic tumor behavior can be realistically modeled using microscopic parameters. As a matter of fact, the rules of division and invasion are the key elements of the CA-based approaches. The hybrid model proposed in [457] incorporates normal cells, tumor cells, necrotic or empty space, and a random network of native microvessels as components of a CA state vector. More recently, Ibrahim-Hashim *et al.* [286] showed that intra-tumoral subpopulations can be defined using cellular adaptive strategies and intra-tumoral Darwinian interactions rather than molecular properties employed in branching clonal evolution models. This multi-scale hybrid mathematical model coupled a discrete CA—as an Agent-Based Model (ABM) for describing explicitly the space and state transitions—and a set of reaction-diffusion PDEs—for spatio-temporal interactions. A reaction-diffusion model describes the change in the distribution and number of tumor cells due to the random movement of tumor cells (diffusion) and proliferation (reaction).

### Brain tumor segmentation based on a CA model

This section presents a novel semi-automatic MR brain tumor image segmentation approach (GTV<sub>CUT</sub>) based on a CA model.



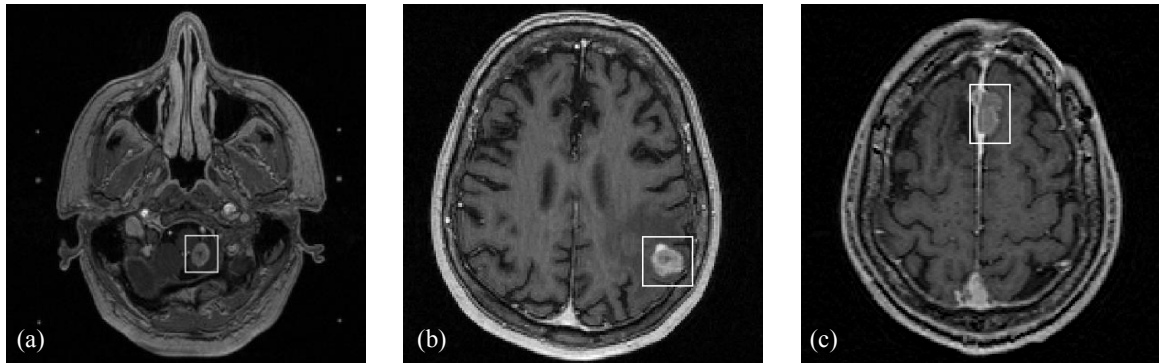


Fig. 4.25 Three instances of input T1w contrast-enhanced MR axial slices with the rectangular bounding area, selected by the user, containing the tumor: (a) moderately hyper-intense enhancing region including a necrotic core; (b) highly hyper-intense enhancing region with inner necrotic material; (c) heterogeneous cancer.

**Patient dataset description** The study was performed on a dataset composed of 25 patients (average age:  $59.84 \pm 9.99$  years), affected by metastatic brain cancers, who underwent Gamma Knife neuro-radiosurgery. Some subjects among the enrolled patients showed multi-metastatic scenarios, so the total number of lesions was 32.

All available T1w FFE CE-MRI series were acquired on a Philips Gyroscan Intera 1.5 T MRI scanner, a few hours before the Gamma Knife treatment and then utilized for the planning phase. Three significant examples of input MR images, representing three brain cancers with different characteristics, are shown in Fig. 4.25. MRI acquisition parameters are reported in Table 4.1.

**GTVcut: brain cancer segmentation for neuro-radiosurgery** The analyzed brain metastatic cancers, wherein necrotic and enhancing tumor tissues are often mixed, exemplify the challenges encountered in radiosurgery scenarios. Thus, interactive segmentation approaches represent a more feasible and safe solution for physician in clinical practice with respect to fully automatic approaches [243]. Since region connectedness is ensured by the CA algorithm, segmentation results are represented by connected-regions with unbroken edges even if some necroses or cysts extend partially outside the tumor. Dealing with connected target regions is very important in radiation therapy treatment planning, because tumor delineation is a critical step that must be performed accurately and effectively. Moreover, since cancer imaging is usually characterized by non-uniform intensity distributions, local image statistics should be considered instead of global ones [27]. CA model is very suitable due to its own local update rules. GTVcut results in a simple but effective application based on a plain cellular automata model, by initializing adaptively the foreground and background

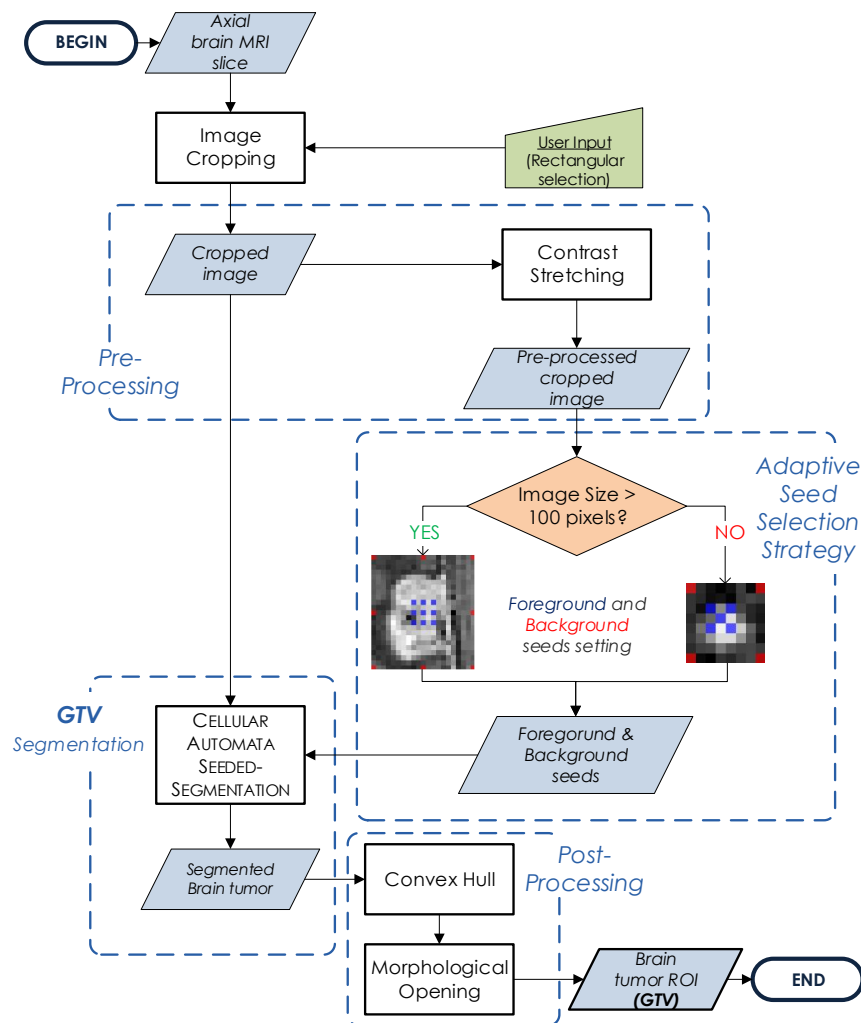


Fig. 4.26 Flow diagram of the GTVCUT segmentation approach. The pipeline can be divided into four main stages: (i) pre-processing to improve segmentation process; (ii) adaptive seed selection to choose suitable foreground and background seed pixels; (iii) GTV segmentation using a CA model; (iv) post-processing to refine the achieved segmentation results.

seeds in a smart fashion. A detailed flow diagram of the proposed brain tumor segmentation approach is shown in Fig. 4.26.

According to the interaction technique employed in “GrabCut” [511], user intervention consists just in dragging a rectangle around the target area, without parameter settings. Unlike “Graph cuts” [65] and “GrowCut” [634] algorithms, wherein the user has to define the foreground and background seeds explicitly by drawing the respective brushing pixels, our approach calculates both foreground and background seeds adaptively by using an adaptive strategy. As a matter of fact, operator-dependency is minimized by lightening the load

involving the user. This results in a very straightforward and usable ROI selection tool, which is generally provided by medical imaging clinical applications [98].

**Pre-processing** After the bounding region selection accomplished by the user, the cropped input images are pre-processed in order to improve the subsequent segmentation procedure. The range of intensity values of the selected part is expanded to enhance the GTV extraction. Thereby, a linear contrast stretching operation is applied. This linear normalization converts input intensity values  $r$  into the output values  $s \in [0, 1]$ , by implying an expansion of the values between  $r_{\min}$  and  $r_{\max}$ , and improving detail discrimination.

**Adaptive seed selection strategy** A strategy for the choice of both foreground (tumor) and background (healthy tissue) seeds was designed *ad hoc*. Especially, according to the flow diagram in Fig. 4.26, seed selection depends on the size of the image cropped by the operator, and then it is adaptive and fully automatic. Let  $\mathcal{I}_{\text{cropped}} : \{1, 2, \dots, M\} \times \{1, 2, \dots, N\} \rightarrow [0, 1]$  be the cropped MR image, whose size and center coordinates are  $M \times N$  and  $C \equiv (C_x, C_y) \equiv (\text{round}(M/2), \text{round}(N/2))$ , respectively. Adaptive seed selection is performed as follows:

- if  $M \times N > 100$ , nine foreground pixels in the central zone and eight background pixels at the border of the cropped image are chosen, such as in Fig. 4.27(a, b, c), to provide significant samples for the CA algorithm initialization;
- otherwise, to avoid the access to pixels outside the cropped image, five foreground pixels in the central zone and four background pixels at the corners of the cropped image are chosen (see Fig. 4.27d).

This discrimination is necessary because very small lesions could be present (5 mm diameter), such as in initial and final tumor slices. Foreground tumor seed-points are set according to the center coordinates of the cropped image, hence the only trivial requirement for the operator is that the rectangular bounding region must be centered on the tumor. This strategy implements a highly robust seed selection procedure.

**GTV segmentation** A two-labeled CA segmentation is the most efficient scheme for classifying a hyper-intense lesion from the healthy part of the brain. As a matter of fact, in contrast-enhanced MR brain scans, metastases have a brighter core than periphery and distinct borders differentiating them from the surrounding healthy brain tissue [18]. Nevertheless, the proposed method can be safely applied to benign and malignant brain tumors, since it relies on contrast medium absorption by tumor tissues. In some instances, either central or adjacent areas of the lesions could comprise necrotic tissue [105, 557] that, for radiotherapeutic

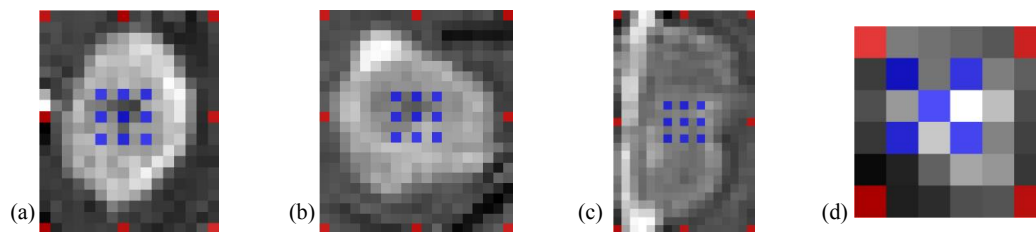


Fig. 4.27 Adaptive seed selection according to the rectangular boundary regions (including the tumor) dragged by the user: (a, b, c) cropped image size greater than 100 pixels (input MR slices in Fig. 4.25(a, b, c), respectively); (d) cropped image size less than or equals to 100 pixels. Foreground and background seeds are shown using transparent blue and red pixels (alpha factor  $\alpha = 0.40$ ), respectively.

purposes, must be considered in the CTV during the treatment planning phase. Our method is able to segment correctly heterogeneous tumors, even when foreground seeds are located in necrotic core proximity. The pseudo-code for the proposed image segmentation GTVCUT approach based on a CA model is given in the Algorithm 4 box. Differently to “GrowCut” [634], the user does not explicitly select seed-points. In the proposed segmentation approach the user has to drag only a bounding rectangle. Inheriting the interaction modality used in “GrabCut” [511], the user is indicating the region of background and is free of any need to select foreground markers because seed selection is accomplished automatically by the adaptive strategy explained previously. Unlike “Tumor-Cut” [243], in which tumor segmentation is performed by a level set evolution on a tumor probability map, the proposed method is based solely on a CA model for the automated GTV extraction. We used both the pixel similarity functions reported in Eqs. (4.10) and (4.11) and an experimental comparison was performed.

**Post-processing** A pair of simple morphological operations [571] is also required to refine the segmentation results achieved by the proposed CA-based approach. Firstly, a convex hull algorithm is used to envelope the segmented lesion into the smallest convex polygon that contains the automatically segmented region as well as to remove possible holes at tumor boundaries. In particular, this operator allows for: (i) hole filling when necroses or cysts are fully included in the tumor, and (ii) whole tumor detection when part of the necrotic material may extend partially outside the tumor. Secondly, a morphological opening with a circular structuring element (1-pixel radius) is employed to smooth the GTV contour shape by unlinking poorly connected pixels.

---

**Algorithm 4** Pseudo-code of the GTVCUT segmentation algorithm based on a CA model.

**Input:** Foreground (tumor) seeds  $\mathcal{F}$ , Background seeds  $\mathcal{B}$ , Cropped MRI axial slice  $\mathcal{I}_{\text{cropped}}$

**Output:** ROI segmentation result  $\mathcal{S}$

**Parameters:**  $\theta_{\mathbf{p}}^{(t)} \in [0, 1]$  is the strength value of the cell  $\mathbf{p}$  at the current time step  $t$   
 $l_{\mathbf{p}} \in \{0, 1, 2\}$  is the label of the cell  $\mathbf{p}$  at the current time step  $t$ , according to:

$$\begin{cases} l_{\mathbf{p}}^{(t)} = 0 : \mathbf{p} \text{ is an unlabeled cell} \\ l_{\mathbf{p}}^{(t)} = 1 : \mathbf{p} \text{ is assigned to the foreground class} \\ l_{\mathbf{p}}^{(t)} = 2 : \mathbf{p} \text{ is assigned to the background class} \end{cases}$$

```

1: for each pixel  $\pi_i \in \mathcal{I}_{\text{cropped}}$  do                                ▷ Initialization of the pixels in  $\mathcal{I}_{\text{cropped}}$ 
2:    $l_{\pi_i}^{(0)} \leftarrow 0; \theta_{\pi_i}^{(0)} \leftarrow 0$ 
3: end for
4: for each foreground seed  $\varphi_i \in \mathcal{F}$  do                            ▷ Initialization of the foreground pixels
5:    $l_{\varphi_i}^{(0)} \leftarrow 1; \theta_{\varphi_i}^{(0)} \leftarrow 1$ 
6: end for
7: for each foreground seed  $\beta_i \in \mathcal{B}$  do                            ▷ Initialization of the background pixels
8:    $l_{\beta_i}^{(0)} \leftarrow 2; \theta_{\beta_i}^{(0)} \leftarrow 1$ 
9: end for
10: while convergence is not achieved do
11:   for each cell  $\mathbf{p} \in \mathcal{I}_{\text{cropped}}$  do                                ▷ Moore 8-neighborhood tries to attack the cell
   ▷ For each neighbor cell  $\mathbf{q} \in \mathcal{N}(\mathbf{p})$  find  $\bar{\mathbf{q}}$  with  $\max \{g(\mathbf{p}, \mathbf{q}) \cdot \theta_{\mathbf{q}}^{(t)}\}$ 
12:    $\bar{\mathbf{q}} \leftarrow \arg \max_{\mathbf{q} \in \mathcal{N}(\mathbf{p})} \{g(\mathbf{p}, \mathbf{q}) \cdot \theta_{\mathbf{q}}^{(t)}\}$ 
13:   if  $g(\mathbf{p}, \bar{\mathbf{q}}) \cdot \theta_{\bar{\mathbf{q}}}^{(t)} > \theta_{\mathbf{p}}^{(t)}$  then
14:      $\theta_{\mathbf{p}}^{(t+1)} \leftarrow g(\mathbf{p}, \bar{\mathbf{q}}) \cdot \theta_{\bar{\mathbf{q}}}^{(t)}; l_{\mathbf{p}}^{(t+1)} \leftarrow l_{\bar{\mathbf{q}}}^{(t)}$                                 ▷ Update the current state
15:   else
16:      $\theta_{\mathbf{p}}^{(t+1)} \leftarrow \theta_{\mathbf{p}}^{(t)}; l_{\mathbf{p}}^{(t+1)} \leftarrow l_{\mathbf{p}}^{(t)}$                                 ▷ Keep the previous state
17:   end if
18:   end for
19: end while
20: for each cell  $\mathbf{p} \in \{\mathcal{I}_{\text{cropped}} : l_{\mathbf{p}}^{\text{final}} = 1\}$  do
21:    $\mathcal{S}_{\mathbf{p}} \leftarrow 1$                                             ▷ ROI segmentation output
22: end for

```

---

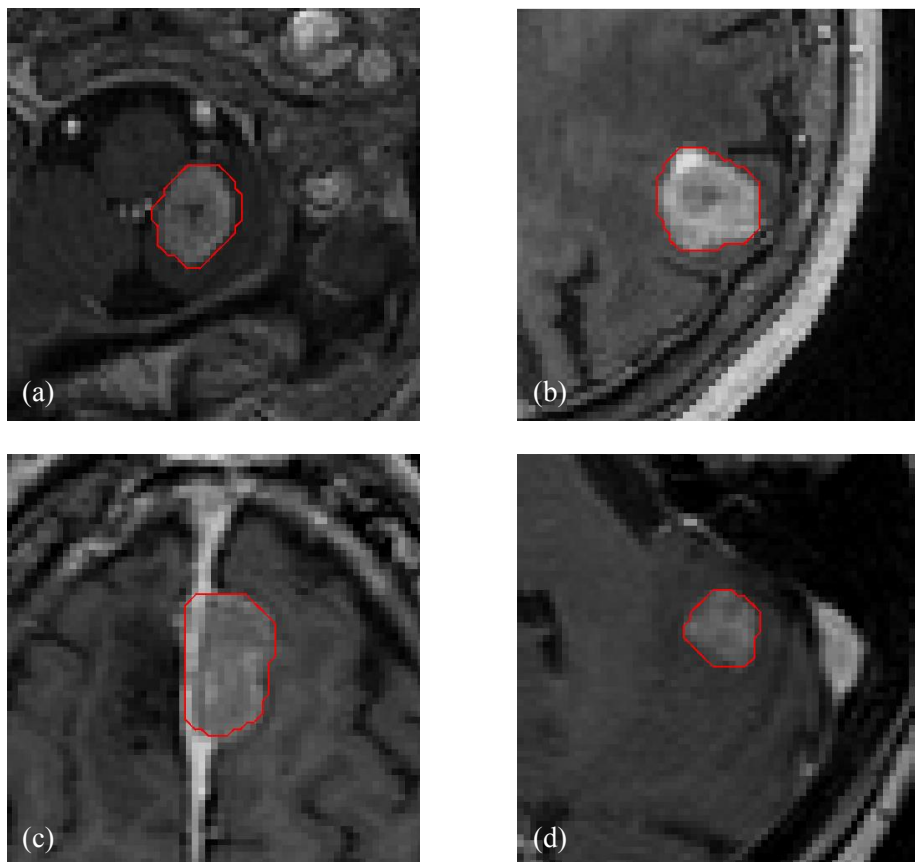


Fig. 4.28 Brain tumor segmentation using GTVCUT in different anatomic areas of human brain: (a) brainstem (Fig. 4.25a); (b) occipital lobe (Fig. 4.25b); (c) temporo-parietal region in proximity to the superior sagittal sinus (Fig. 4.25c); (d) anterior cerebellar lobe. ( $4\times$  zoom factor).

**Experimental results** Fig. 4.28 shows four GTV segmentation results achieved using the proposed GTVCUT approach. It is appreciable how GTVCUT obtains correct results in different brain areas, even with inhomogeneous input MRI data.

A first clinical validation can be performed by a quantitative analysis on tumor volume measurements. Accordingly, the proposed and manual GTVs for the brain cancers (statistical samples), considered in the MRI dataset, were compared. The agreement between the automatic and manual volume measurement can be graphically represented by the Bland-Altman plot in Fig. 4.29, which shows that more than 90% of the volume differences lie within  $\pm 1.96$  standard deviations around the mean difference ( $0.1 \text{ cm}^3$ ). Therefore, the GTVCUT approach achieves reproducible results also considering the different brain tumor sizes and shapes included in the analyzed real MRI dataset.

We developed and tested two different version of the proposed method:

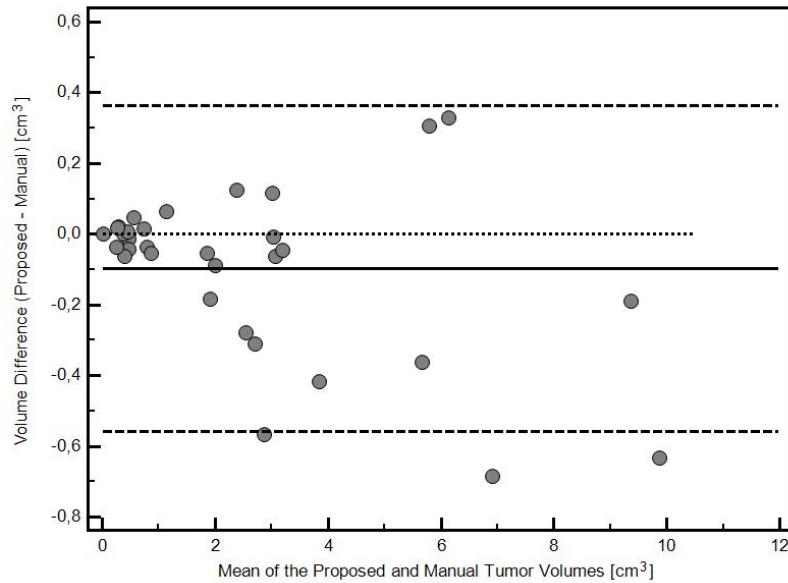


Fig. 4.29 Bland-Altman plot of brain tumor volume measurements: the results yielded by the GTVCUT method against the manual segmentations by an expert neurosurgeon (gold standard). Solid line: mean; dashed lines:  $\pm 1.96$  standard deviations.

- GTVCUT<sub>IFD</sub> implements the pixel similarity function given in Eq. (4.10), i.e., the absolute image feature difference;
- GTVCUT<sub>GM</sub> utilizes the pixel similarity function given in Eq. (4.11), i.e., the gradient magnitude (absolute intensity difference measure).

Moreover, all the evaluation metrics were also calculated on the same MRI datasets using our previous brain tumor segmentation method based on the FCM clustering algorithm [415, 525], which outperformed the most common literature segmentation methods, as it has been shown in [415]. This approach was implemented in MatLab<sup>®</sup> R2014a by exploiting the Image Processing Toolbox<sup>™</sup> and the Fuzzy Logic Toolbox<sup>™</sup>.

Table 4.10 reports the mean and standard deviation values of spatial overlap-based metrics obtained in the experimental segmentation tests on the analyzed 32 brain tumors (concerning the 25 patients undergone neuro-radiosurgery) using the GTVCUT<sub>IFD</sub>, GTVCUT<sub>GM</sub> and FCM-based method. High *DSC* and *JI* values prove the segmentation accuracy and reliability. In addition, *SEN* and *TNR* average values involve the correct detection of the “true” pathological areas as well as the ability of not detecting wrong parts within the segmented tumors. The GTVCUT<sub>GM</sub> segmentation approach achieved slightly better results than GTVCUT<sub>IFD</sub>. Therefore, we chose the pixel similarity function in Eq. (4.11) to be employed in the local transition rule of the GTVCUT computational model. To provide a graphical representation of

Table 4.10 Values of the spatial overlap-based metrics regarding the achieved GTV segmentation results. The results are expressed as average value  $\pm$  standard deviation.

Method	DSC	JI	SEN	TNR	FPR	FNR
GTV <sub>CUT</sub> <sub>IFD</sub>	90.83 $\pm$ 4.12	84.01 $\pm$ 6.64	91.30 $\pm$ 6.96	99.99 $\pm$ 0.01	0.008 $\pm$ 0.009	6.218 $\pm$ 6.455
GTV <sub>CUT</sub> <sub>GM</sub>	90.88 $\pm$ 4.19	84.11 $\pm$ 6.74	91.20 $\pm$ 7.00	99.99 $\pm$ 0.01	0.007 $\pm$ 0.008	6.353 $\pm$ 6.482
FCM [413]	90.88 $\pm$ 4.19	84.11 $\pm$ 6.74	91.20 $\pm$ 7.00	99.99 $\pm$ 0.01	0.007 $\pm$ 0.008	6.353 $\pm$ 6.482

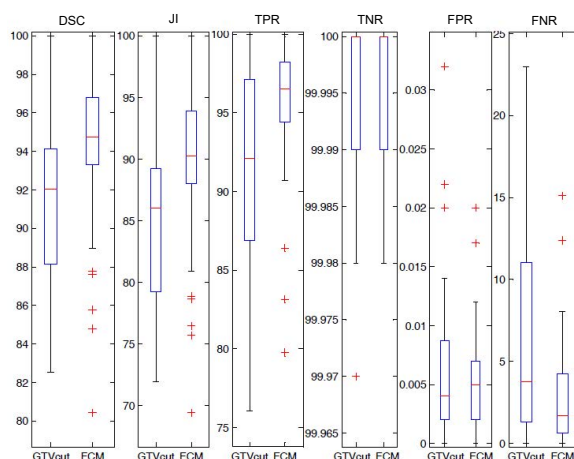


Fig. 4.30 Boxplots of the spatial overlap-based metrics values achieved by the GTV<sub>CUT</sub> approach and the FCM-based method in [413]. The median value (i.e., the second quartile) is represented by a red line dividing the box. Whisker value is 1.5 in all cases.

the statistical distribution of the results, the corresponding boxplots of spatial overlap-based evaluation metrics achieved by GTV<sub>CUT</sub> and the FCM-based approach are also reported in Fig. 4.30. The short width of the interquartile range represented in boxplots implies that values are very concentrated. All index distributions present a small number of outliers, thus demonstrating extremely low variability (whisker value is 1.5 in all cases).

Table 4.11 shows the spatial distance-based metrics mean and standard deviation values achieved by the GTV<sub>CUT</sub><sub>IFD</sub>, GTV<sub>CUT</sub><sub>GM</sub>, and FCM-based segmentation methods. The achieved spatial distance-based indices are consistent with overlap-based metrics. Hence, good performances were obtained also with heterogeneous cancers, since the CA algorithm delineates accurately irregular and complex shapes. The achieved segmentation results are comparable with the FCM-based approach in [413]. Moreover, the distance-based metrics values calculated on the GTV<sub>CUT</sub><sub>GM</sub> segmentations are slightly lower than GTV<sub>CUT</sub><sub>IFD</sub> results, confirming the trend observed in the overlap-based metrics. Fig. 4.31 depicts the boxplots of the spatial distance-based evaluation metrics achieved by GTV<sub>CUT</sub> and our previous FCM-based method in [413].

Even though there are some outliers in *AvgD* and *MaxD* boxplots (the whisker is set to 1.5 in all cases), just a small deviation between the segmentations of the proposed method



Table 4.11 Values of the spatial distance-based metrics regarding the achieved GTV segmentation results. The results are expressed as average value  $\pm$  standard deviation.

Method	AvgD	MaxD	HD
GTV <sub>CUT</sub> <sub>IFD</sub>	$0.492 \pm 0.223$	$1.426 \pm 0.495$	$1.650 \pm 0.453$
GTV <sub>CUT</sub> <sub>GM</sub>	$0.487 \pm 0.223$	$1.414 \pm 0.483$	$1.638 \pm 0.443$
FCM [413]	$0.329 \pm 0.246$	$1.172 \pm 0.576$	$1.389 \pm 0.448$

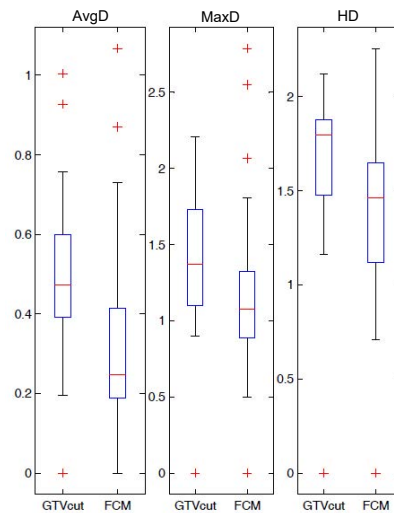


Fig. 4.31 Boxplots of the spatial distance-based metrics values achieved by the GTV<sub>CUT</sub> approach and the FCM-based method in [413]. The median value (i.e., the second quartile) is represented by a red line dividing the box. Whisker value is 1.5 in all cases.

and those of the experienced neurosurgeon is denoted. Considering the surgery treatment planning purpose of the application, referring to MRI spatial resolution (see characteristics in Table 4.1), all distance-based metrics mean values are in the range of 1 mm. This confirms the great validity and accuracy of the proposed segmentation method.

**Conclusions** In this section, an original application of CA for semi-automatic MRI brain GTV segmentation (GTV<sub>CUT</sub>) was presented. The proposed technique can be used for brain tumors segmentation in neuro-radiosurgery treatments. The developed methodology can be integrated in the current clinical practice, by supporting physicians in the target volume delineation task during Gamma Knife treatment planning. The proposed GTV<sub>CUT</sub> segmentation method was tested on an MRI dataset composed of 32 metastatic brain concerning real patients, who underwent Leksell Gamma Knife treatment, by calculating both spatial overlap-based and distance-based metrics. The achieved experimental results are very encouraging and show the effectiveness of the proposed approach.

Considering the used CA-based models, the main experimental finding is that  $\text{GTV-CUT}_{\text{GM}}$ , based on the gradient magnitude defined in Eq. (4.11), obtained slightly better segmentation results than  $\text{GTV-CUT}_{\text{IFD}}$ , which implements the image feature difference in Eq. (4.10). In addition, the experimental results achieved by  $\text{GTV-CUT}$  and by our previous FCM-based approach in [413, 525] are comparable. The  $\text{GTV-CUT}$  segmentation approach works better than FCM-based segmentation when heterogeneous tumors with either diffused inner necrotic material or cysts are processed. Indeed, the employed adaptive seed selection strategy is more robust in these very critical cases with respect to the FCM clustering algorithm that could yield also disjoint output regions. Moreover,  $\text{GTV-CUT}$  does not require any smoothing operations to reduce MRI acquisition noise and no imaging details are missed. These small differences between the  $\text{GTV-CUT}$  and the FCM-based segmentation results, revealed in the experimental trials, can be also due to the different interactive selection tools utilized by the two approaches; the method in [413] employs a free-hand “lasso” tool and  $\text{GTV-CUT}$  uses a draggable bounding rectangle. However, the second ROI selection tool is more usable and less operator-dependent than the first one, by minimizing the user’s degrees of freedom [511]. Furthermore, the overall segmentation approach based on the FCM clustering algorithm in [413] exploited also a sophisticated strategy for necrotic material inclusion, which has to be considered in the resulting GTV.

### 4.2.2 Random Walker

The RW algorithm [221] encodes an image as a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with nodes (vertices)  $v \in \mathcal{V}$  corresponding to the voxels and edges (arcs)  $e \in \mathcal{E}$  are associated to a Gaussian cost function that maps a change in image intensity to the edge weights  $w_{ij}$  defined by means of the following Gaussian function:

$$w_{ij} = e^{-\beta \|g_i - g_j\|^2}, \quad (4.12)$$

where:  $g_i$  and  $g_j$  are the image intensity values at voxels  $v_i$  and  $v_j$ , respectively;  $\beta$  is a free parameter. It is also assumed that the graph  $\mathcal{G}$  is connected and undirected (i.e.,  $w_{ij} = w_{ji}$ ).

The image is then converted into a lattice where some pixels’ classes are known while other pixels are unassigned. In the original method in [221], the known nodes are marked by user input. The segmentation problem is solved by assigning a label to unknown nodes, by finding the minimum energy among all possible graph scenarios to achieve an optimal segmentation. The RW method classifies the nodes into foreground and background classes, considering the probability that a “random walker” starting at a source node, first reaches a node with a pre-assigned label by visiting every voxel. Every unlabeled node is assigned

to the subset which is most likely a RW would reach first starting from that pixel. The RW problem has the same solution as the combinatorial Dirichlet problem, defined by the Dirichlet integral:

$$D[u] = \frac{1}{2} \int_{\Omega} |\nabla u|^2 d\Omega, \quad (4.13)$$

for a field  $u$  and a region  $\Omega$ . This problem can be solved by a harmonic function that satisfies the Laplace equation:

$$\nabla^2 u = 0. \quad (4.14)$$

More specifically, the harmonic function that satisfies the boundary conditions minimizes the Dirichlet integral, since the Laplace equation is the Euler-Lagrange equation for the Dirichlet integral [131].

The combinatorial Laplacian matrix  $\mathbf{L}$  [153] can be defined as:

$$L_{ij} = \begin{cases} \deg(v_i), & \text{if } i = j \\ -w_{ij}, & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes} \\ 0, & \text{otherwise} \end{cases} \quad (4.15)$$

The matrix  $\mathbf{L}$  can be decomposed into four sub-matrices:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_M & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{L}_U \end{bmatrix}, \quad (4.16)$$

where  $\mathbf{L}_U$  is a sub-matrix of edge weights for the unlabeled voxels and  $\mathbf{L}_M$  is a sub-matrix of edge weights for the marked nodes.  $\mathbf{B}$  is a sub-matrix of the probabilities corresponding to the labeled voxels.

Accordingly, the harmonic function can be found by solving the system of linear equations in Eq. (4.17):

$$\mathbf{L}_U \mathbf{X} = -\mathbf{B}^\top \mathbf{M}, \quad (4.17)$$

where:  $\mathbf{M}$  is a Boolean matrix representing the boundary conditions of the seeds;  $\mathbf{X}$  is the probability for each node being a member of the labels. A threshold of 50% is chosen to discriminate the foreground from the background yielding a voxel binary mask  $\mathcal{S}$ , so that:

- $\mathcal{S}_{\mathbf{p}_{\text{fg}}} \leftarrow 1$ , for foreground nodes  $\mathbf{p}_{\text{fg}}$  with probability higher than or equal to 0.5;
- $\mathcal{S}_{\mathbf{p}_{\text{fg}}} \leftarrow 0$ , for background nodes  $\mathbf{p}_{\text{bg}}$  with probability lower than 0.5.

Briefly, a probability map is produced, and a threshold of 50% is chosen to discriminate between foreground and background voxels.

### Brain lesion segmentation on PET images

In medical imaging, especially when dealing with PET, the RW method is able to localize weak boundaries as part of consistent boundaries. To obtain the BTV delineation, the RW parameters were modulated to incorporate PET information:  $g_i$  and  $g_j$  have been replaced with the Standardized Uptake Value (SUV) [584, 585, 122] in the voxels  $i$  and  $j$ :

$$w_{ij} = e^{-\beta(\text{SUV}_i - \text{SUV}_j)^2}, \quad (4.18)$$

where:

$$\text{SUV} = \frac{c_{\mathcal{I}}}{i_{\text{dose}}/W_{\text{pat}}}, \quad (4.19)$$

$c_{\mathcal{I}}$  is the radioactivity activity concentration [kBq/ml] measured from an image  $\mathcal{I}$ ,  $i_{\text{dose}}$  is the decay-corrected amount of injected radiotracer [kBq], and  $W_{\text{pat}}$  is the patient's weight [g].

The SUV is the most common semi-quantitative parameter used to estimate radiotracer accumulation within a lesion in clinical practice [375]. It normalizes the voxel activity considering acquisition time, administered activity, and patient's body weight [696]. Hence, the PET image is converted into a lattice where the SUV of each voxel is assigned to the corresponding graph node  $v \in \mathcal{V}$  and the edge weights  $w_{ij}$  are computed accordingly.

The RW method is very sensitive to the choice of  $\beta$  factor in the Gaussian-like weighting function in Eq. (4.18).  $\beta$  influences how quickly the probability decreases with increasing intensity differences [586]: a high  $\beta$  value reduces the weight of the walker, which weakens the connection between the adjacent voxels and under-estimates the foreground volume, whilst a low  $\beta$  value increases the weight, which over-estimates the target volume.

In addition, to obtain a fully automatic and user-independent method, two crucial improvements have been implemented to overcome two critical issues in the RW methodology:

1. an automatic method to localize starting target and background seeds: the RW algorithm requires a set of pre-labeled seeds, which may be generated by the user, making the RW delineation sensitive to the location of the pre-labeled voxels;
2. a strategy to adaptively determine the appropriate probability threshold rather than fixed one of 50% on the probabilistic output of the RW in order to overcome the operator-dependence of the weighting factor  $\beta$  in Eq. (4.12).

A more detailed explanation of these improvements is reported in [585, 586].

### **Multimodal PET/MRI brain metastasis segmentation**

The study presented in [526] investigated the impact of BTV segmentation, using MET-PET imaging, and the subsequent co-registration with MR images, utilized to delineate the GTV, in stereotactic neuro-radiosurgery therapy. The main goal was to present a novel multimodal PET/MRI automatic segmentation method, combining complementary information, and encourage its use in future Gamma Knife treatments.

To the best of our knowledge, this approach is the first work that presents a co-segmentation method to integrate MET-PET metabolic information with anatomic MRI in stereotactic neuro-radiosurgery treatments. The proposed method improves and combines two computer-assisted and operator-independent methods to segment BTV and GTV from PET [584–586] and MR [413, 525] images, respectively. Each single method has been already validated [413, 584], comparing each of them against the most common literature segmentation methods. The former uses the graph-based RW algorithm [221] and the latter is based on unsupervised FCM clustering [52, 53]. Both segmentation methods are valid operator-independent approaches to identify the BTV and the GTV, respectively, in order to delineate a comprehensive CTV that includes metabolic and morphologic information, useful for treatment planning and patient follow-up. Our method addresses the co-segmentation problem only on multimodal PET/MRI in the brain anatomic district imaging, by exploiting the great sensitivity and specificity of brain MET-PET in distinguishing healthy and pathological tissues. As result, a fully automatic approach was obtained.

The co-segmentation tests on 19 metastatic brain tumors, treated with Leksell Gamma Knife radiosurgery, were retrospectively performed to evaluate the effectiveness of the proposed approach. Overlap-based and spatial distance-based metrics were calculated between PET and MRI segmentation results. Statistics was also considered to quantify correlation. Finally, a qualitative evaluation, based on a Likert score scale, was carried out by three experienced clinicians.

It is worth to note that clinical evaluation is not just a mathematical procedure, based just on tumor area difference or on GTV and BTV union, but a deep decision-making process involving several anatomic and metabolic insights, the specific patient's pathological scenario as well as the physician experience. The CTV identification is actually a critical task accomplished by expert physicians according to tumor volumes (i.e., BTV and GTV), and deciding about the regions to be included or excluded in the planned target volume. This may pave the way for personalized therapy, by integrating molecular imaging and allowing for a diagnosis and therapy that is specialized to the individual metabolism and disease [66].

**Dataset description** PET brain acquisitions without head frame were performed using a time-of-flight PET/CT Discovery 690 by General Electric Medical Systems (Milwaukee, WI, USA). Patients fasted for 4 hours before the PET exam and were intravenous injected with MET. The PET protocol began 10 minutes after the injection. PET images consist of a  $256 \times 256$  pixel matrix of  $1.1719 \times 1.1719 \times 3.27$  mm<sup>3</sup> voxel size.

All available MRI datasets were acquired on a Philips Gyroscan Intera 1.5 T MRI scanner, before treatment, for the planning phase. MR images are T1w FFE CE sequences, using the following MRI acquisition parameters are: TR: 25 ms, TE: 1.808-3.688 ms, matrix size:  $256 \times 256$  pixels, slice thickness: 1.5 mm, slice spacing: 1.5 mm, pixel spacing: 1.0 mm. Therefore, the size of each voxel was  $1.0 \times 1.0 \times 1.50$  mm<sup>3</sup>. Imaging data are encoded in the 16-bit DICOM format. Representative instances of input brain MR and PET image pairs are shown in Fig. 4.32.

**The proposed multimodal PET/MRI segmentation approach** Fig. 4.33 outlines the overall flow diagram of the proposed multimodal PET/MRI segmentation method. For a better understanding, PET and MR image processing pipelines are represented with blue and purple blocks, respectively. It is worth noting the smart combination of the two single modality pipelines (green connections in Fig. 4.33). In particular, PET and MRI segmentation results mutually exploit each other. Firstly, the IUR, obtained on co-registered PET images by the graph-based RW tumor segmentation method, is used for the generation of an area including the tumor ROI on MR brain images. These ROI bounding regions are calculated adaptively on MR images using an LSF method [357] and then utilized by the MRI brain lesion segmentation method based on the FCM clustering technique. Lastly, the MRI GTV masks are combined with PET images, which are processed by the RW algorithm, to influence and refine BTV segmentation on PET images.

All the employed parameters were fixed in the proposed computational model for all the analyzed datasets. These settings are completely hidden to the end-users (clinicians), who have not to select any parameters. Therefore, this does not affect the fully automation of the proposed method. First of all, the IUR, obtained on co-registered PET images using the graph-based BTV segmentation method, is exploited for the automatic generation of a suitable region including the tumor on MR brain images. These bounding regions, computed adaptively on MR images using an LSF method, are then utilized by the GTV segmentation method. Lastly, GTV masks are combined with PET images to influence the BTV segmentation. The presented multimodal segmentation approach yields the tumor boundaries in both PET and MRI modalities, conveying different information, which are not always complementary, considering that enhancement, edema, and necrosis regions are

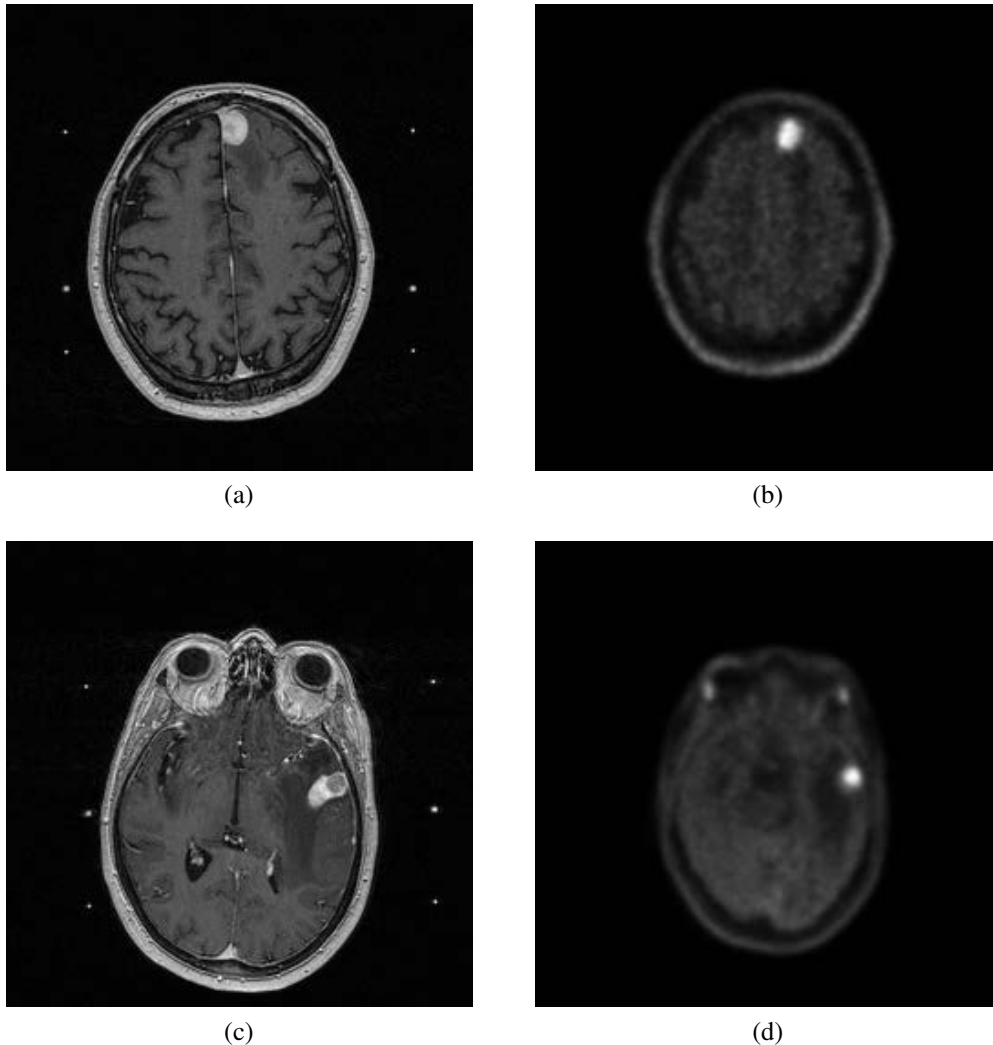


Fig. 4.32 Examples of original input brain MR and PET image pairs concerning patients with brain tumors who underwent Gamma Knife: (a, c) MR slices and (b, d) the nearly respective corresponding PET slices (tumors #6 and #14). Note the different dimensions and the substantial tridimensional misalignment between corresponding MR and PET images. The images are displayed in gray-scale.

imaged differently by the modalities [171]; so the tumor volumes defined on PET and on CT or MRI could be highly different [576].

The main key novelties of the proposed multimodal segmentation approach are:

- presenting a fully automatic method for GTV segmentation, calculated automatically by exploiting the IURs detected on PET images to define a suitable bounding region for the tumor zone with a DRLSE approach [361];

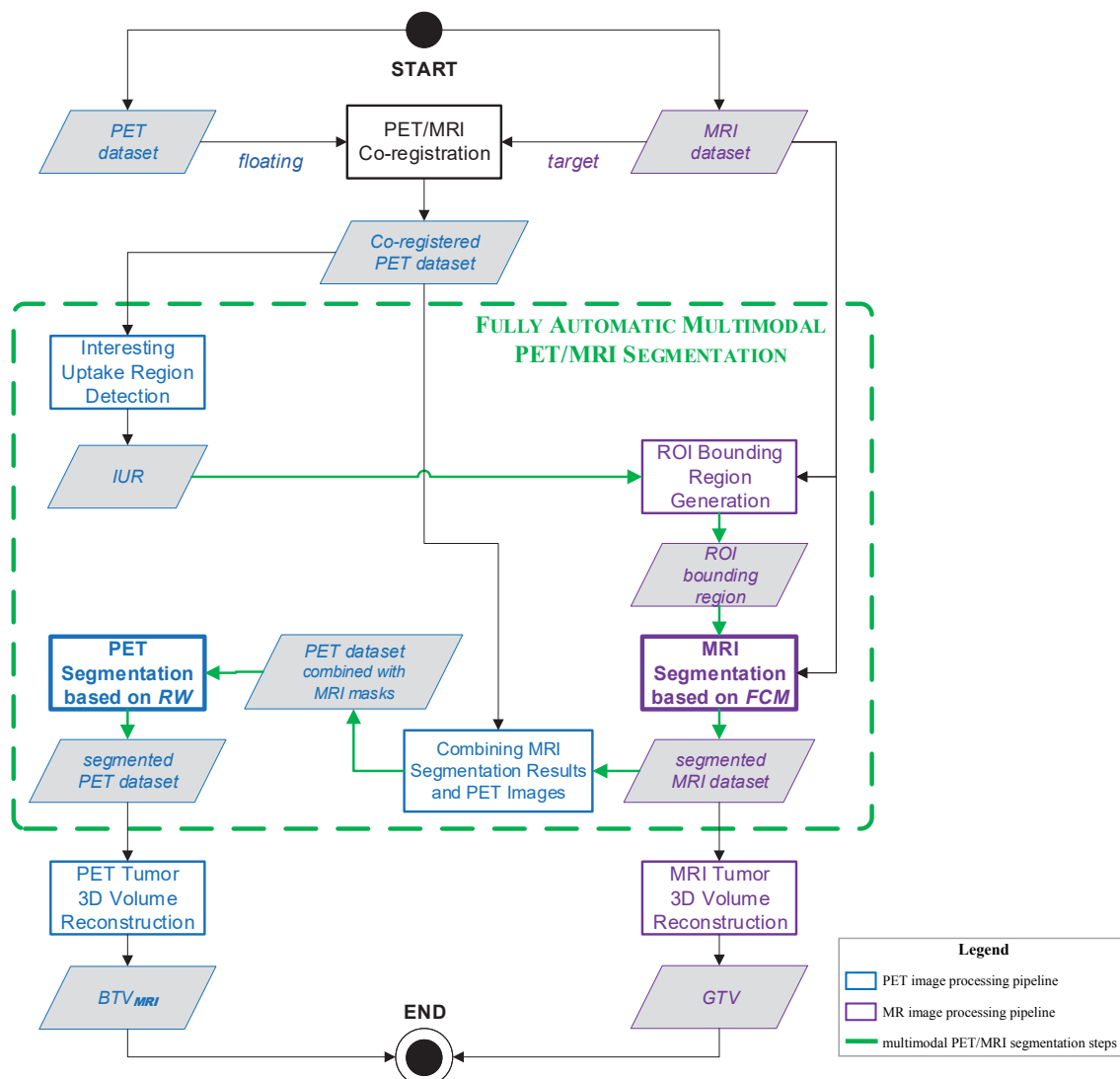


Fig. 4.33 Flow diagram of the proposed fully automatic multimodal PET/MRI segmentation method. The adopted graphical and color notations are explained in the legend box.

- computing the  $BTV_{MRI}$ , a new BTV that can reduce radioactivity spill-in and spill-out effects, between tumor and surrounding tissues, affecting the segmented BTV (according to the morphologic GTV information) and the subsequent  $BTV_{MRI}$  integration for a comprehensive treatment planning.

Thereby, a complete knowledge about the clinical scenario is achieved from both anatomic and metabolic imaging perspectives. In conclusion, the tackled problem and the aim of the proposed multimodal approach are considerably different with respect to joint segmentation approaches. Employing a decision-level fusion, the assumption of a “ground truth” joint volume, defined on fused multimodal imaging data, and the consequent comparison of the



proposed method with related works is not significant, and sometimes misleading. Under these hypotheses, a reference gold standard for the evaluation is not consistent with the purpose of our work. This retrospective study aims to evaluate the feasibility and the clinical value of BTV integration in Gamma Knife treatment planning.

**PET/MRI modalities co-registration** An image co-registration stage is mandatory to bring the different MRI and MET-PET datasets, concerning the same patient, into the same reference space. In this way, it is possible to make quantitative and meaningful comparisons between the brain lesion segmentation results achieved by both MRI and PET image segmentation methods (see Section 5.2.1 for more details on medical image registration). In our multimodality PET/MRI registration, MRI is used as reference image (target) while the PET is the source image (floating), because MRI conveys more anatomic information than PET. Indeed, PET imaging is characterized by weak boundaries and has a lower spatial resolution than MRI. Realignment and reslicing operations are thus required to get a one-to-one mapping between PET and MRI slices. From an algorithmic perspective, image co-registration involves finding parameters (i.e., geometric transformation matrix) that either maximize or minimize some objective function. However an accurate interpolation method is clearly required, by which the floating and target images are sampled when being represented in different spaces [381, 474]. PET/MRI inter-modal 3D registration was performed using SPM 12 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, University College, London, UK), a software package designed for the analysis of brain imaging data sequences [463]. We relied on the SPM tool that is widely used in the neuroimaging community and in the clinical routine, also for Voxel-Based Morphometry (VBM). The registration method used by SPM is based on the work by Collignon *et al.* [121], where the original interpolation method has been changed to give a smoother cost function. The images are also smoothed slightly, by means of their histogram. This makes the cost function as smooth as possible, to give faster convergence and less chance of local minima. A 3D rigid-body model, parameterized by 3 translations and 3 rotations about the different axes, is used by means of voxel-to-voxel affine transformations. This approach is very efficient for brain anatomic district. We chose Normalized Mutual Information (NMI) as the cost function to be optimized. The Mutual Information (MI) registration criterion states that the mutual information of the image intensity values of corresponding voxel pairs is maximal if the images are geometrically aligned. MI is the most intensively investigated criterion for registration of intra-individual human brain images [114, 475]. In addition, PET/MRI registration misalignment can be large with respect to the FOVs, then a criterion invariant to image overlap statistics is very important, such as the

NMI [589]. Since MET-PET and MRI multimodal images are written in a different space, an efficient resampling and interpolation method must be utilized. Although Nearest-Neighbor and Trilinear interpolations are faster, it may be better to use a higher degree approach to achieve more accurate results (i.e., reducing the deviation from an ideal low-pass filter, by avoiding artifacts mostly near the edges) [610]. In particular, we used the 4th Degree B-Spline interpolation.

Two instances of PET/MRI co-registration are shown in Fig. 4.34, where the quality of the achieved registration can be qualitatively appreciated with the checkerboard images (Fig. 4.34(a, d)) as well as fused PET/MR images (Fig. 4.34(b, e)). In addition, the joint histogram, which is a feature space constructed by counting the number of times a combination of gray values occurs contextually on source (MET-PET) and reference (MRI) images, was also computed and plotted (Fig. 4.34(c, f)). Along the axes of each joint histogram the gray values of the two images are represented: from left to right for source (MET-PET) and from bottom to top for reference (MRI). Original and final joint histograms prove the quality of the NMI registration process, because pixel values are mainly redistributed on the histogram diagonal.

**Interesting uptake region detection** The fully automatic multimodal PET/MRI segmentation starts with the automatic identification of the PET slices with the maximum SUV ( $SUV_{max}$ ), relative to each lesion detected by means of the iterative procedure reported in Algorithm 5, on PET dataset previously co-registered against the corresponding MRI. The voxels with a SUV greater than 95% of the  $SUV_{max}$  concerning each tumor are marked as target seeds. The neighborhood of the node with  $SUV_{max}$ , through searching in all 8 neighboring directions, is explored to identify the voxels with a SUV less than 30% of the average of target seed SUVs [584]. In this way, 8 background nodes are identified. Once the foreground and background seeds are automatically localized, the PET lesion is segmented using our enhanced RW method: the probability threshold to discriminate between target and background voxels is obtained slice by slice to follow the whole lesion volume. In this way, the method takes into account the intensity gradient and contrast changes of the metabolic lesion over the entire range of PET slices. Once the BTV has been extracted, the ROI of the PET slice with the  $SUV_{max}$  is propagated to the corresponding MRI dataset. In addition, the range of segmented PET slices with high uptake regions is also provided to the next processing phase, to generate efficiently a ROI bounding region for the tumor on MR images. Such information defines the so-called IUR.

By taking advantage of the great sensitivity and specificity of  $^{11}\text{C}$ -labeled MET radiotracers in the discrimination among malign versus benign tissues, each brain tumor is independently processed by our algorithm. For each patient study, the highest uptake regions

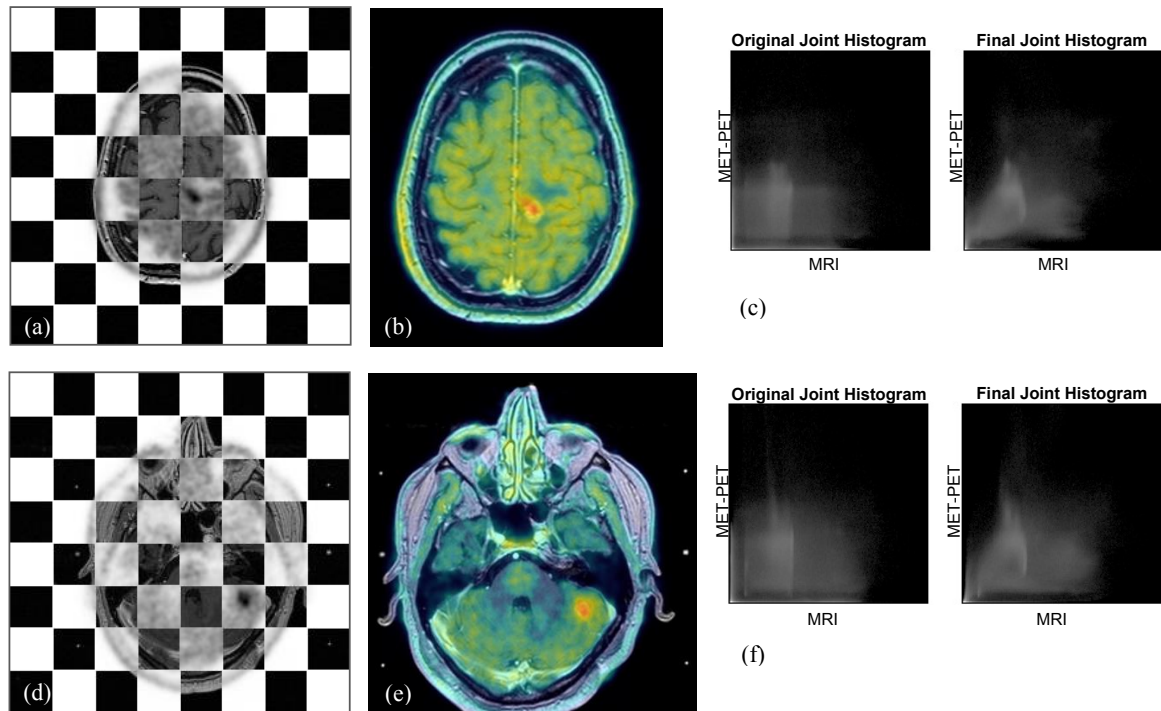


Fig. 4.34 Instances of PET/MRI co-registration using SPM *via* maximization of NMI (tumors #7 and #10): (a, d) checkerboard images of co-registered PET (negative gray-scale) and MRI (normal gray-scale); (b, e) fused images (MRI in gray-scale and PET in Jet colormap); (c, f) corresponding original and final joint histograms (feature space constructed by counting the number of times a combination of gray values occurs contextually on MRI and MET-PET images) for (b) and (e), respectively.

are automatically analyzed. The  $SUV_{\max}$  value is not global, but is relative to each single brain tumor present in the PET study under examination. Therefore, in the case of  $L_{\text{tot}}$  brain tumors a different  $SUV_{\max}(l_i)$  is identified for each lesion  $l_i$ , with  $i \in \{1, 2, \dots, L_{\text{tot}}\}$ .

**ROI bounding region generation** A coarse lesion delineation on MR images is required in order to define automatically a bounding region that encloses the actual GTV on MR slices. Hence user interaction, needed by the methods in [413, 525] through a ROI selection tool, will be completely prevented and avoided. Because of the variability of brain lesions in terms of location and intensity values, the definition of a valid ROI bounding region is not a trivial problem. A bad choice of the ROI bounding region on MR image might affect the whole FCM clustering process. Therefore, this task must be carried out not in a static way (e.g., using a simple morphological dilation or closing with a fixed structuring element [68]), but rather dynamically by processing MRI input data. This adaptive procedure starts on the lesion ROI found on the  $SUV_{\max}$  slice. This IUR is first dilated (using a structuring element

---

**Algorithm 5** Iterative procedure for the IUR detection and BTV delineation on PET images.

---

**Input:** Analyzed PET study

**Output:** List  $\mathcal{L}_{\text{BTV}}$  that stores the coordinates of the voxels included in the segmented  $L_{\text{tot}}$  brain lesions on the processed PET study

```

1:  $\mathcal{L}_{\text{BTV}} \leftarrow \{\}$  ▷ Initialize  $\mathcal{L}_{\text{BTV}}$  as an empty list
2:  $i \leftarrow 1$  ▷ Initialize the iteration counter for the detected tumors
3: do
4:   Find the slice with the highest  $\text{SUV}_{\text{max}}^{(i)}$  in the PET study, except the voxels in  $\mathcal{L}_{\text{BTV}}$ 
5:   Let  $l_i$  be the current lesion with  $\text{SUV}_{\text{max}}^{(i)}$ 
6:    $\text{SUV}_{\text{max}}(l_i) \leftarrow \text{SUV}_{\text{max}}^{(i)}$  ▷ Set the  $\text{SUV}_{\text{max}}(l_i)$  regarding  $l_i$  to  $\text{SUV}_{\text{max}}^{(i)}$ 
7:    $\text{BTV}_i \leftarrow \text{BTV}_{\text{SEGMENTATION}}(l_i)$  ▷ Segment the current  $l_i$  using the RW algorithm
   ▷ Store the segmented  $\text{BTV}_i$  voxel coordinates, by updating the list  $\mathcal{L}_{\text{BTV}}$ 
8:    $\mathcal{L}_{\text{BTV}} \leftarrow \mathcal{L}_{\text{BTV}} \cup \text{GETVOXELCOORDINATES}(\text{BTV}_i)$ 
9:    $i \leftarrow i + 1$  ▷ Increment the iteration counter
   ▷ Check if the  $\text{SUV}_{\text{max}}^{(i)}$  in the  $i$ -th iteration is at least 2.0 (minimum threshold value for considering a candidate region to be a tumor), otherwise stop the loop
10: while  $\text{SUV}_{\text{max}}^{(i)} \geq 2.0$ 

```

---

represented by a disk of 3-pixel radius) and then utilized as a binary step function to initialize an LSF-based method. A rough and over-estimated MRI brain lesion segmentation, based on the DRLSE formulation proposed in [357], is used to generate dynamically a bounding region that includes the brain lesion in the MR slice corresponding with the SUVmax PET slice. Unlike region growing algorithms [3, 524], the DRLSE method allows a controlled and regular LSF evolution, by avoiding leaking in the brain tissue area. In short, Level Set methods represent an active contour as the zero level set of a higher dimensional function, i.e., a time-dependent LSF  $\phi(x, y, t) : \Omega \rightarrow \mathbb{R}$  (where  $\Omega$  is the domain), and determine the evolution of the contour [446].

Three examples of ROI bounding regions delineated, on original MR images, by the DRLSE-based method are shown in Fig. 4.35(a-c). These computations are performed just once for each dataset. In fact, only one bounding region is determined *via* DRLSE-based segmentation and it is extruded along the  $z$  axis onto the other candidate MR slices. For this purpose, the range of PET slices with high uptake regions is also provided by the IUR detection steps, to construct a cylindroid (i.e., a cylinder with irregular-shaped bases) including the whole GTV brain lesion. Since PET and MRI tumor volumes may be distributed along different axial slices, the range of PET active slices is augmented by three slices on upper and lower sides, in order to ensure the total enclosing of the tumor into the cylindroid defined on the MRI dataset. Fig. 4.35d shows a 3D volume rendering of the cylindroid

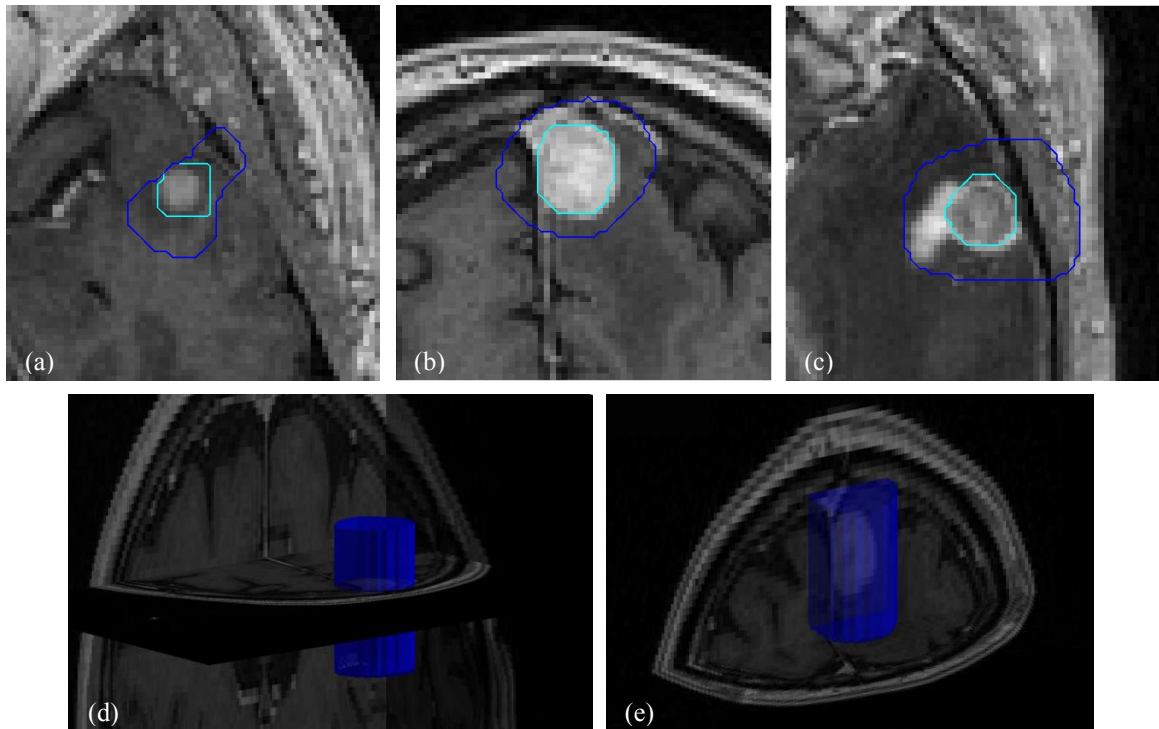


Fig. 4.35 ROI bounding region generation on MR images: (a, b, c) three examples (tumors #5, #6, and #10) of bounding region (blue contour) determined *via* a DRLSE-based method, starting from IUR (cyan contour) imported from the corresponding  $SUV_{max}$  PET slice. Cylindroid extrusion on MRI dataset according to active PET slice range: (d) tridimensional bounding cylindroid for the tumor in the temporal brain lobe, reported in (b), rendered with transparent blue surface (Alpha blending with  $\alpha = 0.70$ ) on 3D orthogonal tri-planar view (middle slice is shown in the axial plane), that includes the whole GTV, as visible in (e).

extruded on the ROI bounding region shown in Fig. 4.35a. The whole GTV is included into the cylinder volume (see Fig. 4.35e).

**MRI segmentation based on the FCM algorithm** Typical MRI protocols, employed in stereotactic radio-surgery treatment planning, include only T1w CE-MRI of the head volume. Therefore, no multiparametric MRI structural data (i.e., T1w, T2w, Proton Density, and FLAIR images) are generally available and brain tumor segmentation approaches based on multispectral MR images cannot be applied to Leksell Gamma Knife treatment planning. These approaches rely on these multispectral MRI for segmenting and distinguishing cancer enhancement region, necrosis or edema [154]. As a matter of fact, tumor detection and segmentation on T1w CE-MRI datasets are based on the hyper-intense enhancement regions. Since sometimes either edema or necrotic material (hypo-intense regions in T1w CE-MR images) could be present in the analyzed tumors, necrosis inclusion in the planned target

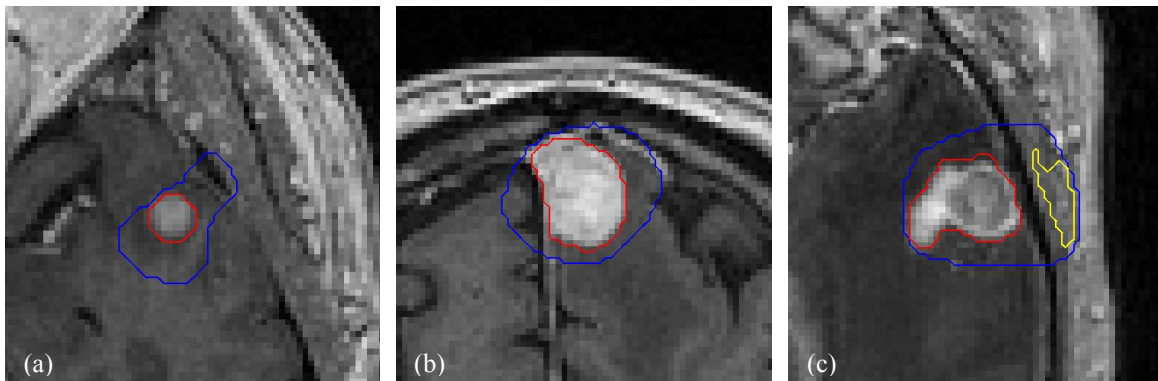


Fig. 4.36 Examples (tumors #5, #6, and #10) of brain GTV segmentation on MR images (red contour), starting from the ROI bounding region (blue contour). A difficult case is shown in (c), where a lengthened hyper-intense region of the head skin is detected by the FCM clustering (yellow contour), but it is removed using the shape-based control.

volume is required for radiosurgery purposes [523, 557]. Please refer to Section 4.1.1 for an extensive description of the method [413, 525].

Unfortunately, due to the automation of the segmentation multi-modal process, the direct control of the human operator is lost during ROI bounding region manual delineation. Some MR images included in the extruded cylindroid, especially the most distant slices from the IUR, might include no brain tumor as well as anatomic parts with pixel values similar to enhancement regions (such as bone or epidermal tissue). For the former case, our MRI brain lesion segmentation method yields an empty ROI. For the latter case, a shape-based control on connected-components belonging to the cluster with the highest intensity pixels is also defined, by taking advantage of the pseudospherical appearance of metastatic brain tumors [18]. When either skull bone or skin are included in the ROI bounding region, since the IUR is located near brain boundary, other high-valued pixel areas may be erroneously segmented by the unsupervised FCM clustering algorithm. These areas are often characterized by a lengthened shape and may be removed by checking eccentricity (that is, the ratio between the foci distance related to an ellipse and its own major axis length) and extent (that is, the ratio between the pixels belonging to the region and the bounding box pixels) of the connected-components. Theoretically suitable and experimentally validated feature values for this refinement step are:  $\text{extent} \geq 0.6$  and  $0.0 \leq \text{eccentricity} < 0.8$ .

Fig. 4.36 shows segmentation results on the images reported in Fig. 4.35. In particular, refinement steps can be appreciated in Fig. 4.36c, where the region composed of the skin highest pixels included in the ROI bounding region is correctly removed by the shape-based controls (yellow boundary).

Although the literature offers 3D approaches for brain lesion segmentation [307], we chose a 2D methodology because 3D FCM should be not well suitable on clusters with more general shapes [281]. The analyzed brain tumors are composed of inhomogeneous cancerous tissue and necrotic material. As a matter of fact, tumor detection and segmentation on T1w CE-MRI datasets are based on the hyper-intense enhancement regions. Sometimes a dark area might be present due to either cystic or necrotic tissue, which must be included into the GTV for radiotherapeutic purposes. These hypo-intense areas may negatively affect the FCM clustering process when 3D MRI data are processed. In these cases, clustering could yield several disconnected volumes and successful post-processing steps are very difficult. On the other hand, performing FCM cluster analysis on the single MRI slices is definitely more reliable and post-processing steps proposed in [413] are able to segment the whole tumor.

**PET segmentation based on the RW algorithm** The RW-based segmentation process on the input PET dataset is updated and repeated, in the same way, on the same PET dataset combined with the corresponding MRI GTV binary masks in order to achieve a new BTV, called as  $BTV_{MRI}$ . In PET imaging, the adaptive probability threshold is computed separately slice-by-slice to follow the whole lesion volume, by taking into account the intensity gradient and contrast changes of the metabolic lesion over the entire range of PET slices. This key feature could be lost using a fully 3D method: the proposed method is then performed in parallel for the slices adjacent to the starting PET slice with the  $SUV_{max}$  [584].

**Combining MRI segmentation results and PET images** Lesion segmentation results on the morphologic brain MRI data are combined with the metabolic information of the co-registered PET images. This PET dataset represents the input of the RW algorithm. In this way, MRI GTV is utilized to combine the superior contrast of PET images with the higher spatial resolution of MR images. Each PET slice is weighted pixel-by-pixel according to the corresponding MRI ROIs. Nevertheless, an assumption that there is a one-to-one correspondence between BTV and GTV is unrealistic. Lesions may have smaller uptake regions compared to GTV, such as shown in Figs. 4.37c and 4.38c. In the same way, the PET lesion can show additional area compared to lesion boundaries in MR images (see Figs. 4.38d 4.39c). For this reason, we incorporate, with extreme caution, the MRI GTV binary mask in the RW-based segmentation procedure. In particular, PET pixels inside the corresponding MR target are multiplied by a gain factor to slightly emphasize probable lesion area; PET pixels outside the corresponding MRI ROIs are weakly underestimated in order to reduce the radioactivity spill-in effect—from background into the lesion—and spill-out effect—from the lesion into background—by taking advantage of the higher spatial resolution of MR images.

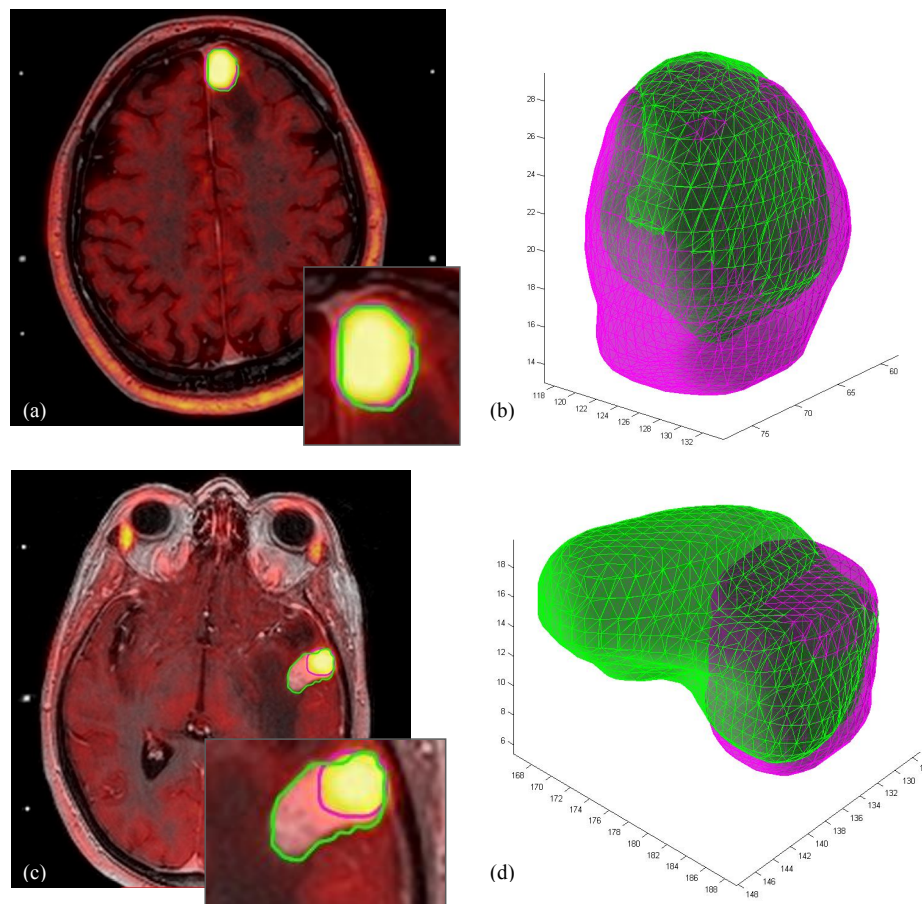


Fig. 4.37 Results achieved by the proposed multimodal PET/MRI segmentation method on the input image pairs (tumors #6 and #14) reported in Fig. 4.32(a, b) and Fig. 4.32(c, d), respectively. The  $BTV_{MRI}$  (magenta contour) and GTV (green contour) are superimposed: (a) boundaries computed on PET and MR images are nearly overlaid; (c) the two ROIs are very different in this case. At the bottom right, the lesion regions are zoomed in both cases. Tridimensional volume reconstructions of brain tumors in (a) and (c) are shown in (b) and (d), respectively. Transparent surfaces are rendered with Alpha blending ( $\alpha = 0.50$ ) to emphasize volume intersections.

**Results** The evaluation of the data presented in this study was performed retrospectively on 19 metastatic brain tumors, treated with Leksell Gamma Knife. Nowadays, the clinical protocol does not usually consider co-registered PET images for Gamma Knife treatment planning. In the proposed multimodal segmentation approach, two different PET segmentation results were computed: (i) BTV, by considering PET images alone, and (ii)  $BTV_{MRI}$ , by considering PET/MRI co-segmentation.

Firstly, volumes measured on PET and MRI data are reported and, for a quantitative assessment, PET/MRI volume difference and centroid distance were calculated for each



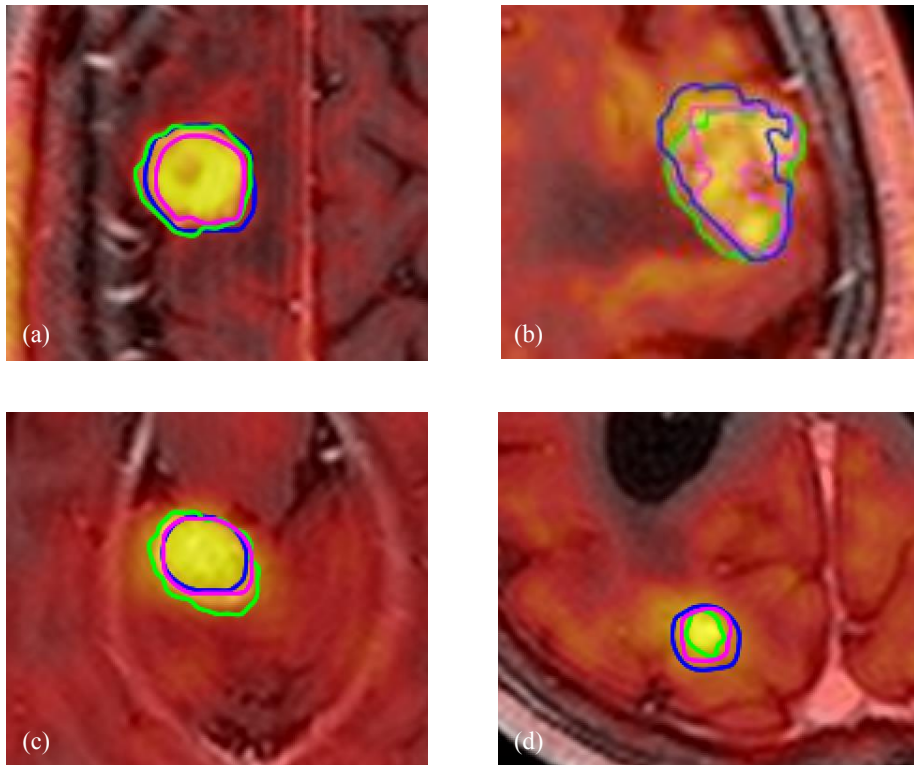


Fig. 4.38 Graphical examples of co-segmentation results achieved by the proposed multi-modal PET/MRI approach, where GTV (green contour), BTV (blue contour), and  $BTV_{MRI}$  (magenta contour) ROIs are overlaid on the corresponding fused PET/MRI slices (displayed in Hot Look-Up Table). Four metastatic tumors in different anatomic regions of human brain are considered: (a) left temporo-parietal (tumor #3), (b) right temporo-parietal (tumor #11), (c) limbic (tumor #15), and (d) occipital (tumor #13).

lesion. Figs. 4.37a and 4.37c show experimental results achieved by the proposed multimodal PET/MRI segmentation method on the input image pairs reported in Figs. 4.32(a, b), and Figs. 4.32(c, d), respectively. In addition, the corresponding GTV and  $BTV_{MRI}$  volume rendering is reported in Figs. 4.37b and 4.37d, respectively. Three other interesting examples obtained by the proposed multimodal segmentation method can be visually and qualitatively evaluated in Fig. 4.38.

It is worth to observe that the margin of enhancement of tumors in MRI datasets is not always strongly correlated with high uptake regions in PET images. Segmentation results of PET images combined with GTV MRI masks are very stable (magenta boundaries in Fig. 4.38), even if GTV is quite different. In Fig. 4.38d, PVE on PET image [583] is reduced by using MRI ROIs.

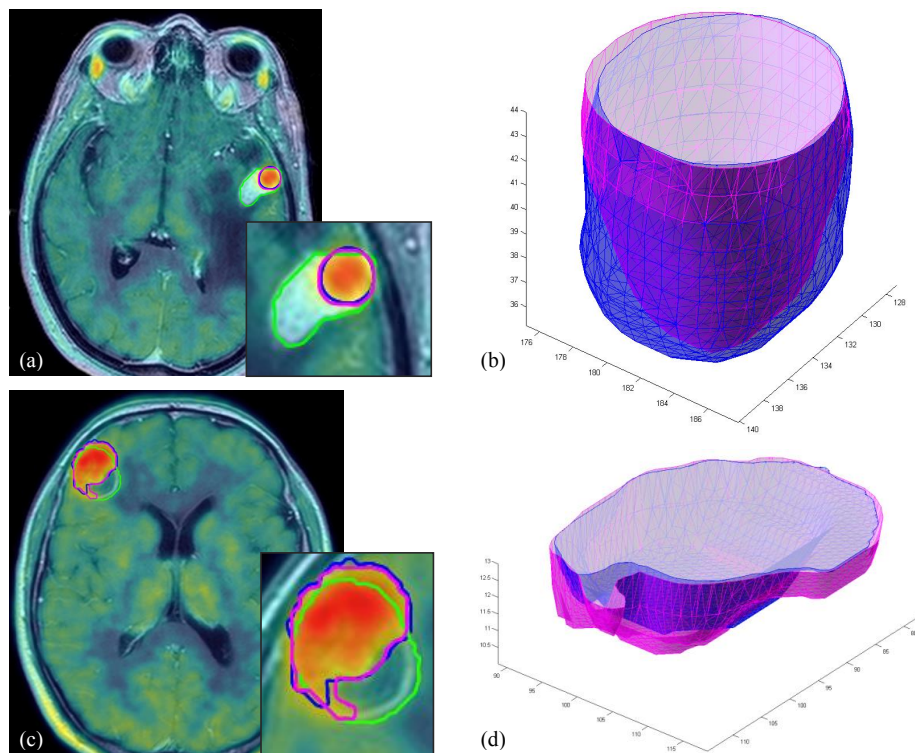


Fig. 4.39 Segmentation results, on brain tumors #14 and #16, obtained by the single modality PET method (BTV) versus those obtained by the multimodal PET/MRI approach after the PET data updating based on the results of the MRI brain lesion segmentation ( $BTV_{MRI}$ ): (a, c) GTV (green contour), BTV (blue contour), and  $BTV_{MRI}$  (magenta contour) ROIs overlaid on the corresponding fused PET/MRI slice (PET and MRI are displayed using Jet and gray-scale LUTs, respectively); (b, d) 3D representation of BTV and  $BTV_{MRI}$  rendered with transparent surfaces (Alpha blending with  $\alpha = 0.40$ ), respectively, cut at the corresponding slice shown in (a) and (c).

**PET and MRI tumor volume measurements** The volumes, calculated for each brain tumor imaged on MRI and PET modalities, are shown in Table 4.12. Absolute average volume difference (defined as  $\text{abs}(V_1 - V_2)/V_2$ ) between BTV,  $BTV_{MRI}$ , and GTV were also calculated—by generalizing the formulation in Eq. (A.1)—in order to evaluate either overestimation or underestimation of PET functional imaging with respect to anatomic MRI data, before and after GTV binary mask combination into PET images. In addition, Euclidean distances between BTV,  $BTV_{MRI}$ , and GTV centroids (centers of mass) were reported.

First of all, mean BTV and GTV values and their average absolute volume difference attest that tumor regions are imaged and represented differently on PET and MRI datasets. Furthermore, high standard deviation in BTV and GTV measurements implies that we are dealing with metastatic lesions characterized by various types and dimensions. On the

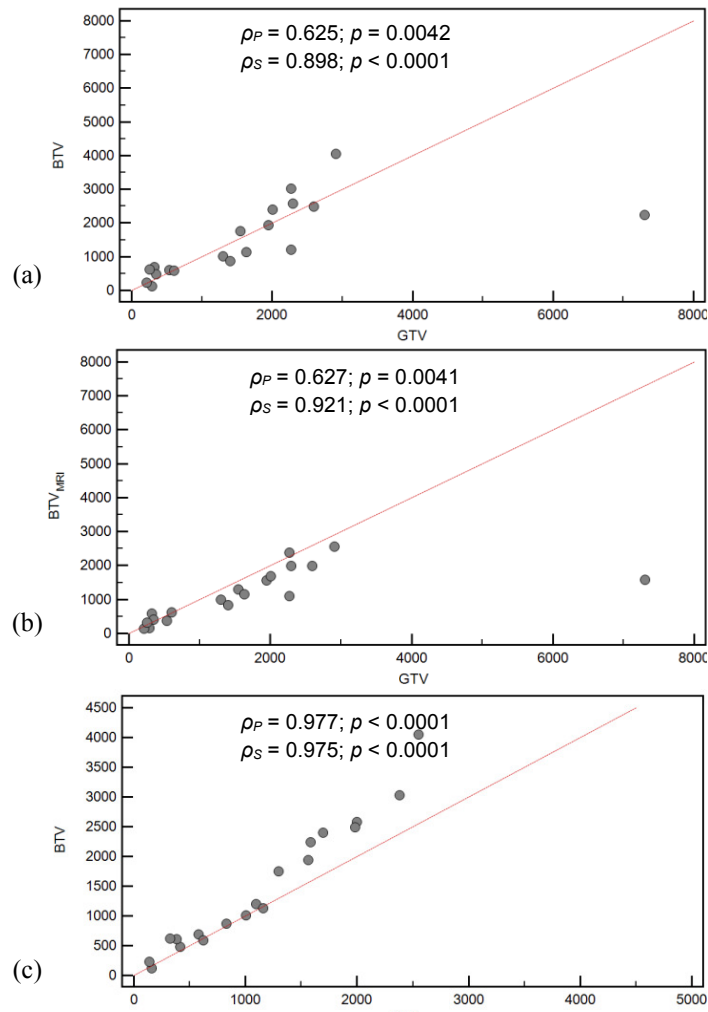


Fig. 4.40 Scatter diagrams of tumor volume measurements on multimodal PET/MRI data: (a) BTV versus GTV; (b)  $BTV_{MRI}$  versus GTV; (c) BTV versus  $BTV_{MRI}$ . Red dashed line is the line of equality. In the upper part of each sub-figure, Pearson and Spearman correlation coefficients with the corresponding significance level ( $p$ -value) are also reported.

other hand,  $BTV_{MRI}$  measurements follow a different trend because  $BTV_{MRI}$  segmentation is influenced by the combination of GTV masks and PET images. Although we slightly integrate MRI ROIs with PET datasets during  $BTV_{MRI}$  computation, this conservative choice frequently could reduce radioactivity spill-in and spill-out effects, by exploiting higher MRI spatial resolution. This is evident in the examples shown in Fig. 4.38. Moreover, the  $BTV_{MRI}$  measurement standard deviation is lower than BTV and GTV ones. This means that multimodal PET/MRI segmentation yields more stable results with respect to single modality segmentations, since both metabolic and morphological imaging are considered. These experimental evidences are corroborated by the achieved absolute volume difference

Table 4.12 Brain tumor volume measurements for each metastatic tumor using MR and PET imaging. Absolute volume difference and centroid distance between segmented GTV, BTV, and  $BTV_{MRI}$  are also reported. Mean value and standard deviation are reported in the last two rows.

Tumor	Absolute Volume Difference [voxels]			Volume Centroid Distance [pixels]		
	BTV vs GTV	$BTV_{MRI}$ vs GTV	BTV vs $BTV_{MRI}$	BTV vs GTV	$BTV_{MRI}$ vs GTV	BTV vs $BTV_{MRI}$
#1	33.61	4.98	27.27	1.805	1.067	1.457
#2	13.45	16.11	35.24	1.523	0.740	0.839
#3	0.36	19.74	24.15	1.250	0.949	0.398
#4	112.69	79.26	18.65	2.591	1.275	0.607
#5	39.88	20.52	16.07	1.446	0.845	0.246
#6	12.60	12.95	29.36	0.974	0.971	0.098
#7	13.41	30.98	59.84	2.248	1.042	1.229
#8	22.49	23.10	0.80	1.830	0.802	1.454
#9	57.34	44.41	23.27	2.238	1.681	0.399
#10	1.17	3.50	4.51	0.697	1.080	0.730
#11	39.44	12.18	58.78	0.647	0.772	0.801
#12	7.62	34.29	63.77	3.118	2.720	0.447
#13	143.14	27.45	90.77	5.660	0.840	0.648
#14	47.07	51.83	9.87	5.753	5.551	0.718
#15	30.91	29.13	2.50	1.162	0.768	0.518
#16	3.79	23.37	25.55	5.941	5.614	0.431
#17	38.11	41.03	4.95	0.972	1.099	0.238
#18	19.42	15.49	41.31	1.414	1.211	0.224
#19	69.37	78.36	41.56	7.141	6.769	0.885
Avg.	37.15	29.93	30.43	2.548	1.884	0.651
Std. dev.	37.59	21.37	24.22	2.018	1.890	0.395

values, where  $BTV_{MRI}$  has about the same difference with GTV and BTV. Volume centroid distances reflect the aforementioned situation, too.

The above observations, based on numerical experimental results, are qualitatively supported by the graphical multimodal PET/MRI segmentation results in Fig. 4.39. To measure the degree of dependence between GTV, BTV, and  $BTV_{MRI}$ , both Pearson's correlation coefficient  $\rho_P$  and Spearman's rank correlation coefficient  $\rho_S$  were calculated for each comparison. Spearman's correlation yields different information with respect to Pearson's correlation, since  $\rho_S$  is less sensitive to strong outliers that are in the tails of both samples. In fact, outlier importance is limited to the value of its rank. As expected, GTV revealed greater correlation to  $BTV_{MRI}$ , especially according to Spearman's  $\rho_S$ . In addition, BTV and  $BTV_{MRI}$  are significantly correlated, using both Pearson and Spearman correlation coefficients. Accordingly, scatter diagrams in Fig. 4.40 represent graphically these numerical results.

**Evaluation of the overall multimodal PET/MRI segmentation method** In this section, an overall evaluation of the proposed multimodal PET/MRI segmentation method is presented. Spatial overlap-based, and distance-based metrics were calculated, to measure the similarity and correlation concerning PET and MRI segmentation approaches. Considering

Table 4.13 Achieved tumor spatial overlap-based metrics for GTV, BTV, and BTV<sub>MRI</sub> segmentations. Total average and standard deviation calculated on 19 brain lesions are also reported.

Tumor	DSC [voxels]		
	BTV vs GTV	BTV <sub>MRI</sub> vs GTV	BTV vs BTV <sub>MRI</sub>
#1	63.97	78.89	76.17
#2	80.55	81.11	80.24
#3	81.33	79.81	83.91
#4	40.79	56.32	79.30
#5	55.42	59.24	73.70
#6	76.43	73.26	86.76
#7	58.29	78.46	62.22
#8	68.31	83.90	80.82
#9	39.22	51.24	83.99
#10	72.59	74.20	84.35
#11	65.74	78.71	68.14
#12	46.79	51.72	69.23
#13	52.57	76.90	68.36
#14	53.38	54.62	87.37
#15	76.66	79.08	89.81
#16	67.80	70.59	86.97
#17	65.55	67.03	85.80
#18	77.30	79.51	82.39
#19	32.85	31.66	78.24
Avg.	61.87	68.75	79.36
Std. dev.	14.64	14.07	7.79

these results, it was possible to argue the complementarity between PET and MR imaging modalities, combined by the proposed multimodal segmentation method.

Table 4.13 shows tumor spatial overlap-based metrics for PET and MRI segmentation methods. The achieved results demonstrated that GTV was, on average, more similar to BTV<sub>MRI</sub> than to BTV. The highest mean values and the lowest standard deviation of the evaluation measures were obtained in BTV and BTV<sub>MRI</sub> comparisons.

Table 4.14 reports the values of spatial distance-based metrics achieved by the proposed multimodal PET/MRI segmentation approach in the experimental trials. Distance-based metrics values agreed with achieved volume-based and overlap-based measurements, showing the same trend.

Table 4.15 shows clinical evaluation results regarding both BTV and BTV<sub>MRI</sub> integration with GTV to determine CTV, according to the explanation in Appendix A.4. In the current clinical practice the CTV is always based on the GTV (defined on MR images). As a consequence, the BTV contribution (defined on PET images) cannot get worse the CTV definition from a clinical perspective, i.e., the Likert score is always greater than 2. This qualitative evaluation performed by three physicians using a five-point Likert scale, defined as: 1) strong worsening in CTV definition; 2) moderate worsening in CTV definition; 3) indifferent, neither enhancement nor worsening in CTV definition; 4) moderate enhancement in CTV definition; 5) strong enhancement in CTV definition.

Table 4.14 Achieved tumor spatial distance-based metrics for GTV, BTV, and BTV<sub>MRI</sub> segmentations. Total average and standard deviation calculated on 19 brain lesions are also reported.

Tumor	HD [pixels]			AvgD [pixels]			MHD [pixels]		
	BTV vs GTV	BTV <sub>MRI</sub> vs GTV	BTV vs BTV <sub>MRI</sub>	BTV vs GTV	BTV <sub>MRI</sub> vs GTV	BTV vs BTV <sub>MRI</sub>	BTV vs GTV	BTV <sub>MRI</sub> vs GTV	BTV vs BTV <sub>MRI</sub>
#1	8.544	6.403	7.616	2.669	1.377	1.756	0.412	0.264	0.339
#2	5.831	4.000	5.831	1.343	1.163	1.259	0.340	0.178	0.195
#3	8.485	7.810	4.472	1.400	1.355	1.145	0.436	0.312	0.138
#4	5.831	3.606	5.000	1.090	0.667	0.778	3.497	3.257	0.277
#5	5.385	3.000	3.606	1.295	0.798	0.742	1.857	1.814	0.107
#6	5.099	4.123	3.000	1.121	1.243	0.940	1.208	1.247	0.029
#7	9.220	4.000	5.831	2.233	0.804	1.332	0.786	0.441	0.421
#8	10.630	3.162	10.630	1.312	0.943	0.815	0.610	0.239	0.524
#9	4.472	4.472	1.414	0.727	0.969	0.204	1.392	0.994	0.377
#10	7.810	6.403	3.162	1.368	1.369	0.689	0.285	0.379	0.286
#11	12.000	6.083	11.662	2.630	1.535	2.262	0.734	0.180	0.214
#12	4.472	5.388	2.236	1.898	1.556	1.167	1.496	1.389	0.249
#13	7.280	2.828	7.280	2.345	0.809	1.711	0.530	0.404	0.296
#14	15.621	14.213	3.606	3.618	3.281	0.810	1.115	1.045	0.233
#15	4.243	3.606	2.236	1.250	1.107	0.643	0.286	0.219	0.147
#16	11.705	11.402	7.616	3.503	3.175	1.906	0.788	0.754	0.094
#17	5.385	4.243	2.236	2.136	1.939	0.769	0.267	0.319	0.087
#18	7.071	5.385	4.472	1.782	1.420	1.356	0.374	0.336	0.062
#19	15.811	16.125	10.000	5.820	5.811	1.486	1.446	1.360	0.195
Avg.	8.15	6.12	5.36	2.081	1.648	1.146	0.940	0.796	0.225
Std. dev.	3.57	3.79	3.04	1.211	1.235	0.515	0.787	0.778	0.131

In  $\sim 26\%$  of cases, the CTV definition was strongly influenced by PET imaging. In more than  $50\%$  of cases, the CTV was strongly or moderately conditioned by metabolic imaging: only in 8/19 cases PET volume did not modify the CTV definition. In more than  $25\%$  of cases, BTV<sub>MRI</sub> enhanced the CTV more accurately than BTV, because radioactivity spill-in and spill-out effects between tumor and surrounding tissues were reduced by the contribution of MRI ROIs in PET segmentation, shrinking false positive uptake areas.

**Discussion and conclusions** The proposed fully automated multimodal segmentation method introduced several novelties and advantages. Firstly, we used two computer-assisted single modality segmentation approaches for PET and MR images, which have been already tested using evaluation metrics and validated by clinicians. Although image co-registration is needed to bring multimodal PET/MR image information into the same reference system, the co-segmentation results are not just derived *via* image registration since anatomic and metabolic are carefully integrated, by relying also on the physicians' expertise, in the proposed multimodal PET/MRI segmentation algorithm. These accurate and operator-independent single modality segmentation methods are certainly more effective and reliable on PET and MRI datasets, with respect to approaches that unify the multimodal information into a single graph.

State-of-the-art methods based on a hyper-graph [32, 248] or subgraphs with inter-subgraph arcs [576] are more sensitive to registration errors, especially when multimodal images are not acquired contextually using a hybrid scanner. In these cases, no significant anatomic and functional changes have to be assumed between the images acquired with

Table 4.15 Clinical value of BTV and BTV<sub>MRI</sub>, for each brain metastasis, in CTV definition.

Tumor	Clinical value	
	BTV integration	BTV <sub>MRI</sub> integration
# 1	4	5
# 2	4	4
# 3	3	3
# 4	4	4
# 5	3	3
# 6	4	5
# 7	4	4
# 8	3	3
# 9	3	3
# 10	3	3
# 11	5	5
# 12	3	3
# 13	3	4
# 14	4	4
# 15	3	3
# 16	4	4
# 17	3	3
# 18	4	5
# 19	4	5

different modalities at the same clinical phase to directly construct the hyper-graph or the interconnected subgraphs. The proposed approach was not tailored specifically for the latest generation hybrid PET/MRI scanners. Obviously, in the case of multimodal PET/MRI scanner [87, 301], the scenario is simplified considering that all the initial registration stages are not necessary, whereas the subsequent steps in the processing pipeline are still valid. Methods based on hyper-graphs generally yield a single target volume on the fused multimodal images. This may compromise the segmentation quality since PET and MRI could convey different information, resulting in a disagreement between BTV and GTV, because enhancement, edema, and necrosis regions are imaged differently by the PET and MRI modalities; so the tumor volumes defined on metabolic PET and on anatomic CT or MRI could be highly different [576]. The proposed multimodal method, keeping all anatomic and metabolic information and enabling a decision-level fusion [124, 556], is more reliable on PET and MRI datasets with respect to approaches that unify the multimodal information into a single graph structure. Result repeatability is ensured because it is able to adapt in different pathological scenarios. This choice allows for a greater awareness on the decision-making process, carried out by the clinical staff that is going to plan the neuro-radiosurgery treatment, that brings to the CTV definition. The CTV is determined by the clinical staff, by taking into account the specific patient's pathological scenario. According to the computed BTV and GTV volumes, the CTV identification is a deep decision-making process involving numerous anatomic and metabolic insights. Therefore, the appropriateness of the multimodal segmentation is justified by the addressed scenario, by allowing the clinicians to carefully consider the possibility to include BTV information into the planned CTV and to determine

a personalized therapy for each single cancer metabolism [66]. Secondly, IURs detected on PET images are not blindly propagated to GTV segmentation algorithm, but properly modified according to MRI data using the developed ROI bounding region generation method. This tailored MRI-driven method, based on LSFs, finds adaptively a valid bounding region in order to eliminate user intervention, required by the MRI brain tumor segmentation method in [413]. On the other hand, in [32], the identified seeds on PET images are propagated to the corresponding anatomic images before segmentation process beginning. Hence, two existing and efficient single modality processing pipelines are combined in a smart fashion. In particular, PET and MRI segmentation results are mutually exploited. Lastly, this fully automated approach is very reproducible and reliable, also due to the brain anatomic district imaging. We focused on the application domain regarding brain tumors that underwent Leksell Gamma Knife stereotactic radiosurgery. In fact, the fully automatic identification of target IURs is feasible because MET-PET datasets for stereotactic neuro-radiosurgery treatment planning include only the brain area avoiding the possible presence of false positives in other anatomic regions. Differently, in total body FDG-PET examinations user interaction to manually identify the target lesion is always needed (normal structures such as brain, heart, bladder, kidneys, and ureters normally have high FDG uptake) [586]. Accordingly, these co-segmentation software implementation platforms must provide manual seeding facilities in addition to automated multimodal segmentation algorithm. Although a one-to-one relationship between two different structural images of the same abnormal region could be reasonable, the assumption of identical lesion contours in both functional and morphologic images is often infeasible in many anatomic regions [32]. For instance, depending on the metabolic characteristics of the cells within a lung tumor [35], it may not take up a radiotracer in all its volume, or uptake region may be larger than the anatomic boundary of the tumor due to cellular activation in nearby tissues. Further issues are introduced by combining more than two structural modalities, i.e., MRI and CT, with PET metabolic imaging. Briefly, different imaging modalities (PET/MRI vs PET/CT), anatomic regions, and pathologies require a customization/adaptation of the method according to the specific clinical scenario. In our study, PET and MRI datasets regarding the same subject were acquired at two different times (MRI is scanned a few days after MET-PET) by two different dedicated scanners. This represents a non-trivial problem in co-segmentation domain, because an efficient and accurate multimodal co-registration is required. On the contrary, state-of-the-art about co-segmentation methods process mostly simultaneous PET/CT or PET/MRI acquisitions. Of course, the potential of hybrid imaging (PET/CT and PET/MRI systems) overcomes image registration issues by giving multimodal images contextually [468, 655], especially in brain imaging [263].



The achieved experimental results showed that GTV and BTV segmentations are statistically correlated (Spearman's rank correlation coefficient: 0.898) but have not higher degree of similarity (average  $DSC$ :  $61.87 \pm 14.64$ ). In fact, GTV and BTV measurements as well as evaluation metrics values (volume-based, overlap-based, and spatial distance-based metrics) corroborated that MRI and MET-PET convey different but complementary imaging information. Furthermore, when lesion sizes are greater than 2-3 times the Full Width at Half Maximum (FWHM) of the point spread function of the PET image resolution reconstructed by the PET imaging system, the underestimation of metabolic activity due to PVE can be assumed to be negligible [578]. In other cases, a recovery coefficient method could be included in our algorithm, such as the approach described in [198]. Nevertheless, the output of the proposed multimodal PET/MRI segmentation method, named  $BTV_{MRI}$ , accurately combines PET and MR imaging.  $BTV_{MRI}$  segmentation could reduce radioactivity spill-in and spill-out effects between tumor and surrounding tissues, taking advantage of the higher spatial resolution of MRI.

In addition, since it is not possible to define a gold standard CTV according to both MRI and PET images without treatment response assessment, the feasibility and the clinical value of BTV integration in Gamma Knife treatment planning were considered. Therefore, a visual and qualitative evaluation was performed by experienced physicians to assess the clinical value of BTV integration in CTV delineation using a five-point Likert scale. In most cases, CTV delineation was strongly or moderately influenced by PET imaging: GTV does not match with the functional area of the tumor, and metabolic imaging must be considered to assist the radiation oncologist in treatment planning. These results agree with [230], where in 74% the expansion of the MET-PET volume was larger than the Gadolinium enhancement area. Table 4 shows that  $BTV_{MRI}$  influenced the CTV more accurately than BTV (in 5/19 of cases): this finding was attributed to the higher spatial resolution of MRI, which shrank small false positive areas in PET target segmentation. Finally, PET imaging should be usually considered by clinicians in order to determine a CTV that takes into account the "active" part of the cancer in addition to anatomic tumor boundaries. According to these findings, it is appropriate to include PET images in stereotactic neuro-radiosurgery treatment planning.

In FDG-PET studies, target lesions must always be specified by an operator, because healthy structures (such as brain, heart, bladder, and kidneys) normally have high radiotracer uptake [586]. We do not claim that our approach is able to segment other types of imaging data. However, there are no general approaches that can work on all acquisition modalities and in all the clinical contexts without any limitations. As a matter of fact, multimodal PET/CT and PET/MRI segmentation techniques in the literature are quite different and have to be designed *ad hoc*. These methods have to be adapted considering the different imaging

acquisition modalities in certain clinical scenarios, as well as the involved body districts and pathologies. As concerning state-of-the-art methods on PET/MRI segmentation, we would like to underline that methods based on unifying graphs, which assume no significant anatomic and functional changes between the images acquired using different modalities to directly construct the hyper-graph [248, 32] or the subgraphs with inter-subgraph arcs [576], generally yield a single target volume on the fused multimodal images. Considering a decision-level fusion (i.e., taking into account the physicians' final decision for CTV definition), rather than a feature-level fusion (i.e., processing and fusing the input imaging data into a unifying data-structure) [124, 556], keeps all the initial information even when PET and MR images are considerably different.

This study showed that the BTV should be used to modify the GTV, by properly considering both metabolic and morphologic information. By using our multimodal methodology, clinicians could be assisted in CTV delineation during stereotactic radiosurgery treatment planning. Therefore, the proposed multimodal segmentation approach can be considered clinically feasible, since it can be integrated in the current clinical practice. After the positive opinion provided by clinicians, we have the rationale to begin clinical trials to evaluate tumor response after Leksell Gamma Knife treatment execution, even when BTV is included in CTV delineation during the planning phase.

# Chapter 5

## Computational Intelligence

### 5.1 Genetic Algorithms

Genetic Algorithms (GAs) represent an Evolutionary Computation technique for global optimization tasks [166]. Taking inspiration from Darwin's theory of biological evolution, GAs search optimal solutions to complex problems by evolving a population  $P$  of randomly created candidate solutions [270]. In the most general formulation, each solution represents an "individual" encoded as a fixed length string of characters taken from a finite alphabet (i.e., the genes); in alternative versions, the individuals codify real-values genes [14]. It is also possible to leverage GAs in other contexts by using more aggregate objects for the individuals, such as image histogram bins (see for instance [252]).

In GAs, the population of candidate solutions evolves in competition under controlled variation. Relying on a quality measure (i.e., the fitness value), each individual competes against the other individuals according to a *selection* process. The individuals characterized by a better fitness value will be selected with higher probability, using one of the available selection strategies, e.g., *roulette wheel selection* [216], *ranking selection* [33], and *tournament selection* [417]. Starting from these individuals, new ones are generated applying both *crossover* and *mutation* genetic operations. These operators allow us to mix the parent characteristics and introduce new genetic material in the population, respectively. Taking advantage of the information accumulation about an unknown search space, GAs are able to converge to optimal solutions, adapting the individuals during the iterations through exploration and exploitation of promising subspaces. Assuming a sufficiently large population, the *schema theorem* proves that GAs asymptotically converge to the optimal solution of any problem under investigation [270].

Given an optimization problem, a GA is initialized with a randomly generated population of individuals corresponding to potential solutions. During each *generation* (i.e., iteration)

the quality of each individual is evaluated using the fitness function. Then, individuals undergo selection, crossover, and mutation to create the next population (see pseudo-code in Algorithm 6).

Different termination criteria can be used: for instance, the GA could stop when a given number of generations  $T_{\max}$  is reached, or when the fitness value of an individual is lower than a fixed threshold value (assuming a minimization problem). In our approach, the algorithm stops when the number of generations  $T_{\max}$  is reached.

As mentioned above, several *selection* mechanisms exist, such as *roulette wheel selection* [216], *ranking selection* [33], or *tournament selection* [417]. Given a population  $P$ , consisting in  $|P|$  individuals, all selection techniques have the same goal, that is, obtaining an intermediate population  $P'$  composed of copies of individuals from the previous population  $P$ . Each individual can be added more than one time at  $P'$ , depending on its fitness value (the better the fitness, the higher the chance to be selected).

Once  $P'$  has been created by means of a selection procedure, the variation operators (i.e., crossover and mutation) are applied to obtain new individuals, named offspring, for the next generation. The *crossover* operator is used to recombine the genetic information from two parent individuals (Parent<sub>1</sub> and Parent<sub>2</sub>). Such operator, which plays a key role in GAs, allows for obtaining offspring solutions with better characteristics with respect to the parents. Crossover is applied to several selected pairs of individuals with a probability value  $p_c$ , named crossover rate. There are different versions of crossover, such as *single-point crossover*, *two-point crossover*, and *uniform crossover*.

In order to arbitrarily alter one or more genes of a selected offspring, we can apply the *mutation* operator, so increasing the variability inside  $P'$ . This simple operator, introducing new genetic material into the population, is able to prevent the convergence of the individuals to local optimal solutions, increasing the probability to reach every point in the search space. The mutation operator is applied with a probability  $p_m$  (i.e., mutation rate), such that some randomly selected genes of each individual are selected and their values are perturbed. Generally, the crossover is executed with high probability values (e.g.,  $p_c = 0.95$ ), conversely the mutation is applied with very low rates (e.g.,  $p_m = 0.05$ ). These parameter settings aim at reflecting the natural evolution process in which the characteristics (genes) of two parents are deeply blended, while mutations has a very low probability to occur.

Finally, the *elitism* strategy could also be exploited, by copying the best individual (or a subset of the best individuals) directly into the next generation without modifying it (them). This mechanism prevents the quality of the best solution from decreasing during the iterations, since the best individual is copied into the next population without undergoing the genetic operators.

**Algorithm 6** General pseudo-code for GAs**Input:** the number of individuals  $|P|$ **Output:** the individual  $C_{\text{best}}$  characterized by the best fitness value

---

```

1:  $t \leftarrow 0$ 
2:  $P^{(0)} \sim \text{RANDOMINITIALIZATION}(|P|)$ 
3: while  $t < T_{\text{max}}$  do
4:    $\text{EVALUATEFITNESS}(P^{(t)})$ 
5:    $P' \leftarrow \{\}$ 
6:   for  $i$  in  $1, 2, \dots, \lfloor \frac{|P|}{2} \rfloor$  do
7:      $\{\text{Parent}_1, \text{Parent}_2\} \leftarrow \text{SELECTIONOPERATOR}(P^{(t)})$ 
8:      $\eta \sim \text{Uniform}(0,1)$ 
9:     if  $\eta < p_c$  then
10:       $\{\text{offspring}_1, \text{offspring}_2\} \leftarrow \text{CROSSOVEROPERATOR}(\text{Parent}_1, \text{Parent}_2)$ 
11:    else
12:       $\{\text{offspring}_1, \text{offspring}_2\} \leftarrow \{\text{Parent}_1, \text{Parent}_2\}$ 
13:    end if
14:    for each  $\text{gene} \in \text{offspring}_1$  do
15:       $\eta \sim \text{Uniform}(0,1)$ 
16:      if  $\eta < p_m$  then
17:         $\text{MUTATIONOPERATOR}(\text{gene})$ 
18:      end if
19:    end for
20:    for each  $\text{gene} \in \text{offspring}_2$  do
21:       $\eta \sim \text{Uniform}(0,1)$ 
22:      if  $\eta < p_m$  then
23:         $\text{MUTATIONOPERATOR}(\text{gene})$ 
24:      end if
25:    end for
26:     $P' \leftarrow P' \cup \{\text{offspring}_1, \text{offspring}_2\}$ 
27:  end for
28:   $P^{(t)} \leftarrow P'$ 
29:   $t \leftarrow t + 1$ 
30: end while
31:  $C_{\text{best}} \leftarrow \text{SELECTBESTSOLUTION}(P^{(t)})$ 

```

---

### 5.1.1 Medical image enhancement

Medical imaging plays a key role in the clinical workflow, thanks to its capability of representing anatomical and physiological features [345, 513], by means of several different principles to measure spatial distributions of physical attributes of the human body, allowing for a better understanding of complex or rare diseases [612]. The effectiveness of such techniques can be reduced by a plethora of phenomena, such as noise and PVE [612], which might affect the measurement processes involved in imaging and data acquisition devices. In addition, computer-aided medical image acquisition procedures generally include reconstruction methods (producing two, three or even four dimensional imaging data), which could cause the appearance of artifacts. Image contrast and details might also be impaired by the procedures used in medical imaging, as well as by the physiological nature of the body part under investigation.

Medical images actually convey an amount of information—mainly related to high image resolution and high pixel depth—that could overwhelm the human vision capabilities in distinguishing among dozens of gray levels [445]. Improvements in the appearance and visual quality of medical images are therefore essential to allow physicians attaining valuable information that would not be immediately observable in the original image, thus assisting them in anomaly detection, diagnosis, and treatment. This kind of diagnosis includes two basic processes: image observation (visual perception), and diagnostic interpretation (cognition) [338]. Errors occurring in these diagnostic and therapeutic decision-making processes may have a significant impact on patient care, by causing possible misdiagnoses. In this context, image enhancement techniques aim at realizing a specific improvement in the quality of a given medical image. The enhanced image is expected to better reveal certain features, compared to their original appearance [143]. In particular, these methods could have a significant clinical impact in the case of not commensurable dynamic range of the actual pictorial data with respect to the displayed ones (i.e., monitor luminance response), as well as when the input image is characterized either by a high level of noise or by low contrast [218, 454]. This also applies to specialized computer screens for diagnostic reporting.

Although the majority of the enhancement techniques are typically applied to generate improved images for a human observer, others are exploited as a pre-processing step to provide enhanced images to further algorithms for computer-assisted analyses [454]. The first category includes techniques devoted to remove noise, enhance contrast, and sharpen the details. The second category, partially overlapped with the former one, includes additional techniques such as edge detection and object segmentation for automated processing [494]. It was shown that a high-contrast medical image could lead to a better interpretation of the different adjacent tissues in the imaged body part [99]. Accordingly, the resulting enhanced

image—in terms of signal intensities of different tissues—can facilitate the automated segmentation, feature extraction, and classification of these tissues.

In the clinical routine, CE-MRI is a diagnostic technique that enables a more precise assessment of the imaged tissues after the administration of a Gadolinium-based contrast medium in patients [579]. MRI is currently the most prominent modality to obtain soft-tissue imaging [70], especially in oncology, since it provides significant improvements—in terms of image contrast and resolution—between lesion and healthy tissue [407]. However, MRI data are affected by acquisition noise [590] and are also prone to imaging artifacts, related to magnetic susceptibility and large intensity inhomogeneities of the static magnetic field (i.e., streaking or shadowing artifacts [46]), especially using high magnetic field strengths. These aspects make MR image enhancement a challenging task devoted to improve the outcome of automatic segmentation.

The existing image enhancement approaches generally attempt to improve the contrast level of the whole image and do not address the issues related to overlapped gray level intensities; as a consequence, neither the region contour sharpness nor the image thresholding results can be improved. In the case of threshold-based image segmentation with two classes (i.e., foreground and background) [426], the input image is assumed to have a bimodal histogram [677]. Thus, an appropriate image enhancement method that yields medical images with a stronger bimodal distribution is required. However, determining the best pre-processing of an image—able to preserve the structural information of the image while enhancing the underlying bimodal distribution of the histogram bins—is a complex task on a multi-modal fitness landscape that mandates the use of global optimization approaches.

### **Related work**

Most of the existing enhancement techniques are empirical or heuristic methods—strongly related to particular types of images—which generally aim at improving the contrast level of images degraded during the acquisition process [100]. As a matter of fact, finding the best gray level mapping that adaptively enhances an input image for each different sample can be considered an optimization problem [155, 459]. Unfortunately, no unifying theory employing a standardized image quality measure is currently available to define a general criterion for image enhancement [429]. In addition, in the case of medical imaging, techniques tailored on specific tasks are necessary to achieve a significant enhancement and, in general, interactive procedures involving considerable human effort are needed to obtain satisfactory results.

In order to achieve objective and reproducible measurements conveying clinically useful information, operator-dependence should be minimized by means of automated methods. Point-wise operations in the spatial (pixel) domain, representing the simplest form of image

processing, are effective solutions since efficiency requirements have also to be met. In the case of image enhancement, they re-map each input gray level into a certain output gray level, according to a global transformation [218]. Thus, such kind of techniques treat images as a whole, without considering specific features of different regions or selectively distinguishing between a collection of contrast enhancement degrees or settings [429]. Histogram Equalization (HE) is the most common image enhancement technique [218], but it is not suitable for medical images due to the obtained over-brightness [202]. Indeed, this method uniformly spreads the input gray level values according to the cumulative density function of the image histogram. However, HE does not preserve the input mean brightness, possibly suffering from over-enhancement and giving rise to artifacts such as washed-out effect [101]. This global transformation applies contrast stretching just on gray levels with the highest frequencies, causing a significant contrast loss for gray levels characterized by lower occurrences in the input histogram [320]. In order to address the issues related to input mean brightness preservation, a modification of the standard HE technique, called Bi-Histogram Equalization (Bi-HE), was introduced [320]. Bi-HE attempts to improve the results achieved by HE, by first splitting the original histogram into two components according to the global mean of the original image, and then separately performing the HE method on the two sub-histograms.

Other traditional global gray level transformations generally used for contrast stretching are formalized as transformation functions of the form  $s = \mathcal{T}(r)$ , where  $\mathcal{T}(\cdot)$  maps an input intensity value  $r$  into an output intensity value  $s$  [218]. Power-law transformation—also called Gamma Transformation (GT)—is a non-linear operation of the form  $\mathcal{T}(r) = cr^\gamma$ , where typically  $c = 1$ . For instance, when the image is predominantly dark, an expansion of the intensity levels is desirable. In such a case, GT with  $\gamma < 1$  yields a brighter image by increasing the number of hyper-intense pixels. On the contrary, by using  $\gamma > 1$ , the GT converts the input gray-scale range into a darker one, by increasing the occurrences of darker pixels. Obviously, the value of  $\gamma$  strongly depends on the medical application. Accordingly, logarithmic and anti-logarithmic transformations make an image much brighter and darker, respectively. Unfortunately, for medical images characterized by low contrast and weak edges at adjacent tissue boundaries, GT may result in merely brighter or darker images, leading to difficulties in the visualization and interpretation of different tissues. Therefore, to adequately enhance contrast, the two different behaviors—corresponding to values  $\gamma > 1$  and  $\gamma < 1$ —should be combined for contextually decreasing the darker pixel gray values and increasing the brighter pixel gray values. This results in a significant improvement of the contrast, by enhancing the edges thanks to the increased gradient magnitude of the image [202]. This kind of contrast stretching can be achieved by using a Sigmoid intensity Transformation (ST), which darkens a wide range of hypo-intense gray levels and brightens a wide range



of hyper-intense gray levels [218]. Such an operation indirectly increases the difference between low and high intensity values, resulting in the overall contrast enhancement of the image [202].

In addition to HE, which automatically yields an image with a uniform histogram, it is possible also to explicitly specify the desired shape of the output histogram. This method, named Histogram Specification (HS), aims at matching the histogram of the gray level intensities of the input image against a desired histogram [218]. Unfortunately, such approaches cannot be applied in the case of image datasets with heterogeneous gray level distributions, since the histogram to be matched should be defined either *a priori* for all the images in the dataset or interactively for each processed image, by separating and shaping the two underlying sub-distributions [669]. It is therefore necessary to employ global search metaheuristics to automatically identify the best solution case-by-case.

The complexity of the enhancement criteria to be met (i.e., the effective contrast stretching combined with image detail preserving) leads to the application of global search metaheuristics that allow for coping with several constraints, which are not generally tractable by means of traditional exhaustive computational approaches [429, 445, 459]. Evolutionary methods have been widely adopted in the image enhancement domain to find the optimal enhancement kernel [429], sequence of filters [328], or input-output mapping transformation [81, 537]. Recently, Hashemi *et al.* [252] proposed a GA that efficiently encodes the histogram by means of the non-zero intensity levels, by employing genetic operators that directly process images to increase the visible details and contrast of low illumination regions, especially in the case of high dynamic ranges. The authors argued that this method yields “natural-looking” images, considering the visual appearance.

In this context, Genetic Programming (GP) [334] was shown to be a powerful framework to select and combine existing algorithms in the most suitable way. Differently to GAs, GP evolves a population of functions, or more generally, computer programs to solve a problem. The solutions in the computer program space can be represented as trees, lines of code, expressions in prefix or postfix notations as well as strings of variable length [85]. For instance, the authors of [57] tackled the video change detection problem (among the frames of video streams) by combining existing algorithms *via* different GP solutions exploiting several fusion schemes. The fitness function was composed of different performance measures regarding change detection evaluation. For what concerns the application of GP in image enhancement, Poli and Cagnoni [477] proposed an approach to yield optimally pseudo-colored images for visualization purposes, aiming at combining multiple gray-scale images (e.g., time-varying images, multi-modal medical images and multi-band satellite images) into a single pseudo-color image. This approach relies on user interactions to determine which

individual should be the winner in tournament selection, so it does not explicitly require a fitness function. As case studies, a pair of brain MRI sequences were fused as well as the motion of the heart on echocardiographic images was synthesized into a single pseudo-color image. GP-based semi-supervised learning was shown to be successful even in the case of noisy labeled data in classification tasks [566]. In the medical domain, GP with geometrics semantic operators [630] was employed to predict the relative position of a slice within a CT image stack when DICOM metadata are unreliable [84].

Other works exploited Swarm Intelligence techniques. The approach presented in [559], called Multi-Objective HE, uses Particle Swarm Optimization (PSO) [316] to enhance the contrast and preserve the brightness at the same time. Draa and Bouaziz [155] employed the same encoding of individuals and histogram mapping strategy described in [252], within an optimization strategy based on the Artificial Bee Colony (ABC) algorithm [310]. However, since ABC natively works in a continuous space, while a discrete representation is used for the solutions (i.e., gray-level mapping), a discretization step is mandatory in the correction operation during the search phase. An alternative approach using the ABC algorithm for image contrast enhancement has been proposed in [100], wherein the optimal values for the parameters of a parametrics image transformation, namely the Incomplete Beta Function, are estimated. Differently to the work described in [155], the optimization procedure is carried out in a continuous search space. Finally, multi objective Bat Optimization and a neuron-based model of Dynamic Stochastic Resonance were combined in [568] for the enhancement of brain MR images.

Image pre-processing methods often include operations at the lowest level of abstraction—wherein both input and output are gray-scale images—employing intensity transformations. These global transformations just deal with the gray levels, without considering the pixel position, and the corresponding relationship with its neighborhood, in the image. The main objective of pre-processing is to improve the raw imaging data, dealing with undesired artifacts or enhancing some image features important for further processing [577, 599].

Conventional pre-processing techniques generally expand a narrow input range of gray levels into a wider range of output levels, in order to improve the image detection performance. Pre-processing tackles the problems related to image degradation correction, whether no assumption about the nature of noise/artifacts can be made as well as when neither *a priori* knowledge about the acquisition device nor the objects of interest are available [577]. Thus, Soft Computing methods could be beneficially exploited—by considering, for instance, metaheuristics [596] or Fuzzy Logic [91]—to find the best solution in a high variety of scenarios. In addition, interactive procedures and trial-and-error parameter tuning are generally required to obtain satisfactory results [429]. GAs offer a very powerful framework

for image enhancement since they implement an unbiased global optimization method for huge search spaces [459]. Considering the Computational Intelligence metaheuristics, GAs are the most suitable framework because of the discrete encoding of the candidate solutions and the combinatorial structure of the search space [607]. More specifically, here we exploited GAs to efficiently encode the individuals by means of their corresponding image histograms.

Biomedical images usually require methods tailored on a specific task [558]. Well-designed preliminary image pre-processing techniques can provide benefits for a better ROI feature detection to be exploited in a downstream image processing pipeline. Especially, these pre-processed data can be fed as a suitable input for an image segmentation algorithm to improve its accuracy.

Naïve smoothing techniques, by using both linear (e.g., average or Gaussian filters) and non-linear (e.g., median filter) operators, may improve the shape of the histogram [504], also also by widening the separation between the two modes in bimodal gray-level histograms [536]. These approaches are generally exploited for noise reduction, even though sharp boundaries might be replaced with a fuzzy region of varying shades of gray. Thus, such a kind of low-pass filtering could affect the actual pictorial content, compromising the clinical effectiveness of the subsequent ROI detection and delineation phases.

In medical imaging, pre-processing steps are valuable for further computer-assisted image analysis, by making anatomical or functional structures more easily detectable. Therefore, pre-processing involving image enhancement operations may be needed to achieve more identifiable and sharpened boundaries for medical image segmentation [612]. A better distinction among adjacent tissues in medical images can be achieved, by means of an adequate contrast performing gray-scale transformations.

It is worth noting that all works above mentioned are focused on consumer electronics or medical applications, to obtain more “visually pleasant” images by mainly increasing the contrast of the whole image. As a matter of fact, the visual appearance is a fundamental feature in quality assessment and recognition tasks, such as in the case of food products [109, 464]. On the contrary, the main key novelty of MedGA consists in better revealing the two underlying sub-distributions occurring in an image sub-region characterized by a roughly bimodal histogram, overcoming the limitations of the state-of-the-art contrast enhancement methods, which could produce false edges and consequently over-segmentation when the input images are affected by noise, as in the case of MRI data [202]. There exist other algorithms, like HS, whose aim is similar to MedGA and consists in matching the histogram of the gray level intensities of the input MR image with a desired output histogram [218]. Unfortunately, such approaches cannot be applied in the case of image datasets with high variability in gray level distributions, since the histogram to be matched should be

defined either *a priori* for the whole dataset or interactively for each processed image, by strengthening and shaping the two underlying sub-distributions. It is therefore necessary to employ global search metaheuristics to automatically identify the best solution, as in the case of MedGA.

Finally, supervised—such as Artificial Neural Networks (ANNs) [561] and SVMs [702]—as well as unsupervised learning approaches—such as Self-Organizing Maps (SOMs) [445]—require a training phase and are generally exploited in image classification and recognition tasks.

### Clinical MRI datasets

MedGA was applied as a pre-processing step in two segmentation tasks concerning challenging clinical MRI dataset from patients affected by: (i) uterine fibroids; (ii) brain metastatic cancers. All the analyzed MRI data are encoded in the 16-bit DICOM format. The MRI acquisition characteristics are reported in Table 5.1.

Table 5.1 MRI acquisition parameters of uterine fibroids and brain metastatic tumor dataset.

Dataset	MRI sequence	TR [ms]	TE [ms]	Matrix size [pixels]	Slice spacing [mm]	Slice thickness [mm]	Pixel spacing [mm]
Uterine fibroids	T1w FSPGR+FS+C	150-260	1.392-1.544	512 × 512	5.0	6.0	0.6641-0.7031
Brain metastases	T1w FFE	25	1.808-3.688	256 × 256	1.5	1.5	1.0

**Uterine fibroids** Eighteen patients affected by symptomatic uterine fibroids who underwent MRgFUS therapy [501] were considered. The total number of the examined fibroids was 29, overall represented on 163 MR slices, since some patients presented a pathological scenario with multiple fibroids. The analyzed images were acquired using a Signa HDxt 1.5 T MRI scanner at two different institutions. These MRI series were acquired after the MRgFUS treatment, executed with the ExAblate 2100 (Insightec Ltd., Carmel, Israel) HIFU equipment. The considered MR slices were scanned using the T1w “Fast Spoiled Gradient Echo + Fat Suppression + Contrast mean” (FSPGR+FS+C) sequence. This MRI protocol is usually employed for NPV assessment, since ablated fibroids appear as hypo-intense areas due to low perfusion of the contrast mean [412]. Sagittal MRI sections were processed, in compliance with the current clinical routine for therapy response assessment [412]. In current clinical practice, the NPV evaluation procedure is fully manual [524]. Two uterine fibroid MR slices are depicted in Figs. 5.1a and 5.1b.

**Brain metastatic tumors** Twenty-seven brain metastases treated using a Leksell Gamma Knife stereotactic neuro-radiosurgical device [354] were processed, for a total of 248 MR

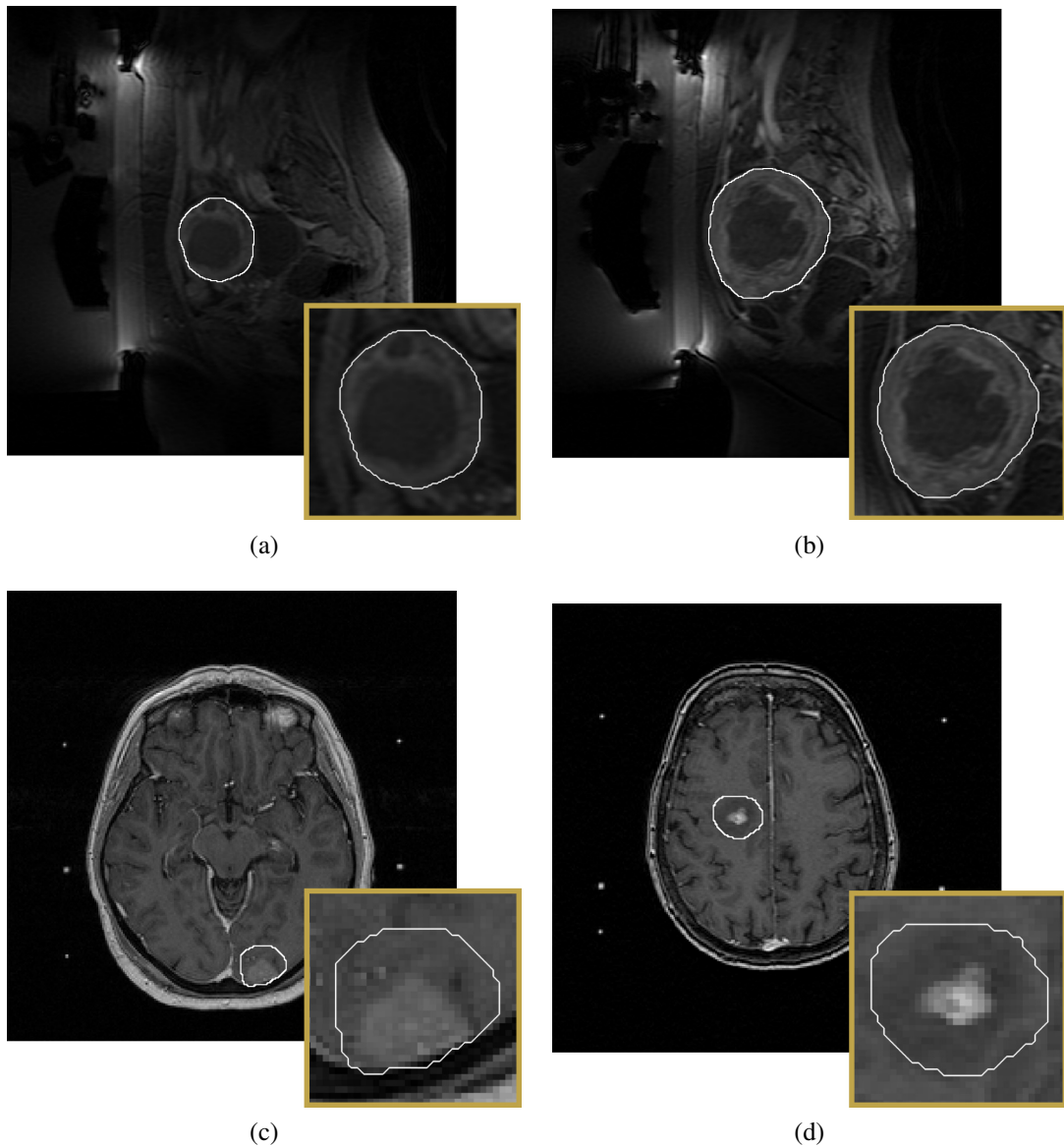


Fig. 5.1 Examples of input MR images: (a, b) uterine fibroid inside the uterus region; (c, d) brain tumor inside a ROI bounding region selected by the healthcare operator. The image regions including the ROIs, defined by the white contour and zoomed at the bottom right of each sub-figure, are characterized by nearly bimodal histograms.

slices. All the available MRI datasets were acquired on a Philips Gyroscan Intera 1.5 T MRI scanner, before treatment, for the planning phase. In current radiation therapy practice, Gamma Knife treatments are planned manually by a neurosurgeon on MRI alone, by typically using T1w FFE CE-MRI sequences [521, 526]. Thanks to the Gadolinium-based contrast agent, brain lesions appear as enhanced hyper-intense zones. Sometimes a dark area might

be present due to either edema or necrotic tissues [521, 523]. Two representative instances of brain tumors are shown in Figs. 5.1c and 5.1d.

### State-of-the-art image enhancement methods

To achieve a fair comparison, we considered the most common computational techniques for image enhancement and pre-processing, namely:

- HE [218], which adjusts pixel intensities for contrast enhancement according to the normalized histogram of the original image  $\mathcal{I}_{\text{orig}}$ . With HE, gray levels are more uniformly distributed on the histogram, by spreading the most frequent intensity values;
- Bi-HE [320]—a modification of the traditional HE—which addresses issues concerning mean brightness preservation;
- GT, which is a non-linear operation using the power-law relationship  $s(r) = cr^\gamma$ , where  $r$  and  $s$  are the input and the output gray-scale values, respectively, and  $c$  is a multiplication constant ( $c = 1$  in the following tests). The parameter  $\gamma$  is set to values greater than 1 (i.e., decoding gamma) to obtain a gamma expansion, or to values smaller than 1 (i.e., encoding gamma) to realize a gamma compression (see Fig. 5.2a). In our tests we considered the values  $\gamma = 0.4$  and  $\gamma = 2.5$ ; higher (lower) values of  $\gamma = 2.5$  ( $\gamma = 0.4$ ) tend to logarithmic (anti-logarithmic) functions, resulting in an excessively bright (dark) output image, unsuitable for medical applications [202];
- ST function (Fig. 5.2b), also called S-shaped curve, which is a global non-linear mapping defined as follows:

$$s(r) = \frac{l_{in}^{(\max)}}{1 + \exp(-\lambda(r - \alpha))}, \quad (5.1)$$

where  $l_{in}^{(\max)} = \max\{\mathcal{L}_{in}\} = \max\{\mathcal{L}'_{in}\}$  is the asymptotic maximum value of the function,  $\alpha = \frac{1}{2} \left( l_{in}^{(\max)} - l_{in}^{(\min)} \right)$  is the midpoint value, and  $\lambda$  defines the function steepness. This transformation stretches the intensity around the level  $\alpha$ , by making the hypo-intense histogram part darker and the hyper-intense histogram part brighter. Thus, the difference between the minimum and maximum gray values as well as the gradient magnitude of the image are increased, obtaining strong edges [202]. In our validation, we used sigmoid functions that allow to consider the entire input dynamic range, by varying the curve slope with the values  $\lambda \in \left\{ \frac{4}{\alpha}, \frac{6}{\alpha}, \frac{8}{\alpha} \right\}$ .

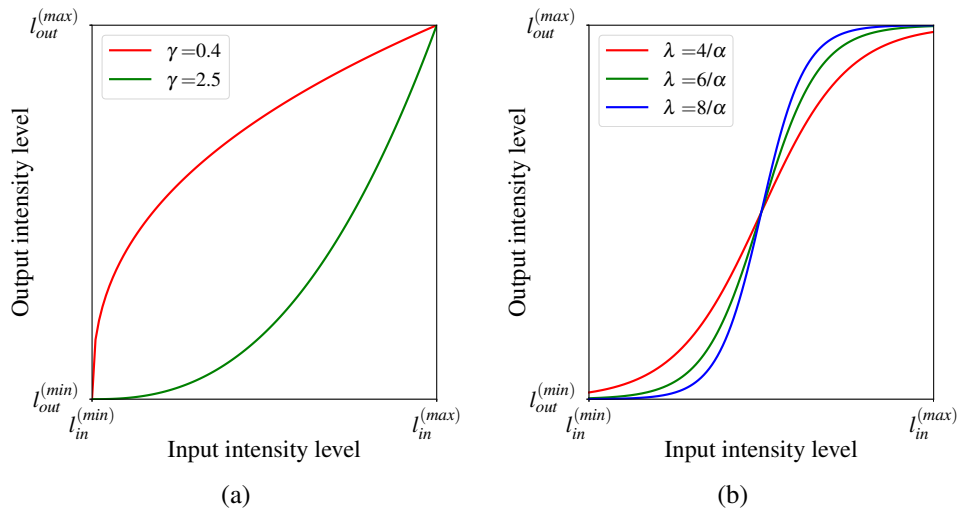


Fig. 5.2 Plots of the implemented global non-linear intensity transformations for image enhancement: (a) Gamma Transformation; (b) Sigmoid intensity Transformation. We report on the  $x$ -axis the input intensity range  $[I_{in}^{(min)}, \dots, I_{in}^{(max)}]$ , and on the  $y$ -axis the output intensity range  $[I_{out}^{(min)}, \dots, I_{out}^{(max)}]$ .

### MR image enhancement

In order to address these issues, we proposed a novel image enhancement technique based on GAs [270], called MedGA [529], specifically aimed at strengthening the sub-distributions in medical images with an underlying bimodal histogram of the gray level intensities. Among the existing soft computing methods for global optimization, GAs represent the most suitable technique because of the discrete structure of the candidate solutions and the intrinsic combinatorial structure of the problem under investigation.

Considering the possible clinical applications, MedGA improves the visual perception of a ROI in MRI data with an underlying bimodal intensity distribution [529]. In addition, MedGA can be used as a pre-processing step, in a framework defined to realize an efficient threshold-based image segmentation with two classes (i.e., binarization), applied to MRI data [527]. Image thresholding approaches performed on CE-MR image regions could considerably benefit from input data pre-processed by MedGA.

The main contributions of MedGA in the context of intelligent clinical systems can be briefly outlined as follows. MedGA acts as an expert system by playing a two-fold role: (i) image enhancement to visually assist physicians during their interactive decision-making tasks, and (ii) improvement of the results in downstream automated processing pipelines for clinically useful measurements. The motivations underlying MedGA rely on the design of an image enhancement technique for medical images with roughly bimodal histograms. Indeed,

our objective is to propose an intelligent model that is well-suited for effectively enhancing such a particular type of medical images. To the best of our knowledge, MedGA is the first work that explicitly deals with the improvement of thresholding-based segmentation results [677]. Therefore, our computational framework can be employed as an intelligent solution in CDSSs. As a matter of fact, MedGA represents an interpretable computational model [86] that allows for the understandability of the results (i.e., the gray level histogram is readable by the user). The compelling issues related to the interpretability of Machine Learning and Computational Intelligence methods in medicine [133] are fundamental for the adoption and the clinical feasibility of a novel CDSS [74]. In addition, thanks to the reliable calibration step of MedGA, no user parameter setting is required. We aim at showing how evolutionary methods can boost the state-of-the-art performance in medical image enhancement, thus fostering GAs as a new concrete support tool for the clinical practice. Such an expert system may have a significant impact in real healthcare environments.

Even though MedGA exploits the same encoding of individuals defined in [252] and [155], its purpose is very different as it is designed to explicitly strengthen the two sub-distributions of medical images characterized by an underlying bimodal histogram. To this aim, we defined a specific fitness function that emphasizes the two Gaussian distributions composing a bimodal histogram. This achievement plays a fundamental role for threshold-based segmentation approaches. As a matter of fact, these approaches strongly rely on the assumption that the bimodal histogram under investigation is composed of two nearly Gaussian distributions with almost equal size and variance [677].

Finally, MedGA differs from the GP-based approaches, in which the final generated solution might have large size [85], even when the GP model is implemented efficiently, representing a limitation that could significantly impair the readability and interpretability of the solutions. Moreover, MedGA does not require any user interaction step, differently to [477] where the user, being directly involved in the tournament selection, controls the evolution of simple programs that enhance and integrate multiple gray-scale images into a single pseudo-color image. Therefore, the selection is based on the output image (i.e., phenotype) rather than the structure and size of the program (i.e., genotype) [477].

**The proposed GA-based image enhancement approach** MedGA requires an input image  $\mathcal{I}_{\text{orig}}$ , with  $M$  rows and  $N$  columns, depicting a ROI included in a bounding region whose pixel values are approximately characterized by an underlying bimodal histogram. Therefore, as a first step, either a computational method or a user must detect a region including the ROI. Then, the input MR image is masked with this bounding region and the whole image is cropped according to the smallest rectangle enclosing the bounding region. Fig. 5.3 outlines



the overall procedure of MedGA, by presenting the initialization phase as well as the flow diagram of the proposed GA for image enhancement. A linear normalization is applied on the initial full range of the masked MR image to balance out the pixel distribution for the following bin rearrangement by means of MedGA. It is worth noting that no additional pre-processing operation (e.g., low-pass or high-pass filtering) is needed by MedGA. The final best solution found by MedGA will emphasize the two underlying bimodal Gaussian distributions occurring in the bimodal MR image, for the subsequent image thresholding phase according to the optimal adaptive threshold  $\theta_{opt}$  computed using the efficient IOTS algorithm [500, 619].

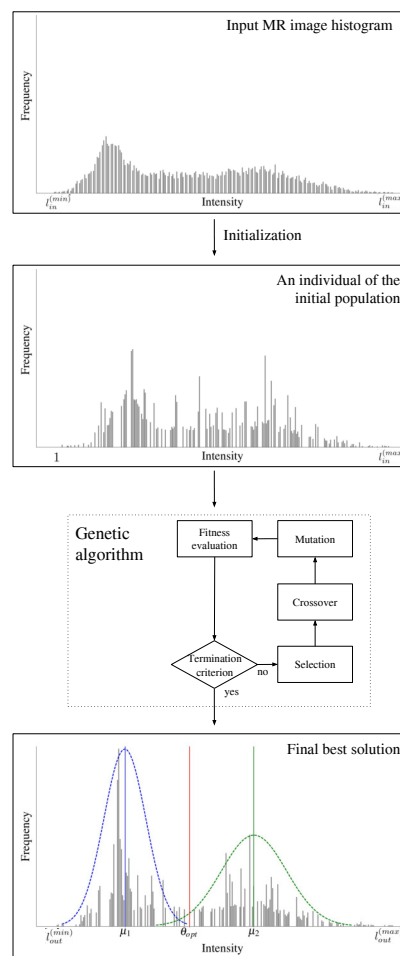


Fig. 5.3 Workflow of MedGA: the individuals are initialized according to the characteristics of the input MR image, and processed by the GA. The final best solution of MedGA strengthens the two underlying distributions in the gray levels intensity characterized by mean values  $\mu_1$  and  $\mu_2$ , and standard deviations  $\sigma_1$  and  $\sigma_2$ , respectively. The two distributions are highlighted in the plot with blue and green dashed lines.

Hereafter, we present MedGA in detail, showing the structure of individuals and genes and explaining the used selection, crossover and mutation operators. Finally, we explain the fitness function that allows for obtaining two well-separated normal distributions starting from the input histogram computed on the original image, while preserving the input mean brightness.

**Individual's structure and gene encoding** MedGA exploits a very efficient and comprehensive individual's structure. The size of each individual is equal to  $n$ , where  $n$  is the number of gray levels (different from 0) of the input MR image.

Let  $l_{in}^{(\min)}$ ,  $l_{in}^{(\max)}$ ,  $l_{out}^{(\min)}$ , and  $l_{out}^{(\max)}$  be the minimum non-zero and maximum gray levels of input and output images, respectively. Assuming that  $l_{in}^{(\min)} \leq l_{out}^{(\min)}$  and  $l_{in}^{(\max)} \geq l_{out}^{(\max)}$ , the linear normalization on  $\mathcal{L}_{in}$  gives rise to the extended range of the non-zero gray levels, that is, the ordered set  $\mathcal{L}'_{in} = [1, \dots, l_{in}^{(\max)}] \subset \mathbb{N}$  (typically,  $l_{in}^{(\min)} > 1$ ). This normalization operation, which employs only values of gray levels already representable in the initial dynamic range, does not alter the image content and allows MedGA to process additional intensity levels with respect to the initial full range  $\mathcal{L}_{in}$ , by considering the variability within the analyzed MRI datasets. During each iteration  $t = 1, 2, \dots, T$ , each individual  $C_i^t = [C_i^t(1), C_i^t(2), \dots, C_i^t(n)]$  (with  $i = 1, \dots, |P|$ ) is defined as a circular array of integer numbers of size  $n$ , where  $n = |\mathcal{L}'_{in}|$  corresponds to the number of different gray levels belonging to  $\mathcal{L}'_{in}$  identified in the input MR image (i.e., the gray levels whose frequency is greater than zero in the input MR image). The  $n$  values are then sorted in ascending order so that the intensity levels  $C_i^t(j)$  (with  $j = 1, \dots, n$ ) codified by the individual can be mapped onto the intensity levels of the input MR image (i.e., the gray level frequencies of the input MR image are assigned to the corresponding intensity levels of the individual).

In order to evaluate the fitness value of the individuals, we apply the following transformation  $\mathcal{T}$  that re-maps each input gray level  $r$  into  $s$ :

$$s = \mathcal{T}(C_i^t(j)) = \mathcal{T}(r), \quad (5.2)$$

where  $r \in \mathcal{L}_{in} = [l_{in}^{(\min)}, \dots, l_{in}^{(\max)}] \subset \mathbb{N}$  and  $s \in \mathcal{L}_{out} = [l_{out}^{(\min)}, \dots, l_{out}^{(\max)}] \subset \mathbb{N}$  are intensity values in the input and output gray-scale ranges, respectively. By so doing, a direct mapping between the gray levels of the original image  $\mathcal{I}_{orig}$  and the output  $\mathcal{I}_{enh}$  is defined: each gray level in the original histogram is replaced with the gray level value contained by the same position in the final best solution  $C_{best} \in P$ .

Since GAs are unbiased global optimization methods [459], as population initialization method we used a simple uniform distribution to create the initial histogram bins. Generally, computational approaches requiring initialization settings should be initialized randomly in

order to increase the convergence speed and the probability to avoid local optima determined by fixed settings. Thus, the following steps are used to initialize the population  $P$ :

1. each individual  $C_i^t \in P$  is randomly initialized by sampling  $n$  integer values from the discrete uniform distribution in  $\mathcal{L}'_{in}$ ;
2. these  $n$  values are then sorted in ascending order allowing for a mapping with the intensity levels of the input MR image.

During the initialization, if an integer value is sampled more than once then the frequency values of the input MR image, corresponding to these gray level intensities, are summed-up and assigned to the same gray level of the individual. After constructing the initial population, the fitness values are calculated for all individuals. In every generation, the size of the population  $P$  is kept constant to  $|P|$ .

**Selection operator** For the selection of individuals, MedGA exploits a *tournament* strategy for three main reasons: (i) the selection pressure can be controlled by setting the tournament size  $k$  (with  $k \ll |P|$ ); (ii) the fitness evaluations are performed only on the  $k$  individuals selected for the tournaments, and not on the whole population; (iii) this technique could be easily implemented on parallel architectures [417].

When a tournament is executed,  $k \ll |P|$  individuals are randomly selected from the population  $P$  and the winner, that is, the individual with the best fitness, is deterministically selected and added to  $P'$ . This procedure is usually performed with reinsertion, that is, the individual can compete to multiple tournaments during a generation. The size of the tournament  $k$  balances the selection pressure: if  $k$  is large, then the individuals characterized by worse fitness have a reduced probability to be selected for the intermediate population  $P'$ . MedGA exploits this selection mechanism since it has three advantages: (i) the fitness evaluations are required only for the individuals selected for the tournaments; (ii) the selection pressure can be increased or decreased simply varying the value of  $k$ ; (iii) this technique is suitable for parallel implementations.

**Crossover and mutation operators** At each iteration of MedGA, a number of individuals properly selected from the current population are inserted into the intermediate population  $P'$ , and modified by means of crossover and mutation operators. A *single-point crossover* operator is applied with a given probability  $p_c$  to the individuals selected by the tournament strategy and belonging to the first intermediate population  $P'$ . MedGA employs the following *single-point crossover* to blend the genetic information of two parents, resulting in a mixing ratio fixed to 0.5 (see scheme in Fig. 5.4):

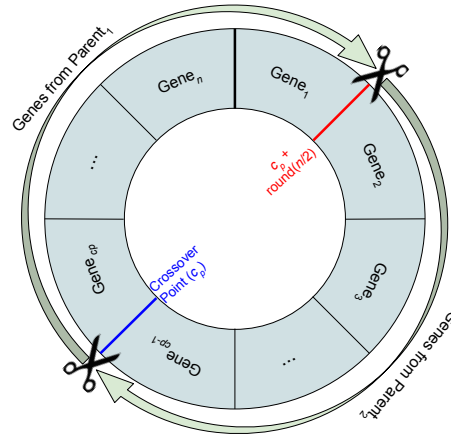


Fig. 5.4 Graphical representation of the single-point crossover strategy used for generating the first offspring from two parents (Parent<sub>1</sub> and Parent<sub>2</sub>). The second offspring is generated changing the Parent<sub>1</sub>'s genes with those of Parent<sub>2</sub> and vice-versa.

1. one crossover point  $c_p$  is randomly selected from the ordered set  $[1, 2, \dots, n]$ . Let  $h_p$  be equal to  $\text{round}(\frac{n}{2})$ ;
2. if  $c_p > \text{round}(\frac{n}{2})$ , the first offspring is generated using the genes  $1, \dots, c_p - h_p - 1$  and  $c_p, \dots, n$  from Parent<sub>1</sub>, and the genes  $c_p - h_p, \dots, c_p - 1$  from Parent<sub>2</sub>. Otherwise, it is generated using the genes  $c_p, \dots, c_p + h_p - 1$  from Parent<sub>1</sub>, and the genes  $1, \dots, c_p - 1$  and  $c_p + h_p, \dots, n$  from Parent<sub>2</sub>;
3. the second offspring is generated by replacing the Parent<sub>1</sub>'s genes with those of Parent<sub>2</sub> and vice-versa.

The *mutation* operator is applied with probability  $p_m$  to each element  $C_i^t(j) \in C_i^t = [l_{out,i}^{(\min)}, \dots, l_{out,i}^{(\max)}]$  of each offspring from the population  $P'$ , where  $l_{out,i}^{(\min)}$  and  $l_{out,i}^{(\max)}$  are the minimum and maximum non-zero gray levels encoded by  $C_i^t$  during the current iteration, respectively. In particular, if the gray level intensity encoded in  $C_i^t(j)$  is smaller than the optimal threshold  $\theta_{opt,i}^t$  evaluated by IOTS at that iteration for the individual  $C_i^t$ , then an integer is randomly sampled from the uniform distribution in  $[l_{out,i}^{(\min)}, \dots, \theta_{opt,i}^t - 1] \subset \mathbb{N}$  to update the value  $C_i^t(j)$ ; otherwise, an integer is randomly sampled from the uniform distribution in  $[\theta_{opt,i}^t, \dots, l_{out,i}^{(\max)}] \subset \mathbb{N}$  to update the value  $C_i^t(j)$ .

Finally, to prevent the quality of the best solution from decreasing during the optimization, MedGA also exploits an *elitism* strategy, so that the best individual from the current population is copied into the next population without undergoing the genetic operators.

**Fitness function** In order to obtain a bimodal histogram separation that yields better results for further automated image processing phases, we propose a fitness function that aims at realizing two well-separated normal distributions with equal distance from the optimal threshold  $\theta_{\text{opt}}$ . For each iteration  $t$  (with  $t = 1, 2, \dots, T$ ) and for each individual  $C_i^t$  (with  $i = 1, 2, \dots, |P|$ ), first the mean values  $\mu_{1,i}^t$  and  $\mu_{2,i}^t$ , concerning the two sub-distributions in the histogram  $\mathcal{H}_{1,i}^t$  and  $\mathcal{H}_{2,i}^t$ , respectively, and the corresponding optimal threshold  $\theta_{\text{opt},i}^t$ , are efficiently computed using the IOTS algorithm [500, 619]. Then, for each individual  $C_i^t$  the following fitness function is calculated:

$$\begin{aligned} \mathcal{F}(C_i^t) &= \tau_1 + \tau_2 + \tau_3, \quad \text{where:} \\ \tau_1 &= |2 \cdot \theta_{\text{opt},i}^t - \mu_{1,i}^t - \mu_{2,i}^t|, \\ \tau_2 &= |\omega_{1,i}^t - 3\sigma_{1,i}^t|, \\ \tau_3 &= |\omega_{2,i}^t - 3\sigma_{2,i}^t|. \end{aligned} \tag{5.3}$$

The terms  $\omega_{1,i}^t = \frac{1}{2}(\theta_{\text{opt},i}^t - \min_{j \in \{1, \dots, n\}} \{C_i^t(j)\})$  and  $\omega_{2,i}^t = \frac{1}{2}(\max_{j \in \{1, \dots, n\}} \{C_i^t(j)\} - \theta_{\text{opt},i}^t)$  correspond to the half width of  $\mathcal{H}_{1,i}^t$  and  $\mathcal{H}_{2,i}^t$ , respectively, while  $\sigma_{1,i}^t$  and  $\sigma_{2,i}^t$  are the standard deviations of  $\mathcal{H}_{1,i}^t$  and  $\mathcal{H}_{2,i}^t$ , respectively. The three terms of the fitness function  $\mathcal{F}(\cdot)$  cooperate together to achieve the desired image enhancement:  $\tau_1$  aims at maintaining the mean values  $\mu_{1,i}^t$  and  $\mu_{2,i}^t$  equidistant from the yielded optimal threshold  $\theta_{\text{opt},i}^t$ , while  $\tau_2$  and  $\tau_3$  are designed to force the sub-histograms  $\mathcal{H}_{1,i}^t$  and  $\mathcal{H}_{2,i}^t$ , respectively, to approximate normal distributions. Especially, we exploited the empirical 3- $\sigma$  rule, which states that approximately 99.73% of the values lie within  $3\sigma$  according to  $\Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.9973$ , where  $\mu$ ,  $\sigma$ , and  $X$  represent the mean, the standard deviation and an observation from a normally distributed random variable, respectively.

Two examples of image enhancement results, achieved by MedGA on a uterine fibroid and on a brain tumor, are shown in Figs. 5.5 and 5.6, respectively. In the case of uterine fibroids, the proposed method enhances the input MR image by making fibroid regions more uniform and with sharper edges in terms of both visual human perception and automated image segmentation. The histogram in Fig. 5.5d highlights that the output image is characterized by a more defined bimodal distribution compared to the initial image (Fig. 5.5b), which presents approximately a trimodal gray level distribution. In the case of brain tumors, MedGA enhances the underlying bimodal distribution related to contrast-enhancing tumoral tissue and brain healthy tissues on CE-MR images. This visual achievement is endorsed by the histogram of the enhanced image (see Fig. 5.6d) that shows two more distinct peaks with respect to the initial gray level distribution (see Fig. 5.6b).

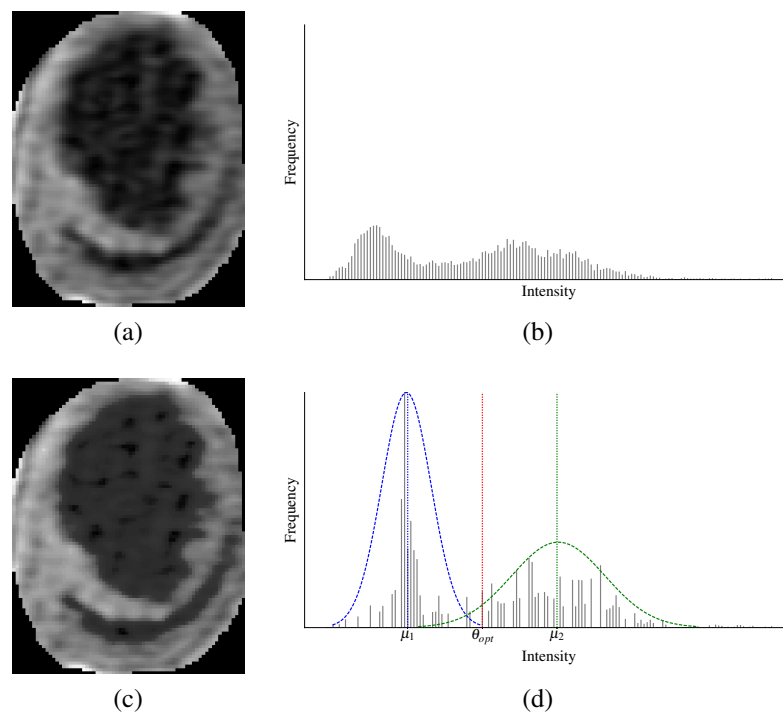


Fig. 5.5 Enhanced image obtained by MedGA on an example of uterine fibroid (size:  $89 \times 70$  pixels): (a) normalized input image using linear contrast stretching on the initial full range of the masked MR image; (c) resulting image after the application of the pre-processing using MedGA. The histograms corresponding to the sub-images in (a) and (c) are shown in (b) and (d), respectively. The final histogram emphasizes the two underlying distributions in the gray levels intensity characterized by mean values  $\mu_1$  and  $\mu_2$ , and standard deviations  $\sigma_1$  and  $\sigma_2$ , respectively. The two distributions are highlighted in the plot with blue and green dashed lines.

Although [155, 252] used the same individual structures, both the fitness function and the aim of the GA are very different. They defined two similar fitness functions that combine the number of edge pixels and the intensity of such pixels. The goal of these studies is the image contrast enhancement for consumer electronics products. The proposed method aims to enhance the image thresholding results on MRI data, generating two well-distinct normal distributions based on the input histogram, where the final best solution emphasizes the underlying bimodal distribution of the input histogram.

**MedGA parameter setting and calibration** To analyze the performances of MedGA and identify the best settings for the image enhancement problem, we considered a calibration set consisting of 80 medical images randomly selected from the available uterine fibroid MRI dataset, and we varied the settings of MedGA used throughout the optimization process,

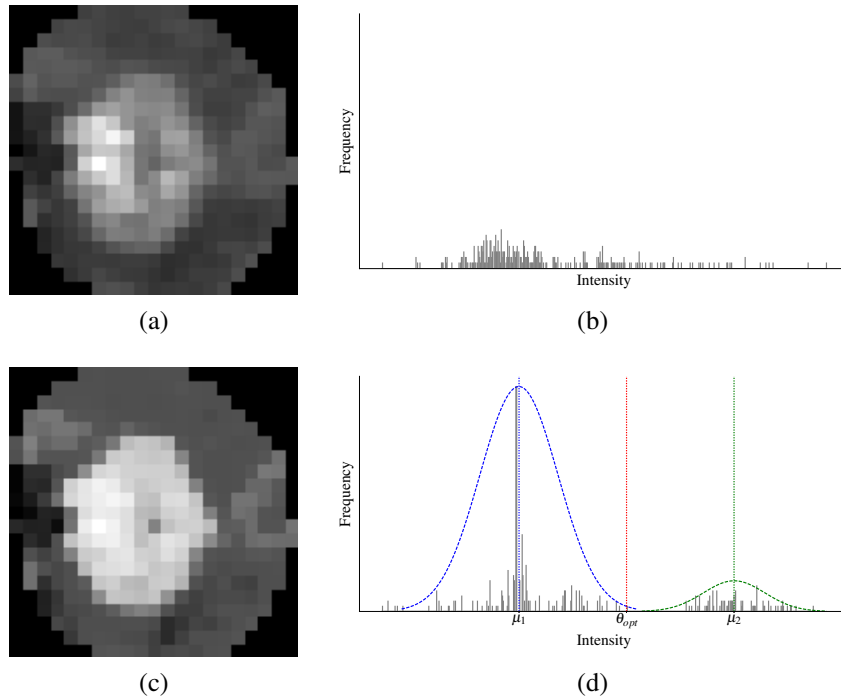


Fig. 5.6 Enhanced image obtained by MedGA on an example of brain tumor (size:  $21 \times 21$  pixels): (a) normalized input image using linear contrast stretching on the initial full range of the masked MR image; (c) resulting image after the application of the pre-processing using MedGA. The histograms corresponding to the sub-images in (a) and (c) are shown in (b) and (d), respectively. The final histogram emphasizes the two underlying distributions in the gray levels intensity characterized by mean values  $\mu_1$  and  $\mu_2$  and standard deviations  $\sigma_1$  and  $\sigma_2$ , respectively. The two distributions are highlighted in the plot with blue and green dashed lines.

that is: (i) the size of the population  $|P| \in \{50, 100, 150, 200\}$ ; (ii) the crossover probability  $p_c \in \{0.8, 0.85, 0.9, 0.95, 1.0\}$ ; (iii) the mutation probability  $p_m \in \{0.01, 0.05, 0.1, 0.2\}$ ; (iv) the size of the tournament selection strategy  $k \in \{5, 10, 15, 20\}$ . In all tests MedGA was run for  $T_{\max} = 100$  iterations. Each MedGA execution was performed by varying one setting at a time, for a total of 320 different settings tested and a total number of  $320 \times 80 = 25,600$  MedGA executions.

The results of these tests (data not shown) highlighted that, for each value of  $|P|$ , the best settings in terms of fitness values achieved are:

1.  $|P| = 50, p_c = 0.85, p_m = 0.01, k = 15$ ;
2.  $|P| = 100, p_c = 0.9, p_m = 0.01, k = 20$ ;
3.  $|P| = 150, p_c = 0.85, p_m = 0.01, k = 20$ ;

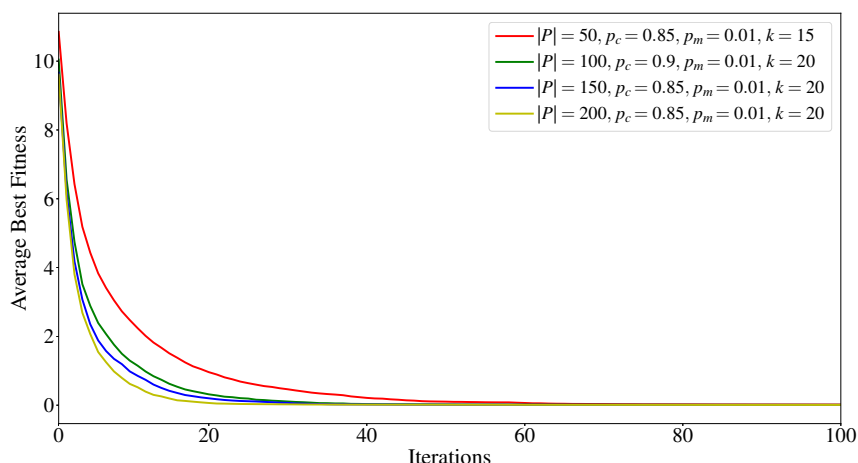


Fig. 5.7 Comparison of the ABF achieved by MedGA with the best parameterizations found for each value of  $|P|$  tested here. The average was computed over the results of the optimization of 80 MR images.

4.  $|P| = 200, p_c = 0.85, p_m = 0.01, k = 20$ .

Figure 5.7 reports the comparison of the performances achieved by MedGA with these settings, where the Average Best Fitness (ABF) was computed by taking into account, at each iteration of MedGA, the fitness value of the best individuals over the 80 optimization processes. It is clear from the plot that, despite the final ABF values are comparable in all settings, the convergence speed increases with the size of the population, as well as the running time required by MedGA; therefore, to choose the best settings, we analyzed the computational performances concerning the 4 tests listed above.

Considering that an MRI series related to a single patient contains on average 10 slices with ROI fibroids, we tested the clinical feasibility of MedGA by calculating the total execution time for enhancing 10 randomly chosen MR images. For what concerns the tests 1-4 described above, the executions lasted on average 672.12 s, 1290.15 s, 1987.8 s, and 2669.74 s, respectively, for the optimization of the same batch of 10 images running a single core of the Intel® Xeon® E5-2440 CPU with 2.40 GHz clock frequency. On the other hand, by exploiting the 6 cores of the same CPU to execute the parallel version of MedGA, we achieved up to  $3.6\times$  speed-up with respect to the sequential version. The results achieved using the parallel version of MedGA confirm the importance of HPC solutions in the field of real healthcare environment to obtain clinically feasible outcomes, that is, enhancing MR images in reasonable time for medical imaging practice.

By considering both the performance of MedGA in terms of ABF and the running time required to process 80 images, we selected the parameter settings  $|P| = 100, p_c = 0.9$ ,



Table 5.2 Values of the image enhancement evaluation metrics, achieved by MedGA and the classical image pre-processing approaches, calculated on the MRI dataset with 18 patients affected by uterine fibroids and expressed as mean and standard deviation values. Numbers in bold indicate the best values.

Method	PSNR		#DE		AMBE		SSIM	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HE	30.994	1.949	<b>975.465</b>	475.951	0.085	0.029	0.859	0.044
Bi-HE	31.880	2.046	907.177	415.703	0.038	0.020	0.907	0.032
GT $\gamma = 0.4$	30.194	2.170	717.555	401.186	0.212	0.019	0.823	0.024
GT $\gamma = 2.5$	29.952	2.127	965.012	380.967	0.261	0.012	0.586	0.075
ST $\lambda = 4/\alpha$	33.971	1.874	872.594	396.016	0.040	0.014	0.880	0.023
ST $\lambda = 6/\alpha$	32.286	1.975	869.032	378.277	0.060	0.021	0.715	0.056
ST $\lambda = 8/\alpha$	31.353	2.029	841.420	348.674	0.073	0.025	0.613	0.070
MedGA	<b>37.366</b>	2.347	866.604	409.604	<b>0.033</b>	0.011	<b>0.928</b>	0.025

Table 5.3 Values of the image enhancement evaluation metrics, achieved by MedGA and the classical image pre-processing approaches, calculated on the MRI dataset composed of 27 brain tumors and expressed as mean and standard deviation values. Numbers in bold indicate the best values.

Method	PSNR		#DE		AMBE		SSIM	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HE	34.215	1.447	38.779	21.287	0.124	0.049	0.756	0.112
Bi-HE	36.758	1.718	44.923	26.252	<b>0.042</b>	0.024	<b>0.932</b>	0.021
GT $\gamma = 0.4$	33.193	0.841	21.271	18.467	0.229	0.020	0.713	0.065
GT $\gamma = 2.5$	33.520	1.113	<b>45.119</b>	27.444	0.229	0.028	0.457	0.096
ST $\lambda = 4/\alpha$	36.812	0.923	43.574	27.307	0.055	0.019	0.848	0.048
ST $\lambda = 6/\alpha$	35.270	0.942	43.779	26.492	0.079	0.028	0.645	0.105
ST $\lambda = 8/\alpha$	34.435	0.991	44.072	26.508	0.090	0.034	0.543	0.121
MedGA	<b>37.751</b>	1.990	43.534	22.598	0.079	0.039	0.881	0.053

$p_m = 0.01$ ,  $k = 20$  as the best trade-off characterized by a good convergence speed and an adequate running time (for this specific application), and we exploited this configuration for all tests reported and discussed in what follows.

**Results** Tables 5.2 and 5.3 show the image pre-processing results achieved by each method on the uterine fibroid and brain tumor MRI datasets, respectively, using the metrics reported in Appendix A.2. In both MR image analysis applications, MedGA remarkably achieves the highest Peak Signal-to-Noise Ratio (*PSNR*) mean values with respect to the state-of-the-art methods, generally involving the highest signal quality.

Considering the results in Table 5.2, HE over-enhances the processed uterine fibroid MR images, as denoted by the highest mean number of detected edges (#DE) value [79], while Bi-HE allows for the preservation of the mean brightness, as also indicated by the lower mean value of Absolute Mean Brightness Error (*AMBE*) [26, 101]. For what concerns the other techniques, on the one hand, GT with  $\gamma = 0.4$  yields better results compared to GT

with  $\gamma = 2.5$ , especially in the case of the Structural Similarity Index (*SSIM*) metrics [651]; on the other hand, all metrics related to the tested ST functions show that their performances decrease as the value of  $\lambda$  increases. As it can be observed in Figs. 5.10 and 5.11, MedGA strengthens the ROI edges by enhancing details and features useful for image binarization; this result confirms, from a qualitative perspective, the quantitative results presented above. From an overall view of the metrics values, we can claim that the approaches obtaining the highest values of the #DE measure (i.e., HE and GT with  $\gamma = 2.5$ ) could imply a considerable over-enhancement of the output image, according to the other image quality metrics.

The results on brain tumor MRI data reported in Table 5.3 show a slightly different trend, also due to the small size of the pre-processed cropped sub-images. As a first evidence, both GTs do not preserve the input mean brightness considering the *AMBE* measure. Interestingly, GT with  $\gamma = 2.5$  and  $\gamma = 0.4$  achieve the highest and the lowest #DE values, respectively. Bi-HE strongly improves the enhancement metrics obtained by HE, by generally reporting the best results. Consistently with the metrics calculated on uterine fibroid MRI data, all the results concerning ST functions get worse when the value of  $\lambda$  increases. The highest *SSIM* mean value is achieved by Bi-HE, revealing the best structural information, even though MedGA obtains the best signal quality in terms of *PSNR* mean values.

These findings are also corroborated by a visual inspection of Figs. 5.10 and 5.11, where the enhanced images using HE and GT with  $\gamma = 2.5$  present an inadequate appearance for image observation and interpretation. Overall, these results highlight that MedGA generally outperforms the conventional image enhancement approaches considering signal and perceived image quality, while preserving the input mean brightness.

**Conclusion and future work** Differently to the state-of-the-art methods that aim at improving the contrast level of the whole image to obtain “visually pleasant” images, MedGA is the first work that explicitly addresses the challenging issues pertaining to the enhancement of medical images characterized by a nearly bimodal gray level histogram distribution. Our intelligent image enhancement system can visually assist physicians during their interactive decision-making tasks as well as improve the results in downstream automated processing pipelines for clinically useful measurements [513].

MedGA deals with the practical problems regarding the interpretability of Machine Learning and Computational Intelligence methods in medicine [74]. As a matter of fact, the final best solution (i.e., the output gray level histogram) and the enhanced image are understandable by physicians. The efficient encoding of the individuals—taking inspiration by [252]—coupled with effective HPC solutions, allows for a clinically feasible computational framework. By defining a specific fitness function to emphasize the two Gaussian distribu-

tions composing a bimodal histogram, while the existing approaches based on Evolutionary Computation or Swarm Intelligence techniques were conceived for a different purpose, i.e., improving the perceived visual information in terms of image contrast. Thanks to this *ad hoc* fitness function (differently to the GP-based image enhancement method in [477], where no fitness function is defined because the user interactively selects the best solution) and the robustness achieved by means of the calibration step for GA's parameter setting, MedGA does not require any user interaction step. Moreover, MedGA differs from the GP-based approaches, in which the final generated solution could have large size [85], so heavily affecting the readability and interpretability of the provided solutions. The main impact of this contribution consists in showing how an effective method based on evolutionary computation can outperform the existing methods in medical image enhancement.

As a future extension, in the case of large size images (e.g.,  $1000 \times 1000$  pixels), Graphics Processing Units (GPUs) can represent an enabling technology for real-time radiology applications [168], since the running time of histograms computation can be considerably reduced by using a parallel implementation [549].

Considering the achieved results in terms of #DE (see Tables 5.2 and 5.3), MedGA's performance could be further improved—in terms of contrast—by integrating a novel component in the fitness function, which explicitly relies on the number of detected edges. Since this additional component would have a different purpose and a different magnitude, a multi-objective optimization method should be taken into account. In particular, MedGA could be extended by means of an effective evolutionary approach, such as Non-dominated Sorting Genetic Algorithm (NSGA-III) [146], to simultaneously optimize both conflicting objectives, which consist in maximizing the number of edges while minimizing the distance between the optimal threshold and the two normal distributions.

### 5.1.2 Computational framework to improve MR image thresholding

Medical image segmentation concerns both detection and delineation of anatomical or physiological structures from the background, distinguishing among the different components included in the image [37]. This important task allows for the extraction of clinically useful information and features in medical image analysis [296, 648]. Accordingly, computer-assisted approaches enable quantitative imaging [159], aiming at accurate and objective measurements from digital images regarding a ROI [214, 680]. Indeed, image segmentation is still one of the most challenging research areas especially in medical image analysis [159]. Accurately delineating the ROIs is a critical task, since manual segmentation procedures are time-expensive, error-prone, and operator-dependent (i.e., not ensuring result repeatability).

In Pattern Recognition, among low-level intensity-based techniques, which are widely adopted in scenarios with real-time constraints, the most basic unsupervised image segmentation approach is global thresholding that essentially reduces to a pixel classification problem [218] (see Section 3.1.1 for the theoretical prerequisites). In particular, image binarization classifies the input pictorial data into exactly two classes (i.e., foreground and background), given a threshold intensity value [504]. This global threshold value is efficiently computed by operating on the image histogram alone. Unfortunately, image binarization techniques work properly only for input images with a bimodal histogram [677]. In practice, different types of regions in an image could overlap, thus affecting the bimodality conditions of the gray level histogram. As a matter of fact, the histogram modes semantically correspond to different types of regions. Image pre-processing can definitely improve the result accuracy achieved by computer-assisted segmentation methods [504], by sharpening the peaks of the two sub-distributions. The resulting histogram is supposed to be thus more strongly bimodal, dealing also with blurred region contours and the related Mach band effect pertaining to edge-detection in the human visual system [137, 338]. As a matter of fact, in radiology, this phenomenon is accentuated in the edges of adjacent regions that slightly differ in terms of gray level intensities [495].

No existing pre-processing technique addresses the issues related to medical image enhancement for subsequent binarization using adaptive thresholding [677]. Literature methods may be inadequate when dealing with low contrast images [358], producing false edges and under-/over-segmentation when input images are affected by noise, as in the case of MRI data [202]. In order to overcome these limitations and automatically determine a suitable optimal threshold, this computational framework exploits MedGA [529], for improving bimodal histogram separation [529]. MedGA is designed *ad hoc* to help basic thresholding techniques by improving the threshold selection between the two underlying distributions in CE-MR images with a nearly bimodal histogram.

Therefore, in [527], we used MedGA as a pre-processing step, before performing threshold-based two-class image segmentation of MR images characterized by a nearly bimodal histogram distribution. As case study, MedGA was applied to two processing pipelines, based on the efficient IOTS algorithm [500, 619], concerning two different clinical contexts that require MR image analysis:

- uterine fibroid segmentation in MRgFUS [412];
- brain metastatic cancer segmentation in neuro-radiosurgery therapy [401].

In both cases, a precise and reliable and reproducible segmentation is mandatory. Unfortunately, these tasks are generally carried out manually by experienced physicians. The

achieved image enhancement results were quantitatively evaluated using image enhancement and medical image segmentation metrics.

**MR image segmentation using adaptive thresholding** Image enhancement techniques can facilitate the user interpretation of an image, as well as improve the automated interpretation. Therefore, we use MR image segmentation as an important processing goal [582]. After the application of MedGA, the enhanced images are segmented using the IOTS algorithm [500, 619], which is the most straightforward automated segmentation approach.

Differently to existing image enhancement techniques, we aimed at designing an appropriate pre-processing method that is able to reliably enhance a particular type of medical images for further automated image processing phase. In our case, MedGA tackles the problems related to medical images with a roughly bimodal histogram, by strengthening the two underlying sub-distributions. The main goal of MedGA is to yield pre-processed medical images well-suited for classic threshold-based segmentation techniques, in order to improve ROI delineation. By so doing, the computational load—required for achieving accurate segmentation results—is transferred from the pixel classification stage to the pre-processing phase by means of an effective Soft Computing technique.

Fig. 5.8 shows the overall flow diagram of the proposed computational framework that integrates MedGA for the thresholding-based segmentation of bimodal MR images. We developed two slightly different post-processing pipelines to refine the results achieved by this efficient adaptive thresholding technique. These post-processing steps, here applied to perform uterine fibroid and brain tumor segmentation, are described in what follows. MedGA is able to enhance images in segmentation tasks involving both hyper- and hypo-intense ROIs in CE-MR images, also dealing with data unbalanceness (i.e., the number of foreground pixels is either much higher or lower than the number of background pixels).

**Uterine fibroid segmentation** Firstly, uterus region delineation is required. This task can be accomplished manually by the user or automatically by means of computational methods to reduce operator-dependency, as described in [415]. ROIs are represented by tissues with low contrast mean absorption (i.e., NPV), thus the pixels with lower values with respect to the achieved threshold are yielded in the binarized MR image. However, segmentation approaches have to take into account NPV inhomogeneities, due to sonication spots during the MRgFUS treatment.

The used post-processing refinement steps are the following (Fig. 5.8a):

1. morphological opening with a circular structuring element (2-pixel radius) to separate possible loosely connected hypo-intense regions;

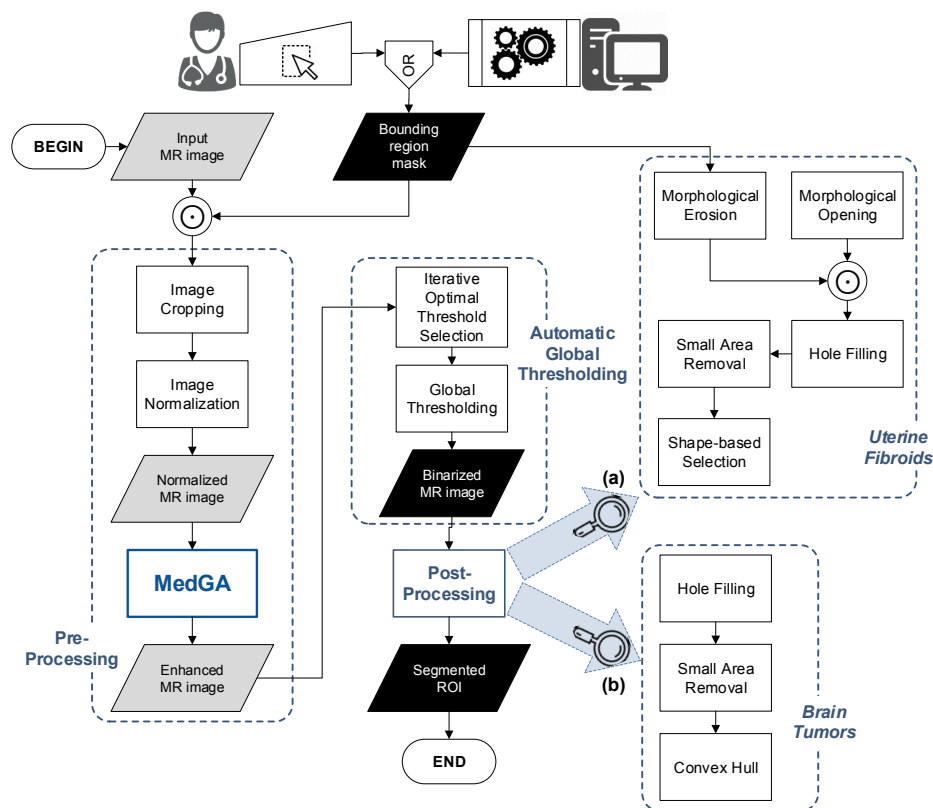


Fig. 5.8 Flow diagram of the proposed computational framework that integrates MedGA as a pre-processing step for MR image segmentation based on the efficient IOTS algorithm [500, 619]. Note that two slightly different post-processing pipelines were developed for (a) uterine fibroids and (b) brain tumors. Gray and black data blocks denote MRI gray-scale images and binary masks, respectively.

2. some regions at the boundary of the uterus bounding region mask could present similar intensity values to gray levels characterizing fibroid regions, so being included in the thresholding output. To eliminate this ambiguity, it is appropriate to apply a morphological erosion (with a circular structuring element of 5-pixel radius) to the ROI binary mask, and then the logical pixel-by-pixel product (i.e., Hadamard multiplication) with the image resulting from the previous step is performed;
3. a hole filling algorithm is necessary to deal with possible holes in fibroid regions also due to non-uniform distribution of ablated tissue caused by sonication spots;
4. segmentation is further improved through a connected-component labeling based operation by removing objects that are smaller than a certain area (i.e., 120 pixels) and characterized by similar intensity with respect to the fibroids to be treated, because there may be regions or artifacts caused by very small dark areas;

5. some lengthened connected-component with sufficiently large area could be present (i.e., due to other anatomical structures or to acquisition artifacts). Fibroids, in fact, present a spherical or semi-spherical shape [632] that can be denoted by means of the parameters of the various connected-components. This connected-component based selection considers the eccentricity and the extent of the detected regions. Specifically, experimental reference values to discern fibroids from the rest of connected-components are:  $0.3 \leq \text{extent} < 0.8$  and  $0.0 \leq \text{eccentricity} < 0.8$ , according to [415]. Lastly, any connected-component, which has passed the shape-based control and whose centroid distance is more than a given upper limit (i.e.,  $\frac{\sqrt{M^2+N^2}}{3}$ ) from the MR image center, is removed.

**Brain tumor segmentation** The accurate and reproducible measurement of tumor size and its changes over time is crucial for diagnosis, treatment planning, as well as monitoring of response to oncologic therapy for brain tumors [401]. As a preliminary step, the user has to interactively select a bounding region that includes the tumor zone (by means of a free-hand “lasso” tool). Since the areas to segment are enhancement regions, the pixels that have higher intensities than the threshold are selected during the image binarization phase. Brain metastatic cancers may contain also necrotic material, so these inner necrotic cores could affect the achieved enhancement region segmentation. Therefore, some refinement steps are useful to cope with this situation.

The used post-processing pipeline is described in the following (see Fig. 5.8b):

1. hole filling algorithm to consider also necrotic areas;
2. adaptive post-processing steps based on the size of the input image, consisting in small area removal (considering 4-connectivity) with minimum threshold equal to 30 pixels on images with size greater than 300 pixels, or 10 pixels otherwise;
3. to allow also for large bounding regions, shape-based selection is applied in the case of at least two connected-components, according to:  $\text{extent} \geq 0.6$  and  $0.0 \leq \text{eccentricity} < 0.8$  (see [526]). However, when a single connected-component is present, these controls are avoided;
4. brain metastases have a pseudo-spherical appearance [18], therefore a convex hull algorithm is employed to envelope the segmented lesion into the smallest convex polygon containing this region [707].

The thresholding-based segmentation results achieved on the MR images pre-processed by MedGA in Figs. 5.5c and 5.6c are shown in Figs. 5.9a and 5.9b, respectively.

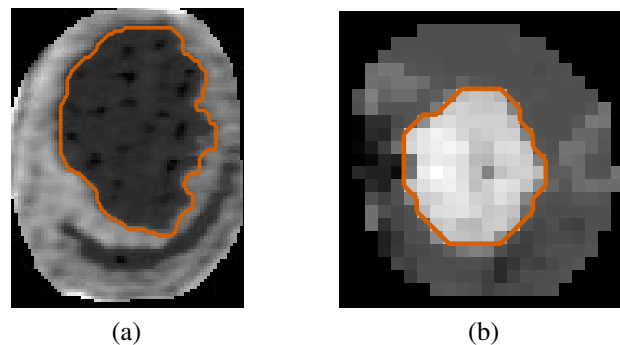


Fig. 5.9 Segmentation results achieved by the IOTS algorithm on the MR images pre-processed by MedGA: (a) uterine fibroid delineation on the image given in Fig. 5.5c; (b) brain tumor delineation on the image given in 5.6c.

**Results** This section presents the experimental results achieved by MedGA, according to the evaluation metrics defined in Appendix A.1. In order to achieve a comprehensive comparison between MedGA and the other pre-processing techniques listed above, we exploited the entire set of MRI data consisting in 18 patients affected by uterine fibroids and 27 brain metastatic cancers. Figs. 5.10 and 5.11 show two examples of uterine fibroid and brain tumor MR images, respectively, which were pre-processed by means of the comparison methods considered in this work and segmented using the processing pipelines described above. Fig. 5.12 shows two pairs of examples concerning tridimensional reconstructions of the ROIs, i.e., uterine fibroids and brain tumors, allowing us to display their actual locations in the whole uterus and brain (represented by means of a transparent red surface), respectively. This visualization allows for an intuitive and comprehensive representation of complex data [646].

The quantitative segmentation results achieved by using the pipeline in Fig. 5.8a, employing the different pre-processing approaches, on the analyzed MRI dataset composed of 18 patients affected by uterine fibroids are depicted in the boxplots in Fig. 5.13, reporting both overlap-based and distance-based metrics values. Analogously, the boxplots concerning the segmentation results, achieved by using the pipeline in Fig. 5.8b on the analyzed MRI dataset consisting in 27 brain metastases, are shown in Fig. 5.14.

In the literature, it has been shown that a *DSC* above 70% is generally regarded as a satisfactory level of agreement between two segmentations (i.e., manual and automated delineations) in clinical applications [705]. Since the MR image segmentation methods above described obtain a *DSC* appreciably higher than 70% regardless the pre-processing technique, we can consider that the processing pipelines in Figs. 5.8 are clinically valuable,



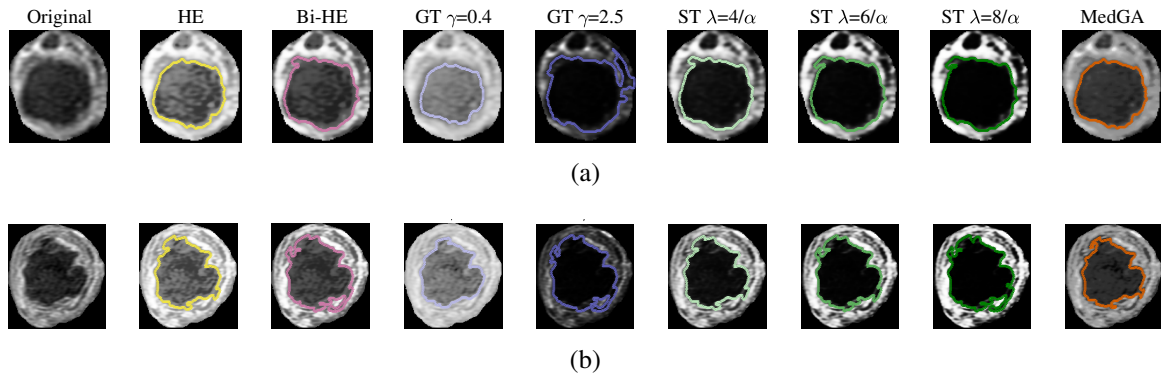


Fig. 5.10 Segmentation results on the uterine fibroids shown in Figs. 5.1a and 5.1b, achieved by the processing pipeline presented in Fig. 5.8a exploiting the state-of-the-art image pre-processing approaches (namely: HE, Bi-HE, GT  $\gamma = 2.5$ , GT  $\gamma = 0.4$ , ST  $\lambda = 4/\alpha$ , ST  $\lambda = 8/\alpha$ , ST  $\lambda = 6/\alpha$ ) and MedGA.

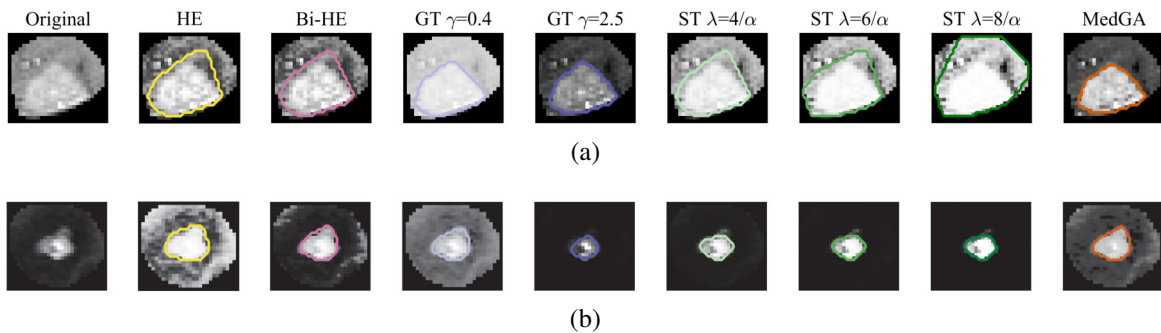


Fig. 5.11 Segmentation results on the the brain tumors in Figs. 5.1c and 5.1d, achieved by the processing pipeline in Fig. 5.8b by exploiting the state-of-the-art image pre-processing approaches (namely: HE, Bi-HE, GT  $\gamma = 2.5$ , GT  $\gamma = 0.4$ , ST  $\lambda = 4/\alpha$ , ST  $\lambda = 8/\alpha$ , ST  $\lambda = 6/\alpha$ ) and MedGA.

so allowing for a fair comparison on segmentation performance among the state-of-the-art pre-processing algorithms.

In both cases, the segmentation results concerning the images pre-processed using MedGA achieved the highest mean and median *DSC* values, with low standard deviation. GT with  $\gamma = 0.4$  and Bi-HE obtained the second best performances for uterine fibroid and brain tumor MR image segmentation, respectively. So, we can claim that MedGA shows the highest accuracy and reliability in the two considered MRI analysis tasks. This evidence is also confirmed by the boxplots, where the distributions for MedGA present significantly less than 10% outliers in all the overlap-based metrics, thus evidencing extremely low statistical dispersion. As a matter of fact, MedGA is the only technique that significantly supports the IOTS algorithm in both dark (i.e., uterine fibroid NPV) and bright (i.e., brain tumor

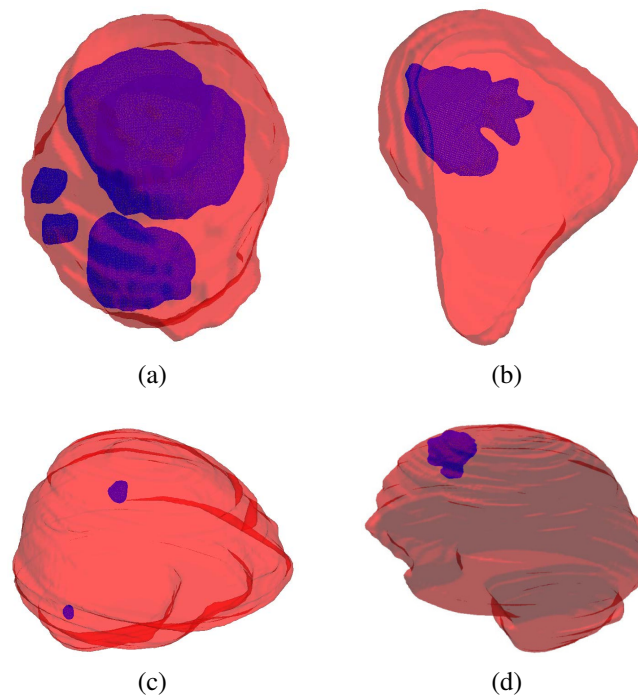


Fig. 5.12 Tridimensional reconstruction of the segmented ROIs (blue volumetrics models) in their real location with respect to the enclosing organ (transparent red surface): (a, b) uterine fibroids within the uterus, segmented using the method in [415]; (c, d) brain tumors within the whole brain, achieved using a skull stripping algorithm. Transparent surfaces are rendered with alpha blending ( $\alpha = 0.40$ ).

enhancement region) ROI extraction. In agreement with the image enhancement results discussed in Section 5.1.1, GT with  $\gamma = 0.4$  considerably outperforms GT with  $\gamma = 2.5$ . The decreasing trend, related to ST when the value of  $\lambda$  increases, is also confirmed. ST with  $\lambda = 4/\alpha$  achieved good results in both cases. Brain tumor MR images pre-processed by means of HE achieved low *DSC* values, but better results are obtained on uterine fibroid MR segmentation with respect to Bi-HE. Overall, the achieved segmentation performance shows the great accuracy and reliability of the proposed evolutionary method. Considering *SEN* and *SPC*, MedGA yielded the best trade-off between these two often conflicting measures that should be always considered and combined together. These metrics reveal that the other techniques could involve over-estimation or under-estimation of the segmented regions.

The achieved spatial distance-based indices are consistent with overlap-based metrics, also observing the corresponding boxplots shown in Figs. 5.13 and 5.14, respectively. Hence, MedGA allows also for accurate results in terms of distance between the automated and the manual boundaries. Note that generally the boxplots pertaining to MedGA results present the lowest statistical dispersion (in terms of box width and number of outliers), so implying a

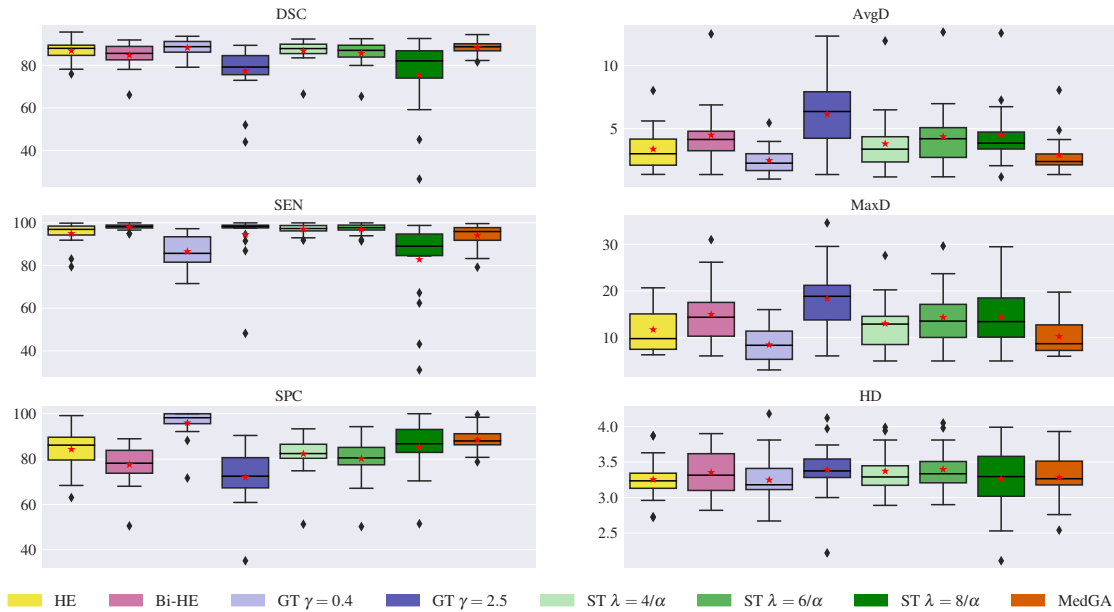


Fig. 5.13 Boxplots of the overlap-based and distance-based metrics (left and right columns, respectively) obtained on the MRI dataset composed of 18 patients affected by uterine fibroids who underwent MRgFUS treatment. The lower and the upper bounds of each boxplot represent the first and third quartiles of the statistical distribution, respectively. The median (i.e., the second quartile) and the mean values are represented by a black solid line and a red star, respectively. Whisker value is 1.5 in all cases, and outliers are displayed as black diamonds.

lower standard deviation with respect to the conventional techniques. Therefore, the use of MedGA as a pre-processing step allows for considerably robust and reliable segmentation results. These experimental findings are endorsed by the segmentation examples shown in Figs. 5.10 and 5.11.

**Conclusion** We proposed MedGA [529, 527], a novel Soft Computing method based on Evolutionary Computation, designed *ad hoc* for enhancing MR image segmentation results of MRI data characterized by an underlying bimodal histogram. This image pre-processing approach exploits a GA that aims at emphasizing bimodal histogram separation, and consequently optimize the subsequent segmentation phase based on the efficient IOTS algorithm [500, 619].

In Pattern Recognition, among the low-level intensity-based techniques, the most straightforward unsupervised image segmentation technique is global thresholding [677]. In order to overcome the limitations related to the assumptions underlying threshold selection methods and automatically determine a suitable optimal threshold, MedGA tackles this

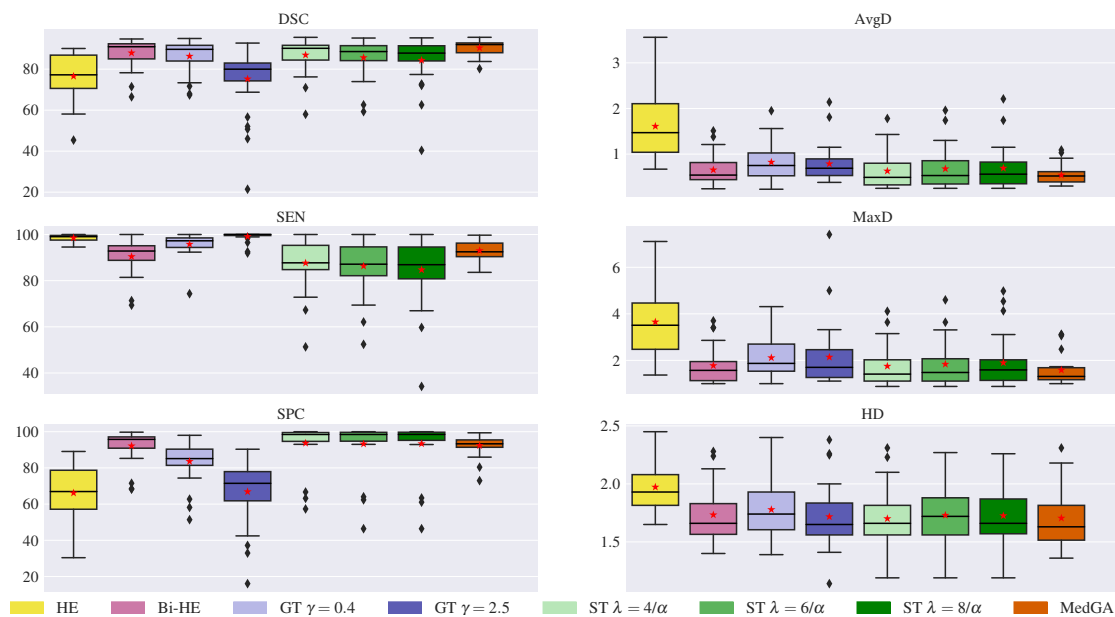


Fig. 5.14 Boxplots of the overlap-based and distance-based metrics (left and right columns, respectively) obtained on the MRI dataset composed of 27 brain metastatic tumors underwent stereotactic neuro-radiosurgery. The lower and the upper bounds of each boxplot represent the first and third quartiles of the statistical distribution, respectively. The median (i.e., the second quartile) and the mean values are represented by a black solid line and a red star, respectively. Whisker value is 1.5 in all cases, and outliers are displayed as black diamonds.

complex problem, introducing a fitness function tailored to better separate the two underlying sub-distributions of the gray level intensities. Unlike the traditional image enhancement techniques that, generally, improve the contrast level of the whole image, MedGA focuses on MR image sub-regions characterized by a roughly bimodal histogram, making it valuable in clinical contexts, especially involving CE-MRI analysis. As a matter of fact, the Soft Computing approaches presented in [155, 252] explicitly consider in the fitness function both the number of edge pixels and the intensity of these pixels, thus achieving high #DE values that would consistently lead to over-enhanced images, possibly yielding also inaccurate ROI segmentations.

MedGA was integrated into two processing pipelines concerning different clinical contexts that require MR image analysis: (i) uterine fibroid segmentation in MRgFUS treatments, and (ii) brain metastatic cancer segmentation in neuro-radiosurgery therapy. Overall, the MedGA pre-processing outperformed, in terms of image quality and segmentation accuracy, the conventional image enhancement approaches. According to the achieved experimental results, MedGA was shown to be an appropriate and reliable solution when employed as a medical image pre-processing method. Even considering the statistical dispersion of the

segmentation evaluation metrics, MedGA achieved the most robust and repeatable segmentation results. More generally, in addition to the clinical contexts investigated here, the application of MedGA can be extended also to real-world problems involving the analysis of images characterized by an underlying bimodal histogram, as in the case of bright-field and fluorescence microscopy imaging [402].

MedGA has been currently used only for off-line image analysis. Our Python-coded implementation requires approximately two minutes for each image (running on a computational platform equipped with a 6-core Intel<sup>®</sup> Xeon<sup>®</sup> E5-2440 CPU at 2.40 GHz, 16 GB RAM, and CentOS 7 operating system), using the selected GA parameters. In the case of tomography image stack analysis, we achieved a sublinear speed-up with respect to the number of the available cores, by developing a Master-Slave version to distribute on multiple cores the computations pertaining to different slices [469]. An additional speed-up could be definitely obtained by porting the current Python code into a faster compiled programming language (e.g., C/C++) [44]. Consequently, leveraging efficient programming languages and HPC paradigms, MedGA may become a clinically feasible solution as a pre-processing step in real-time radiology applications.

In the near future, we plan to apply MedGA as an image pre-processing phase also in other clinical contexts requiring MR image analysis and segmentation to provide useful insights for differential diagnosis and prognosis, such as in the case of breast cancer [10, 253] and meningiomas [241], also for differentiating tumor grade.

## 5.2 Particle Swarm Optimization

Swarm Intelligence studies the collective behavior of decentralized, self-organized natural or artificial systems [163]. These models consist typically in a population of simple agents interacting locally each other and with their environment. The agents follow very simple rules and, although there is no centralized control structure dictating how individual agents should behave, local interactions among such agents—often affected by a certain degree of stochasticity—lead to a complex intelligent emergent global behavior, with effects that would not have been expected by each individual.

PSO is a population-based stochastic optimization algorithm, introduced in 1995 by Kennedy and Eberhart [316], which searches an optimal solution in the computable search space [161]. This global optimization technique, which is inspired by the collective movement of bird flocks and fish schools, results in a metaheuristic for solving non-linear optimization problems [606]. Generally, such problems cannot be solved exactly by an explicit method because in practice the mathematical expression of the objective function, also called fitness

function  $\mathcal{F} : \Omega \rightarrow \mathbb{R}$ , is not available, where  $\Omega$  represents the search space. Instead of computing the optimum's position  $\mathbf{x}^* \in \Omega$ , a sufficiently good not necessarily optimal point  $\hat{\mathbf{x}}^* \in \Omega$ , called sub-optimal solution, is obtained by a metaheuristic in the given search space.

PSO can be also seen as an evolutionary technique which, in contrast to GAs [270] and traditional ESs [29] that use the competitive characteristics of biological survival, exploits cooperative and social aspects. Starting from a widely diffused population (swarm), individual components (particles) tend to move through the search space, eventually clustering in regions where minima are identified. Briefly, PSO simulates natural movement evolution for searching a solution with higher quality. This population-based algorithm is suitable for optimization problems with solutions encoded as real-valued vectors (i.e.,  $\mathbf{x}_i \in \mathbb{R}^D$ , with  $i = 1, 2, \dots, N$ ) [316].

In the traditional PSO formulation, a *swarm* of  $N$  individuals (*particles*) moves inside a bounded  $D$ -dimensional search space to the aim of identifying the optimal solution, collecting and sharing information about the best position found, with respect to a user-defined fitness function, and cooperating to identify the optimal solution for a complex problem. Typically, the particles are randomly initialized in the search space with a uniform distribution. The movement of particles in the (bounded) search space is influenced by two different attractors: (i) the social component that leads particles towards the (current) best position  $\mathbf{g} \in \mathbb{R}^D$  identified by the whole swarm (or by a specified particle's neighborhood), and (ii) the cognitive component that drives each particle towards its own best position  $\mathbf{b}_i \in \mathbb{R}^D$  found so far. Social and cognitive attractions are weighted by means of two constants, named the social ( $c_{\text{soc}} \in [0, 2]$ ) and cognitive ( $c_{\text{cog}} \in [0, 2]$ ) factors, respectively. In order to prevent a chaotic exploration of the search space, the movement of particles is further weighted by using an inertia factor  $w$ . Since a deterministic movement of particles would probably lead to local optima, both attractors are multiplied by two vectors  $\mathbf{r}_1$  and  $\mathbf{r}_2$  of random numbers, uniformly sampled from the interval  $(0, 1)$ . In addition, to avoid a chaotic behavior of the swarm, the change of particle velocity is attenuated by an inertia weight  $w \in \mathbb{R}^+$ .

Formally, the velocity of the  $i$ -th particle at iteration  $t$  is updated as follows:

$$\begin{aligned} \mathbf{v}_i(t) = & w \cdot \mathbf{v}_i(t-1) + \\ & + c_{\text{soc}} \cdot \mathbf{r}_1 \odot (\mathbf{x}_i(t-1) - \mathbf{g}(t-1)) + \\ & + c_{\text{cog}} \cdot \mathbf{r}_2 \odot (\mathbf{x}_i(t-1) - \mathbf{b}_i(t-1)), \end{aligned} \quad (5.4)$$

where  $\odot$  denotes the component-wise multiplication operator (i.e., Hadamard product). Afterwards, the positions of particles are updated by calculating:

$$\mathbf{x}_i(t) = \mathbf{x}_i(t-1) + \mathbf{v}_i(t), \text{ for all } i = 1, \dots, N. \quad (5.5)$$

The quality of the positions within the search space encoded by each particle is evaluated by means of a *fitness function*  $\mathcal{F}(\cdot)$ . The hyper-surface described by the fitness values over the set of feasible solutions is called the fitness landscape. The movement of particles is the result of the two attractors, which are determined according to the fitness function  $\mathcal{F}$  defined for the problem under investigation: during each iteration of the PSO,  $\mathbf{b}_i$  and  $\mathbf{g}$  are updated according to the new positions. Summarizing, the  $i$ -th particle is able to memorize its own best position  $\mathbf{b}_i$  from the past, thus creating a kind of “nostalgia” to return there. When the fitness function is to be maximized, if  $\mathcal{F}(\mathbf{x}_i(t))$ , then  $\mathbf{b}_i(t) \leftarrow \mathbf{x}_i(t)$ , otherwise the previous individual best position  $\mathbf{b}_i(t-1)$  is kept. On the other hand, the particle  $\mathbf{x}_i$  has also a social behavior, because it follows the swarm in its global best position  $\mathbf{g}(t) \leftarrow \arg \max_{\forall \mathbf{b}_i(t)} \{\mathcal{F}(\mathbf{b}_i(t))\}$ .

This search procedure may bring some particles to regions outside the feasible space of solutions for the optimization problem. To avoid this situation, the search space is bounded (according to the domain knowledge) and some specific boundary conditions are exploited to keep the particles inside the feasible space of solutions. For instance, the so-called *damping* boundary conditions could be considered, in which a random bounce on the limit of the search space is simulated to let the particle go back to the feasible space of solutions [673]. The boundaries of the  $d$ -th dimension of the search space are denoted by  $\beta_{\min_d}$  and  $\beta_{\max_d}$ , for each  $d = 1, \dots, D$ , with  $\beta_{\min_d}, \beta_{\max_d} \in \mathbb{R}$ . The particle velocity is usually clamped to a maximum value  $v_{\max_d} \in \mathbb{R}^+$  along each  $d$ -th dimension of the search space (with  $d = 1, \dots, D$ ), to avoid divergence during the optimization process [476]. The velocity of particles may diverge during the optimization process; for this reason, its value is usually clamped to a maximum value  $v_{\max_d} \in \mathbb{R}^+$  along each  $d$ -th dimension of the search space, with  $d = 1, \dots, D$  [478].

Note that the values used for the  $N$ ,  $c_{\text{soc}}$ ,  $c_{\text{cog}}$ ,  $w$  parameters, and the vector of the allowed maximum velocity values  $\mathbf{v}_{\max} : \mathbf{v}_i \in [-\mathbf{v}_{\max}, \mathbf{v}_{\max}]$ , typically set by the user and strongly problem-dependent, have a deep impact on the performance of PSO [478], both in terms of convergence speed and quality of the best solution found. PSO performance strongly depends on a proper choice of the aforementioned settings (i.e.,  $N$ ,  $c_{\text{soc}}$ ,  $c_{\text{cog}}$ ,  $w$ , the maximum number of iterations  $T_{\max}$ ) [162, 618], and the vector of the allowed maximum velocity values  $\mathbf{v}_{\max}$ ; in addition, even the initial distribution of particles can largely affect the results [89]. Unfortunately, these settings are problem-dependent, so that a proper setting for PSO can be established only by having an in-depth knowledge of the shape and roughness of the fitness landscape. These parameters are typically tuned by the user and may heavily influence the convergence speed and quality of the best solution found [478]. In general, the identification of the best settings is a complex, lengthy and time-consuming meta-problem. Hence, an intense research has been devoted to self-tuning, settings-free, or adaptive variations of PSO and other evolutionary and Swarm Intelligence based algorithms. TRIBES is a PSO version

able to automatically modify the particles' behavior, as well as the topology of the swarm, resulting in a settings-free version of the algorithm [117]. Other examples are represented by the Parameter-Less Evolutionary Search that dynamically selects the settings by exploiting the statistical properties of the population [452]. More recently, Nobile *et al.* proposed in [440, 438] a novel PSO variant that exploits a set of fuzzy rules [691] to dynamically change, for each particle of the swarm, the values of  $w$ ,  $c_{soc}$ , and  $c_{cog}$ . In this way, PSO's *reactive* individuals are turned into *proactive* optimizing agents [440]. The resulting version of the PSO, called Fuzzy Self-Tuning PSO (FST-PSO) [438], achieved promising performance on benchmark functions [604] as well as in the parameter estimation of biochemical systems [441]. Finally, in order to avoid particles to get trapped in local optima, reboot strategies for PSO that reinitialize particle positions may be useful [580].

In addition to the basic PSO algorithm, several modifications are possible to enhance the achieved results:

- the best position in a neighborhood of each particle  $\mathbf{x}_i$  could be considered, named  $\mathbf{b}'_i$ , instead of the global best position  $\mathbf{b}$  [592]. Therefore,  $\mathbf{b}'_i$  replaces  $\mathbf{g}$  in the updating rules in Eq. (5.4);
- Evolutionary Computation may also be integrated, giving rise to a hybrid PSO. Genetic operators, such as mutation and crossover, can be used to preserve exploration capability in the various iterations, especially in the later stage of the evolution process [603]. Such a case could occur even at the early stage for a particle that is very close to the global best position  $\mathbf{g}$ , and the velocity will tend to zero. To avoid premature convergence and stagnation, after the particle positions updating, pairs of particles are selected for crossover with probability  $p_c$ , which is a random number generated uniformly in the interval  $[0, 1]$ . For each pair, two child particles are generated by a crossover rule and replace their parents, maintaining the population size constant to  $N$  [643]. The authors of [371] proposed the following crossover rule for parents  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , with  $i \neq j$ :

$$\begin{cases} \mathbf{x}'_i = p_c \mathbf{x}_i + (1 - p_c) \mathbf{x}_j \\ \mathbf{x}'_j = p_c \mathbf{x}_j + (1 - p_c) \mathbf{x}_i \end{cases} \quad (5.6)$$

The velocities  $\mathbf{v}_i$  and  $\mathbf{v}_j$  are also updated from the velocities of the parents  $\mathbf{v}'_i = \mathbf{v}'_i \cdot \mathbf{v}$ ,  $\mathbf{v}'_j = \mathbf{v}'_j \cdot \mathbf{v}$ , where:  $\mathbf{v} = (\mathbf{v}_i + \mathbf{v}_j) / \|\mathbf{v}_i + \mathbf{v}_j\|$ ;

- grouping the particles into sub-populations is a further alternative. Any clustering method can be used to perform this subdivision. Another random number  $p_{spc} \in [0, 1]$



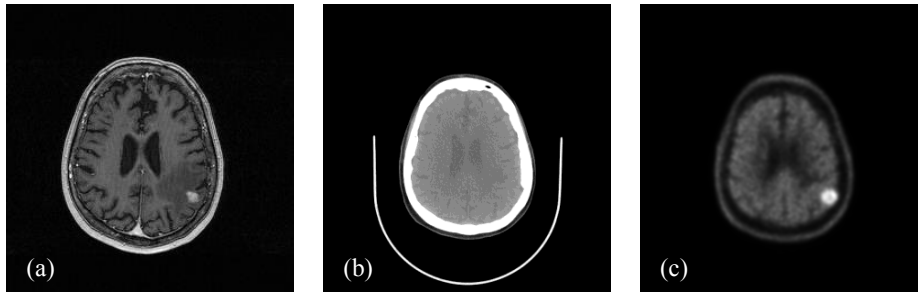


Fig. 5.15 Examples of different medical image modalities concerning the same subject affected by brain tumor: (a) MRI; (b) CT; (c) PET.

is specified to represent the probability of intra-population crossover. Crossover among different sub-populations occurs with probability  $1 - p_{sp_c}$ ;

- a constriction factor  $\chi$  was also introduced in Eq. (5.4) to control the movement of  $\mathbf{x}_i$ , by balancing both convergence and explosive particle movements [118]. The  $\chi$  coefficient must hold the following conditions:

$$\chi = \frac{2\kappa}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|}, \varphi = c_{\text{cog}} + c_{\text{soc}}, \varphi > 4, \kappa = [0, 1]. \quad (5.7)$$

### 5.2.1 Multimodal biomedical image registration using PSO

Image co-registration is a fundamental task in medical imaging because it enables to integrate different images into a single representation (i.e., the same reference system), allowing physicians and researchers to access at this complementary information more easily and accurately [449, 505]. Fused image data can improve medical diagnosis, surgery planning and simulation, as well as intra-operative navigation. The choice of the modality depends strongly on the medical task. In radiotherapy planning, for instance, dose calculation is based on the CT data, while tumor delineation is often better performed on the corresponding MRI scans, especially in soft-tissue imaging (e.g., brain or prostate) [381]. The same anatomical district of the human body is often imaged with different modalities (Fig. 5.15). In particular, morphologic images, which define the anatomy of organs or pathological tissues, can be integrated with functional ones, which describe the cellular physiology or the metabolism (i.e., PET imaging). These images are used in a complementary fashion to gain additional insights into biological phenomena and pathologies.

Even though combining appropriately different modalities is definitely clinically useful, multimodal images even concerning the same subject generally differ by local geometrical differences. Therefore, in order to perform quantitative and precise evaluations on biomedical

imaging data, such images have to be mapped into the same coordinate system by means of the alignment process, named biomedical image co-registration. Medical image registration algorithms can be conveniently applied in: (i) intra-modality matching, for patient follow-up (e.g., tumor response assessment over the time); (ii) inter-modality matching, for comparisons and quantitative measurements concerning images acquired with different modalities. The aim of this section is to provide an overview on biomedical image registration using PSO.

### Theoretical background on biomedical image registration

An image co-registration stage is mandatory to integrate and quantitatively compare biomedical imaging data originating from different modalities. As a matter of fact, image registration is able to bring the different medical datasets, concerning the same patient, into the same space. In this way, it will be possible to make quantitative and meaningful comparisons between the ROIs, such as organs and tumors, obtained by processing several image series from different scanners and modalities [474].

Registration approaches can be principally distinguished in feature-based and intensity-based schemes [505]:

- feature-based techniques find correspondences between geometrical image features (e.g., points, lines, contours, surfaces) or landmarks (external or anatomical). Either image features or landmarks are first extracted from the input images, and then a transformation is established according to the correspondence between the found features. However, segmenting and finding correspondences are very difficult tasks;
- intensity-based techniques directly exploit image intensities to compute the transformation that maximizes a similarity metrics by searching in a certain space of transformations and comparing intensity patterns. The main advantage of these schemes is that explicit image segmentation or feature extraction is not required.

Intensity-based registration techniques are the most suitable for applying Swarm Intelligence approaches. Fig. 5.16 shows the overall flow diagram of the biomedical image registration process, by using an intensity-based technique.

From an algorithmic perspective, image co-registration involves finding the parameters, i.e., a geometric transformation matrix  $\mathbf{T}$ , which either maximize or minimize some objective function, also called fitness function  $\mathcal{F}$ . Therefore, image co-registration can be modeled as an iterative procedure by successive refinements. In each iteration, the current estimate of the transformation  $\hat{\mathbf{T}}$  is used to calculate a similarity measure. Afterwards, the optimization algorithm makes another (hopefully better) estimate of the transformation, evaluates the

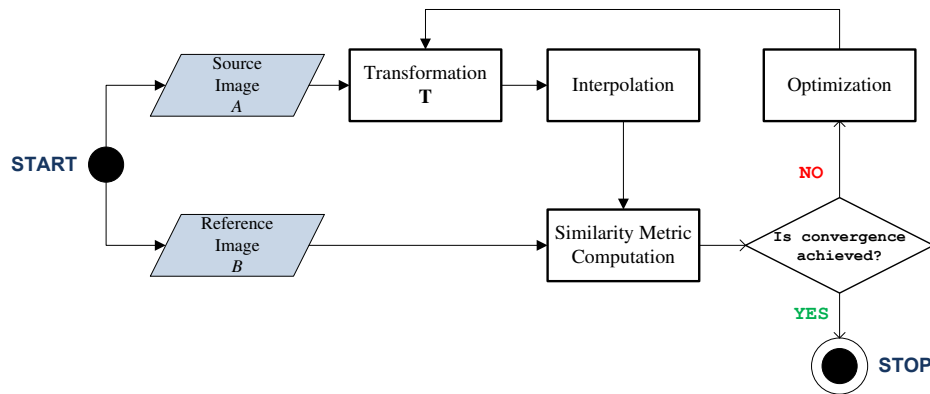


Fig. 5.16 Flow diagram of the biomedical image registration process, by using an intensity-based technique.

similarity measure again and continues to iterate until the convergence condition is achieved (i.e., no transformation can improve the value of the similarity metrics above a preset tolerance threshold  $\epsilon$  or the number of possible iterations  $T_{\max}$  is achieved). The multimodal registration process starts from two images  $A$  and  $B$  probably characterized by different FOVs:  $A$  is used as source image (floating) while  $B$  is the reference image (target). The reference imaging modality is usually selected according to the higher spatial resolution and image content. For instance, in PET/MRI registration, MR images are chosen as target images because they convey more anatomical information than PET images.

The geometric transformation  $T$  is defined only in the region of overlap of the image FOVs, and must take into account image sampling and resolution. It is also important to note that the images are discrete. The discretization is determined by the sampling grids, which are different when FOVs are not the same [267]. Even if the images  $A$  and  $B$  have exactly the same sampling grid, the grid points do not normally coincide in the volume of overlap and an interpolation step is therefore necessary. Realignment and reslicing operations are mandatory to get a one-to-one mapping between different modality slices. Especially, in iterative registration algorithms, an accurate interpolation method is required [381, 474] (e.g., B-Spline interpolation [610]). In intensity-based registration techniques, the three core components are: (i) the search space (i.e., parameter space); (ii) the similarity metrics (i.e., objective function); (iii) the search strategy (i.e., optimization algorithm).

**Search space** The search space  $\Omega$  is the set of potential transformations used to align the images. Each point in the parameter space corresponds to a different estimate of the transformation. Accordingly, the parameter space can be thought as a high-dimensionality

function in which the value of each location corresponds to the value of the similarity measure for that transformation estimate. Geometric transformations may be rigid, affine, and elastic.

Three dimensional (3D) rigid-body registration has six degrees of freedom, regarding:

- translations along the three axes of the reference system  $x$ ,  $y$ , and  $z$ , denoted by the displacements  $t_x$ ,  $t_y$ , and  $t_z$ , respectively;
- rotations around the  $x$ ,  $y$ , and  $z$  axes, denoted by the angles  $\alpha$ ,  $\beta$ , and  $\gamma$ , respectively.

In 3D rigid-body registration, the mapping of the coordinates  $\mathbf{p}_A = (x, y, z)^\top$  into the transformed coordinates  $\mathbf{p}'_A = (x', y', z')^\top$  can be suitably formulated as a matrix multiplication in homogeneous coordinates with the geometric transformation matrix  $\mathbf{T}$ . Some registration algorithms increase the number of degrees of freedom by allowing for anisotropic scaling (giving nine degrees of freedom) and shears (giving twelve degrees of freedom). A transformation that includes scaling and shearing as well as the rigid-body parameters is referred to as affine, and has the important characteristics that it can be described in matrix form and that all points, straight lines and planes are preserved. Rigid-body registration is widely used in medical applications where the structures of interest are either bony or are enclosed in bones (e.g., head, neck, pelvis, leg or spine), but the errors are likely to be larger. The use of an affine transformation rather than a rigid-body one does not significantly increase the applicability of image registration, as there are not many organs that only stretch or shear. Tissues usually deform in more complicated ways. However, errors introduced by the scanner may occur, resulting in scaling or skew terms, and affine transformations are sometimes used to overcome these problems [267]. For most organs in the body, many more degrees of freedom are necessary to describe the tissue deformation with adequate accuracy, thus elastic or non-rigid methods are required to cope with local differences between the images. However, for global elastic transformation, the number of parameters to be optimized is generally too large (often many thousands) to be feasible in practice. Therefore, two-step intensity-based registration approaches are typically used. In the first step, the global affine medical image registration is used to establish a one-to-one mapping between the two images to be registered. Afterwards, the images are registered up to small local elastic deformation.

**Similarity metrics** The similarity metrics is an indicator that quantifies the degree of closeness between features or intensity values of two images. The Sum of Squared intensity Differences (SSD), correlation coefficient, ratio image uniformity are often utilized in intra-modality registration [267]. Because of the similarity of the intensities in the images being registered, these subtraction, correlation and ratio techniques are pretty intuitive. With

inter-modality registration, the situation is quite different: there is, in general, no simple relationship between the intensities in the images  $A$  and  $B$  [474].

MI, denoted by  $I(A, B)$ , is an information theoretic concept for estimating the degree of dependence of the random variables  $A$  and  $B$ , with marginal probability distributions  $p_A(a)$  and  $p_B(b)$ , by measuring the distance between the joint distribution  $p_{AB}(a, b)$  and the distribution associated to the case of complete statistical independence  $p_A(a) \cdot p_B(b)$  [381, 639]:

$$I(A, B) = \sum_{a,b} p_{AB}(a, b) \cdot \log \frac{p_{AB}(a, b)}{p_A(a) \cdot p_B(b)}. \quad (5.8)$$

Let the random variables  $A$  and  $B$  represent the image intensity values  $a$  and  $b$  concerning the pairs of voxels in the two images to be registered, respectively. Estimations for the joint distribution and the marginal distributions can be simply obtained by normalizing the joint and marginal histograms of the overlapping parts of both images. In general, the input images are also smoothed slightly, by means of their histogram. This makes the cost function (i.e., similarity metrics) as smooth as possible to give faster convergence and less chance of being trapped in local minima. The intensities  $a$  and  $b$  are related through the geometric transformation  $\mathbf{T}$ . The MI registration criterion states that the images  $A$  and  $B$  are geometrically aligned by the transformation  $\mathbf{T}$  for which  $I(A, B)$  is maximal. Therefore, the objective of intensity-based registration is to find an estimation of the transformation  $\mathbf{T}$  that best aligns the source image  $A$  against the reference image  $B$ :  $\hat{\mathbf{T}} = \arg \max_{\mathbf{T} \in \Omega} \{I(\mathbf{T}(A), B)\}$ .

The results presented by Maes *et al.* [381] proved that sub-voxel registration differences with respect to the stereotactic reference solution can be obtained for CT/MRI and PET/MRI matching without using any prior knowledge about the gray-value content of both images and the correspondence between them. As a matter of fact, MI is the most intensively investigated criterion for registration of intra-individual human brain images [449]. NMI is also frequently used as the cost function to be optimized [589]. Especially, when misalignment can be large with respect to the imaged FOVs, a criterion invariant to image overlap statistics should be used:

$$I_N(A, B) = \frac{H(A) + H(B)}{H(A, B)}, \quad (5.9)$$

where  $H(\cdot)$  and  $H(\cdot, \cdot)$  are the marginal and the joint entropies, respectively.

**Optimization of the similarity metrics** Intensity-based registration techniques determine the registration transformation  $\mathbf{T}$  by optimizing a certain voxel similarity measure. Unfortunately, parameter spaces for image registration are frequently not so simple. There are often multiple optima within the parameter space, and the registration can fail if the optimization algorithm converges to the wrong optimum. Some of these optima may be very small, caused

either by interpolation artifacts or a local good match between features or intensities [267]. As explained previously, these small optima can often be removed from the parameter space by smoothing the images before the registration.

Local methods, such as Powell's direction set method [483], Nelder-Mead simplex algorithm, conjugate gradient, Levenberg-Marquardt algorithm [50], are usually employed in image registration [382]. Because many similarity metrics (i.e., functions of transformation parameters) are generally irregular and rough, especially in multimodal image registration, local methods are more accurate when the initial orientation is very close to the transformation that yields the best registration [643]. These strategies are also susceptible to premature convergence to local optima, especially for non-smooth functions. An approach to address this issue is to apply multiresolution techniques, whereby images are aligned at increasing resolutions with initial orientations from the previous (lower) resolution registration result. However, these hierarchical methods frequently become trapped in local optima, as the global optimum may not be present in lower resolutions [267]. Global optimization is often required for dealing with the most general and tricky situations. Such global approaches include simulated annealing, tabu search, GAs, and ESs. Efficiency is the primary reason that local techniques are preferred for registration. Efficient global optimization may gain acceptance if a significant improvement in accuracy can be demonstrated.

For intensity-based registration, the problem is even more complicated. The desired optimum when registering images using voxel similarity measures is not often the global optimum, but it is one among the local optima [267]. A solution to this problem is to start the algorithm in the proximity of the correct optimum, which is within the portion of the parameter space in which the algorithm is more likely to converge to the correct optimum than the incorrect global one. In practical terms, this requires that the starting estimate of the registration transformation is reasonably close to the correct solution.

### **Multimodal medical image registration approaches based on PSO**

In this section, multimodal medical image registration approaches that use PSO as searching strategy are described and critically reviewed according to our work in [528]. First, the introduction of an initial orientation term in the PSO formulation is explained. Then, the different PSO modifications are described.

**Introduction of an initial position in the standard PSO formulation** In a large amount of the optimization techniques, such as PSO, standard test functions are employed for benchmark testing and continuous optimization algorithms assessment. In these cases, the initial set of parameters has little importance. However, in many practical applications,  $a$

*priori* knowledge of the characteristics of  $\mathbf{x}^*$  could be available for which  $\mathcal{F}(\mathbf{x}^*)$  is the global optimum [643]. This situation is certainly true in biomedical image registration, since the users of clinical imaging systems are generally skilled physicians. These clinicians can supply a trustworthy indication of the correct orientation, by choosing an accurate initial transformation. Although co-registration is required because of both medical image complexity and human subjectivity or error, registration algorithms can definitely benefit from an accurate initial guess.

This may also occur with PSO, as  $\mathbf{v}_{\max}$  and constriction coefficient  $\chi$  only prevent particle straying from the region of feasible solutions. However, if the particles were arranged according the user's initial orientation, while swarming, they may have a higher probability of discovering a region that contains  $\mathbf{x}^*$ . Briefly, in addition to the swarming effect around the current global best  $\mathbf{g}$  and each particle recollection of its personal best  $\mathbf{b}_i$ , the initial orientation  $\mathbf{x}_{\text{init}}$  can be also introduced into the velocity  $\mathbf{v}_i$  of each particle [643]. The formulation in Eq. (5.4) is modified into Eq. (5.10):

$$\begin{aligned} \mathbf{v}_i(t) = & w(t) \cdot \mathbf{v}_i(t-1) + c_{\text{soc}} \cdot \mathbf{r}_1 \odot (\mathbf{x}_i(t-1) - \mathbf{g}(t-1)) + \\ & + c_{\text{cog}} \cdot \mathbf{r}_2 \odot (\mathbf{x}_i(t-1) - \mathbf{b}_i(t-1)) + c_{\text{ret}} \cdot \mathbf{r}_3 \odot (\mathbf{x}_{\text{init}} - \mathbf{x}_i(t-1)), \end{aligned} \quad (5.10)$$

where  $c_{\text{ret}}$  is the acceleration constant for returning to the initial orientation and  $\mathbf{r}_3$  is a uniformly distributed random number in  $[0, 1]$ . Stochastic and evolutionary global optimization techniques, including PSO, can generally discover the promising region or the “basin of attraction” in the search landscape. However, they typically exhibit slow convergence to the global optimum  $\mathbf{x}^*$ , even though the complexity of similarity metrics computation needs fast convergence. For these reasons, a local method is applied to the best point in the promising region found by the PSO [643]. Powell's direction set algorithm [483] is well-suited for this purpose, because it does not require any derivative computation [382].

**Versions of medical image registration approaches using PSO** Several literature works addressed the issues related to biomedical image registration using PSO. In addition to the “traditional” plain PSO, other versions, including initial orientation knowledge and/or ESs, have been proposed in the literature. These more advanced registration approaches aim at balancing out exploration and exploitation, avoiding premature convergence.

In [643], three main variants of biomedical image registration based on PSO are described and compared:

- (i) hybrid PSO with crossover operators for positions and velocities updating [371]. However, for normal velocity updates, Eq. (5.4) with the constriction factor  $\chi$  is exploited. The convergence criteria are  $T_{\text{NoImprove}} = 20$  iterations in which there is no

improvement in  $\mathcal{F}(\mathbf{x})$ , or reaching the maximum number of iterations  $T_{\max}$ . After the PSO convergence, the Powell's local optimization method [483] is applied to the best point in the swarm;

- (ii) hybrid PSO with crossover and sub-populations, analogous to the previous one except that five sub-populations were initially determined with the  $K$ -Means clustering algorithm [368]. Additionally, after the convergence, the Powell's method is applied to the best points in each sub-population, resulting in the final registration transformation;
- (iii) PSO with constriction coefficient and relaxed convergence criteria is the most "controlled" of the three techniques. A "loose" local optimization through the Powell's method is applied to  $\mathbf{x}_{\text{init}}$ , resulting in  $\mathbf{x}'_{\text{init}}$  around which particles are generated. PSO with the constriction factor  $\chi$  is applied, the velocities are updated according to Eq. (5.10) and the convergence criteria are more relaxed. If during some iteration  $\mathcal{F}(\mathbf{x}_i) < \mathcal{F}(\mathbf{g})$ , then  $\mathbf{g}$  is set to  $\mathbf{x}_i$ , but if  $\|\mathbf{x}_i - \mathbf{g}\| < \varepsilon$  the iteration is still considered to be a non-improving iteration. Convergence is faster, as the iteration counter is not reset to zero for improving points very close to  $\mathbf{g}$ . The Powell's local optimization is applied to  $\mathbf{g}$  after the convergence (a function value change less than 0.005).

Although the constriction coefficient prevents the particles from straying out of the space of feasible solutions, the particles have a greater probability of being drawn out of local optima by the additional term. The authors of [643] argued that the modifications were designed *ad hoc* for image registration and this term improved registration accuracy. In other applications, however, there may be no prior knowledge of the location of the global optimum. In these cases, the last version may prevent particles from moving towards the global optimum, and Eq. (5.4) should be used for the velocity update. If a feasible region, wherein the correct transformation likely lies, cannot be identified, then the other PSO hybridizations (i.e., the use of crossover and sub-populations) are recommended. Some knowledge of the correct orientation might greatly improve the search. Both the initial orientation term and the constriction coefficient prevent the search from straying too far from the global optimum. Hybridization with ES operators appears to improve accuracy by diversifying particle locations. As shown in the last version, convergence criteria during the global search can be relaxed, as local optimization can find the global optimum if a particle is sufficiently close to it.

The authors of [553] investigated four different plain PSO versions for the registration of the images, without integrating other optimization approaches: (i) standard PSO corresponds to the standard PSO introduced in 2007 by Bratton and Kennedy [67]; (ii) standard PSO with variable inertia weight represents a modification of the earlier described standard PSO based



on [643], where the inertia weight  $w(t)$  monotonically decreases during the iterations; (iii) standard PSO with initial orientation, relaxed convergence criteria and constriction coefficient is another alteration presented by Wachowiak *et al.* [643]. This version includes the initial orientation of the volumes to one another, according to Eq. (5.10); (iv) PSO with constriction coefficient, relaxed convergence criteria and variable influence of the initial orientation is a modification of the previous version. Since the involvement of the initial orientation may prevent the convergence of the swarm, the influence of the component  $\mathbf{x}_{\text{init}}$  in Eq. (5.10) should decrease in each iteration:

$$\begin{cases} c_{\text{soc}}(t) = c_{\text{soc}}(t-1) - \Delta\varphi \\ c_{\text{ret}}(t) = c_{\text{ret}}(t-1) - \Delta\varphi \\ \varphi = c_{\text{cog}} + c_{\text{soc}} + c_{\text{ret}} \\ \Delta\varphi = \frac{c_{\text{ret}_{\text{min}}} - c_{\text{ret}_{\text{max}}}}{T_{\text{max}}} \xrightarrow{c_{\text{ret}_{\text{min}}}=0} \Delta\varphi = \frac{c_{\text{ret}_{\text{max}}}}{T_{\text{max}}} \end{cases} \quad (5.11)$$

Chen *et al.* [102] investigated an extension of the PSO to a Hybrid Particle Swarm Optimizer (HPSO), by integrating two methods from the GAs into the standard PSO. The authors argued that classical PSO is convenient for 2D-2D registration, but is less efficient for 3D-3D alignment. They solved this problem by using a hybrid algorithm. The chosen optimization metrics is again MI. An alternative non-linear 2D-2D affine registration technique for MR and CT modality images of human brain slices was presented in [141], using a correlation function as objective function. Both GA and PSO schemes were considered in a multiresolution domain (using the Haar wavelet transform) to decrease the sensitivity of the registration procedure to local maxima and achieve an idea of the initial orientation of the images to be registered.

Finally, in [1], a hybrid approach is proposed, which uses a modified MI by including the spatial image information into the computation of the similarity metrics by means of a linear combination of image pixel intensity and image Gradient Vector Flow (GVF) [672] intensity. The values of the weighing factors are determined using numerical experimentation. By so doing, statistical and spatial information can be integrated.

## Discussion

Regarding possible experimental findings that can be drawn from the state-of-the-art works, an accurate and fair comparison among the different literature approaches is not so straightforward. Indeed, different medical datasets were used in experimental trials, due to the unavailability of public medical benchmarks for intensity-based image registration. In ad-

dition, all the approaches used a small number of images for 2D or 3D registration tests, making the experimental findings much less significant.

Swarm Intelligence techniques have been shown to be effective and powerful in a wide variety of computer science areas. Depending on the problem nature, continuous or discrete search spaces may be properly defined. Accordingly, the different Swarm Intelligence approaches, such as PSO, may provide more efficient solution encoding either in continuous or discrete optimization problems. However, although each optimization technique was first designed for a particular purpose, the majority of evolutionary algorithms were adapted from continuous to discrete search space, and vice-versa. Other Swarm Intelligence techniques, such as the Bat Algorithm (BA) [679] and the ABC algorithm [310], could be useful to improve the PSO performance in biomedical image registration. In this context, the elastic registration [187] is still a challenging problem because of the thousands of parameters to be optimized [326]. However, to the best of my knowledge, no literature work has addressed yet this challenging issue using Swarm Intelligence techniques.

In conclusion, although more accurate and robust comparisons must be performed with other global and local optimization paradigms, PSO achieves encouraging results in biomedical image registration. This approach deserves certainly further study and represents a promising open research topic for multimodal medical image registration.

# Chapter 6

## Deep Neural Networks

### 6.1 Convolutional Neural Networks

In the latest years, DNNs have been exploited to learn a hierarchy of increasingly complex features from the processed data, enabling multiple levels of abstraction [350]. This kind of models allows for raw data processing differently to conventional Machine Learning classifiers that are built on top of hand-engineered features, such as Local Binary Patterns (LBP) [443], Scale-Invariant Feature Transform (SIFT) [372], and Histogram of Oriented Gradients (HOG) [139].

In traditional multi-layer networks, each neuron is densely or fully connected to every neuron of the subsequent layer. However, in Pattern Recognition exploiting only local sub-structural information could be useful. Inspired by the idea of the neurocognitron [195, 196], it was shown that pixels that are close together in the image (e.g., adjacent pixels) tend to be strongly correlated and can represent meaningful features such as edges, while pixels that are far apart in the image tend to be weakly correlated or uncorrelated.

CNNs are deep feed-forward neural network architectures that are suitable to work on data structured with grid-like topology, such as images, videos, and time-series. CNNs are effective in training and generalization much better than common neural networks. As a matter of fact, efficient neural architectures are a hot topic in computer science engineering, also considering hardware devices with limited resources [125].

Many standard feature representations used in Computer Vision problems are based upon local features within the image. In order to reflect this powerful property, a CNN architecture takes advantage of sparse connectivity: the local sub-structure within an image is captured by constraining each neuron to depend only on a spatially local subset of the variables of the previous layer. The set of neurons in the input layer that influences the activation of a neuron is known as the neuron's receptive field. This property needs for less parameters to

learn, so allowing for a more efficient training. In such a practical scenario, leveraging GPUs [168, 570] has been representing an important technological enabling factor in effective training [364, 560].

The main components of a CNN [337, 351] are:

- *convolutional layers*, defining linear convolution operations organized in feature maps. Each unit is connected to local patches in the feature maps of the previous layer by means of a set of weights (i.e., filter bank) allowing for weight-sharing;
- *non-linear activation layers*, applied after each convolutional layer or fully-connected layer, allowing for the learning of multimodal functions. Considering the universal approximation theorem—firstly proved by Cybenko [134] for sigmoidal activation function and then extended by Hornik [272] to an arbitrary bounded and non-constant activation functions—feed-forward neural networks with a single hidden layer of finite size can learn any continuous function on a compact subset of  $\mathbb{R}^n$ . In addition to the traditional non-linear activation layer (e.g., the sigmoid and the hyperbolic tangent), which could give rise to the vanishing gradient phenomenon [48]. Thereby, relying on a biological basis [239], piecewise linear activation functions—such as the Rectified Linear Unit (ReLU) [430]—have been introduced in DNNs. ReLU efficiently allows for scale-invariance and sparse activation. ReLU variants are ReLU the Leaky ReLU [378], and the Parametric ReLU (PReLU) [257];
- *pooling layers*, denoting an operation that replaces the output of the net at a certain location with a summary statistics of the neighborhood's outputs [62]. Examples are max pooling and average pooling.

The convolutional and pooling layers in CNNs are directly inspired by the classic notions of simple cells and complex cells in visual neuroscience [283]. Deep neural networks exploit the property that many natural signals are compositional hierarchies (i.e., higher-level features are obtained by composing lower-level ones).

In supervised learning, the training phase relies on the minimization of an objective function, also called loss function  $\mathcal{L}(\cdot)$ . As long as the modules are relatively smooth functions of their inputs and of their internal weights, the gradients can be computed using the backpropagation algorithm [515]. In this context, the backpropagation procedure aims at computing the gradient of an objective function with respect to the weights of a multi-layer architecture. This algorithm applies the chain rule for derivatives. Regarding the difficulties encountered during DNN training, Haeffele and Vidal [236] provided the conditions to guarantee that local minima are globally optimal and that a local descent strategy can reach

a global minima from any initialization. In particular, these conditions require both the network output and the regularization to be positively homogeneous functions of the network parameters, with the regularization being designed to control the network size. Multi-layer architectures can be trained by the Stochastic Gradient Descent (SGD) [61], especially when the optimization problem is convex or pseudo-convex [277]. Alternatives are adaptive methods, such as AdaGrad [157], ADADELTA [697], and Adam [322].

Another novelty in DNN training is the dropout technique, which randomly drop neurons from the neural network during training and, thus, prevents the co-adaptation of features [581]. Initially, dropout was proposed as an implicit regularization method, in [34, 88] it was demonstrated the equivalence between dropout and a fully deterministic model for Matrix Factorization in which the factors are regularized by the sum of the product of squared Euclidean norms of the columns.

However, in medical imaging it is also important to achieve high classification rates on potentially perturbed test data [364]. The analysis of state-of-the-art deep classifiers' robustness to perturbation at test time is therefore an important step for validating the models' reliability to unexpected (possibly adversarial) nuisances that might occur when deployed in uncontrolled environments [180], such as clinical scenarios.

Considering the most recent computational methods in medical image segmentation, along with traditional Pattern Recognition techniques [523], significant advances have been proposed in CNN-based architectures. For instance, to overcome the limitations related to accurate image annotations, DeepCut [491] relies on weak bounding box labeling [511, 521]. Moreover, Wang *et al.* [647] proposed an interactive segmentation framework by incorporating CNNs into a bounding box and scribble-based segmentation pipeline. This method aims at learning features for a CNN-based classifier from bounding box annotations.

Among the architectures devised for biomedical image segmentation [255, 305], U-Net [506] showed to be a noticeably successful solution, thanks to the combination of a contracting path, for coarse-grained context detection, and a symmetric expanding path, for fine-grained localization. This fully CNN is capable of stable training with reduced samples. The authors of V-Net [418] extended U-Net for volumetric medical image segmentation, by introducing also a different loss function based on the *DSC*. Oktay *et al.* [444] presented an Attention Gate (AG) model for medical imaging, which aims at focusing on target structures or organs. AGs were introduced into the standard U-Net, so defining Attention U-Net, which achieved high performance in multi-class image segmentation without relying on multi-stage cascaded CNNs. Recently MS-D Net [460] was shown to yield better segmentation results in biomedical images than U-Net [506] and SegNet [30], by creating dense connections among features at different scales obtained by means of dilated convolutions. By so doing,

features at different scales can be contextually extracted using fewer parameters than full CNNs. Finally, also image-to-image translation approaches—e.g., pix2pix [289], leveraging conditional adversarial neural networks—were exploited for image segmentation.

### **6.1.1 Cross-dataset generalization abilities of CNNs: Prostate zonal segmentation of multi-centric MRI datasets**

Even though CNNs achieved state-of-the-art performance for automatic medical image segmentation, they have not yet shown sufficiently accurate and reliable results for clinical use [647]. As a matter of fact, CNN-based architectures are limited by the lack of image-specific adaptation and the lack of generalizability to previously unseen object classes.

In the specific case of prostate MR image analysis, MRI plays a decisive role in PCa diagnosis and disease monitoring (even in an advanced status [450]), revealing the internal prostatic anatomy, prostatic margins, and PCa extent [635]. According to the zonal compartment system proposed by McNeal, the prostate WG can be partitioned into the CG and PZ [554]. In prostate imaging, T2w MRI serves as the principal sequence [547], thanks to its high resolution that enables to differentiate the hyper-intense PZ and hypo-intense CG in young male subjects [269].

Considering prostate zonal segmentation, an improved PCa diagnosis requires a reliable and automatic zonal segmentation method, since manual delineation is time-consuming and operator-dependent [428, 519]. Moreover, in clinical practice, the generalization ability among multi-institutional prostate MRI datasets is an unresolved issue due to large anatomical inter-subject variability and the lack of a standardized pixel intensity representation for MRI (such as in the case of CT-based radiodensity measurements expressed in HUs) [327]. Hence, we aim at automatically segmenting the prostate zones on three multi-institutional T2w MRI datasets to evaluate the generalization ability of CNN-based architectures. This task is challenging because images coming from multi-institutional datasets are characterized by different contrasts, visual consistencies, and heterogeneous characteristics [629].

#### **Multi-institutional MRI datasets**

We segmented the CG and PZ from the WG on three completely different multi-parametric prostate MRI datasets, namely:

- #1 dataset (21 patients/193 MR slices with prostate), acquired with a whole body Philips Achieva 3.0 T MRI scanner at the Cannizzaro Hospital (Catania, Italy) [519]. MRI parameters: matrix size =  $288 \times 288$  pixels; slice thickness = 3.0 mm; inter-slice

spacing = 4 mm; pixel spacing = 0.625 mm; number of slices per image series (including slices without prostate) = 18. Average patient age:  $65.57 \pm 6.42$  years;

#2 Initiative for Collaborative Computer Vision Benchmarking (I2CVB) dataset (19 patients/503 MR slices with prostate), acquired with a whole body Siemens TIM 3.0 T MRI scanner at the Hospital Center Regional University of Dijon-Bourgogne (Dijon, France) [355]. MRI parameters: matrix size  $\in \{308 \times 384, 336 \times 448, 360 \times 448, 368 \times 448\}$  pixels; slice thickness = 1.25 mm; inter-slice spacing = 1.0 mm; pixel spacing  $\in \{0.676, 0.721, 0.881, 0.789\}$  mm; number of slices per image series = 64. Average patient age:  $64.36 \pm 9.69$  years;

#3 National Cancer Institute – International Symposium on Biomedical Imaging (NCI-ISBI) 2013 Automated Segmentation of Prostate Structures Challenge dataset (40 patients/555 MR slices with prostate) *via* The Cancer Imaging Archive (TCIA) [485], acquired with a whole body Siemens TIM 3.0 T MRI scanner at Radboud University Medical Center (Nijmegen, The Netherlands) [59]. The prostate structures were manually delineated by five experts. MRI parameters: matrix size  $\in \{256 \times 256, 320 \times 320, 384 \times 384\}$  pixels; slice thickness  $\in \{3.0, 4.0\}$  mm; inter-slice spacing  $\in \{3.6, 4.0\}$  mm; pixel spacing  $\in \{0.500, 0.600, 0.625\}$  mm; number of slices per image series ranging from 15 to 24. Average patient age:  $63.90 \pm 7.17$  years.

All the analyzed MR images are encoded in the 16-bit DICOM format. It is worth noting that even MR images from the same dataset have intra-dataset variations (such as the matrix size, slice thickness, and number of slices). Furthermore, inter-rater variability for the CG and PZ annotations exists, as different physicians delineated them. For clinical feasibility [269], we analyzed only axial T2w MR slices—the most commonly used sequence for prostate zonal segmentation—among the available sequences. In our multi-centric study, we conducted the following seven experiments resulting from all possible training/testing conditions:

- Individual dataset #1, #2, #3: training and testing on dataset #1 (#2, #3, respectively) alone in 4-fold cross-validation, and testing also on whole datasets #2 and #3 (#1 and #3, #1 and #2, respectively) separately for each round;
- Mixed dataset #1/#2, #2/#3, #1/#3: training and testing on both datasets #1 and #2 (#2 and #3, #1 and #3, respectively) in 4-fold cross-validation, and testing also on whole dataset #3 (#1, #2, respectively) separately for each round;
- Mixed dataset #1/#2/#3: training and testing on whole datasets #1, #2, and #3 in 4-fold cross-validation.

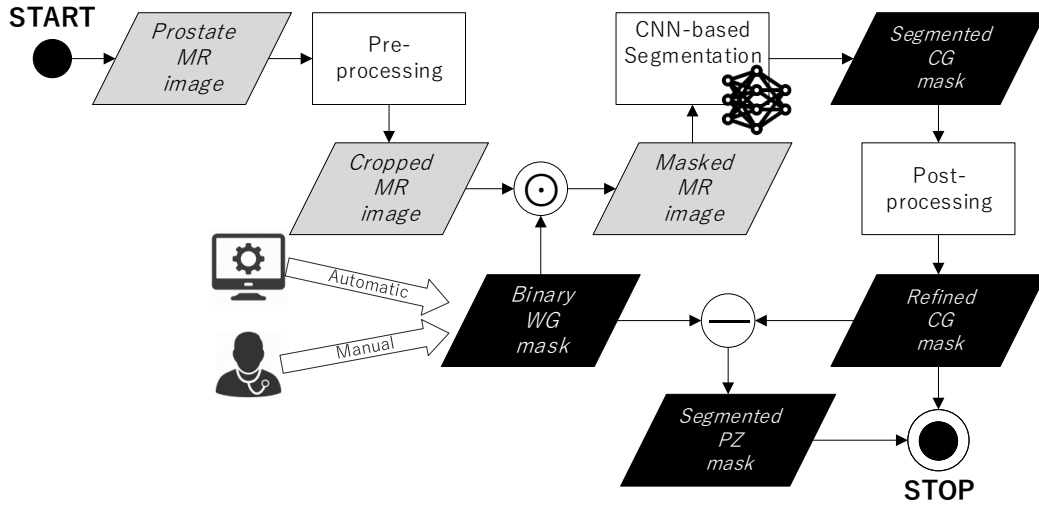


Fig. 6.1 Workflow of the overall CNN-based prostate zonal segmentation approach. The gray and black data blocks denote gray-scale images and binary masks, respectively.

For the 4-fold cross-validation, we partitioned the datasets #1, #2, and #3 into 4 folds by using the following patient indices:  $\{[1, \dots, 5], [6, \dots, 10], [11, \dots, 15], [16, \dots, 21]\}$ ,  $\{[1, \dots, 5], [6, \dots, 10], [11, \dots, 15], [16, \dots, 19]\}$ , and  $\{[1, \dots, 10], [11, \dots, 20], [21, \dots, 30], [31, \dots, 40]\}$ , respectively. Finally, the results from the different cross-validation rounds were averaged to obtain a final descriptive value. We chose a 4-fold cross-validation scheme to conduct reliable and fair training/testing phases, according to the number of patients provided by each dataset, as well as to calculate the evaluation metrics on a statistically significant test set. Our experimental settings were designed to evaluate the CNN's generalization abilities among different MRI acquisition options, such as different devices and functioning parameters. For instance, both intra-/inter-scanner evaluations can be carried out, as dataset #1's scanner is different from those of datasets #2 and #3.

Our study adopts a selective delineation approach to focus on internal prostatic anatomy: the CG and PZ, denoted by  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$ , respectively. Let the entire image and the WG region be  $\mathcal{I}_{\Omega}$  and  $\mathcal{R}_{WG}$ , respectively, the following relationships can be defined:

$$\mathcal{I}_{\Omega} = \mathcal{R}_{WG} \cup \mathcal{R}_{bg} \text{ and } \mathcal{R}_{WG} \cap \mathcal{R}_{bg} = \emptyset, \quad (6.1)$$

where  $\mathcal{R}_{bg}$  represents background pixels. Relying on [487, 635],  $\mathcal{R}_{PZ}$  was obtained by subtracting  $\mathcal{R}_{CG}$  from  $\mathcal{R}_{WG}$  meeting the constraints:

$$\mathcal{R}_{WG} = \mathcal{R}_{CG} \cup \mathcal{R}_{PZ} \text{ and } \mathcal{R}_{CG} \cap \mathcal{R}_{PZ} = \emptyset. \quad (6.2)$$

The overall prostate zonal segmentation method is outlined in Fig. 6.1.



### Quantitative comparison of the state-of-the-art CNN-based architectures

In the preliminary study in [517], we automatically segmented the CG and PZ using Deep Learning to evaluate the generalization ability of CNNs on two different MRI prostate datasets—namely, the datasets #1 and #2—by means of a 4-fold cross-validation scheme. However, this is challenging since multi-centric datasets are generally characterized by different contrast, visual consistencies, and image characteristics. Therefore, prostate zones on T2w MR images were manually annotated for supervised learning, and then automatically segmented using a mixed scheme by (i) training on either each individual dataset or both datasets and (ii) testing on both datasets, using CNN-based architectures: SegNet [30], U-Net [506], and pix2pix [289]. In such a context, we compared the segmentation performances achieved with/without pre-training [602].

**Pre-processing** To fit the image resolution of the dataset #1, we center-cropped the images of the dataset #2 and resized them to  $288 \times 288$  pixels. Furthermore, the images of these datasets were masked using the corresponding prostate binary masks to omit the background and only focus on extracting the CG and PZ from the WG. This operation can be performed either by an automated method [519] or previously provided manual WG segmentation [355]. For better training, we randomly cropped the input images from  $288 \times 288$  to  $256 \times 256$  pixels and horizontally flipped them.

**Investigated CNN-based Architectures** During the training of all architectures, the  $\mathcal{L}_{DSC}$  loss function (i.e., a continuous version of *DSC*) was used [418]:

$$\mathcal{L}_{DSC} = -\frac{2 \sum_{i=1}^N s_i \cdot r_i}{\sum_{i=1}^N s_i + \sum_{i=1}^N r_i}, \quad (6.3)$$

where  $s_i$  and  $r_i$  represent the continuous values of the prediction map (i.e., the result of the final layer of the CNN) and the ground truth at the  $i$ -th pixel ( $N$  is the total number of pixels to be classified).

**SegNet** is a CNN architecture for semantic pixel-wise segmentation [30]. It consists of an encoder/decoder network followed by a pixel-wise classification layer. Since our classification task involves only one class, the soft-max operation and ReLU activation function at the final layer were removed for stable training.

We implemented SegNet using PyTorch. During training, we used the SGD [61] with a learning rate of 0.01, momentum of 0.9, weight decay of  $5 \times 10^{-4}$ , and batch size of 8. It

was trained for 50 epochs and the learning rate was multiplied by 0.2 at the 20-th and 40-th epochs.

**U-Net** is a fully CNN capable of stable training with a reduced number of samples [506], combining pooling operators with up-sampling operations. The general architecture is an encoder-decoder with skip connections between mirrored layers in the encoder/decoder stacks. By so doing, high resolution features from the contracting path are combined with the up-sampled output for better localization. We utilized four scaling operations.

U-Net was implemented using Keras on top of TensorFlow. We used SGD with a learning rate of 0.01, momentum of 0.9, weight decay of  $5 \times 10^{-4}$ , and batch size of 4. Training was executed for 50 epochs, multiplying the learning rate by 0.2 at the 20-th, and 40-th epochs.

**pix2pix** is an image-to-image translation method with conditional adversarial networks [289]. As a generator, U-Net is used to translate the original image into the segmented one [506], preserving the highest level of abstraction. The generator and discriminator include 8 and 5 scaling operations, respectively.

We implemented pix2pix on PyTorch. Adam was used as an optimizer with a learning rate of 0.0002 and 0.01 for the discriminator and generator, respectively. The learning rate for generator was multiplied by 0.1 every 20 epochs. It was trained for 50 epochs with a batch size of 12.

**Post-processing** Two simple morphological steps were applied on the obtained CG binary masks to smooth boundaries and avoid disconnected regions:

- a hole filling algorithm on the segmented  $\mathcal{R}_{CG}$  to remove possible holes in the predicted map;
- a small area removal operation to delete connected components with area less than  $\lfloor |\mathcal{R}_{WG}|/8 \rfloor$  pixels, where  $|\mathcal{R}_{WG}|$  denotes the number of the pixels contained in the WG segmentation. This criterion effectively adapts according to the different dimensions of the  $\mathcal{R}_{WG}$  (ranging from the apical to the basal prostate slices).

**Influence of pre-training** In medical imaging, due to the lack of training data, ensuring CNN's proper training convergence is difficult from scratch. Therefore, pre-training the models on a different application and then fine-tuning them is common [602].

To evaluate cross-dataset generalization ability *via* pre-training, we compared the performances of the three CNN-based architectures with/without pre-training on a similar

application. We used a relatively large dataset of 50 manually segmented examples from the PROMISE12 challenge [365]. Since this competition focuses only on WG segmentation without providing prostate zonal labeling, we pre-trained the architectures on WG segmentation. To adjust this dataset to our experimental setup, the images of this dataset were resized from  $512 \times 512$  to  $288 \times 288$  pixels and randomly cropped to  $256 \times 256$  pixels; because our task only focuses on slices with prostate, we also omitted initial/final slices without prostate, so the number of slices for each sample was fixed to 25.

**Results** This section explains how the three CNN-based architectures segmented the prostate zones, evaluating their cross-dataset generalization ability.

Table 6.1 shows the 4-fold cross-validation results obtained in the different experimental conditions. When training and testing are both performed on the dataset #1, U-Net outperforms the other architectures on both CG and PZ segmentation; however, it experiences problems with testing on the dataset #2 due to the limited number of training images in the dataset #1. In such a case, pix2pix generalizes better thanks to its internal generative model. When trained on the dataset #2 alone, U-Net yields the most accurate results both in intra- and cross-dataset testing. This probably derives from the dataset #2's relatively larger training data as well as U-Net's good generalization ability when sufficient data are available. Moreover, SegNet reveals rather unstable results, especially when trained on a limited amount of data. Finally, when trained on the mixed dataset, all three architectures—especially U-Net—achieve good results on both datasets without losing accuracy compared to training on the same dataset alone. Therefore, using mixed MRI datasets during training can considerably improve the performance in cross-dataset generalization towards other clinical applications. Comparing the CG and PZ segmentation, when tested on the dataset #1, the results on the PZ are generally more accurate, except when trained on the dataset #2 alone; however, for the dataset #2, segmentations on the CG are generally more accurate.

Fine-tuning after pre-training sometimes leads to slightly better results than training from scratch, when trained only on a single dataset. However, its influence is generally negligible or rather negative, when trained on the mixed dataset. This modest impact is probably due to the ineffective data size for pre-training. For a visual assessment, two examples (one for each dataset) are shown in Fig. 6.2. Relying on the gold standards in Figs. 6.2(d, h), it can be seen that U-Net generally achieves more accurate results compared with SegNet and pix2pix. This finding confirms the trend revealed by the *DSC* values in Table 6.1.

Table 6.1 Prostate zonal segmentation results of the three CNN-based architectures in 4-fold cross-validation assessed by the *DSC* metrics (presented as the average and standard deviation). The experimental results are calculated on the different setups of (i) training on either each individual dataset or both datasets and (ii) testing on both datasets. Numbers in bold indicate the highest *DSC* values for each prostate region (i.e., CG and PZ) among all architectures with/without pre-training (PT).

	Architecture	Zone	Testing on Dataset #1		Testing on Dataset #2	
			Average	Std. Dev.	Average	Std. Dev.
Training on Dataset #1	SegNet (w/o PT)	CG	80.20	3.28	74.48	5.82
		PZ	80.66	11.51	59.57	12.68
	SegNet (w/ PT)	CG	83.38	3.22	72.75	2.80
		PZ	87.39	3.90	66.20	5.64
	U-Net (w/o PT)	CG	84.33	2.37	74.18	3.77
		PZ	88.98	2.98	66.63	1.93
	U-Net (w/ PT)	CG	<b>86.88</b>	1.60	70.11	5.31
		PZ	<b>90.38</b>	3.38	58.89	7.06
	pix2pix (w/o PT)	CG	82.35	2.09	<b>76.61</b>	2.17
		PZ	87.09	2.72	73.20	2.62
pix2pix (w/ PT)	CG	80.38	2.81	76.19	5.77	
	PZ	83.53	5.65	<b>73.73</b>	2.40	
Training on Dataset #2	SegNet (w/o PT)	CG	76.04	2.05	87.07	2.41
		PZ	77.25	3.09	82.45	1.77
	SegNet (w/ PT)	CG	77.99	2.15	87.75	2.83
		PZ	76.51	2.70	82.26	2.09
	U-Net (w/o PT)	CG	78.88	0.88	88.21	2.10
		PZ	74.52	1.85	<b>83.03</b>	2.46
	U-Net (w/ PT)	CG	<b>79.82</b>	1.11	<b>88.66</b>	2.28
		PZ	<b>74.56</b>	5.12	82.48	2.47
	pix2pix (w/o PT)	CG	77.90	0.73	86.95	2.93
		PZ	66.09	3.07	81.33	0.90
pix2pix (w/ PT)	CG	77.21	1.02	85.94	4.31	
	PZ	67.39	5.04	80.07	0.84	
Training on Mixed Dataset	SegNet (w/o PT)	CG	84.28	3.12	87.92	2.80
		PZ	87.74	1.66	82.21	0.79
	SegNet (w/ PT)	CG	86.08	1.92	87.78	2.75
		PZ	89.53	3.28	82.39	1.50
	U-Net (w/o PT)	CG	<b>86.34</b>	2.10	<b>88.12</b>	2.34
		PZ	90.74	2.40	<b>83.04</b>	2.30
	U-Net (w/ PT)	CG	85.82	1.98	87.42	1.89
		PZ	<b>91.44</b>	2.15	82.17	2.11
	pix2pix (w/o PT)	CG	83.07	3.39	86.39	3.16
		PZ	83.53	2.36	80.40	1.80
pix2pix (w/ PT)	CG	82.08	4.37	85.96	5.40	
	PZ	83.04	4.20	80.60	1.49	

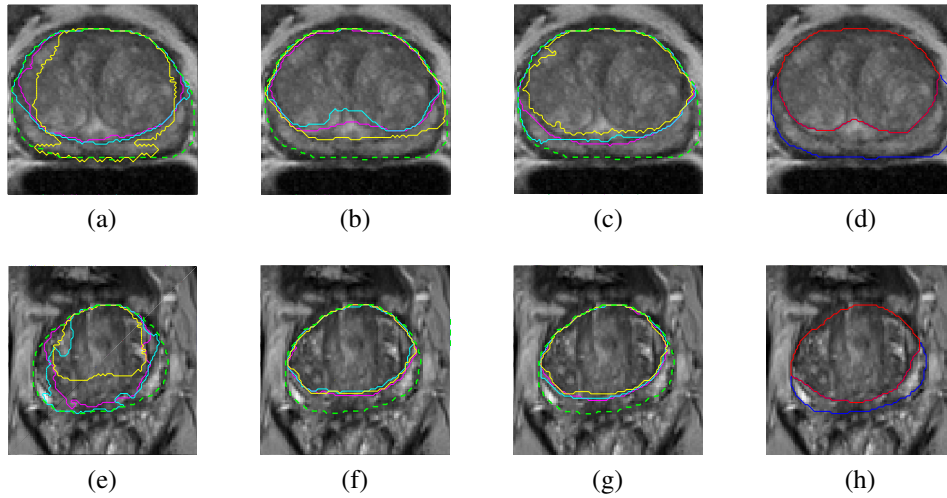


Fig. 6.2 Examples of prostate zonal segmentation in pre-training/fine-tuning. The first row concerns testing on dataset #1, trained on: (a) dataset #1; (b) dataset #2; (c) mixed dataset. The second row concerns testing on dataset #2, trained on: (e) dataset #1; (f) dataset #2; (g) mixed dataset. The  $\mathcal{R}_{CG}$  segmentation results are represented with magenta, cyan, and yellow solid contours for SegNet, U-Net, and pix2pix, respectively. The dashed green line denotes the  $\mathcal{R}_{WG}$  boundary. The last column (sub-figures (d, h)) shows the gold standard for  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$  with red and blue lines, respectively. The images are zoomed with a  $4\times$  factor.

**Discussion and conclusions** Our preliminary results show that CNN-based architectures can segment prostate zones on two different MRI datasets to some extent, leading to valuable clinical insights; CNNs suffer when training and testing are performed on different MRI datasets acquired by different devices and protocols, but this can be mitigated by training the CNNs on multiple datasets, even without pre-training. Generally, considering different experimental training and testing conditions, U-Net outperforms SegNet and pix2pix thanks to its good generalization ability. Furthermore, this study suggests that significant performance improvement *via* fine-tuning may require a remarkably large dataset for pre-training.

As future developments, we plan to improve the results by refining the predicted binary masks for better smoothness and continuity, avoiding disconnected segments; furthermore, we should enhance the output delineations considering the three-dimensional spatial information among slices.

### USE-Net: Incorporating SE blocks into U-Net

In this research context, we proposed a novel CNN, called USE-Net [516], which incorporates Squeeze-and-Excitation (SE) blocks [280] into U-Net after every Encoder (Enc USE-Net) or Encoder-Decoder block (Enc-Dec USE-Net). The rationale behind the design of USE-

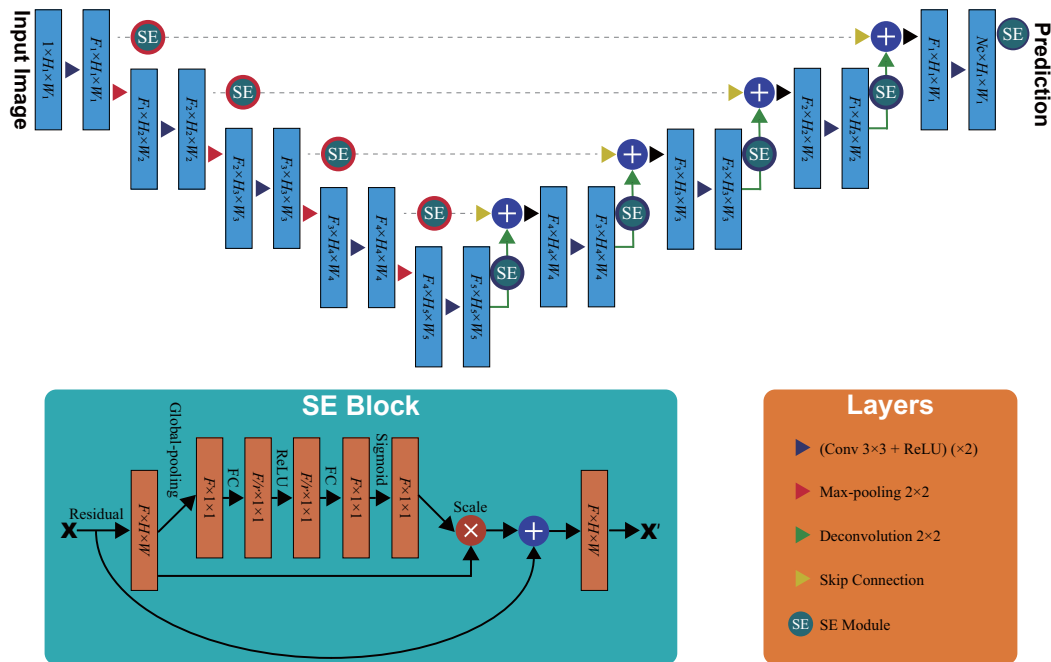


Fig. 6.3 Scheme of the proposed USE-Net architecture: Enc USE-Net has only 4 (red-contoured) SE blocks after every encoder, whilst Enc-Dec USE-Net has 9 SE blocks integrated after every encoder/decoder (represented with red/blue contours, respectively).

Net is to exploit adaptive channel-wise feature recalibration to boost the generalization performance. Considering the three datasets above mentioned, this study adopted a mixed scheme for cross- and intra-dataset generalization: (i) training on either each individual dataset or pair of datasets, and (ii) testing on all three datasets with all possible training/testing combinations. To the best of our knowledge, we tackle for the first time CNN-based prostate zonal segmentation on T2w MRI alone. We compared USE-Net against three CNN-based architectures using both spatial overlap-/distance-based metrics: U-Net, pix2pix, and Mixed-Scale Dense Network (MS-D Net) [460], along with a semi-automatic continuous max-flow model [487].

**Contributions** Our main contributions are:

- **Prostate zonal segmentation:** our novel Enc-Dec USE-Net achieves accurate CG and PZ segmentation results on T2w MR images, remarkably outperforming the other competitor methods when trained on all datasets used for testing in multi-institutional scenarios.

- **Cross-dataset generalization:** this first cross-dataset study among three medical datasets shows that training on the union of multi-institutional datasets generally outperforms training on each dataset during testing, realizing both intra-/cross-dataset generalization—thus, we may train CNNs by feeding samples from multiple different datasets for improving the performance.
- **Deep Learning for medical imaging:** this research reveals that SE blocks provide excellent intra-dataset generalization in multi-institutional scenarios, when the testing is performed on samples from the datasets used during the training. Therefore, adaptive mechanisms (e.g., feature recalibration in CNNs) may be a valuable solution in medical imaging applications involving multi-institutional settings.

We propose to introduce SE blocks [280] following every Encoder (Enc USE-Net) or Encoder-Decoder (Enc-Dec USE-Net) of U-Net [506], as shown in Fig. 6.3. As pointed out before, U-Net allows for a multi-resolution decomposition/composition technique [597], by combining encoders/decoders with skip connections between them [682]; in our implementation, encoders and decoders consist of four pooling operators that capture the context and up-sampling operators that conduct precise localization, respectively.

We introduce SE blocks to enhance image segmentation, expecting an increased representational power from modeling the channel-wise dependencies of convolutional features [280]. These blocks were originally envisioned for image classification using adaptive feature recalibration to boost informative features and suppress weak ones at minimal computational burden. Enc USE-Net and Enc-Dec USE-Net are investigated to evaluate the effect of strengthened feature recalibration.

SE blocks can be formally described as follows:

**Squeeze** Let  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_F]$  be an input feature map, where  $\mathbf{u}_f \in \mathbb{R}^{H \times W}$  is a single channel with size  $H \times W$ . Through spatial dimensions  $H \times W$ , a global average pooling layer generates channel-wise statistics  $\mathbf{z} \in \mathbb{R}^F$ , whose  $f$ -th element is given by:

$$z_f = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W [\mathbf{u}_f]_{i,j}. \quad (6.4)$$

**Excitation** To limit model complexity and boost generalization, two fully-connected layers and the ReLU [430] function  $\delta$  transform  $\mathbf{z}$  with a sigmoid activation function  $\sigma(\cdot)$ :

$$\mathbf{s} = \sigma(g(\mathbf{z}, \mathbf{W})) = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{z})), \quad (6.5)$$

where  $\mathbf{W}_1 \in \mathbb{R}^{\frac{F}{r} \times F}$ ,  $\mathbf{W}_2 \in \mathbb{R}^{F \times \frac{F}{r}}$ , and  $r$  is the reduction ratio. In order to obtain an adaptive recalibration that ignores less important channels and emphasizes important ones (allowing for non-mutual exclusivity among multiple channels, differently from one-hot encoding),  $\mathbf{U}$  is rescaled into  $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_F]$  by applying Eq. (6.6):

$$\tilde{\mathbf{x}}_f = \mathbf{F}_{\text{scale}}(\mathbf{u}_f, s_f) = s_f \cdot \mathbf{u}_f, \text{ for } f = 1, 2, \dots, F, \quad (6.6)$$

where  $\mathbf{F}_{\text{scale}}(\mathbf{u}_f, s_f)$  represents the channel-wise multiplication between the feature map  $\mathbf{u}_f \in \mathbb{R}^{H \times W}$  and the scalar  $s_f \in [0, 1]$ .

**Pre-processing** To fit the image resolution of dataset #1, we either center-cropped or zero-padded the images of datasets #2 and #3 to resize them to  $288 \times 288$  pixels. Afterwards, all images in the three datasets were masked using the corresponding prostate binary masks to omit the background and only focus on extracting the CG and PZ from the WG. This operation can be performed either by an automated method [519] or previously provided manual WG segmentation [355]. As a simple form of Data Augmentation (DA), we randomly cropped the input images from  $288 \times 288$  to  $256 \times 256$  pixels and horizontally flipped them.

**Post-processing** Two efficient morphological operations were applied on the obtained  $\mathcal{R}_{CG}$  binary masks to smooth boundaries and deal with disconnected regions:

- a hole filling algorithm on the segmented  $\mathcal{R}_{CG}$  to remove possible holes in a predicted map;
- a small area removal operation dealing with connected components smaller than  $\lfloor |\mathcal{R}_{WG}|/8 \rfloor$  pixels, where  $|\mathcal{R}_{WG}|$  denotes the number of pixels contained in WG segmentation. This adaptive criterion takes into account the different sizes of  $\mathcal{R}_{WG}$  (ranging from the apical to the basal prostate slices).

**Comparison against the state-of-the-art methods** After our preliminary study in [517], we discarded SegNet [30] due to the poor performance in this particular task. Therefore, we compared USE-Net against three supervised CNN-based architectures (i.e., U-Net [506], pix2pix [289], and MS-D Net [460]) and the unsupervised continuous max-flow model proposed in [487]. Therefore, in this comparative analysis we introduced also a very recent CNN (i.e., MS-D Net [460]) and a very powerful state-of-the-art method tailored for prostate zonal segmentation [487]. All the investigated CNN-based architectures were trained using the  $\mathcal{L}_{DSC}$  loss function (i.e., a continuous version of the  $DSC$ ) [418] in Eq. (6.3).



**USE-Net and U-Net** Using four scaling operations, U-Net and USE-Net were implemented on Keras with TensorFlow backend. We used the SGD method [61] with a learning rate of 0.01, momentum of 0.9, weight decay of  $5 \times 10^{-4}$ , and batch size of 4. Training was executed for 50 epochs, multiplying the learning rate by 0.2 at the 20-th and 40-th epochs.

**pix2pix** This image-to-image translation method with conditional adversarial networks was used to translate the original image into the segmented one [289]. The generator and discriminator (both U-Nets in our implementation) include eight and five scaling operations, respectively. We developed pix2pix on PyTorch. Adam [322] was used as an optimizer with a learning rate of 0.01 for the generator—which was multiplied by 0.1 every 20 epochs—and  $2 \times 10^{-4}$  for the discriminator. Training was executed for 50 epochs with a batch size of 12.

**MS-D Net** This dilated convolution-based method, characterized by densely connected feature maps, is designed to capture features at various image scales [460]. It was implemented on PyTorch with a depth of 100 and width of 1. We used Adam [322] with a learning rate of  $1 \times 10^{-3}$  and trained it for 100 epochs with a batch size of 12.

**Continuous max-flow model** This model [487] exploits duality-based convex relaxed optimization [688] to achieve better numerical stability (i.e., convergence) than classic graph cut-based methods [192]. This semi-automatic approach simultaneously segments both  $\mathcal{R}_{WG}$  and  $\mathcal{R}_{CG}$  under the constraints given in Eq. (6.2), relying on user intervention. The initialization procedure consists in two closed surfaces defined by a thin-plate spline interpolating 10-12 control points interactively selected by the user (considering both the axial and sagittal views). These 3D partitions estimate the intensity probability density functions associated with three sub-regions of background, CG, and PZ. This allows for defining the region appearance models for global optimization-based multi-region segmentation [688].

Since the supervised CNN-based architectures rely on the gold standard  $\mathcal{R}_{WG}$  for zonal segmentation, we apply the continuous max-flow method on CG for single-region segmentation for fair comparison. Moreover, in our tests, a very accurate slice-by-slice  $\mathcal{R}_{CG}$  initialization is provided by eroding the gold standard CG with a circular structuring element (radius: 6 pixels).

**Experimental results** This section shows how the CNN-based architectures and the continuous max-flow model segmented the prostate zones, through the evaluation of their cross-dataset generalization ability.

Table 6.2 shows the 4-fold cross-validation results, as assessed by the *DSC* metrics, obtained under different training/testing conditions. For visual comparison, the Kiviat diagrams [149] for each CNN-based architecture are also displayed in Fig. 6.4. Here, we can observe the impact of leaving dataset #3 out of the training set and, at the same time, using it as test set: the corresponding spokes III, VI, and XII generally show lower performance, probably due to the peculiar image characteristics of dataset #3 (comprising the highest number of patients) that are not learned during the training phase on datasets #1/#2. In general, Enc USE-Net performs similarly to U-Net, which stably yields satisfactory results. More interestingly, Enc USE-Net obtains considerably better results when trained/tested on multiple datasets. Enc-Dec USE-Net (characterized by a higher number of SE blocks with respect to Enc USE-Net) consistently and remarkably outperforms the other methods on both CG and PZ segmentation when trained on all the investigated datasets, also performing well when trained and tested on the same datasets.

In order to visualize the achieved results, example images segmented by each method are compared in Fig. 6.5 under the three-dataset training condition. The critical difference diagram (Fig. 6.6) using the Bonferroni-Dunn’s *post hoc* test also confirms this trend, considering *DSC* values for every round of the 4-fold cross-validation.

However, as shown in Fig. 6.7, Enc-Dec USE-Net shows less powerful cross-dataset generalization when trained and tested on different datasets, achieving slightly lower average performance than Enc USE-Net (considering all training/testing combinations). This implies that the SE blocks’ adaptive feature recalibration—boosting informative features and suppressing weak ones—provides excellent intra-dataset generalization when testing is performed on multiple datasets used during training (i.e., when training samples from every testing dataset are fed to the model).

On the contrary, pix2pix achieves good generalization when trained and tested on different datasets, especially under mixed-dataset training conditions, thanks to its internal generative model. MS-D Net generally works better in single dataset scenarios, using a limited amount of training samples, according to [460]. The unsupervised continuous max-flow model achieves comparable results to the supervised ones only when trained and tested on different datasets. However, this semi-automatic approach is outperformed by the supervised methods when trained and tested on the same datasets, as it underestimates  $\mathcal{R}_{CG}$ .

The results also reveal that training on multi-institutional datasets generally outperforms training on each dataset during testing on any dataset/zone, realizing both intra-/cross-dataset generalization. For instance, training on datasets #1 and #2 generally outperforms training on dataset #1 during testing on all datasets #1, #2, and #3, without losing accuracy.

Table 6.2 Prostate zonal segmentation results of the CNN-based architectures and the unsupervised continuous max-flow model (proposed by Qiu *et al.* [487]) in 4-fold cross-validation assessed by *DSC* (presented as the mean value  $\pm$  standard deviation). The supervised experimental results are calculated under the different seven conditions described in Section 6.1.1. Numbers in bold indicate the best *DSC* values (the higher the better) for each prostate region (i.e.,  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$ ) among all architectures.

	Method	Testing on dataset #1		Testing on dataset #2		Testing on dataset #3	
		CG	PZ	CG	PZ	CG	PZ
Training on dataset #1	MS-D Net	<b>84.3</b> $\pm$ 1.6	<b>86.7</b> $\pm$ 1.6	77.3 $\pm$ 4.1	65.6 $\pm$ 11.2	66.2 $\pm$ 3.7	<b>50.9</b> $\pm$ 1.2
	pix2pix	81.9 $\pm$ 2.2	85.9 $\pm$ 5.0	77.2 $\pm$ 2.9	73.7 $\pm$ 4.3	52.5 $\pm$ 3.2	47.1 $\pm$ 1.3
	U-Net	78.6 $\pm$ 4.1	78.3 $\pm$ 7.6	77.4 $\pm$ 5.4	<b>75.3</b> $\pm$ 1.4	73.6 $\pm$ 6.2	50.9 $\pm$ 1.5
	Enc USE-Net	79.3 $\pm$ 3.5	77.7 $\pm$ 2.7	<b>81.3</b> $\pm$ 1.5	74.7 $\pm$ 1.8	<b>75.0</b> $\pm$ 4.2	50.3 $\pm$ 1.2
	Enc-Dec USE-Net	78.8 $\pm$ 2.9	79.4 $\pm$ 7.9	76.9 $\pm$ 5.5	72.7 $\pm$ 1.7	63.7 $\pm$ 14.6	46.3 $\pm$ 1.8
Training on dataset #2	MS-D Net	78.7 $\pm$ 1.1	70.0 $\pm$ 4.4	86.8 $\pm$ 3.7	81.1 $\pm$ 0.5	83.2 $\pm$ 1.0	54.6 $\pm$ 0.8
	pix2pix	78.3 $\pm$ 0.9	67.3 $\pm$ 3.2	87.1 $\pm$ 2.9	81.8 $\pm$ 1.0	80.0 $\pm$ 2.5	51.1 $\pm$ 1.5
	U-Net	78.6 $\pm$ 1.0	70.9 $\pm$ 3.2	87.7 $\pm$ 2.0	82.4 $\pm$ 2.4	<b>83.8</b> $\pm$ 1.8	<b>54.9</b> $\pm$ 1.8
	Enc USE-Net	<b>78.8</b> $\pm$ 1.4	<b>72.3</b> $\pm$ 5.6	87.4 $\pm$ 2.5	82.6 $\pm$ 2.1	82.9 $\pm$ 2.5	54.5 $\pm$ 2.0
	Enc-Dec USE-Net	77.5 $\pm$ 2.1	70.6 $\pm$ 5.5	<b>87.8</b> $\pm$ 2.7	<b>82.8</b> $\pm$ 1.9	82.7 $\pm$ 1.5	53.6 $\pm$ 1.0
Training on dataset #3	MS-D Net	<b>81.2</b> $\pm$ 1.3	<b>73.3</b> $\pm$ 3.7	82.5 $\pm$ 1.9	<b>74.7</b> $\pm$ 2.0	91.6 $\pm$ 1.1	71.4 $\pm$ 5.6
	pix2pix	79.1 $\pm$ 5.6	64.6 $\pm$ 22.1	81.2 $\pm$ 4.2	66.6 $\pm$ 19.1	89.4 $\pm$ 4.8	62.8 $\pm$ 10.0
	U-Net	75.9 $\pm$ 3.4	63.3 $\pm$ 5.0	82.1 $\pm$ 2.9	66.6 $\pm$ 8.4	<b>91.7</b> $\pm$ 2.4	76.1 $\pm$ 4.1
	Enc USE-Net	77.3 $\pm$ 3.6	64.7 $\pm$ 6.4	<b>82.7</b> $\pm$ 4.3	66.7 $\pm$ 15.9	91.5 $\pm$ 3.2	74.0 $\pm$ 7.8
	Enc-Dec USE-Net	76.1 $\pm$ 4.2	58.9 $\pm$ 13.7	81.8 $\pm$ 4.8	67.6 $\pm$ 13.2	90.7 $\pm$ 3.1	<b>76.6</b> $\pm$ 7.8
Training on datasets #1/#2	MS-D Net	84.4 $\pm$ 3.1	86.5 $\pm$ 2.7	86.4 $\pm$ 2.8	81.2 $\pm$ 1.3	81.7 $\pm$ 2.3	54.9 $\pm$ 2.5
	pix2pix	<b>83.8</b> $\pm$ 2.6	84.8 $\pm$ 3.1	<b>87.1</b> $\pm$ 2.7	81.0 $\pm$ 0.4	<b>82.1</b> $\pm$ 2.5	54.0 $\pm$ 1.8
	U-Net	82.6 $\pm$ 3.3	90.0 $\pm$ 2.7	86.4 $\pm$ 2.0	82.2 $\pm$ 2.7	81.8 $\pm$ 2.1	55.3 $\pm$ 2.5
	Enc USE-Net	81.7 $\pm$ 5.4	90.0 $\pm$ 2.1	87.0 $\pm$ 2.1	82.2 $\pm$ 1.8	80.8 $\pm$ 2.7	<b>55.8</b> $\pm$ 1.7
	Enc-Dec USE-Net	85.4 $\pm$ 1.8	<b>87.7</b> $\pm$ 2.5	85.9 $\pm$ 2.1	<b>82.9</b> $\pm$ 1.4	81.1 $\pm$ 2.7	55.1 $\pm$ 2.1
Training on datasets #1/#3	MS-D Net	<b>85.4</b> $\pm$ 1.8	87.7 $\pm$ 2.5	80.9 $\pm$ 2.7	72.6 $\pm$ 3.7	91.0 $\pm$ 2.9	72.2 $\pm$ 1.9
	pix2pix	85.2 $\pm$ 1.6	86.8 $\pm$ 2.4	<b>82.7</b> $\pm$ 1.9	<b>75.7</b> $\pm$ 3.6	91.5 $\pm$ 1.9	71.0 $\pm$ 3.6
	U-Net	84.8 $\pm$ 0.4	90.4 $\pm$ 2.8	82.1 $\pm$ 2.9	72.5 $\pm$ 4.5	<b>92.6</b> $\pm$ 1.5	78.9 $\pm$ 4.0
	Enc USE-Net	83.8 $\pm$ 1.4	<b>91.1</b> $\pm$ 1.4	81.6 $\pm$ 3.7	71.9 $\pm$ 8.1	92.5 $\pm$ 1.9	79.6 $\pm$ 2.1
	Enc-Dec USE-Net	83.3 $\pm$ 3.2	90.4 $\pm$ 2.1	81.5 $\pm$ 4.6	71.8 $\pm$ 6.7	92.2 $\pm$ 2.4	<b>80.8</b> $\pm$ 1.8
Training on datasets #2/#3	MS-D Net	81.0 $\pm$ 1.3	72.5 $\pm$ 5.6	86.2 $\pm$ 2.5	77.4 $\pm$ 4.7	91.7 $\pm$ 0.9	69.8 $\pm$ 3.6
	pix2pix	<b>81.1</b> $\pm$ 1.1	<b>73.4</b> $\pm$ 3.3	87.4 $\pm$ 2.2	79.6 $\pm$ 5.7	92.0 $\pm$ 1.3	71.3 $\pm$ 3.4
	U-Net	79.2 $\pm$ 2.0	65.7 $\pm$ 6.3	88.1 $\pm$ 2.9	81.4 $\pm$ 2.6	92.9 $\pm$ 1.1	<b>77.6</b> $\pm$ 3.0
	Enc USE-Net	79.8 $\pm$ 1.8	70.3 $\pm$ 7.6	<b>88.5</b> $\pm$ 2.4	<b>82.0</b> $\pm$ 3.2	92.8 $\pm$ 1.0	76.3 $\pm$ 2.7
	Enc-Dec USE-Net	79.4 $\pm$ 2.5	67.4 $\pm$ 8.9	88.2 $\pm$ 2.9	82.0 $\pm$ 4.1	<b>93.7</b> $\pm$ 0.6	76.1 $\pm$ 3.4
Training on datasets #1/#2/#3	MS-D Net	84.8 $\pm$ 4.5	83.6 $\pm$ 6.9	86.8 $\pm$ 2.6	78.6 $\pm$ 5.1	91.1 $\pm$ 1.0	69.4 $\pm$ 4.5
	pix2pix	85.5 $\pm$ 2.6	87.6 $\pm$ 3.5	87.5 $\pm$ 2.0	80.9 $\pm$ 5.3	91.8 $\pm$ 1.3	69.7 $\pm$ 4.8
	U-Net	84.6 $\pm$ 1.9	90.5 $\pm$ 3.0	86.6 $\pm$ 2.0	80.9 $\pm$ 3.3	92.9 $\pm$ 1.1	77.2 $\pm$ 2.0
	Enc USE-Net	84.8 $\pm$ 2.3	91.1 $\pm$ 2.5	87.4 $\pm$ 1.8	81.4 $\pm$ 4.4	93.2 $\pm$ 0.7	79.1 $\pm$ 3.5
	Enc-Dec USE-Net	<b>87.1</b> $\pm$ 3.6	<b>91.9</b> $\pm$ 2.1	<b>88.6</b> $\pm$ 1.5	<b>83.1</b> $\pm$ 2.9	<b>93.7</b> $\pm$ 1.0	<b>80.1</b> $\pm$ 5.5
None	Qiu <i>et al.</i> [487]	78.0 $\pm$ 4.9	75.3 $\pm$ 6.4	71.0 $\pm$ 7.0	77.3 $\pm$ 2.6	82.1 $\pm$ 1.5	61.9 $\pm$ 4.6

Therefore, training schemes with mixed MRI datasets can achieve reliable and excellent performance, potentially useful for other clinical applications. Comparing the CG and PZ segmentation, the results on the CG are generally more accurate, except when trained and tested on dataset #1; this could be due to intra- and inter-scanner generalization, since dataset #1's scanner is different from those of datasets #2 and #3.

Hence, Enc-Dec USE-Net obtained high performance also in terms of difference between the automated and the manual boundaries.

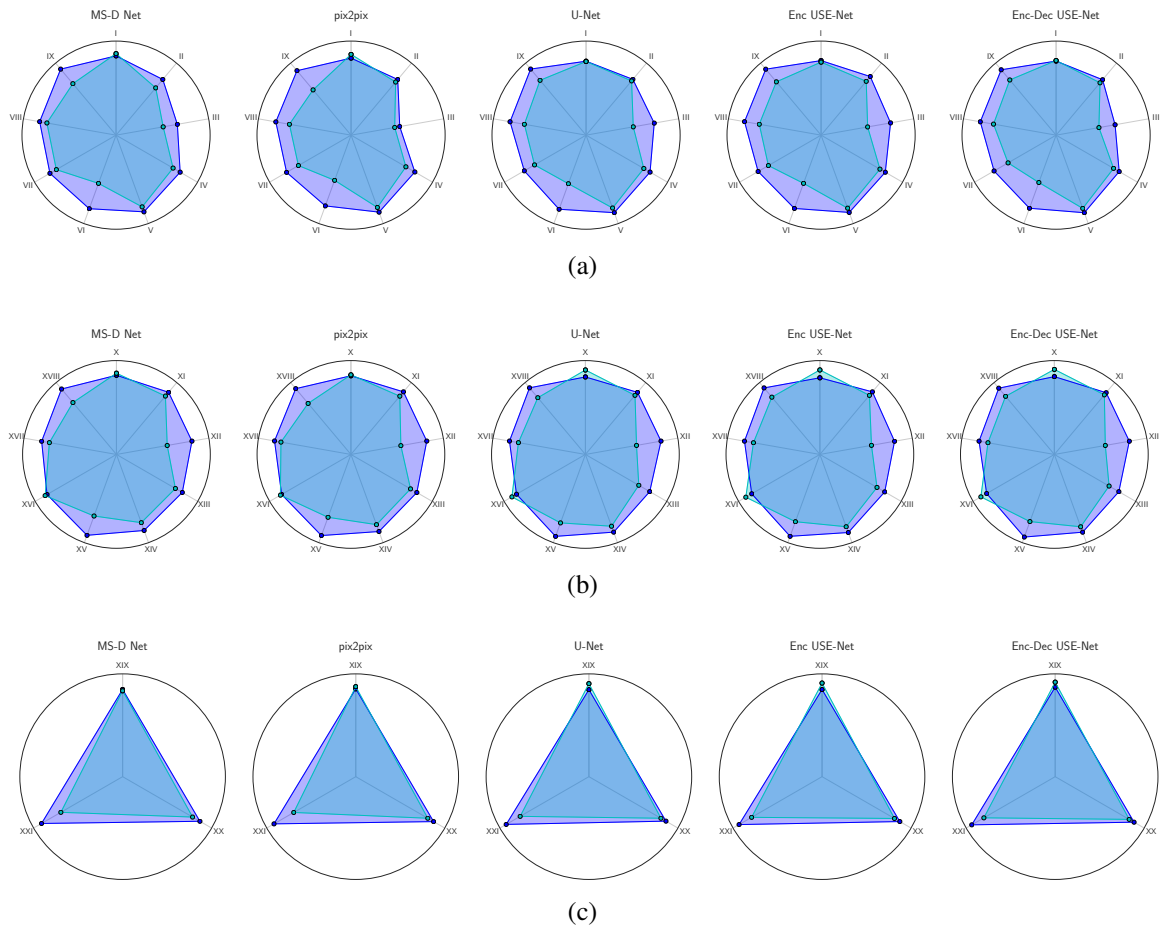


Fig. 6.4 Kiviatt diagrams showing the  $DSC$  values achieved by each method under different conditions.  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$  results are denoted by blue and cyan colors, respectively. Each variable represents a “training-set  $\rightarrow$  test-set” condition as follows: (a) one-dataset training: I) #1  $\rightarrow$  #1; II) #1  $\rightarrow$  #2; III) #1  $\rightarrow$  #3; IV) #2  $\rightarrow$  #1; V) #2  $\rightarrow$  #2; VI) #2  $\rightarrow$  #3; VII) #3  $\rightarrow$  #1; VIII) #3  $\rightarrow$  #2; IX) #3  $\rightarrow$  #3. (b) two-dataset training: X) #1/#2  $\rightarrow$  #1; XI) #1/#2  $\rightarrow$  #2; XII) #1/#2  $\rightarrow$  #3; XIII) #1/#3  $\rightarrow$  #1; XIV) #1/#3  $\rightarrow$  #2; XV) #1/#3  $\rightarrow$  #3; XVI) #2/#3  $\rightarrow$  #1; XVII) #2/#3  $\rightarrow$  #2; XVIII) #2/#3  $\rightarrow$  #3. (c) three-dataset training: XIX) #1/#2/#3  $\rightarrow$  #1; XX) #1/#2/#3  $\rightarrow$  #2; XXI) #1/#2/#3  $\rightarrow$  #3.

**Discussion and conclusions** The novel CNN architecture introduced in this work, Enc-Dec USE-Net, achieved accurate prostate zonal segmentation results when trained on the union of the available datasets in the case of multi-institutional studies—significantly outperforming the competitor CNN-based architectures, thanks to the integration of SE blocks [280] into U-Net [506]. This also derives from the presented cross-dataset generalization approach among three prostate MRI datasets, collected by three different institutions, aiming at segmenting  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$ ; Enc-Dec USE-Net’s segmentation performance considerably improved when trained on multiple datasets with respect to individual training conditions. Since the training

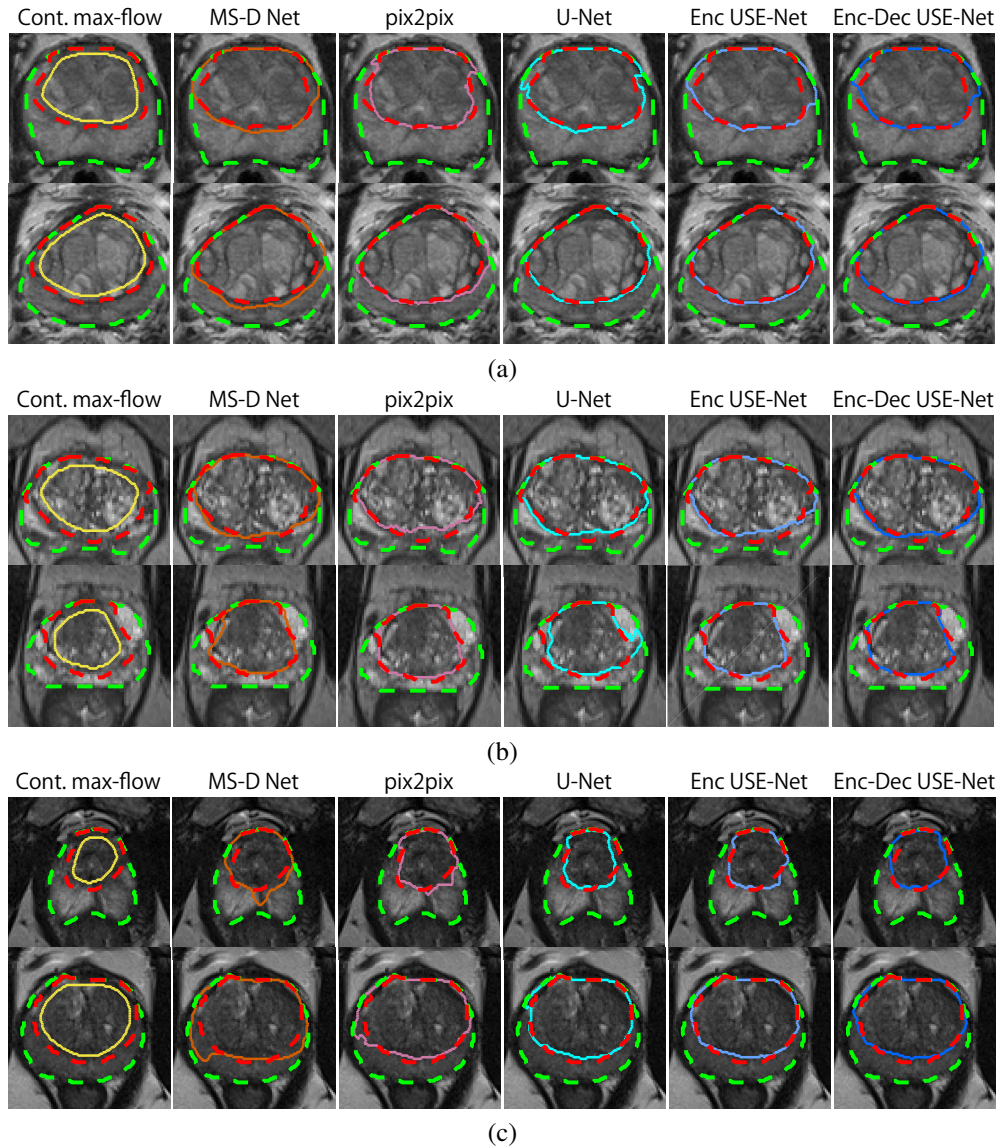


Fig. 6.5 Segmentation results obtained by the six investigated methods (under the three-dataset training condition) on two different images for each dataset: (a) #1; (b) #2; (c) #3. Automatic  $\mathcal{R}_{CG}$  segmentations (solid lines) are compared against the corresponding gold standards (dashed red line).  $\mathcal{R}_{PZ}$  segmentations can be obtained from  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{WG}$  (dashed green line) the constraints according to Eq. (6.2).

on multi-institutional datasets analyzed in this work achieved good intra-/cross-dataset generalization, CNNs could be trained on multiple datasets with different devices/protocols to obtain better outcomes in clinically feasible applications. Moreover, our research also implies that state-of-the-art CNN architectures properly combined with innovative concepts, such as feature recalibration provided by the SE blocks [280], allow for excellent intra-dataset generalization when tested on samples coming from the datasets used for the training

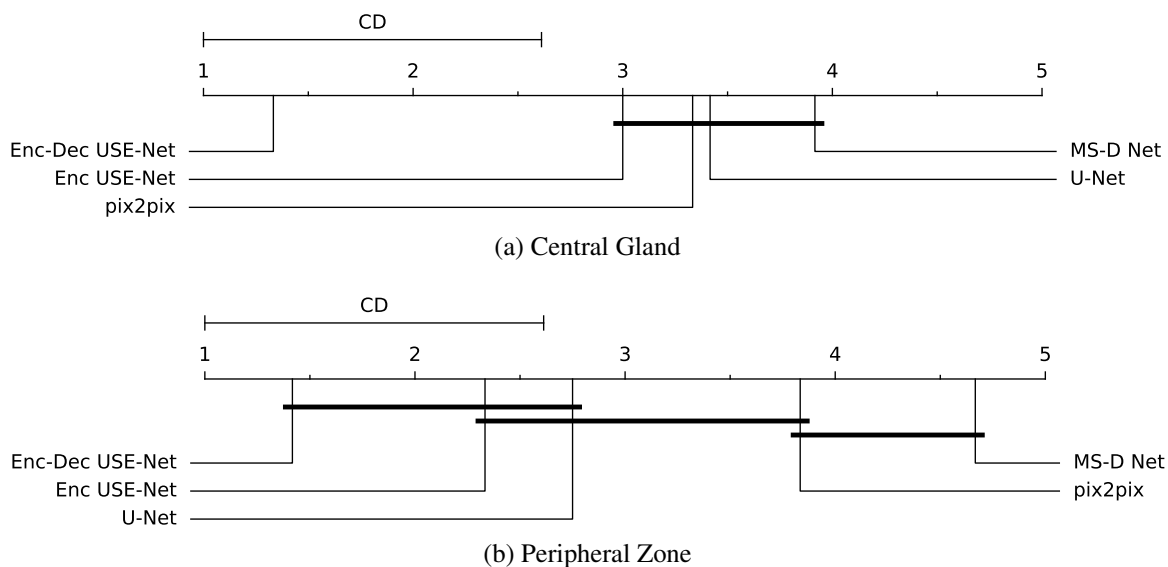


Fig. 6.6 Critical Difference (CD) diagram comparing the  $DSC$  values achieved by all the investigated CNN-based architectures using the Bonferroni-Dunn's *post hoc* test with 95% confidence level for the three-dataset training conditions. Bold lines indicate groups of methods whose performance difference was not statistically significant.

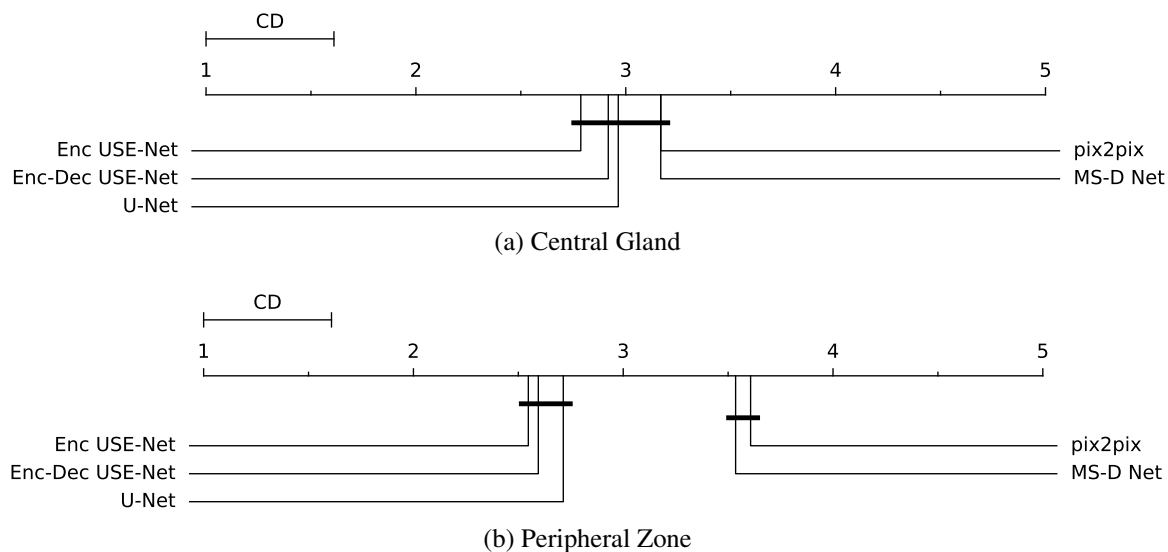


Fig. 6.7 Critical Difference (CD) diagram comparing the  $DSC$  values achieved by all the investigated CNN-based architectures using the Bonferroni-Dunn's *post hoc* test with 95% confidence level considering all training/testing combinations. Bold lines indicate groups of methods whose performance difference was not statistically significant.

phase. Therefore, such adaptive mechanisms may be a valuable solution in medical imaging applications involving multi-institutional settings.

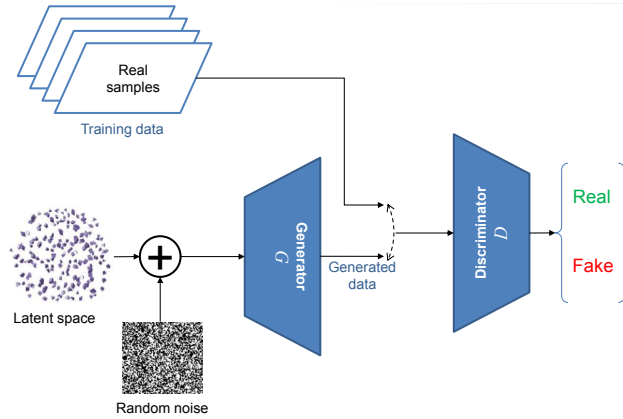


Fig. 6.8 Conceptual scheme of GANs.

As future developments, we will refine the output images considering the 3D spatial information among the prostate MR slices. Finally, for better cross-dataset generalization, we plan to use domain adaptation *via* transfer learning by maximizing the distribution similarity [629]. In this context, Generative Adversarial Networks (GANs) [219, 246] and Variational Auto-Encoders (VAEs) [323] represent useful solutions.

## 6.2 Generative Adversarial Networks

GANs introduce a framework for unsupervised learning aiming at estimating generative models *via* an adversarial process by simultaneously training two models [219]. Originally proposed by Goodfellow *et al.* in 2014 [219], GANs have shown remarkable results in image generation [704]. Let  $p_{\text{data}}$  be a generating distribution over data  $\mathbf{x}$ . The generator  $G(\mathbf{z}; \theta_g)$  is a mapping to data space that takes a prior on input noise variables  $p_{\mathbf{z}}(\mathbf{z})$ , where  $G$  is a neural network with parameters  $\theta_g$ . Similarly, the discriminator  $D(\mathbf{x}; \theta_d)$  is a neural network with parameters  $\theta_d$  that takes either real data or synthetic data and outputs a single scalar probability that  $\mathbf{x}$  came from the real data. The discriminator  $D$  maximizes the probability of classifying both training examples and samples from  $G$  correctly while the generator  $G$  minimizes the likelihood (Fig. 6.8); it is formulated as a minimax two-player game with value function  $\mathcal{V}(G, D)$ :

$$\min_G \max_D \mathcal{V}(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))] \quad (6.7)$$

This can be reformulated as the minimization of the Jensen-Shannon (JS) divergence between the distribution  $p_{\text{data}}$  and another distribution  $p_g$  derived from  $p_{\mathbf{z}}$  and  $G$ .

Along with classic Image Processing methods, CNNs have recently revolutionized medical image analysis [560], including MRI segmentation [255, 305]. However, CNN training demands large-scale annotated training data that are laborious to obtain in the medical domain [496]. Unfortunately, obtaining such massive medical data is challenging; consequently, better training requires intensive DA techniques, such as geometric/intensity transformations of original images [418, 506].

However, those reconstructed images intrinsically resemble the original ones, leading to limited performance improvement in terms of generalization abilities; thus, generating realistic (similar to the real image distribution) but completely new images is essential. In this context, GAN-based DA has excellently performed in general computer vision tasks. It attributes to GAN's good generalization ability from matching the generated distribution from noise variables to the real one with a sharp value function. Especially, Shrivastava *et al.* (SimGAN) outperformed the state-of-the-art with a relative 21% improvement in eye-gaze estimation [562].

### 6.2.1 GAN-based synthetic brain MR image generation

The generation of realistic medical images completely different from the original samples is worth to investigate. Our work in [246] aims to generate synthetic multi-sequence brain MR images using GANs, which is essential in medical imaging to increase diagnostic reliability, such as *via* DA in computer-assisted diagnosis as well as physician training and teaching (Fig. 6.9) [484]. However, this is extremely challenging—MR images are characterized by low contrast, strong visual consistency in brain anatomy, and intra-sequence variability. Our novel GAN-based approach for medical DA adopts Deep Convolutional GAN (DCGAN) [489] and Wasserstein GAN (WGAN) [25] to generate realistic images, and an expert physician validates them *via* the visual Turing test (described in Appendix A.5) [540].

We mainly addressed two research questions:

- **GAN selection:** Which GAN architecture is well-suited for realistic medical image generation?
- **Medical DA:** How can we handle MR images with specific intra-sequence variability?

Our main contributions are as follows:



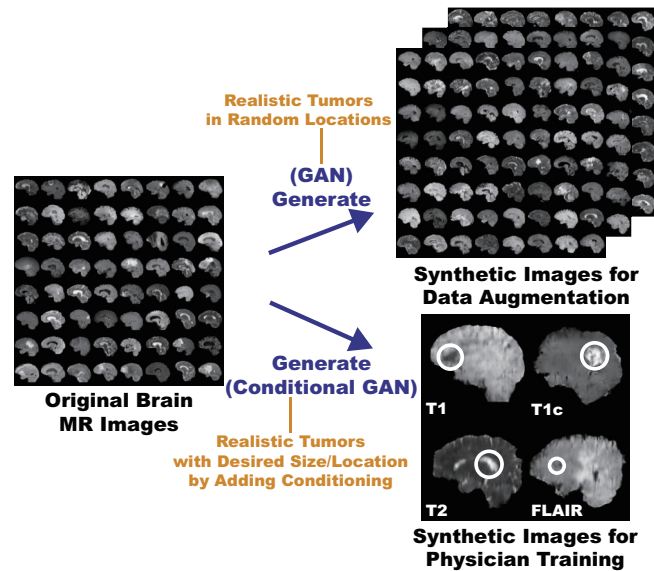


Fig. 6.9 Potential applications of the proposed GAN-based synthetic brain MR image generation: (1) DA for better diagnostic accuracy by generating random realistic images giving insights in classification; (2) physician training for better understanding various diseases to prevent misdiagnosis by generating desired realistic pathological images.

- **MR image generation:** This research shows that WGAN can generate realistic multi-sequence brain MR images, possibly leading to valuable clinical applications: DA and physician training.
- **Medical image generation:** This research provides how to exploit medical images with intrinsic intra-sequence variability towards GAN-based DA for medical imaging.

GANs generate highly realistic images, without a well-defined objective function associated with difficult training accompanying oscillations and mode collapse—i.e., a common failure case where the generator learns with extremely low variety. Whereas VAEs [322], the other most used deep generative models, have an objective likelihood function to optimize, and could so generate blurred samples because of the injected noise and imperfect reconstruction [405].

Therefore, many medical imaging researchers have begun to use GANs recently, such as in image super-resolution [383], anomaly detection [551], and estimating CT images from the corresponding MR images [436]. As GANs allow for adding conditioning on the class labels and images, they often use such conditional GANs [419] to produce desired images, while it makes learning robust latent spaces difficult.

Differently from a very recent work of GANs for biological image synthesis (fluorescence microscopy) [447], to the best of our knowledge, this is the first GAN-based realistic

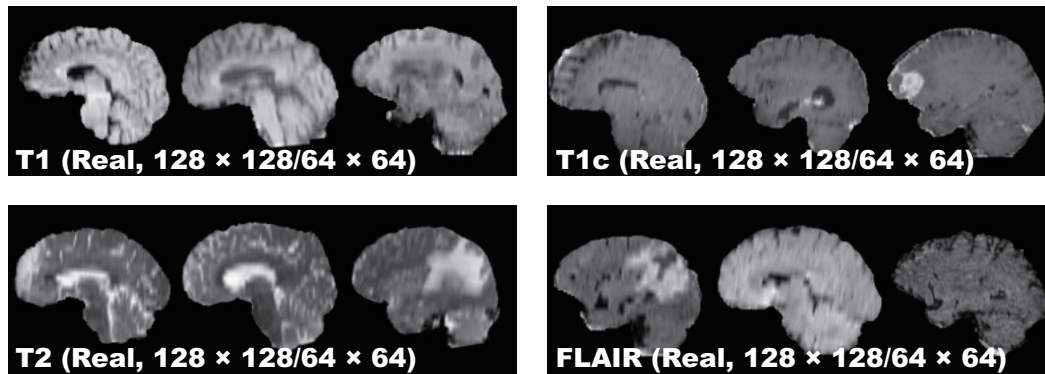


Fig. 6.10 Examples of real MR images used for training the GANs: the resized sagittal multi-sequence brain MRI scans of patients with HGG on the BRATS 2016 training dataset [404].

brain tumor MR image generation approach aimed at DA and physician training. Instead of reconstructing real brain MR images themselves with respect to geometry/intensity, a completely different approach—generating novel realistic images using GANs—may become a clinical breakthrough.

Towards clinical applications utilizing realistic brain MR images, we generate synthetic brain MR images from the original samples using GANs. Here, we compare the two most used GANs, namely DCGAN and WGAN, to find a well-suited GAN between them for medical image generation—it must avoid mode collapse and generate realistic MR images with high resolution.

### The BRATS 2016 dataset

This study exploits a dataset of multi-sequence brain MR images to train GANs with sufficient data and resolution, which was originally produced for the Multimodal BRATS challenge [404]. In particular, the BRATS 2016 training dataset contains 220 High-Grade Glioma (HGG) and 54 Low-Grade Glioma (LGG) cases, with T1w (T1), T1w-CE (T1c), T2w (T2), and FLAIR sequences—they were skull stripped and resampled to isotropic  $1\text{mm} \times 1\text{mm} \times 1\text{mm}$  resolution with image dimension  $240 \times 240 \times 155$ ; among the different sectional planes, we use the sagittal multi-sequence scans of patients with HGG to show that our GANs can generate a complete view of the whole brain anatomy (allowing for visual consistency among the different brain lobes), including also severe tumors for clinical purpose.

### Proposed GAN-based image generation approach

This section describes the details of the GAN-based MR image generation.

**Pre-processing** We select the slices from #80 to #149 among the whole 240 slices to omit initial/final slices, since they convey a negligible amount of useful information and could affect the training. The images are resized to both  $64 \times 64$  and  $128 \times 128$  from  $240 \times 155$  for better GAN training (DCGAN architecture results in stable training on  $64 \times 64$  [489], and so  $128 \times 128$  is reasonably a high-resolution). Fig. 6.10 shows some real MR images used for training; each sequence contains 15,400 images with  $220 \text{ patients} \times 70 \text{ slices}$  (61,600 in total).

**GAN-based MR image generation** DCGAN and WGAN generate six types of images as follows:

- T1 sequence ( $128 \times 128$ ) from the real T1;
- T1c sequence ( $128 \times 128$ ) from the real T1c;
- T2 sequence ( $128 \times 128$ ) from the real T2;
- FLAIR sequence ( $128 \times 128$ ) from the real FLAIR;
- Concat sequence ( $128 \times 128$ ) from concatenating the real T1, T1c, T2, and FLAIR (i.e., feeding the model with samples from all the MRI sequences);
- Concat sequence ( $64 \times 64$ ) from concatenating the real T1, T1c, T2, and FLAIR.

Concat sequence refers to a new ensemble sequence for an alternative DA, containing the features of all four sequences. We also generate  $64 \times 64$  Concat images to compare the generation performance in terms of image size.

**DCGAN.** DCGAN [489] is a standard GAN [219] with a convolutional architecture for unsupervised learning; this generative model uses up-convolutions interleaved with ReLU non-linearity and batch-normalization.

**DCGAN implementation details.** We use the same DCGAN architecture [489] with no tanh in the generator, ELU as the discriminator, all filters of size  $4 \times 4$ , and a half channel size for DCGAN training. A batch size of 64 and Adam optimizer with  $2.0 \times 10^{-4}$  learning rate were implemented.

**WGAN.** WGAN [25] is an alternative to traditional GAN training, as the JS divergence is limited, such as when it is discontinuous; this novel GAN achieves stable learning with less mode collapse by replacing it to the Earth Mover (EM) distance (a.k.a. the Wasserstein-1 metrics):

$$W(p_g, p_r) = \inf_{p \in \Pi(p_g, p_r)} \mathbb{E}_{(\mathbf{x}, \mathbf{x}') \sim p} \|\mathbf{x} - \mathbf{x}'\|, \quad (6.8)$$

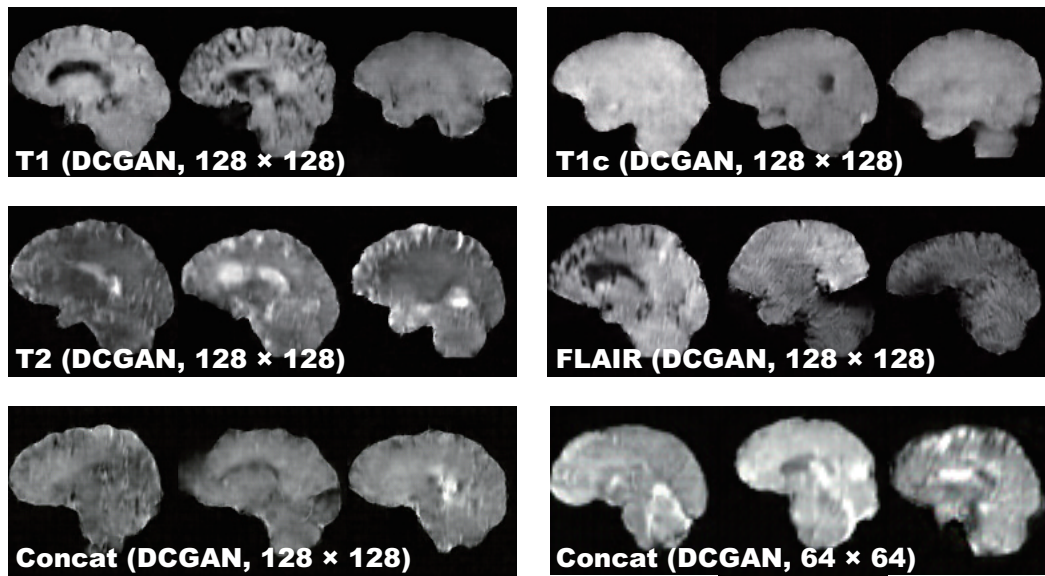


Fig. 6.11 Examples of synthetic MR images yielded by DCGAN.

where  $\Pi(p_g, p_r)$  is the set of all joint distributions  $p$  whose marginals are  $p_g$  and  $p_r$ , respectively. In other words,  $p$  implies how much mass must be transported from one distribution to another. This distance intuitively indicates the cost of the optimal transport plan.

**WGAN implementation details.** We used the same DCGAN architecture [489] for WGAN training. A batch size of 64 and Root Mean Square Propagation (RMSprop) optimizer with  $5.0 \times 10^{-5}$  learning rate were implemented.

**Clinical validation using the visual Turing test** In the visual Turing test (see Appendix A.5 for more details), an expert physician was asked to constantly classify a random selection of 50 real/50 synthetic MR images as real or synthetic shown in a random order for each GAN/sequence, without previous training stages revealing which is real/synthetic; Concat images were classified together with real T1, T1c, T2, and FLAIR images in equal proportion.

## Results

This section shows how DCGAN and WGAN generate synthetic brain MR images. The results include instances of synthetic images and their quantitative evaluation of the realism by an expert physician. The training took about 2 (1) hours to train each  $128 \times 128$  ( $64 \times 64$ ) sequence on an NVIDIA (Santa Clara, CA, USA) GeForce GTX 980 GPU, increasingly learning realistic features.

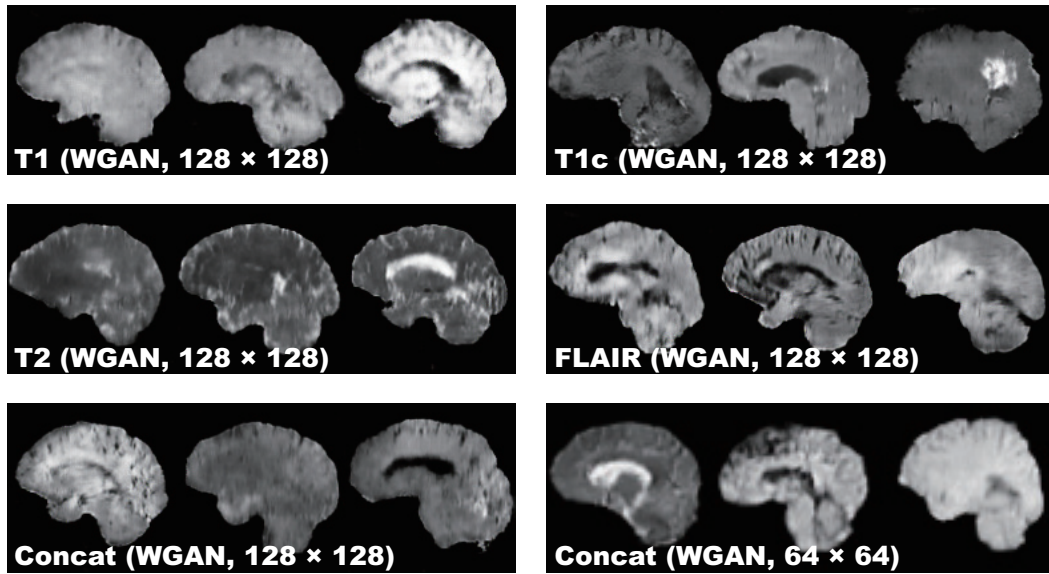


Fig. 6.12 Examples of synthetic MR images yielded by WGAN.

**MR images generated by DCGAN** Fig. 6.11 illustrates examples of synthetic images by DCGAN. The images look similar to the real samples. Concat images combine appearances and patterns from all the four sequences used in training. Since DCGAN's value function could be unstable, it often generates hyper-intense T1-like images analogous to mode collapse for  $64 \times 64$  Concat images, while sharing the same hyper-parameters with  $128 \times 128$ .

**MR images generated by WGAN** Fig. 6.12 shows the example output of WGAN in each sequence. Outperforming remarkably DCGAN, WGAN successfully captures the sequence-specific texture and the appearance of the tumors while maintaining the realism of the original brain MR images. As expected,  $128 \times 128$  Concat images tend to have more messy and unrealistic artifacts than  $64 \times 64$  Concat ones, especially around the boundaries of the brain, due to the introduction of unexpected intensity patterns.

**Visual Turing test results** Table 6.3 shows the confusion matrix concerning the visual Turing test. Even the expert physician found classifying real and synthetic images challenging, especially in lower resolution due to their less detailed appearances unfamiliar in clinical routine, even for highly hyper-intense  $64 \times 64$  Concat images by DCGAN; distinguishing Concat images was easier compared to the case of T1, T1c, T2, and FLAIR images because the physician often felt odd from the artificial sequence. WGAN succeeded to deceive the physician significantly better than DCGAN for all the MRI sequences except FLAIR images (62% to 54%).

Table 6.3 Visual Turing test results by a physician for classifying real vs synthetic images. It should be noted that proximity to 50% of accuracy indicates superior performance (chance = 50%).

	Accuracy (%)	Real Selected as Real	Real as Synt	Synt as Real	Synt as Synt
T1 (DCGAN, $128 \times 128$ )	70	26	24	6	44
T1c (DCGAN, $128 \times 128$ )	71	24	26	3	47
T2 (DCGAN, $128 \times 128$ )	64	22	28	8	42
FLAIR (DCGAN, $128 \times 128$ )	54	12	38	8	42
Concat (DCGAN, $128 \times 128$ )	77	34	16	7	43
Concat (DCGAN, $64 \times 64$ )	54	13	37	9	41
T1 (WGAN, $128 \times 128$ )	64	20	30	6	44
T1c (WGAN, $128 \times 128$ )	55	13	37	8	42
T2 (WGAN, $128 \times 128$ )	58	19	31	11	39
FLAIR (WGAN, $128 \times 128$ )	62	16	34	4	46
Concat (WGAN, $128 \times 128$ )	66	31	19	15	35
Concat (WGAN, $64 \times 64$ )	53	18	32	15	35

## Conclusion

Our preliminary results show that GANs, especially WGAN, can generate  $128 \times 128$  realistic multi-sequence brain MR images that even an expert physician is unable to accurately distinguish from the real, leading to valuable clinical applications, such as DA and physician training. This attributes to WGAN's good generalization ability with a sharp value function. In this context, DCGAN might be unsuitable due to both the inferior realism and mode collapse in terms of intensity. We only use the slices of interest in training to obtain desired MR images and generate both original/Concat sequence images for DA in medical imaging.

This study confirms the synthetic image quality by the human expert evaluation, but a more objective computational evaluation for GANs should also follow, such as Classifier Two-Sample Tests (C2ST) [370], which assesses whether two samples are drawn from the same distribution. Currently this work uses sagittal MR images alone, so we will generate coronal and transverse images in the near future. As this research uniformly selects middle slices in pre-processing, better data generation demands developing a classifier to only select brain MRI slices with/without tumors.

Towards DA, while realistic images give more insights on geometry/intensity transformations in classification, more realistic images do not always assure better DA, so we have to find suitable image resolutions and sequences; that is why we generate both high-resolution images and Concat images, yet they looked more unrealistic for the physician. For physician training, generating desired realistic tumors by adding conditioning requires exploring extensively the latent spaces of GANs.

Overall, our novel GAN-based realistic brain MR image generation approach sheds light on diagnostic and prognostic medical applications; future studies on these applications are needed to confirm our encouraging results.

### 6.2.2 PGGAN-based DA for MRI brain tumor detection

As reported in the previous section, the classical images transformed geometric/intensity have intrinsically a similar distribution with respect to the original ones, leading to limited performance improvement; thus, generating realistic (i.e., similar to the real image distribution) but completely new samples is essential to fill the real image distribution uncovered by the original dataset [246]. In this context, GAN-based DA is promising, as it has shown excellent performance in computer vision, revealing good generalization ability. Especially, SimGAN outperformed the state-of-the-art with 21% improvement in eye-gaze estimation [562].

Also in medical imaging, realistic retinal image and CT image generation have been tackled using adversarial learning [108, 130]; a very recent study reported performance improvement with synthetic training data in CNN-based liver lesion classification, using a small number of  $64 \times 64$  CT images for GAN training [194]. However, GAN-based image generation using MRI, the most effective modality for soft-tissue acquisition, has not yet been reported due to the difficulties from low-contrast MR images, strong anatomical consistency, and intra-sequence variability; in our previous work [246], we generated  $64 \times 64/128 \times 128$  MR images using conventional GANs and even an expert physician failed to accurately distinguish between the real/synthetic images.

According to the results achieved in [246], we aimed at generating highly-realistic and original-sized  $256 \times 256$  images, while maintaining clear tumor/non-tumor features using GANs [247]. Therefore, our goal is to generate GAN-based synthetic T1w-CE brain MR images—the most commonly used sequence in tumor detection thanks to its high-contrast [413]—for CNN-based tumor detection. This computer-assisted brain tumor MRI analysis task is clinically valuable for better diagnosis, prognosis, and treatment [401].

Generating  $256 \times 256$  images is extremely challenging: (i) GAN training is unstable with high-resolution inputs and severe artifacts appear due to strong consistency in brain anatomy; (ii) brain tumors vary in size, location, shape, and contrast. However, it is beneficial, because most CNN architectures adopt around  $256 \times 256$  input sizes (e.g., InceptionResNetV2 [600]:  $299 \times 299$ , ResNet-50 [258]:  $224 \times 224$ ) and we can achieve better results with original-sized image augmentation—towards this, we use Progressive Growing of GANs (PGGANs), a multi-stage generative training method [313]. Moreover, an expert physician evaluates the generated images' realism and tumor/non-tumor features *via* the visual Turing test [540]. Us-

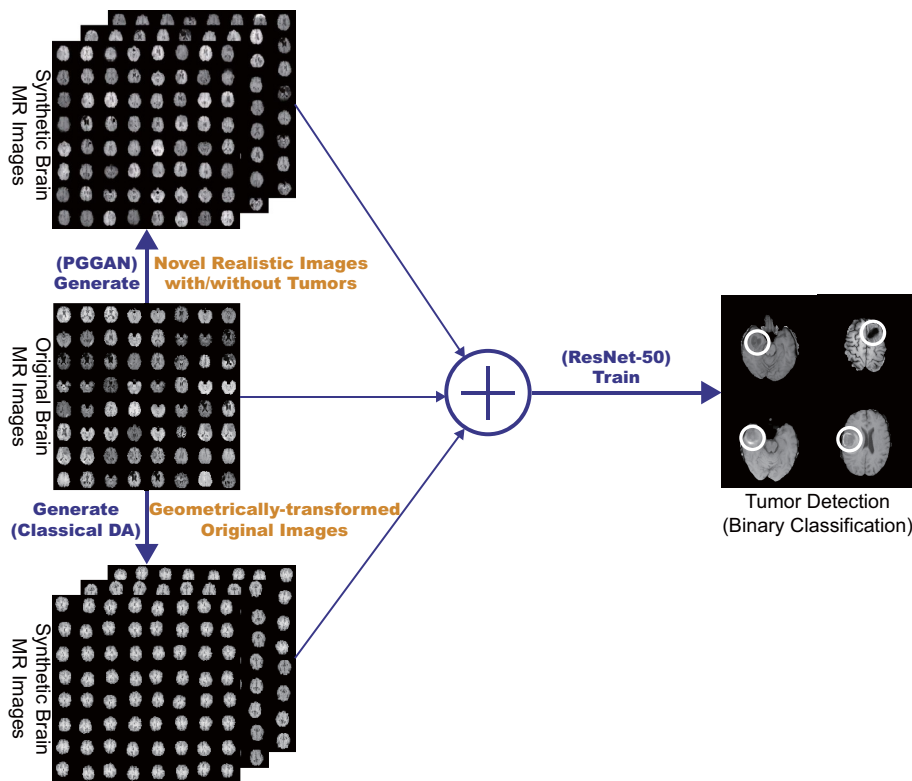


Fig. 6.13 PGGAN-based DA for improved tumor detection: the PGGANs method generates a number of realistic brain tumor/non-tumor MR images and the binary classifier uses them as additional training data.

ing the synthetic images, our novel PGGAN-based DA approach achieves better performance in CNN-based tumor detection, when combined with classical DA (Fig. 6.13).

Recently, researchers applied GANs to medical imaging, mainly for image-to-image translation, such as segmentation [678], super-resolution [383], and cross-modality translation [436]. Since GANs allow for adding conditional dependency on the input information (e.g., category, image, and text), they used such conditional GANs [419] to produce the desired corresponding images. However, GAN-based research on generating large-scale synthetic training images is limited, while the biggest challenge in this field is handling small datasets.

Differently to a very recent DA work for  $64 \times 64$  CT liver lesion ROI classification [194], to the best of our knowledge, this is the first GAN-based whole MR image augmentation approach. This work also firstly uses PGGANs to generate  $256 \times 256$  medical images. Along with classical transformations of real images, a completely different approach—generating novel realistic images using PGGANs—may become a clinical breakthrough.

Our main contributions are as follows:



- **MR image generation:** This research explains how to exploit MRI data to generate realistic and original-sized  $256 \times 256$  whole brain MR images using PGGANs, while maintaining clear tumor/non-tumor features.
- **MR image augmentation:** This study shows encouraging results on PGGAN-based DA, when combined with classical DA, for better tumor detection and other medical imaging tasks.

In the context of the high-resolution images generation, several multi-stage generative training methods have been proposed, such as the StackGAN architecture [699]. Composite GAN exploits multiple generators to separately generate different parts of an image [341]. The PGGANs method adopts multiple training procedures from low resolution to high to incrementally generate a realistic image [313].

### Dataset preparation

We exploited again the BRATS 2016 dataset [404]. We selected the slices from #30 to #130 among the whole 155 slices to omit initial/final slices, since they convey a negligible amount of useful information and negatively affect the training of both PGGANs [313] and ResNet-50 [258]. For tumor detection, our whole dataset (220 patients) was divided into: (i) a training set (154 patients); (ii) a validation set (44 patients); (iii) a test set (22 patients). Only the training set is used for the PGGAN training to be fair. Since tumor/non-tumor annotations were based on 3D volumes, these labels are often incorrect/ambiguous on 2D slices; so, we discard (i) tumor images tagged as non-tumor, (ii) non-tumor images tagged as tumor, (iii) unclear boundary images, and (iv) too small/big images; after all, our datasets consist of:

- training set (5,036 tumor/3,853 non-tumor images);
- validation set (793 tumor/640 non-tumor images);
- test set (1,575 tumor/1,082 non-tumor images).

The images from the training set are zero-padded to reach a power of 2,  $256 \times 256$  from  $240 \times 240$  pixels for better PGGAN training. Fig. 6.14 shows examples of real MR images.

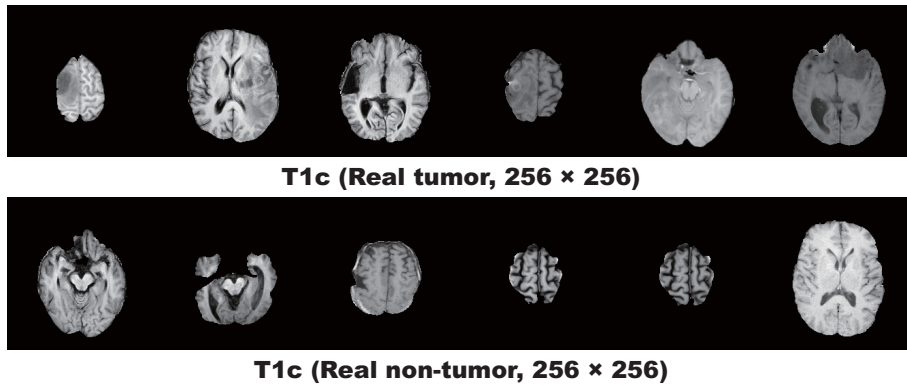


Fig. 6.14 Examples of real  $256 \times 256$  MR images used for PGGAN training.

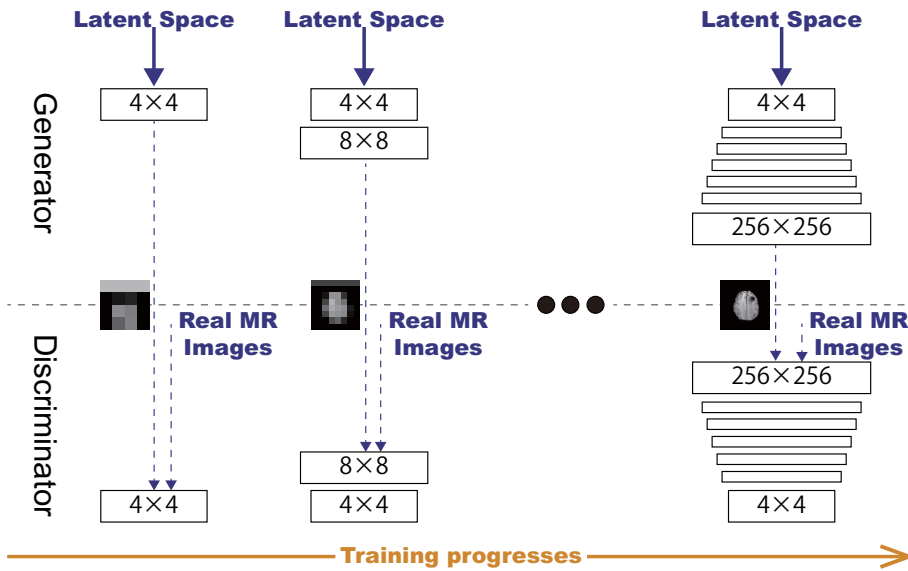


Fig. 6.15 PGGANs architecture for synthetic  $256 \times 256$  MR image generation.

### PGGAN-based image generation

**PGGANs** is a novel training method for GANs with progressively growing generator and discriminator [313]: starting from low resolution, newly added layers model fine-grained details as training progresses. As Fig. 6.15 shows, we adopt PGGANs to generate highly-realistic and original-sized  $256 \times 256$  brain MR images; tumor/non-tumor images are separately trained and generated.

**PGGAN implementation details** We used the PGGAN architecture with the Wasserstein loss using gradient penalty [231]. Training lasts for 100 epochs with a batch size of 16 and  $1.0 \times 10^{-3}$  learning rate for Adam optimizer.

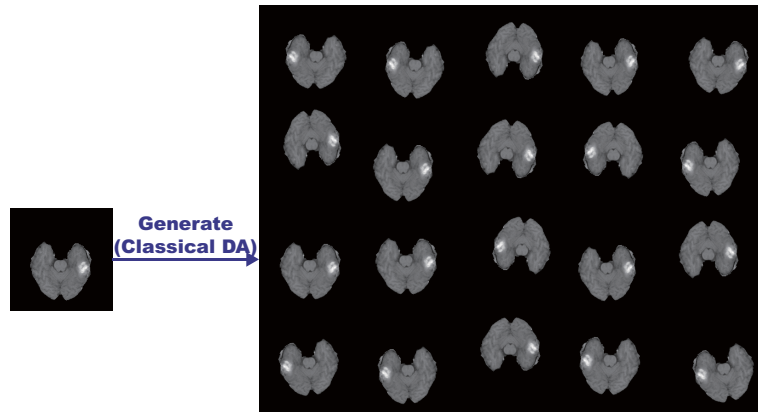


Fig. 6.16 Examples of real MR image and its geometrically-transformed synthetic images.

### Tumor detection using ResNet-50

**Pre-processing** To fit ResNet-50's input size, we center-crop the whole images from  $240 \times 240$  to  $224 \times 224$  pixels.

**ResNet-50** is a residual learning-based CNN with 50 layers [258]: unlike conventional learning unreferenced functions, it reformulates the layers as learning residual functions for sustainable and easy training. We adopt ResNet-50 to detect tumors in brain MR images, i.e., the binary classification of images with/without tumors.

To confirm the effect of PGGAN-based DA, the following classification results are compared: (i) without DA, (ii) with 200,000 classical DA (100,000 for each class), (iii) with 200,000 PGGAN-based DA, and (iv) with both 200,000 classical DA and 200,000 PGGAN-based DA; the classical DA adopts a random combination of horizontal/vertical flipping, rotation up to 10 degrees, width/height shift up to 8%, shearing up to 8%, zooming up to 8%, and constant filling of points outside the input boundaries (Fig. 6.16). For better DA, highly-unrealistic PGGAN-generated images are manually discarded.

**ResNet-50 implementation details** We use the ResNet-50 architecture pre-trained on ImageNet with a dropout of 0.5 before the final softmax layer, along with a batch size of 192,  $1.0 \times 10^{-3}$  learning rate for Adam optimizer, and early stopping of 10 epochs.

### Clinical validation using the visual Turing test

To quantitatively evaluate (i) how realistic the PGGAN-based synthetic images are, (ii) how obvious the synthetic images' tumor/non-tumor features are, we supply, in a random order, to an expert physician a random selection of:

- 50 real tumor images;
- 50 real non-tumor images;
- 50 synthetic tumor images;
- 50 synthetic non-tumor images.

Then, the physician is asked to constantly classify them as both (i) real/synthetic and (ii) tumor/non-tumor, without previous training stages revealing which is real/synthetic and tumor/non-tumor; here, we only show successful cases of synthetic images, as we can discard failed cases for better DA. The so-called visual Turing test [540] is used to probe the human ability to identify attributes and relationships in images, also in evaluating the visual quality of GAN-generated images [562]. Similarly, this applies to medical images in clinical environments [108, 194], wherein physicians' expertise is critical.

**Visualization using t-SNE** To visually analyze the distribution of both (i) real/synthetic and (ii) tumor/non-tumor images, we use t-Distributed Stochastic Neighbor Embedding (t-SNE) [625] on a random selection of:

- 300 real non-tumor images;
- 300 geometrically-transformed non-tumor images;
- 300 PGGAN-generated non-tumor images;
- 300 real tumor images;
- 300 geometrically-transformed tumor images;
- 300 PGGAN-generated tumor images.

Only 300 images per each category are selected for better visualization. t-SNE is a Machine Learning algorithm for dimensionality reduction to represent high-dimensional data into a lower-dimensional (2D/3D) space. It non-linearly adapts to input data using perplexity, which balances between the data's local and global aspects.

**t-SNE implementation details** We used t-SNE with a perplexity of 100 for 1,000 iterations to obtain a 2D visual representation.

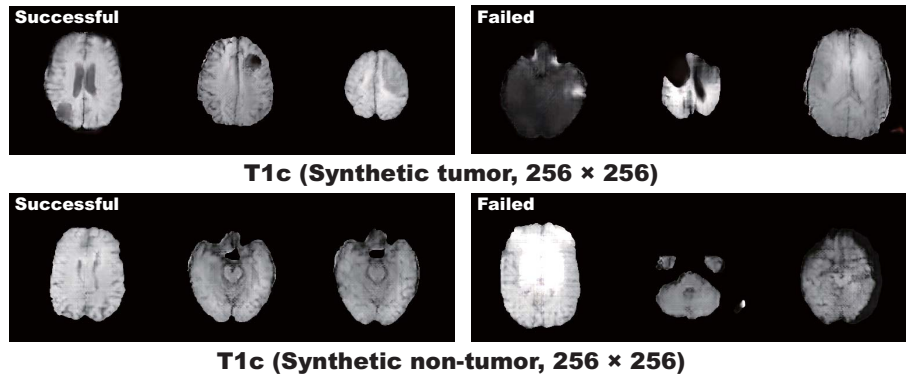


Fig. 6.17 Examples of synthetic MR images yielded by PGGANs: (a) successful cases; (b) failed cases.

Table 6.4 Binary classification results for detecting brain tumors with/without DA.

Experimental condition	Accuracy	Sensitivity	Specificity
ResNet-50 (w/o DA)	90.06%	85.27%	97.04%
ResNet-50 (w/ 200k classical DA)	90.70%	88.70%	93.62%
ResNet-50 (w/ 200k PGGAN-based DA)	62.02%	<b>99.94%</b>	6.84%
ResNet-50 (w/ 200k classical DA + 200k PGGAN-based DA)	<b>91.08%</b>	86.60%	<b>97.60%</b>

**Results** This section shows how PGGANs generates synthetic brain MR images. The results include instances of synthetic images, their quantitative evaluation by an expert physician, and their influence on tumor detection.

**MR images generated by PGGANs** Fig. 6.17 illustrates examples of synthetic tumor/non-tumor images by PGGANs. In our visual confirmation, for about 75% of cases, PGGANs successfully captures the T1c-specific texture and tumor appearance while maintaining the realism of the original brain MR images; however, for about 25% of cases, the generated images lack clear tumor/non-tumor features or contain unrealistic features, such as hyper-intensity, gray contours, and odd artifacts.

**Tumor detection results** Table 6.4 shows the classification results for detecting brain tumors with/without DA techniques. As expected, the test accuracy improves by 0.64% with the additional 200,000 geometrically-transformed images for training. When only the PGGAN-based DA is applied, the test accuracy decreases drastically with almost 100% of sensitivity and 6.84% of specificity, because the classifier recognizes the synthetic images' prevailed unrealistic features as tumors, similarly to anomaly detection.

However, surprisingly, when it is combined with the classical DA, the accuracy increases by 1.02% with higher sensitivity and specificity; this could occur because the PGGAN-based

Table 6.5 Visual Turing test results by a physician for classifying Real (R) vs Synthetic (S) images and Tumor (T) vs Non-tumor (N) images.

Real/Synthetic Classification	R as R	R as S	S as R	S as S
78.5%	58	42	1	99
Tumor/Non-tumor Classification	T as T	T as N	N as T	N as N
90.5%	82	18 (R : 5, S : 13)	1 (S : 1)	99

DA fills the real image distribution uncovered by the original dataset, while the classical DA provides the robustness on training for most cases.

**Visual Turing test results** Table 6.5 shows the confusion matrix for the visual Turing test. Differently from our previous work on GAN-based  $64 \times 64/128 \times 128$  MR image generation, the expert physician easily recognizes  $256 \times 256$  synthetic images [246], while tending also to classify real images as synthetic; this can be attributed to high resolution associated with more difficult training and detailed appearance, making artifacts stand out, which is coherent to the ResNet-50’s low tumor detection accuracy with only the PGGAN-based DA. Generally, the physician’s tumor/non-tumor classification accuracy is high and the synthetic images successfully capture tumor/non-tumor features. However, unlike non-tumor images, the expert recognizes a considerable number of tumor images as non-tumor, especially on the synthetic images; this results from the remaining real images’ ambiguous annotation, which is amplified in the synthetic images trained on them.

**t-SNE results** As presented in Fig. 6.18, tumor/non-tumor images’ distribution shows a tendency that non-tumor images locate from top left to bottom right and tumor images locate from top right to center, while the distinction is unclear with partial overlaps. Classical DA covers a wide range, including zones without any real/GAN-generated images, but tumor/non-tumor images often overlap there. Meanwhile, PGGAN-generated images concentrate differently from real images, while showing more frequent overlaps than the real ones; this probably derives from those synthetic images with unsatisfactory realism and tumor/non-tumor features.

**Conclusion** Our preliminary results show that PGGANs can generate original-sized  $256 \times 256$  realistic brain MR images and achieve higher performance in tumor detection, when combined with classical DA. This occurs because PGGANs’ multi-stage image generation obtains good generalization and synthesizes images with the real image distribution unfilled

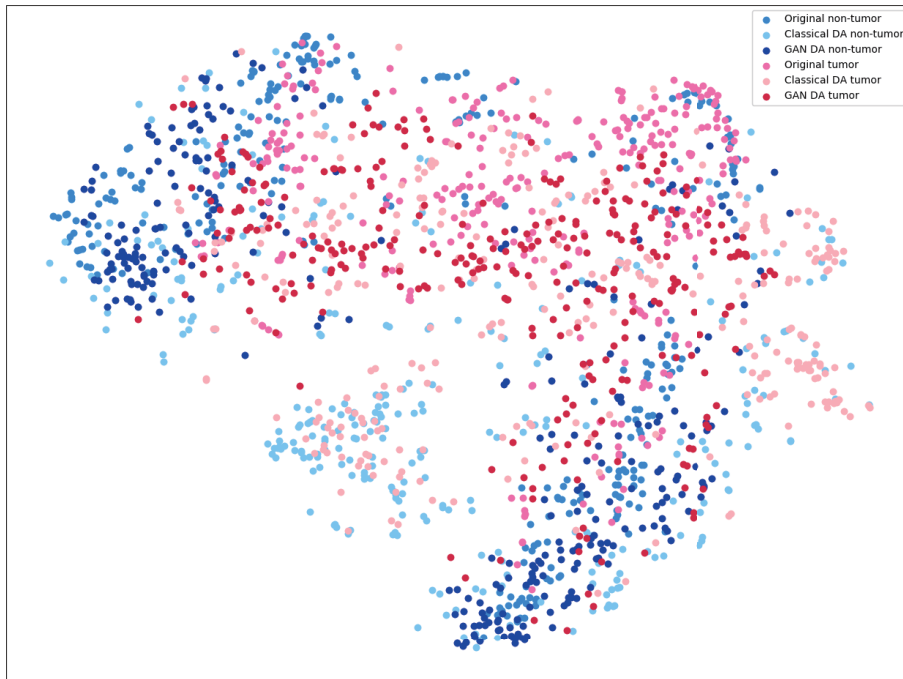


Fig. 6.18 t-SNE result on six categories, with 300 images per each category: (a) real tumor/non-tumor images; (b) geometrically-transformed tumor/non-tumor images; (c) PGGAN-generated tumor/non-tumor images.

by the original dataset. However, considering the visual Turing test and t-SNE results, yet unsatisfactory realism with high resolution strongly limits DA performance, so we plan to (i) discard unrealistic images or (ii) generate only realistic images, and then (iii) refine synthetic images more similar to the real image distribution.

For (i), classifier two-sample tests [370], assessing whether two samples are drawn from the same distribution, can help discard images not from the real image distribution, as manual removal is demanding. Regarding (ii), we can map an input random vector onto each training image [551] and generate images with suitable vectors, to control the divergence of the generated images; virtual adversarial training can be also integrated to control the output distribution. Lastly, (iii) can be achieved by GAN-based image-to-image translation, such as CycleGAN [704] or the UNsupervised Image-to-image Translation (UNIT) framework [366], considering SimGAN's remarkable performance improvement after refinement [562]. At the time of writing, taking inspiration from the SimGAN's rationale [562], we are refining the MR images synthesized by the PGGAN-based simulator by exploiting a refiner Neural Network. Moreover, we should further avoid real images with ambiguous/inaccurate annotation for better tumor detection. Finally, conditional GANs [419] could be exploited to generate labeled synthetic images from arbitrary segmentation masks [499], thus allowing us to

synthesize brain MR images representing tumors in specific positions. By so doing, the generated samples may be employed in DA for medical image segmentation tasks.

Overall, our novel PGGAN-based DA approach sheds light on diagnostic and prognostic medical applications, not limited to tumor detection; future studies are needed to extend these early encouraging results.



# Chapter 7

## Conclusion

### 7.1 Final remarks

The research activities presented in this thesis have been motivated by the need to develop novel and effective methods for the analysis of biomedical images, considering the unique challenges imposed by this kind of data [214, 403]. Recalling the conceptual scheme in Fig. 1.1 (**Chapter 1**), the appropriate combination of different computational techniques can shed light on new discoveries in healthcare and biology (**Chapter 2**). The key contributions of the computational methods proposed in this thesis are presented and discussed in the following.

Classical Image Processing methods (**Chapter 3**) are efficient and still effective, especially in specific situations. Despite their simplicity, region-based and morphological approaches (Sections 3.1 and 3.2, respectively) worked properly when the input images hold the assumptions underlying the algorithms. Global thresholding [415, 525] as well as the combination of region splitting-and-merging with region growing [524] yielded accurate uterine fibroid MR image segmentation results for MRgFUS treatment response evaluation [518]. In clonogenic assays, CHT was suitable for fully automatic well detection in multi-well plates and local thresholding was used for calculating the area covered by the cell colony [411]. The watershed transform was valuable for cell nuclei detection and counting, when dealing with overlapping cells in high-throughput microscopy imaging experiments [530].

Pattern Recognition techniques (**Chapter 4**), with particular focus on unsupervised fuzzy clustering (Section 4.1) and graph-based methods (Section 4.2), were exploited for multi-spectral and multimodal medical imaging data processing. Especially, the FCM clustering algorithm was applied to several clinical contexts, including brain tumor MR image segmentation [413] and necrosis extraction [523] as well as prostate gland segmentation on multispectral MR images [519]. GTVCUT, a CA-based GTV delineation approach, was proposed by exploiting a transition function (based on gradient magnitude) that solves the

shortest path problem [520, 521]. The combination of fuzzy clustering with the RW algorithm can provide valuable solutions for multimodal image segmentation, such as in the case of brain metastasis delineation on co-registered PET/MR images for neuro-radiosurgery treatment planning [526].

Computational Intelligence (**Chapter 5**) was shown to be able to considerably increase the performance of traditional Image Processing techniques. A novel Evolutionary Computation method (Section 5.1) for medical image enhancement, called MedGA [529], was proposed to improve the threshold-based segmentation results [527]. As a matter of fact, MedGA aimed at overcoming the limitations of thresholding approaches that strongly rely on the assumption that the histogram under investigation is bimodal. Moreover, multimodal image co-registration approaches using PSO (Section 5.2) were critically reviewed and analyzed [528]. This global optimization metaheuristic may considerably improve the reliability and robustness of the multimodal medical image co-registration process with respect to local-based search methods.

In the latest years, DNNs (**Chapter 6**) have been exploited in biomedical image analysis achieving state-of-the-art performance. In this context, the generalization abilities of CNNs (Section 6.1) in medical image segmentation for multi-institutional studies was investigated [517]. The generalization abilities of CNNs in medical image segmentation for multi-institutional studies are addressed by conceiving a novel architecture, called USE-Net [516], which integrates adaptive feature recalibration by incorporating Squeeze-and-Excitation blocks into U-Net. Regarding the generation of realistic medical images, the GANs (Section 6.2) were exploited to synthesize completely new brain tumor MR slices [246]. Then, in order to yield real-sized MR images, the PGGANs method achieved promising results in data augmentation for brain tumor MRI detection [247].

Therefore, the synergies emerging from the combination of diverse computer science areas are valuable to develop accurate and comprehensive approaches in clinical and laboratory routine. These computer-assisted biomedical image analysis methods could be beneficial for the definition of imaging biomarkers as well as for quantitative medicine and biology.

Finally, the community of computer scientists in biomedical imaging cannot ignore the interpretability of the results yielded by advanced computational models [86]. As a matter of fact, these compelling issues must be concretely addressed for the adoption of CDSSs in real environments, allowing for the physicians' awareness during complex clinical decision-making tasks [74].

## 7.2 Future work

Quantitative cancer studies represent the immediate future in clinical practice [205, 680]. The proper combination of different imaging modalities provides an in-depth understanding of molecular processes that can guide to personalized medicine [171].

The achievements related to quantitative imaging lead to valuable applications in radiomics [5, 214, 345, 346] and radiogenomics [470, 508, 531, 539] research, so conveying clinically useful information. Especially, radiogenomics studies can provide important information about tumor heterogeneity as well as treatment response [69], by integrating genomic information [689]. In these scenarios, haplotype assembly can be valuable for personalized therapy by relying on Single Nucleotide Polymorphisms (SNPs) [605, 607]. Therefore, in the near future the main objective of biomedical research is moving towards the acquisition and the integration of heterogeneous datasets [344, 556] (i.e., imaging modalities and high-throughput technologies) to gain insights allowing for precision medicine [66].

From the technological perspective, HPC can be an enabling factor for feasible computational solutions in clinical and laboratory practice. GPUs are used today in a wide range of applications in computational biology [439] and medical image analysis [168, 570]. As a matter of fact, GPUs are pervasive, energy-efficient and can dramatically accelerate parallel computing. In the field of biomedical imaging, GPUs are in some cases crucial for enabling practical use of computationally demanding algorithms. For instance, GPUs enabled the training of DNNs in reasonable time [560, 602]. In this context, our research group developed a first GPU-powered version of the online training algorithm for SOMs [329, 330]. We aim at exploiting this efficient method in medical image classification based on textural and structural features [445]. To this end, efficient and reproducible feature computation is a hot-topic in radiomics research [23]. Leveraging GPUs for efficient imaging feature extraction is effective to speed-up the feature extraction [215, 621]. We are currently improving the computational performance of Haralick's feature extraction based on the GLCM [250, 251], by assessing the accuracy relying on state-of-the-art tools for radiation therapy [145] and radiographic phenotyping [627]. With reference to real-time medical diagnosis tools implemented on reconfigurable logic devices, FPGAs can provide higher throughputs with respect to the corresponding software applications [640]. Therefore, FPGA-based solutions could be exploited to accelerate the developed algorithms on specialized processing nodes [189], implementing Systems-on-a-Chip (SoCs) for specific processing tasks [123]. For instance, geometric operations (i.e., reflections, rotations, translations, and uniform scaling) could be parallelized by means of co-processing hardware architectures natively supporting Clifford Algebra [190].

Finally, to address the issues regarding the curse of dimensionality, effective Dimensionality Reduction (DR) techniques, which allow also for online performance, could be required [111]. The Growing Curvilinear Component Analysis (GCCA) [113] enables a non-linear distance preserving reduction technique [112], by means of a self-organized incremental neural network architecture, leading to applications in life sciences [39]. In the case of competing objectives, multi-objective optimization techniques, such as those based on GAs [146, 332], can be exploited for feature selection in complex problems [445].

# Bibliography

- [1] Abdel-Basset, M., Fakhry, A. E., El-Henawy, I., Qiu, T., and Sangaiah, A. K. (2017). Feature and intensity based medical image registration using particle swarm optimization. *J. Med. Syst.*, 41(12):197.
- [2] Abousamra, S., Adar, S., Elia, N., and Shilkrot, R. (2018). Localization and tracking in 4D fluorescence microscopy imagery. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 2290–2298. IEEE.
- [3] Adams, R. and Bischof, L. (1994). Seeded region growing. *IEEE Trans. Pattern Anal.*, 16(6):641–647.
- [4] Adler Jr, J. R., Chang, S. D., Murphy, M. J., Doty, J., Geis, P., and Hancock, S. L. (1997). The Cyberknife: a frameless robotic system for radiosurgery. *Stereotact. Funct. Neurosurg.*, 69(1-4):124–128.
- [5] Aerts, H. J., Velazquez, E. R., Leijenaar, R. T., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., et al. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.*, 5:4006.
- [6] Agnello, L., Comelli, A., Ardizzone, E., and Vitabile, S. (2016). Unsupervised tissue classification of brain MR images for voxel-based morphometry analysis. *Int. J. Imaging Syst. Technol.*, 26(2):136–150.
- [7] Agnello, L., Militello, C., Gagliardo, C., and Vitabile, S. (2015). Radial basis function interpolation for referenceless thermometry enhancement. In *Advances in Neural Networks: Computational and Theoretical Issues*, volume 37 of *Smart Innovation, Systems and Technologies*, pages 195–206. Springer.
- [8] Agu, E., Pedersen, P., Strong, D., Tulu, B., He, Q., Wang, L., and Li, Y. (2013). The smartphone as a medical device: assessing enablers, benefits and challenges. In *Proc. International Workshop of Internet-of-Things Networking and Control (IoT-NC)*, pages 48–52. IEEE.
- [9] Ahmed, H. U., Kirkham, A., Arya, M., Illing, R., Freeman, A., Allen, C., and Emberton, M. (2009). Is it time to consider a role for MRI before prostate biopsy? *Nat. Rev. Clin. Oncol.*, 6(4):197.
- [10] Al-Najdawi, N., Biltawi, M., and Tedmori, S. (2015). Mammogram image visual enhancement, mass segmentation and classification. *Appl. Soft Comput.*, 35:175–185.

- [11] Alaa, A. M., Moon, K. H., Hsu, W., and van der Schaar, M. (2016). ConfidentCare: a clinical decision support system for personalized breast cancer screening. *IEEE Trans. Multimed.*, 18(10):1942–1955.
- [12] AlBadawy, E. A., Saha, A., and Mazurowski, M. A. (2018). Deep learning for segmentation of brain tumors: impact of cross-institutional training and testing. *Med. Phys.*, 45(3):1150–1158.
- [13] Alepis, E. and Lambrinidis, C. (2013). M-health: supporting automated diagnosis and electronic health records. *SpringerPlus*, 2(1):103.
- [14] Ali, M. Z., Awad, N. H., Suganthan, P. N., Shatnawi, A. M., and Reynolds, R. G. (2018). An improved class of real-coded genetic algorithms for numerical optimization. *Neurocomputing*, 275:155–166.
- [15] Allen, P. D., Graham, J., Williamson, D. C., and Hutchinson, C. E. (2006). Differential segmentation of the prostate in MR images using combined 3D shape modelling and voxel classification. In *Proc. International Symposium on Biomedical Imaging (ISBI): From Nano to Macro*, pages 410–413. IEEE.
- [16] Allmendinger, R., Ashlock, D., and Auephanwiriyakul, S. (2018). Computational intelligence techniques in bioinformatics and bioengineering [Guest Editorial]. *IEEE Comput. Intell. Mag.*, 13(3):10–11.
- [17] Altman, M., Kavanaugh, J., Wooten, H., Green, O., DeWees, T., Gay, H., Thorstad, W., Li, H., and Mutic, S. (2015). A framework for automated contour quality assurance in radiation therapy including adaptive techniques. *Phys. Med. Biol.*, 60(13):5199.
- [18] Ambrosini, R. D., Wang, P., and O’dell, W. G. (2010). Computer-aided detection of metastatic brain tumors using automated three-dimensional template matching. *J. Magn. Res. Imaging*, 31(1):85–93.
- [19] Andrews, D. W., Scott, C. B., Sperduto, P. W., Flanders, A. E., Gaspar, L. E., Schell, M. C., Werner-Wasik, M., Demas, W., Ryu, J., Bahary, J.-P., et al. (2004). Whole brain radiation therapy with or without stereotactic radiosurgery boost for patients with one to three brain metastases: phase III results of the RTOG 9508 randomised trial. *Lancet*, 363(9422):1665–1672.
- [20] Angelini, E. D., Clatz, O., Mandonnet, E., Konukoglu, E., Capelle, L., and Duffau, H. (2007). Glioma dynamics and computational models: a review of segmentation, registration, and in silico growth algorithms and their clinical applications. *Curr. Med. Imaging Rev.*, 3(4):262–276.
- [21] Angermueller, C., Pärnamaa, T., Parts, L., and Stegle, O. (2016). Deep learning for computational biology. *Mol. Syst. Biol.*, 12(7):878.
- [22] Antila, K., Nieminen, H. J., Sequeiros, R. B., and Ehnholm, G. (2014). Automatic segmentation for detecting uterine fibroid regions treated with MR-guided high intensity focused ultrasound (MR-HIFU). *Med. Phys.*, 41(7).

- [23] Apte, A. P., Iyer, A., Crispin-Ortuzar, M., Pandya, R., van Dijk, L. V., Spezi, E., Thor, M., Um, H., Veeraraghavan, H., Oh, J. H., et al. (2018). Extension of CERR for computational radiomics: a comprehensive MATLAB platform for reproducible radiomics research. *Med. Phys.*, 45(8):3713–3720.
- [24] Aram, S., Shirvani, R. A., Pasero, E., and Chouikha, M. F. (2016). Implantable medical devices; networking security survey. *J. Internet Serv. Inf. Secur.*, 6(3):40–60.
- [25] Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein generative adversarial networks. In *Proc. International Conference on Machine Learning (ICML)*, pages 214–223.
- [26] Arriaga-Garcia, E. F., Sanchez-Yanez, R. E., and Garcia-Hernandez, M. G. (2014). Image enhancement using bi-histogram equalization with adaptive sigmoid functions. In *Proc. International Conference on Electronics, Communications and Computers (CONI-ELECOMP)*, pages 28–34. IEEE.
- [27] Aslian, H., Sadeghi, M., Mahdavi, S. R., Mofrad, F. B., Astarakee, M., Khaledi, N., and Fadavi, P. (2013). Magnetic resonance imaging-based target volume delineation in radiation therapy treatment planning for brain tumors using localized region-based active contour. *Int. J. Radiat. Oncol. Biol. Phys.*, 87(1):195–201.
- [28] Atherton, T. J. and Kerbyson, D. J. (1999). Size invariant circle detection. *Image Vis. Comput.*, 17(11):795–803.
- [29] Bäck, T. and Schwefel, H. (1993). An overview of evolutionary algorithms for parameter optimization. *Evol. Comput.*, 1(1):1–23.
- [30] Badrinarayanan, V., Kendall, A., and Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(12):2481–2495.
- [31] Bagci, U., Kramer-Marek, G., and Mollura, D. J. (2013a). Automated computer quantification of breast cancer in small-animal models using PET-guided MR image co-segmentation. *EJNMMI Res.*, 3(1):49.
- [32] Bagci, U., Udupa, J. K., Mendhiratta, N., Foster, B., Xu, Z., Yao, J., Chen, X., and Mollura, D. J. (2013b). Joint segmentation of anatomical and functional images: applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images. *Med. Image Anal.*, 17(8):929–945.
- [33] Baker, J. E. (1985). Adaptive selection methods for genetic algorithms. In *Proc. International Conference on Genetic Algorithms and their applications*, pages 101–111, Hillsdale, NJ, USA. L. Erlbaum Associates Inc.
- [34] Baldi, P. and Sadowski, P. J. (2013). Understanding dropout. In *Proc. Advances in Neural Information Processing Systems (NIPS)*, pages 2814–2822.
- [35] Ballangan, C., Wang, X., Fulham, M., Eberl, S., and Feng, D. D. (2013). Lung tumor segmentation in PET images using graph cuts. *Comput. Methods Programs Biomed.*, 109(3):260–268.

- [36] Bandini, S., Mauri, G., and Serra, R. (2001). Cellular automata: From a theoretical parallel computational model to its application to complex systems. *Parallel Comput.*, 27(5):539–553.
- [37] Bankman, I. N. (2009). Part II - Segmentation. In Bankman, I. N., editor, *Handbook of Medical Image Processing and Analysis*, pages 71–72. Academic Press, Burlington, MA, USA, 2 edition.
- [38] Barber, P. R., Vojnovic, B., Kelly, J., Mayes, C. R., Boulton, P., Woodcock, M., and Joiner, M. C. (2001). Automated counting of mammalian cell colonies. *Phys. Med. Biol.*, 46(1):63.
- [39] Barbiero, P., Bertotti, A., Ciravegna, G., Cirrincione, G., Pasero, E., and Piccolo, E. (2019). Supervised gene identification in colorectal cancer. In *Quantifying and Processing Biomedical and Behavioral Signals*, volume 103 of *Smart Innovation, Systems and Technologies*. Springer.
- [40] Barentsz, J. O., Weinreb, J. C., Verma, S., Thoeny, H. C., Tempany, C. M., Shtern, F., Padhani, A. R., Margolis, D., Macura, K. J., Haider, M. A., et al. (2016). Synopsis of the PI-RADS v2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. *Eur. Urol.*, 69(1):41–49.
- [41] Bataineh, B., Abdullah, S. N. H. S., and Omar, K. (2011). An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. *Pattern Recognit. Lett.*, 32(14):1805–1813.
- [42] Bauer, S., Nolte, L.-P., and Reyes, M. (2011). Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 6893 of *LNCS*, pages 354–361. Springer.
- [43] Beavis, A., Gibbs, P., Dealey, R., and Whitton, V. (1998). Radiotherapy treatment planning of brain tumours using MRI alone. *Br. J. Radiol.*, 71(845):544–548.
- [44] Behnel, S., Bradshaw, R., Citro, C., Dalcin, L., Seljebotn, D. S., and Smith, K. (2011). Cython: the best of both worlds. *Comput. Sci. Eng.*, 13(2):31–39.
- [45] Bellman, R. (1998). On a routing problem. *Q. Appl. Math.*, 71(845):544–548.
- [46] Bellon, E. M., Haacke, E. M., Coleman, P. E., Sacco, D. C., Steiger, D. A., and Gangarosa, R. E. (1986). MR artifacts: a review. *Am. J. Roentgenol.*, 147(6):1271–1281.
- [47] Ben-Zadok, N., Riklin-Raviv, T., and Kiryati, N. (2009). Interactive level set segmentation for image-guided therapy. In *Proc. International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 1079–1082. IEEE.
- [48] Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.*, 5(2):157–166.
- [49] Bernard, R., Kanduser, M., Pernus, F., et al. (2001). Model-based automated detection of mammalian cell colonies. *Phys. Med. Biol.*, 46(11):3061.



- [50] Bernon, J., Boudousq, V., Rohmer, J., Fourcade, M., Zanca, M., Rossi, M., and Mariano-Goulart, D. (2001). A comparative study of Powell's and Downhill Simplex algorithms for a fast multimodal surface matching in brain imaging. *Comput. Med. Imaging Graph.*, 25(4):287–297.
- [51] Beucher, S. and Meyer, F. (1993). The morphological approach to segmentation: the watershed transformation. In *Mathematical Morphology in Image Processing*, volume 34, pages 433–481. Marcel Dekker Inc., New York, NY, USA.
- [52] Bezdek, J. C. (1981). Objective function clustering. In *Pattern recognition with fuzzy objective function algorithms*, pages 43–93. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1 edition.
- [53] Bezdek, J. C., Ehrlich, R., and Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.*, 10(2-3):191–203.
- [54] Bezdek, J. C., Hall, L., and Clarke, L. (1993). Review of MR image segmentation techniques using pattern recognition. *Med. Phys.*, 20(4):1033–1048.
- [55] Bhandari, A. K., Kumar, A., Chaudhary, S., and Singh, G. (2016). A novel color image multilevel thresholding based segmentation using nature inspired optimization algorithms. *Expert Syst. Appl.*, 63:112–133.
- [56] Bhattacharjee, K., Naskar, N., Roy, S., and Das, S. (2018). A survey of cellular automata: Types, dynamics, non-uniformity and applications. *Nat. Comput.*, pages 1–29. (In Press).
- [57] Bianco, S., Ciocca, G., and Schettini, R. (2017). Combination of video change detection algorithms by genetic programming. *IEEE Trans. Evol. Comput.*, 21(6):914–928.
- [58] Bidgood Jr, W. D. and Horii, S. C. (1992). Introduction to the ACR-NEMA DICOM standard. *Radiographics*, 12(2):345–355.
- [59] Bloch, N., Madabhushi, A., Huisman, H., Freymann, J., Kirby, J., Grauer, M., et al. (2015). NCI-ISBI 2013 Challenge: automated Segmentation of Prostate Structures. The Cancer Imaging Archive. Online; Accessed on July 30, 2018.
- [60] Borgefors, G. (1986). Distance transformations in digital images. *Comput. Vis. Graph. Image Process.*, 34(3):344–371.
- [61] Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proc. COMPSTAT'2010*, pages 177–186. Springer.
- [62] Boureau, Y.-L., Bach, F., LeCun, Y., and Ponce, J. (2010). Learning mid-level features for recognition. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2559–2566. IEEE.
- [63] Boykov, Y. and Jolly, M.-P. (2000). Interactive organ segmentation using graph cuts. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 1935 of LNCS, pages 276–286. Springer.

- [64] Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239.
- [65] Boykov, Y. Y. and Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *Proc. International Conference on Computer Vision (ICCV)*, volume 1, pages 105–112. IEEE.
- [66] Brady, M., Highnam, R., Irving, B., and Schnabel, J. A. (2016). Oncological image analysis. *Med. Image Anal.*, 33:7–12.
- [67] Bratton, D. and Kennedy, J. (2007). Defining a standard for particle swarm optimization. In *Proc. 2007 IEEE Swarm Intelligence Symposium (SIS)*, pages 120–127. IEEE.
- [68] Breen, E. J., Jones, R., and Talbot, H. (2000). Mathematical morphology: a useful set of tools for image analysis. *Stat. Comput.*, 10(2):105–120.
- [69] Brindle, K. (2008). New approaches for imaging tumour responses to treatment. *Nat. Rev. Cancer*, 8(2):94.
- [70] Brown, R. W., Cheng, Y. C. N., Haacke, E. M., Thompson, M. R., and Venkatesan, R. (2014). *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. John Wiley & Sons.
- [71] Buddemeyer, E. (1975). The physics of diagnostic ultrasound. *Radiol. Clin. North Am.*, 13(3):391–402.
- [72] Burger, M., Müller, J., Papoutsellis, E., and Schönlieb, C.-B. (2014). Total variation regularization in measurement and image space for pet reconstruction. *Inv. Prob.*, 30(10):105003.
- [73] Buttram Jr, V. C. and Reiter, R. C. (1981). Uterine leiomyomata: etiology, symptomatology, and management. *Fertil. Steril.*, 36(4):433.
- [74] Cabitza, F., Rasoini, R., and Gensini, G. F. (2017). Unintended consequences of machine learning in medicine. *J. Am. Med. Assoc.*, 318(6):517–518.
- [75] Caivano, R., Cirillo, P., Balestra, A., Lotumolo, A., Fortunato, G., Macarini, L., Zandolino, A., Vita, G., and Cammarota, A. (2012). Prostate cancer in magnetic resonance imaging: diagnostic utilities of spectroscopic sequences. *J. Med. Imaging Radiat. Oncol.*, 56(6):606–616.
- [76] Calatroni, L., van Gennip, Y., Schönlieb, C.-B., Rowland, H. M., and Flenner, A. (2017). Graph clustering, variational image segmentation methods and Hough transform scale detection for object measurement in images. *J. Math. Imaging Vis.*, 57(2):269–291.
- [77] Cannella, V., Gambino, O., Pirrone, R., and Vitabile, S. (2009). GUI usability in medical imaging. In *Proc. International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*, pages 778–782.
- [78] Cannon, R. L., Dave, J. V., and Bezdek, J. C. (1986). Efficient implementation of the fuzzy c-means clustering algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-8(2):248–255.

- [79] Canny, J. (1986). A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-8(6):679–698.
- [80] Cao, Y., Sundgren, P. C., Tsien, C. I., Chenevert, T. T., and Junck, L. (2006). Physiologic and metabolic magnetic resonance imaging in gliomas. *J. Clin. Oncol.*, 24(8):1228–1235.
- [81] Carbonaro, A. and Zingaretti, P. (1999). A comprehensive approach to image-contrast enhancement. In *Proc. International Conference on Image Analysis and Processing (ICIAP)*, pages 241–246. IEEE.
- [82] Cárdenes, R., de Luis-García, R., and Bach-Cuadra, M. (2009). A multidimensional segmentation evaluation for medical image data. *Comput. Methods Programs Biomed.*, 96(2):108–124.
- [83] Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.*, 7(10):R100.
- [84] Castelli, M., Trujillo, L., Vanneschi, L., and Popovič, A. (2016). Prediction of relative position of CT slices using a computational intelligence system. *Appl. Soft Comput.*, 46:537–542.
- [85] Castelli, M., Vanneschi, L., and Silva, S. (2014). Prediction of the Unified Parkinson’s Disease Rating Scale assessment using a genetic programming system with geometric semantic genetic operators. *Expert Syst. Appl.*, 41(10):4608–4616.
- [86] Castelvechhi, D. (2016). Can we open the black box of AI? *Nat. News*, 538(7623):20.
- [87] Catana, C., Procissi, D., Wu, Y., Judenhofer, M. S., Qi, J., Pichler, B. J., Jacobs, R. E., and Cherry, S. R. (2008). Simultaneous in vivo positron emission tomography and magnetic resonance imaging. *Proc. Natl. Acad. Sci.*, 105(10):3705–3710.
- [88] Cavazza, J., Lane, C., Haeffele, B. D., Murino, V., and Vidal, R. (2017). An analysis of dropout for matrix factorization. *arXiv preprint arXiv:1710.03487*.
- [89] Cazzaniga, P., Nobile, M., and Besozzi, D. (2015). The impact of particles initialization in PSO: parameter estimation as a case in point. In *Proc. 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–8, Niagara Falls, Canada.
- [90] Celery Project. (2018). Celery Distributed Task Queue. <http://www.celeryproject.org/>. Online; accessed 20 October 2018.
- [91] Chaira, T. (2014). An improved medical image enhancement scheme using Type II fuzzy set. *Appl. Soft Comput.*, 25:293–308.
- [92] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Trans. Image Process.*, 10(2):266–277.
- [93] Chan, V. and Perlas, A. (2011). Basics of ultrasound imaging. In Narouze, S., editor, *Atlas of ultrasound-guided procedures in interventional pain management*, pages 13–19. Springer.

- [94] Chang, Y., Chen, R., Yang, Q., Gao, X., Xu, C., Lu, J., and Sun, Y. (2017). Peripheral zone volume ratio (PZ-ratio) is relevant with biopsy results and can increase the accuracy of current diagnostic modality. *Oncotarget*, 8(21):34836.
- [95] Chang, Y.-L. and Li, X. (1994). Adaptive image region-growing. *IEEE Trans. Image Process.*, 3(6):868–872.
- [96] Chapman, A. and Ter Haar, G. (2007). Thermal ablation of uterine fibroids using MR-guided focused ultrasound—a truly non-invasive treatment modality. *Eur. Radiol.*, 17(10):2505–2511.
- [97] Chaudhury, K. N., Sage, D., and Unser, M. (2011). Fast  $o(1)$  bilateral filtering using trigonometric range kernels. *IEEE Trans Image Process*, 20(12):3376–3382.
- [98] Chen, C., Abdelnour-Nocera, J., Wells, S., and Pan, N. (2009a). Usability practice in medical imaging application development. In *Proc. Symposium of the Austrian HCI and Usability Engineering Group*, volume 5889 of LNCS, pages 405–415. Springer.
- [99] Chen, C. M., Chen, C. C., Wu, M. C., Horng, G., Wu, H. C., H., S. H., and Ho, H. Y. (2015). Automatic contrast enhancement of brain MR images using hierarchical correlation histogram analysis. *J. Med. Biol. Eng.*, 35(6):724–734.
- [100] Chen, J., Yu, W., Tian, J., Chen, L., and Zhou, Z. (2018). Image contrast enhancement using an artificial bee colony algorithm. *Swarm Evol. Comput.*, 38:287–294.
- [101] Chen, S. D. and Ramli, A. R. (2003). Minimum mean brightness error bi-histogram equalization in contrast enhancement. *IEEE Trans. Consumer Electron.*, 49(4):1310–1319.
- [102] Chen, Y.-W., Mimori, A., and Lin, C.-L. (2009b). Hybrid particle swarm optimization for 3-D image registration. In *Proc. International Conference on Image Processing (ICIP)*, pages 1753–1756. IEEE.
- [103] Chiang, P.-J., Tseng, M.-J., He, Z.-S., and Li, C.-H. (2015). Automated counting of bacterial colonies by image analysis. *J. Microbiol. Methods*, 108:74–82.
- [104] Chilali, O., Ouzzane, A., Diaf, M., and Betrouni, N. (2014). A survey of prostate modeling for image analysis. *Comput. Biol. Med.*, 53:190–202.
- [105] Chin, L. S., Ma, L., and DiBiase, S. (2001). Radiation necrosis following Gamma Knife surgery: a case-controlled comparison of treatment parameters and long-term clinical follow up. *J. Neurosurg.*, 94(6):899–904.
- [106] Choi, Y. J., Kim, J. K., Kim, N., Kim, K. W., Choi, E. K., and Cho, K.-S. (2007). Functional MR imaging of prostate cancer. *Radiographics*, 27(1):63–75.
- [107] Chow, C. K. and Kaneko, T. (1972). Automatic boundary detection of the left ventricle from cineangiograms. *Comput. Biomed. Res.*, 5(4):388–410.
- [108] Chuquicusma, M. J. M., Hussein, S., Burt, J., and Bagci, U. (2017). How to fool radiologists with generative adversarial networks? A visual Turing test for lung cancer diagnosis. *arXiv preprint arXiv:1710.09762v1*.

- [109] Ciocca, G., Napoletano, P., and Schettini, R. (2017). Food recognition: a new dataset, experiments, and results. *IEEE J. Biomed. Health Inform.*, 21(3):588–598.
- [110] Ciresan, D., Giusti, A., Gambardella, L. M., and Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. In *Proc. Advances in Neural Information Processing Systems (NIPS)*, pages 2843–2851.
- [111] Cirrincione, G., Hérault, J., and Randazzo, V. (2015). The on-line curvilinear component analysis (onCCA) for real-time data reduction. In *Proc. International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.
- [112] Cirrincione, G., Randazzo, V., and Pasero, E. (2018a). Growing Curvilinear Component Analysis (GCCA) for dimensionality reduction of nonstationary data. In *Multidisciplinary Approaches to Neural Computing*, volume 69 of *Smart Innovation, Systems and Technologies*, pages 151–160. Springer.
- [113] Cirrincione, G., Randazzo, V., and Pasero, E. (2018b). The Growing Curvilinear Component Analysis (GCCA) neural network. *Neural Netw.*, 103:108–117.
- [114] Cizžek, J., Herholz, K., Vollmar, S., Schrader, R., Klein, J., and Heiss, W.-D. (2004). Fast and robust registration of PET and MR images of human brain. *NeuroImage*, 22(1):434–442.
- [115] Clark, T., Zhang, J., Baig, S., Wong, A., Haider, M. A., and Khalvati, F. (2017). Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks. *J. Med. Imaging*, 4(4):041307.
- [116] Clarke, M. L., Burton, R. L., Hill, A. N., Litorja, M., Nahm, M. H., and Hwang, J. (2010). Low-cost, high-throughput, automated counting of bacterial colonies. *Cytom. Part A*, 77(8):790–797.
- [117] Clerc, M. (2006). Particle swarm optimization. In *International Scientific and Technical Encyclopaedia*. Wiley, Hoboken, NJ, USA.
- [118] Clerc, M. and Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Trans. Evol. Comput.*, 6(1):58–73.
- [119] Cline, H. E., Schenek, J., Hynynen, K., and Watkins, R. D. (1992). MR-guided focused ultrasound surgery. *J. Comput. Assist. Tomogr.*, 16:956–956.
- [120] Coelho, L. P. (2013). Mahotas: Open source software for scriptable computer vision. *J. Open. Res. Softw.*, 1(1):e3.
- [121] Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., and Marchal, G. (1995). Automated multi-modality image registration based on information theory. In *Proc. 14th International Conference on Information Processing in Medical Imaging (IPMI)*, pages 263–274.
- [122] Comelli, A., Stefano, A., Russo, G., Sabini, M. G., Ippolito, M., Bignardi, S., Petrucci, G., and Yezzi, A. (2018). A smart and operator independent system to delineate tumours in Positron Emission Tomography scans. *Comput. Biol. Med.*, 102:1–15.

- [123] Cong, J., Sarkar, V., Reinman, G., and Bui, A. (2011). Customizable domain-specific computing. *IEEE Des. Test Comput.*, 28(2):6–15.
- [124] Conti, V., Militello, C., Sorbello, F., and Vitabile, S. (2010). A frequency-based approach for features fusion in fingerprint and iris multimodal biometric identification systems. *IEEE Trans. Syst. Man. Cybern. Part C*, 40(4):384–395.
- [125] Conti, V., Rundo, L., Militello, C., Mauri, G., and Vitabile, S. (2017). Resource-efficient hardware implementation of a neural-based node for automatic fingerprint classification. *J. Wirel. Mob. Netw. Ubiquit. Comput. Depend. Appl.*, 8:19–36.
- [126] Cootes, T. F., Hill, A., Taylor, C. J., and Haslam, J. (1994). Use of active shape models for locating structures in medical images. *Image Vis. Comput.*, 12(6):355–366.
- [127] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models-their training and application. *Comput. Vis. Image Underst.*, 61(1):38–59.
- [128] Cootes, T. F., Twining, C. J., Petrovic, V. S., Schestowitz, R., and Taylor, C. J. (2005). Groupwise construction of appearance models using piece-wise affine deformations. In *Proc. British Machine Vision Conference (BMVC)*, volume 2, pages 879–888.
- [129] Cosío, F. A. (2008). Automatic initialization of an active shape model of the prostate. *Med. Image Anal.*, 12(4):469–483.
- [130] Costa, P., Galdran, A., Meyer, M. I., Niemeijer, M., Abràmoff, M., Mendonça, A. M., and Campilho, A. (2018). End-to-end adversarial retinal image synthesis. *IEEE Trans. Med. Imaging*, 37(3):781–791.
- [131] Courant, R. and Hilbert, D. (1989). *Methods of Mathematical Physics: Partial Differential Equations*, volume 2. John Wiley & Sons, Inc., Hoboken, NJ, USA, 1 edition.
- [132] Cramer, S. F. and Patel, A. (1990). The frequency of uterine leiomyomas. *Am. J. Clin. Pathol.*, 94(4):435–438.
- [133] Criminisi, A. (2016). Machine learning for medical images analysis. *Med. Image Anal.*, 33:91–93.
- [134] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.*, 2(4):303–314.
- [135] Czerwinski, R. N., Jones, D. L., and O’Brien, W. D. (1998). Line and boundary detection in speckle images. *IEEE Trans. Image Process.*, 7(12):1700–1714.
- [136] Czerwinski, R. N., Jones, D. L., and O’Brien, W. D. (1999). Detection of lines and boundaries in speckle images-application to medical ultrasound. *IEEE Trans. Med. Imaging*, 18(2):126–136.
- [137] Daffner, R. H. (1980). Visual illusions in computed tomography: phenomena related to Mach effect. *Am. J. Roentgenol.*, 134(2):261–264.
- [138] Dahle, J., Kakar, M., Steen, H. B., and Kaalhus, O. (2004). Automated counting of mammalian cell colonies by means of a flat bed scanner and image processing. *Cytom. Part A*, 60(2):182–188.

- [139] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 886–893. IEEE.
- [140] Dao, D., Fraser, A. N., Hung, J., Ljosa, V., Singh, S., and Carpenter, A. E. (2016). CellProfiler Analyst: interactive data exploration, analysis and classification of large biological image sets. *Bioinformatics*, 32(20):3210–3212.
- [141] Das, A. and Bhattacharya, M. (2011). Affine-based registration of CT and MR modality images of human brain using multiresolution approaches: comparative study on genetic algorithm and particle swarm optimization. *Neural Comput. Appl.*, 20(2):223–237.
- [142] Davenport, J., Bezdek, J., and Hathaway, R. (1988). Parameter estimation for finite mixture distributions. *Comput. Math. Appl.*, 15(10):819–828.
- [143] de Araujo, A. F., Constantinou, C. E., and Tavares, J. M. R. S. (2014). New artificial life model for image enhancement. *Expert Syst. Appl.*, 41(13):5892–5906.
- [144] De los Reyes, J. C. and Schönlieb, C.-B. (2013). Image denoising: Learning the noise model via nonsmooth pde-constrained optimization. *Inv. Prob. Imaging*, 7(4):1183–1214.
- [145] Deasy, J. O., Blanco, A. I., and Clark, V. H. (2003). CERR: a computational environment for radiotherapy research. *Med. Phys.*, 30(5):979–985.
- [146] Deb, K. and Jain, H. (2014). An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: solving problems with box constraints. *IEEE Trans. Evol. Comput.*, 18(4):577–601.
- [147] Delbeke, D. (1999). Oncological applications of FDG PET imaging: brain tumors, colorectal cancer lymphoma and melanoma. *J. Nucl. Med.*, 40(4):591–603.
- [148] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, pages 1–38.
- [149] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.*, 7(Jan):1–30.
- [150] Dhermain, F. G., Hau, P., Lanfermann, H., Jacobs, A. H., and van den Bent, M. J. (2010). Advanced MRI and PET imaging for assessment of treatment response in patients with gliomas. *Lancet Neurol.*, 9(9):906–920.
- [151] Di Stefano, L. and Bulgarelli, A. (1999). A simple and efficient connected components labeling algorithm. In *Proc. International Conference on Image Analysis and Processing (ICIAP)*, pages 322–327. IEEE.
- [152] Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.
- [153] Dodziuk, J. (1984). Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Am. Math. Soc.*, 284(2):787–794.

- [154] Dou, W., Ruan, S., Chen, Y., Bloyet, D., and Constans, J.-M. (2007). A framework of fuzzy information fusion for the segmentation of brain tumor tissues on MR images. *Image Vis. Comput.*, 25(2):164–171.
- [155] Draa, A. and Bouaziz, A. (2014). An artificial bee colony algorithm for image contrast enhancement. *Swarm Evol. Comput.*, 16:69–84.
- [156] Drias, Y., Kechid, S., and Pasi, G. (2016). Bee swarm optimization for medical Web information foraging. *J. Med. Syst.*, 40(2):40.
- [157] Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.*, 12(Jul):2121–2159.
- [158] Duda, R. O. and Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15.
- [159] Duncan, J. S. and Ayache, N. (2000). Medical image analysis: progress over two decades and the challenges ahead. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(1):85–106.
- [160] Earnest IV, F., Kelly, P., Scheithauer, B., Kall, B., Cascino, T., Ehman, R., Forbes, G., and Axley, P. (1988). Cerebral astrocytomas: histopathologic correlation of MR and CT contrast enhancement with stereotactic biopsy. *Radiology*, 166(3):823–827.
- [161] Eberhart, R. and Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Proc. International Symposium on Micro Machine and Human Science (MHS)*, pages 39–43. IEEE.
- [162] Eberhart, R. C. and Shi, Y. (2000). Comparing inertia weights and constriction factors in particle swarm optimization. In *Proc. Congress on Evolutionary Computation (CEC)*, pages 84–88.
- [163] Eberhart, R. C., Shi, Y., and Kennedy, J. (2001). *Swarm Intelligence*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 1 edition.
- [164] Edelstein, W., Glover, G., Hardy, C., and Redington, R. (1986). The intrinsic signal-to-noise ratio in NMR imaging. *Magn. Reson. Med.*, 3(4):604–618.
- [165] Egorov, V., van Raalte, H., and Sarvazyan, A. P. (2010). Vaginal tactile imaging. *IEEE Trans. Biomed. Eng.*, 57(7):1736–1744.
- [166] Eiben, A. E. and Smith, J. (2015). From evolutionary computation to the evolution of things. *Nature*, 521(7553):476–482.
- [167] Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancy, J., Arbuck, S., Gwyther, S., Mooney, M., et al. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer*, 45(2):228–247.
- [168] Eklund, A., Dufort, P., Forsberg, D., and LaConte, S. M. (2013). Medical image processing on the GPU—past, present and future. *Med. Image Anal.*, 17(8):1073–1094.



- [169] El Naqa, I., Yang, D., Apte, A., Khullar, D., Mutic, S., Zheng, J., Bradley, J. D., Grigsby, P., and Deasy, J. O. (2007). Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning a. *Med. Phys.*, 34(12):4738–4749.
- [170] Eliceiri, K. W., Berthold, M. R., Goldberg, I. G., Ibáñez, L., Manjunath, B. S., Martone, M. E., Murphy, R. F., Peng, H., Plant, A. L., Roysam, B., et al. (2012). Biological imaging software tools. *Nat. Methods*, 9(7):697.
- [171] Evanko, D. (2008). Two pictures are better than one. *Nat. Methods*, 5(5):377.
- [172] Evans, P. M. (2008). Anatomical imaging for radiotherapy. *Phys. Med. Biol.*, 53(12):R151.
- [173] Even, A. J., van der Stoep, J., Zegers, C. M., Reymen, B., Troost, E. G., Lambin, P., and van Elmpt, W. (2015). PET-based dose painting in non-small cell lung cancer: comparing uniform dose escalation with boosting hypoxic and metabolically active sub-volumes. *Radiother. Oncol.*, 116(2):281–286.
- [174] Fabijańska, A. (2016). A novel approach for quantification of time–intensity curves in a DCE-MRI image series with an application to prostate cancer. *Comput. Biol. Med.*, 73:119–130.
- [175] Faísca, N. P., Kouramas, K. I., Saraiva, P. M., Rustem, B., and Pistikopoulos, E. N. (2008). A multi-parametric programming approach for constrained dynamic programming problems. *Optim. Lett.*, 2(2):267–280.
- [176] Faísca, N. P., Saraiva, P. M., Rustem, B., and Pistikopoulos, E. N. (2009). A multi-parametric programming approach for multilevel hierarchical and decentralised optimisation problems. *Comput. Manag. Sci.*, 6(4):377–397.
- [177] Fallahi, A., Pooyan, M., Ghanaati, H., Oghabian, M. A., Khotanlou, H., Shakiba, M., Jalali, A. H., and Firouznia, K. (2011). Uterine segmentation and volume measurement in uterine fibroid patients’ MRI using fuzzy c-mean algorithm and morphological operations. *Iran. J. Radiol.*, 8(3):150.
- [178] Fallahi, A., Pooyan, M., Khotanlou, H., Hashemi, H., Firouznia, K., and Oghabian, M. A. (2010). Uterine fibroid segmentation on multiplan MRI using FCM, MPFCM and morphological operations. In *Proc. 2nd International Conference on Computer Engineering and Technology (ICCET)*, volume 7, pages V7–1. IEEE.
- [179] Faruquzzaman, A. B. M., Paiker, N. R., Arafat, J., Karim, Z., and Ali, M. A. (2008). Object segmentation based on split and merge algorithm. In *Proc. IEEE Region 10 Conference TENCN*, pages 1–4. IEEE.
- [180] Fawzi, A., Moosavi Dezfooli, S. M., and Frossard, P. (2017). The robustness of deep networks—a geometric perspective. *IEEE Signal Process. Mag.*, 34(6).
- [181] Felzenszwalb, P. F. and Huttenlocher, D. P. (2012). Distance transforms of sampled functions. *Theory Comput.*, 8(1):415–428.

- [182] Fenster, A. and Chiu, B. (2005). Evaluation of segmentation algorithms for medical imaging. In *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7186–7189. IEEE.
- [183] Ferrari, R. J. (2013). Off-line determination of the optimal number of iterations of the robust anisotropic diffusion filter applied to denoising of brain mr images. *Med. Biol. Eng. Comput.*, 51(1-2):71–88.
- [184] Firbank, M., Coulthard, A., Harrison, R., and Williams, E. (1999). A comparison of two methods for measuring the signal to noise ratio on MR images. *Phys. Med. Biol.*, 44(12):N261.
- [185] Fletcher-Heath, L. M., Hall, L. O., Goldgof, D. B., and Murtagh, F. R. (2001). Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artif. Intell. Med.*, 21(1-3):43–63.
- [186] Ford, L. R. and Fulkerson, D. R. (1956). Maximal flow through a network. *Can. J. Math.*, 8(3):399–404.
- [187] Fornefett, M., Rohr, K., and Stiehl, H. S. (2001). Radial basis functions with compact support for elastic registration of medical images. *Image Vis. Comput.*, 19(1-2):87–96.
- [188] Frahm, J., Merboldt, K.-D., and Hänicke, W. (1993). Functional MRI of human brain activation at high spatial resolution. *Magn. Reson. Med.*, 29(1):139–144.
- [189] Franchini, S., Gentile, A., Sorbello, F., Vassallo, G., and Vitabile, S. (2013). Design and implementation of an embedded coprocessor with native support for 5D, quadruple-based Clifford algebra. *IEEE Trans. Comput.*, 62(12):2366–2381.
- [190] Franchini, S., Gentile, A., Sorbello, F., Vassallo, G., and Vitabile, S. (2015). ConformalALU: a conformal geometric algebra coprocessor for medical image processing. *IEEE Trans. Comput.*, 64(4):955–970.
- [191] Franken, N. A. P., Rodermond, H. M., Stap, J., Haveman, J., and Van Bree, C. (2006). Clonogenic assay of cells in vitro. *Nat. Prot.*, 1(5):2315.
- [192] Freedman, D. and Zhang, T. (2005). Interactive graph cut based segmentation with shape priors. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 755–762. IEEE.
- [193] Freshney, R. I. (2016). *Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications*. Wiley-Blackwell, New York, NY, USA, 7 edition.
- [194] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., and Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321:321–331.
- [195] Fukushima, K. (1980). Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.*, 36(4):193–202.

- [196] Fukushima, K., Miyake, S., and Ito, T. (1983). Neocognitron: a neural network model for a mechanism of visual pattern recognition. *IEEE Trans. Syst. Man Cybern.*, SMC-13(5):826–834.
- [197] Futterer, J. J., Heijmink, S. W., Scheenen, T. W., Veltman, J., Huisman, H. J., Vos, P., de Kaa, C. A. H.-V., Witjes, J. A., Krabbe, P. F., Heerschap, A., et al. (2006). Prostate cancer localization with dynamic contrast-enhanced MR imaging and proton MR spectroscopic imaging. *Radiology*, 241(2):449–458.
- [198] Gallivanone, F., Stefano, A., Grosso, E., Canevari, C., Gianolli, L., Messa, C., Gilardi, M. C., and Castiglioni, I. (2011). PVE correction in PET-CT whole-body oncological studies from PVE-affected images. *IEEE Trans. Nucl. Sci.*, 58(3):736–747.
- [199] Gambhir, S. S., Czernin, J., Schwimmer, J., Silverman, D. H., Coleman, R. E., and Phelps, M. E. (2001). A tabulated summary of the FDG PET literature. *J. Nucl. Med.*, 42(5 suppl):1S–93S.
- [200] Gambino, O., Rundo, L., Cannella, V., Vitabile, S., and Pirrone, R. (2018). A framework for data-driven adaptive GUI generation based on DICOM. *J. Biomed. Inform.*, 88:37–52.
- [201] Gambino, O., Vitabile, S., Re, G. L., La Tona, G., Librizzi, S., Pirrone, R., Ardizzone, E., and Midiri, M. (2010). Automatic volumetric liver segmentation using texture based region growing. In *Proc. 2010 International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*, pages 146–152. IEEE.
- [202] Gandhamal, A., Talbar, S., Gajre, S., Hani, A. F. M., and Kumar, D. (2017). Local gray level S-curve transformation—a generalized contrast enhancement technique for medical images. *Comput. Biol. Med.*, 83:120–133.
- [203] Gao, Y., Sandhu, R., Fichtinger, G., and Tannenbaum, A. R. (2010). A coupled global registration and segmentation framework with application to magnetic resonance prostate imagery. *IEEE Trans. Med. Imaging*, 29(10):1781–1794.
- [204] Garcia, D., Spruill, L. S., Irshad, A., Wood, J., Kepecs, D., and Klauber-DeMore, N. (2018). The value of a second opinion for breast cancer patients referred to a National Cancer Institute (NCI)-designated cancer center with a multidisciplinary breast tumor board. *Ann. Surg. Oncol.*, pages 1–5.
- [205] Gatenby, R. A., Grove, O., and Gillies, R. J. (2013). Quantitative imaging in cancer evolution and ecology. *Radiology*, 269(1):8–14.
- [206] Geissmann, Q. (2013). OpenCFU, a new free and open-source software to count cell colonies and other circular objects. *PLoS One*, 8(2):e54072.
- [207] Geman, D., Geman, S., Hallonquist, N., and Younes, L. (2015). Visual Turing test for computer vision systems. *Proc. Natl. Acad. Sci.*, 112(12):3618–3623.
- [208] Georgescu, W., Wikswo, J. P., and Quaranta, V. (2011). CellAnimation: an open source MATLAB framework for microscopy assays. *Bioinformatics*, 28(1):138–139.

- [209] Georgiadis, P., Cavouras, D., Kalatzis, I., Daskalakis, A., Kagadis, G. C., Sifaki, K., Malamas, M., Nikiforidis, G., and Solomou, E. (2008). Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features. *Comput. Methods Programs Biomed*, 89(1):24–32.
- [210] Gerber, D. E. and Chan, T. A. (2008). Recent advances in radiation therapy. *Am. Fam. Physician*, 78(11):1254–1262.
- [211] Geyer, L. L., Schoepf, U. J., Meinel, F. G., Nance Jr, J. W., Bastarrika, G., Leipsic, J. A., Paul, N. S., Rengo, M., Laghi, A., and De Cecco, C. N. (2015). State of the art: iterative CT reconstruction techniques. *Radiology*, 276(2):339–357.
- [212] Ghita, O. and Whelan, P. F. (2002). Computational approach for edge linking. *J. Electron. Imaging*, 11(4):479–486.
- [213] Ghose, S., Oliver, A., Martí, R., Lladó, X., Vilanova, J. C., Freixenet, J., et al. (2012). A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images. *Comput. Meth. Prog. Biomed.*, 108(1):262–287.
- [214] Gillies, R. J., Kinahan, P. E., and Hricak, H. (2015). Radiomics: images are more than pictures, they are data. *Radiology*, 278(2):563–577.
- [215] Gipp, M., Marcus, G., Harder, N., Suratane, A., Rohr, K., König, R., and Männer, R. (2012). Haralick’s texture features computation accelerated by GPUs for biological applications. In *Modeling, Simulation and Optimization of Complex Processes*, pages 127–137. Springer.
- [216] Golberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Boston, MA, USA.
- [217] Gomes, V. M., Pereira, A. C., Saraiva, P. M., and Reis, M. S. (2012). Development of generalized platforms for the analysis of complex datasets. *Qual. Reliab. Eng. Int.*, 28(5):508–523.
- [218] Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*. Prentice Hall Press, Upper Saddle River, NJ, USA, 3 edition.
- [219] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. In *Proc. Advances in Neural Information Processing Systems (NIPS)*, pages 2672–2680. <http://papers.nips.cc/paper/5423-generative-adversarial-nets>.
- [220] Govindaraj, V. and Murugan, P. R. (2014). A complete automated algorithm for segmentation of tissues and identification of tumor region in T1, T2, and FLAIR brain images using optimization and clustering techniques. *Int. J. Imaging Syst. Technol.*, 24(4):313–325.
- [221] Grady, L. (2006). Random walks for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11):1768–1783.

- [222] Grah, J. S., Harrington, J. A., Koh, S. B., Pike, J. A., Schreiner, A., Burger, M., Schönlieb, C.-B., and Reichelt, S. (2017). Mathematical imaging methods for mitosis analysis in live-cell phase contrast microscopy. *Methods*, 115:91–99.
- [223] Grau, V., Mewes, A., Alcaniz, M., Kikinis, R., and Warfield, S. K. (2004). Improved watershed transform for medical image segmentation using prior information. *IEEE Trans. Med. Imaging*, 23(4):447–458.
- [224] Gray, J. E., Feldman, H., Reti, S., Markson, L., Lu, X., Davis, R. B., and Safran, C. A. (2011). Using Digital Crumbs from an Electronic Health Record to identify, study and improve health care teams. In *AMIA Annu. Symp. Proc.*, volume 2011, page 491. American Medical Informatics Association.
- [225] Greene, G. M., Hitchon, P. W., Schelper, R. L., Yuh, W., and Dyste, G. N. (1989). Diagnostic yield in CT-guided stereotactic biopsy of gliomas. *J. Neurosurg.*, 71(4):494–497.
- [226] Greig, D. M., Porteous, B. T., and Seheult, A. H. (1989). Exact maximum a posteriori estimation for binary images. *J. Royal Stat. Soc. Series B*, 51(2):271–279.
- [227] Gross, M. W., Weber, W. A., Feldmann, H. J., Bartenstein, P., Schwaiger, M., and Molls, M. (1998). The value of F-18-fluorodeoxyglucose PET for the 3-D radiation treatment planning of malignant gliomas. *Int. J. Radiat. Oncol. Biol. Phys.*, 41(5):989–995.
- [228] Grosu, A.-L. and Weber, W. A. (2010). PET for radiation treatment planning of brain tumours. *Radiother. Oncol.*, 96(3):325–327.
- [229] Grosu, A. L., Weber, W. A., Franz, M., Stärk, S., Piert, M., Thamm, R., Gumprecht, H., Schwaiger, M., Molls, M., and Nieder, C. (2005a). Reirradiation of recurrent high-grade gliomas using amino acid PET (SPECT)/CT/MRI image fusion to determine gross tumor volume for stereotactic fractionated radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 63(2):511–519.
- [230] Grosu, A.-L., Weber, W. A., Riedel, E., Jeremic, B., Nieder, C., Franz, M., Gumprecht, H., Jaeger, R., Schwaiger, M., and Molls, M. (2005b). L-(methyl-11C) methionine positron emission tomography for target delineation in resected high-grade gliomas before radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 63(1):64–74.
- [231] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. (2017). Improved training of Wasserstein GANs. In *Proc. Advances in Neural Information Processing Systems (NIPS)*, pages 5769–5779.
- [232] Guo, J. (2018). Smartphone-powered electrochemical biosensing dongle for emerging medical IoTs application. *IEEE Trans. Ind. Inform.*, 14(6):2592–2597.
- [233] Guo, Y., Gao, Y., and Shen, D. (2016). Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Trans. Med. Imaging*, 35(4):1077–1089.

- [234] Guyon, J.-P., Foskey, M., Kim, J., Firat, Z., Davis, B., Haneke, K., and Aylward, S. R. (2003). VETOT, volume estimation and tracking over time: framework and validation. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2879 of *LNCS*, pages 142–149. Springer.
- [235] Guzmán, C., Bagga, M., Kaur, A., Westermarck, J., and Abankwa, D. (2014). ColonyArea: an ImageJ plugin to automatically quantify colony formation in clonogenic assays. *PLoS One*, 9(3):e92444.
- [236] Haeffele, B. D. and Vidal, R. (2017). Global optimality in neural network training. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7331–7339. IEEE.
- [237] Haffner, J., Potiron, E., Bouyé, S., Puech, P., Leroy, X., Lemaitre, L., and Villers, A. (2009). Peripheral zone prostate cancers: location and intraprostatic patterns of spread at histopathology. *Prostate*, 69(3):276–282.
- [238] Hagmann, P., Jonasson, L., Maeder, P., Thiran, J.-P., Wedeen, V. J., and Meuli, R. (2006). Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics*, 26(suppl\_1):S205–S223.
- [239] Hahnloser, R. H., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., and Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*, 405(6789):947.
- [240] Haider, M. A., Van Der Kwast, T. H., Tanguay, J., Evans, A. J., Hashmi, A.-T., Lockwood, G., and Trachtenberg, J. (2007). Combined T2-weighted and diffusion-weighted MRI for localization of prostate cancer. *AJR Am. J. Roentgenol.*, 189(2):323–328.
- [241] Hale, A. T., Wang, L., Strother, M. K., and Chambless, L. B. (2018). Differentiating meningioma grade by imaging features on magnetic resonance imaging. *J. Clin. Neurosci.*, 48:71–75.
- [242] Hall, L. O., Bensaid, A. M., Clarke, L. P., Velthuizen, R. P., Silbiger, M. S., and Bezdek, J. C. (1992). A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. *IEEE Trans. Neural Netw.*, 3(5):672–682.
- [243] Hamamci, A., Kucuk, N., Karaman, K., Engin, K., and Unal, G. (2012). Tumor-Cut: segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications. *IEEE Trans. Med. Imaging*, 31(3):790–804.
- [244] Hamamci, A., Unal, G., Kucuk, N., and Engin, K. (2010). Cellular automata segmentation of brain tumors on post contrast MR images. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 6363 of *LNCS*, pages 137–146. Springer.
- [245] Hambrock, T., Vos, P. C., Hulsbergen-van de Kaa, C. A., Barentsz, J. O., and Huisman, H. J. (2013). Prostate cancer: computer-aided diagnosis with multiparametric 3-T MR imaging—effect on observer performance. *Radiology*, 266(2):521–530.

- [246] Han, C., Hayashi, H., Rundo, L., Araki, R., Shimoda, W., Muramatsu, S., et al. (2018a). GAN-based synthetic brain MR image generation. In *Proc. International Symposium on Biomedical Imaging (ISBI)*, pages 734–738. IEEE.
- [247] Han, C., Rundo, L., Araki, R., Furukawa, Y., Mauri, G., Nakayama, H., and Hayashi, H. (2018b). Infinite brain MR images: PGGAN-based data augmentation for tumor detection. In *28th Italian Workshop on Neural Networks (WIRN) 2018, Vietri sul Mare (SA), Italy, June 13-15, 2018*, Smart Innovation, Systems and Technologies. Springer. (Accepted manuscript).
- [248] Han, D., Bayouth, J., Song, Q., Taurani, A., Sonka, M., Buatti, J., and Wu, X. (2011). Globally optimal tumor segmentation in PET-CT images: a graph-based co-segmentation method. In *Proc. International Conference on Information Processing in Medical Imaging (IPMI)*, volume 6801 of *LNCS*, pages 245–256. Springer.
- [249] Harada, T., Harada, K., and Ueyama, Y. (2012). The enhancement of tumor radioreponse by combined treatment with cepharanthine is accompanied by the inhibition of DNA damage repair and the induction of apoptosis in oral squamous cell carcinoma. *Int. J. Oncol.*, 41(2):565–572.
- [250] Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proc. IEEE*, 67(5):786–804.
- [251] Haralick, R. M., Shanmugam, K., et al. (1973). Textural features for image classification. *IEEE Trans. Syst. Man Cybern.*, SMC-3(6):610–621.
- [252] Hashemi, S., Kiani, S., Noroozi, N., and Moghaddam, M. E. (2010). An image contrast enhancement method based on genetic algorithm. *Pattern Recognit. Lett.*, 31(13):1816–1824.
- [253] Hassanien, A. E., Moftah, H. M., Azar, A. T., and Shoman, M. (2014). MRI breast cancer diagnosis hybrid approach using adaptive ant-based segmentation and multilayer perceptron neural networks classifier. *Appl. Soft Comput.*, 14(Part A):62–71.
- [254] Hatipoglu, N. and Bilgin, G. (2017). Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships. *Med. Biol. Eng. Comput.*, 55(10):1829–1848.
- [255] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., et al. (2017). Brain tumor segmentation with deep neural networks. *Med. Image Anal.*, 35:18–31.
- [256] Hayer, A. and Meyer, T. (2010). High-content imaging. *Nat. Biotech.*, 28(5):424.
- [257] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In *Proc. International Conference on Computer Vision (ICCV)*, pages 1026–1034.
- [258] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.

- [259] Heidenreich, A., Aus, G., Bolla, M., Joniau, S., Matveev, V. B., Schmid, H. P., and Zattoni, F. (2008). EAU guidelines on prostate cancer. *Eur. Urol.*, 53(1):68–80.
- [260] Heijmink, S. W., Futterer, J. J., Hambroek, T., Takahashi, S., Scheenen, T. W., Huisman, H. J., et al. (2007). Prostate cancer: body-array versus endorectal coil MR imaging at 3 T—comparison of image quality, localization, and staging performance. *Radiology*, 244(1):184–195.
- [261] Heimann, T. and Meinzer, H.-P. (2009). Statistical shape models for 3D medical image segmentation: a review. *Med. Image Anal.*, 13(4):543–563.
- [262] Heiss, P., Mayer, S., Herz, M., Wester, H.-J., Schwaiger, M., and Senekowitsch-Schmidtke, R. (1999). Investigation of transport mechanism and uptake kinetics of O-(2-[18F] fluoroethyl)-L-tyrosine in vitro and in vivo. *J. Nucl. Med.*, 40(8):1367.
- [263] Heiss, W.-D. (2009). The potential of PET/MR for brain imaging. *Eur. J. Nucl. Med. Mol. Imaging*, 36(1):105–112.
- [264] Held, M., Schmitz, M. H., Fischer, B., Walter, T., Neumann, B., Olma, M. H., Peter, M., Ellenberg, J., and Gerlich, D. W. (2010). CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nat. Methods*, 7(9):747.
- [265] Hellier, P. and Barillot, C. (2003). Coupling dense and landmark-based approaches for nonrigid registration. *IEEE Trans. Med. Imaging*, 22(2):217–227.
- [266] Hemanth Jude, D. and Anitha, J. (2012). Image pre-processing and feature extraction techniques for magnetic resonance brain image analysis. In *Computer Applications for Communication, Networking, and Digital Contents*, pages 349–356. Springer.
- [267] Hill, D. L., Batchelor, P. G., Holden, M., and J., H. D. (2001). Medical image registration. *Phys. Med. Biol.*, 46(3):R1–R45.
- [268] Hiramatsu, Y., Hotta, K., Imanishi, A., Matsuda, M., Terai, K., Liu, D., Zhang, D., Song, Y., Zhang, C., Huang, H., et al. (2018). Cell image segmentation by integrating multiple CNNs. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 2205–2211. IEEE.
- [269] Hoeks, C. M., Barentsz, J. O., Hambroek, T., Yakar, D., Somford, D. M., Heijmink, S. W., et al. (2011). Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology*, 261(1):46–66.
- [270] Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. MIT Press, Cambridge, MA, USA.
- [271] Honda, H. (1978). Description of cellular patterns by Dirichlet domains: the two-dimensional case. *J. Theor. Biol.*, 72(3):523–543.
- [272] Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Netw.*, 4(2):251–257.



- [273] Horowitz, S. L. and Pavlidis, T. (1976). Picture segmentation by a tree traversal algorithm. *J. ACM*, 23(2):368–388.
- [274] Hoskin, P. J. (2015). Hypoxia dose painting in prostate and cervix cancer. *Acta Oncol.*, 54(9):1259–1262.
- [275] Hoskins, P. R., Martin, K., and Thrush, A. (2010). *Diagnostic Ultrasound: Physics and Equipment*. Cambridge University Press, Cambridge, UK, 2 edition.
- [276] Hough, P. V. C. (1962). Method and means for recognizing complex patterns. US Patent 3,069,654.
- [277] Hovhannisyan, V., Parpas, P., and Zafeiriou, S. (2016). MAGMA: Multilevel accelerated gradient mirror descent algorithm for large-scale convex composite minimization. *SIAM J. Imaging Sci.*, 9(4):1829–1857.
- [278] Hricak, H., Dooks, G. C., McNeal, J. E., Mark, A. S., Marotti, M., Avallone, A., Pelzer, M., Proctor, E. C., and Tanagho, E. A. (1987). MR imaging of the prostate gland: normal anatomy. *AJR Am. J. Roentgenol.*, 148(1):51–58.
- [279] Hsieh, J. et al. (2009). *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE, Bellingham, WA, USA.
- [280] Hu, J., Shen, L., and Sun, G. (2018). Squeeze-and-excitation networks. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. arXiv preprint arXiv:1709.01507.
- [281] Huang, H.-C., Chuang, Y.-Y., and Chen, C.-S. (2012). Multiple kernel fuzzy clustering. *IEEE Trans. Fuzzy Syst.*, 20(1):120–134.
- [282] Huang, Y.-L., Jao, Y.-L., Hsieh, T.-Y., and Chung, C.-W. (2008). Adaptive automatic segmentation of HEp-2 cells in indirect immunofluorescence images. In *Proc. International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC)*, pages 418–422. IEEE.
- [283] Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J. Physiol.*, 160(1):106–154.
- [284] Hung, K., Lee, C., and Choy, S.-O. (2015). Ubiquitous health monitoring: integration of wearable sensors, novel sensing techniques, and body sensor networks. In *Mobile health*, pages 319–342. Springer.
- [285] Iakovidis, D. K. and Koulaouzidis, A. (2015). Software for enhanced video capsule endoscopy: challenges for essential progress. *Nat. Rev. Gastroenterol. Hepatol.*, 12(3):172.
- [286] Ibrahim-Hashim, A., Robertson-Tessi, M., Enrizues-Navas, P., Damaghi, M., Balagurunathan, Y., Wojtkowiak, J. W., Russell, S., Yoonseok, K., Lloyd, M. C., Bui, M. M., et al. (2017). Defining cancer subpopulations by adaptive strategies rather than molecular properties provides novel insights into intratumoral evolution. *Cancer Res.*, pages canres–2844.

- [287] Illingworth, J. and Kittler, J. (1987). The adaptive Hough transform. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI 9(5):690–698.
- [288] Ismail, S. Z. and Bilgen, M. (2013). Use of depth dependent attenuation characteristics of thermal energy deposition to increase spatial uniformity of HIFU ablation in large uterine fibroids. *Med. Instrum.*, 1(1):1.
- [289] Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2016). Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004*.
- [290] Jaccard, P. (1901). Distribution de la florine alpine dans la bassin de dranses et dans quelques regions voisines. *Bulletin de la Societe Vaudoise des Sciences Naturelles*, 37:241–272.
- [291] Jain, A. K. (2002). *Fundamentals of Digital Image Processing*. Prentice Hall Press, Upper Saddle River, NJ, USA, 1 edition.
- [292] Jain, A. K. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1 edition.
- [293] Jensen, R. L. (2009). Brain tumor hypoxia: tumorigenesis, angiogenesis, imaging, pseudoprogession, and as a therapeutic target. *J. Neurooncol.*, 92(3):317–335.
- [294] Jia, H., Xia, Y., Song, Y., Cai, W., Fulham, M., and Feng, D. D. (2018). Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging. *Neurocomputing*, 275:1358–1369.
- [295] Jiang, W., Baker, M. L., Wu, Q., Bajaj, C., and Chiu, W. (2003). Applications of a bilateral denoising filter in biological electron microscopy. *J. Struct. Biol.*, 144(1-2):114–122.
- [296] João, A., Gambaruto, A. M., Tiago, J., and Sequeira, A. (2016). Computational advances applied to medical image processing: an update. *Open Access Bioinform.*, 2016(8):1–15.
- [297] Joe, B. N., Fukui, M. B., Meltzer, C. C., Huang, Q.-s., Day, R. S., Greer, P. J., and Bozik, M. E. (1999). Brain tumor volume measurement: comparison of manual and semiautomated methods. *Radiology*, 212(3):811–816.
- [298] Johnson, P. C., Hunt, S. J., and Drayer, B. P. (1989). Human cerebral gliomas: correlation of postmortem MR imaging and neuropathologic findings. *Radiology*, 170(1):211–217.
- [299] Jolesz, F. A. (2009). MRI-guided focused ultrasound surgery. *Annu. Rev. Med.*, 60:417–430.
- [300] Jones, Z. W., Leander, R., Quaranta, V., Harris, L. A., and Tyson, D. R. (2018). A drift-diffusion checkpoint model predicts a highly variable and growth-factor-sensitive portion of the cell cycle G1 phase. *PLoS One*, 13(2):e0192087.
- [301] Judenhofer, M. S., Wehrl, H. F., Newport, D. F., Catana, C., Siegel, S. B., Becker, M., Thielscher, A., Kneilling, M., Lichy, M. P., Eichner, M., et al. (2008). Simultaneous PET-MRI: a new approach for functional and morphological imaging. *Nat. Med.*, 14(4):459.

- [302] Kaggle (2018). 2018 Data Science Bowl. <https://www.kaggle.com/c/data-science-bowl-2018>. Online; accessed 30 April 2018.
- [303] Kalender, W. A. (1994). Technical foundations of spiral CT. *Semin. Ultrasound CT MRI*, 15(2):81–89.
- [304] Kaliman, S., Jayachandran, C., Rehfeldt, F., and Smith, A.-S. (2016). Limits of applicability of the Voronoi tessellation determined by centers of cell nuclei to epithelium morphology. *Front. Physiol.*, 7:551.
- [305] Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., et al. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.*, 36:61–78.
- [306] Kanade, T., Yin, Z., Bise, R., Huh, S., Eom, S., Sandbothe, M. F., and Chen, M. (2011). Cell image analysis: algorithms, system and applications. In *Proc. Workshop on Applications of Computer Vision (WACV)*, pages 374–381. IEEE.
- [307] Kanas, V. G., Zacharaki, E. I., Davatzikos, C., Sgarbas, K. N., and Megalooikonomou, V. (2015). A low cost approach for brain tumor segmentation based on intensity modeling and 3D random walker. *Biomed. Signal Process. Control*, 22:19–30.
- [308] Kansal, A. R., Torquato, S., Harsh IV, G., Chiocca, E., and Deisboeck, T. (2000). Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. *J. Theor. Biol.*, 203(4):367–382.
- [309] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. (2002). An efficient k-means clustering algorithm: analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):881–892.
- [310] Karaboga, D. and Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J. Glob. Optim.*, 39(3):459–471.
- [311] Kari, J. (2005). Theory of cellular automata: a survey. *Theor. Comput. Sci.*, 334(1–3):3–33.
- [312] Kari, L. and Rozenberg, G. (2008). The many facets of natural computing. *Commun. ACM*, 51(10):72–83.
- [313] Karras, T., Aila, T., Laine, S., and Lehtinen, J. (2018). Progressive growing of GANs for improved quality, stability, and variation. In *Proc. International Conference on Learning Representations (ICLR)*, *arXiv preprint arXiv:1710.10196v3*.
- [314] Kauffmann, C. and Piché, N. (2010). Seeded ND medical image segmentation by cellular automaton on GPU. *Int. J. Comput. Assist. Radiol. Surg.*, 5(3):251–262.
- [315] Kaya, I. E., Pehlivanlı, A. Ç., Sekizkardeş, E. G., and Ibrikci, T. (2017). PCA based clustering for brain tumor segmentation of T1w MRI images. *Comput. Methods Programs Biomed.*, 140:19–28.

- [316] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proc. IEEE International Conference on Neural Networks (ICNN)*, volume 4, pages 1942–1948.
- [317] Khoo, V. and Joon, D. (2006). New developments in MRI for target volume delineation in radiotherapy. *Br. J. Radiol.*, 79(special Issue 1):S2–S15.
- [318] Khotanlou, H., Fallahi, A., Oghabian, M. A., and Pooyan, M. (2014). Segmentation of uterine fibroid on MR images based on Chan-Vese level set method and shape prior model. *Biomed. Eng. Appl. Basis Commun.*, 26(02):1450030.
- [319] Kim, C. K. and Park, B. K. (2008). Update of prostate magnetic resonance imaging at 3 T. *J. Comput. Assist. Tomogr.*, 32(2):163–172.
- [320] Kim, Y. T. (1997). Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Trans. Consumer Electron.*, 43(1):1–8.
- [321] Kimmel, R., Kiryati, N., and Bruckstein, A. M. (1996). Sub-pixel distance maps and weighted distance transforms. *J. Math. Imaging Vis.*, 6(2-3):223–233.
- [322] Kingma, D. P. and Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [323] Kingma, D. P. and Welling, M. (2014). Auto-encoding variational Bayes. In *Proc. International Conference on Learning Representations (ICLR)*. arXiv preprint arXiv:1312.6114.
- [324] Kirby, R. S. and Gilling, R. S. (2011). *Fast Facts: Benign Prostatic Hyperplasia*. Health Press Limited, Abingdon, UK, 7 edition.
- [325] Kittler, J. and Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognit.*, 19(1):41–47.
- [326] Klein, S., Staring, M., Murphy, K., Viergever, M. A., and Pluim, J. P. (2010). elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med Imaging*, 29(1):196–205.
- [327] Klein, S., van der Heide, U. A., Lips, I. M., van Vulpen, M., Staring, M., and Pluim, J. P. W. (2008). Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.*, 35(4):1407–1417.
- [328] Kohmura, H. and Wakahara, T. (2006). Determining optimal filters for binarization of degraded characters in color using genetic algorithms. In *Proc. International Conference on Pattern Recognition (ICPR)*, volume 3, pages 661–664. IEEE.
- [329] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, 43(1):59–69.
- [330] Kohonen, T. (1990). The self-organizing map. *Proc. IEEE*, 78(9):1464–1480.
- [331] Kom, G., Tiedeu, A., and Kom, M. (2007). Automated detection of masses in mammograms by local adaptive thresholding. *Comput. Biol. Med.*, 37(1):37–48.

- [332] Konak, A., Coit, D. W., and Smith, A. E. (2006). Multi-objective optimization using genetic algorithms: a tutorial. *Reliab. Eng. Syst. Saf.*, 91(9):992–1007.
- [333] Kondziolka, D., Patel, A., Lunsford, L. D., Kassam, A., and Flickinger, J. C. (1999). Stereotactic radiosurgery plus whole brain radiotherapy versus radiotherapy alone for patients with multiple brain metastases. *Int. J. Radiat. Oncol. Biol. Phys.*, 45(2):427–434.
- [334] Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA, USA, 1 edition.
- [335] Kraus, O. Z., Ba, J. L., and Frey, B. J. (2016). Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics*, 32(12):i52–i59.
- [336] Kraus, O. Z., Grys, B. T., Ba, J., Chong, Y., Frey, B. J., Boone, C., and Andrews, B. J. (2017). Automated analysis of high-content microscopy data with deep learning. *Mol. Syst. Biol.*, 13(4):924.
- [337] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Proc. Advances Neural Information Processing Systems (NIPS)*, pages 1097–1105.
- [338] Krupinski, E. A. (2010). Current perspectives in medical image perception. *Atten. Percept. Psychophys.*, 72(5):1205–1217.
- [339] Kurhanewicz, J., Vigneron, D., Carroll, P., and Coakley, F. (2008). Multiparametric magnetic resonance imaging in prostate cancer: present and future. *Curr. Opin. Urol.*, 18(1):71.
- [340] Kurita, T., Otsu, N., and Abdelmalek, N. (1992). Maximum likelihood thresholding based on population mixture models. *Pattern Recognit.*, 25(10):1231–1240.
- [341] Kwak, H. and Zhang, B. (2016). Generating images part by part with composite generative adversarial networks. *arXiv preprint arXiv:1607.05387*.
- [342] Kwok, S. M., Chandrasekhar, R., Attikiouzel, Y., and Rickard, M. T. (2004). Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. *IEEE Trans. Med. Imaging*, 23(9):1129–1140.
- [343] Lagendijk, J. J., Raaymakers, B. W., Van den Berg, C. A., Moerland, M. A., Philippens, M. E., and Van Vulpen, M. (2014). MR guidance in radiotherapy. *Phys. Med. Biol.*, 59(21):R349.
- [344] Lahat, D., Adali, T., and Jutten, C. (2015). Multimodal data fusion: an overview of methods, challenges, and prospects. *Proc. IEEE*, 103(9):1449–1477.
- [345] Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., de Jong, E. E., van Timmeren, J., Sanduleanu, S., Larue, R. T., Even, A. J., Jochems, A., et al. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.*, 14(12):749.

- [346] Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., Zegers, C. M., Gillies, R., Boellard, R., Dekker, A., et al. (2012). Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer*, 48(4):441–446.
- [347] Law, M. Y. and Liu, B. (2009). DICOM-RT and its utilization in radiation therapy. *Radiographics*, 29(3):655–667.
- [348] Le Bihan, D. (2013). Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. *Radiology*, 268(2):318–322.
- [349] Le Rhun, E., Dhermain, F., Vogin, G., Reyns, N., and Metellus, P. (2016). Radionecrosis after stereotactic radiotherapy for brain metastases. *Expert Rev. Neurother.*, 16(8):903–914.
- [350] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436.
- [351] LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., and Jackel, L. D. (1990). Handwritten digit recognition with a back-propagation network. In *Proc. Advances Neural Information Processing Systems (NIPS)*, pages 396–404.
- [352] Leibfarth, S., Eckert, F., Welz, S., Siegel, C., Schmidt, H., Schwenzer, N., Zips, D., and Thorwarth, D. (2015). Automatic delineation of tumor volumes by co-segmentation of combined PET/MR data. *Phys. Med. Biol.*, 60(14):5399.
- [353] Leksell, L. (1949). A stereotaxic apparatus for intracerebral surgery. *Acta Chir. Scand.*, 99:229–233.
- [354] Leksell, L. (1951). The stereotaxic method and radiosurgery of the brain. *Acta Chir. Scand.*, 102(4):316–319.
- [355] Lemaître, G., Martí, R., Freixenet, J., Vilanova, J. C., Walker, P. M., and Meriaudeau, F. (2015). Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: a review. *Comput. Biol. Med.*, 60:8–31.
- [356] Levivier, M., Wikler Jr, D., Massager, N., David, P., Devriendt, D., Lorenzoni, J., Pirotte, B., Desmedt, F., Simon Jr, S., Goldman, S., et al. (2002). The integration of metabolic imaging in stereotactic procedures including radiosurgery: a review. *J. Neurosurg.*, 97:542–550.
- [357] Li, C., Xu, C., Gui, C., and Fox, M. D. (2010). Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. Image Process.*, 19(12):3243.
- [358] Li, C., Yang, Y., Xiao, L., Li, Y., Zhou, Y., and Zhao, J. (2016a). A novel image enhancement method using fuzzy Sure entropy. *Neurocomputing*, 215:196–211.
- [359] Li, K., Miller, E. D., Chen, M., Kanade, T., Weiss, L. E., and Campbell, P. G. (2008). Computer vision tracking of stemness. In *Proceedings of the 5th International Symposium on Biomedical Imaging (ISBI): From Nano to Macro*, pages 847–850. IEEE.
- [360] Li, Y., Jia, F., and Qin, J. (2016b). Brain tumor segmentation from multimodal magnetic resonance images via sparse representation. *Artif. Intell. Med.*, 73:1–13.

- [361] Li, Y.-I. and Shen, Y. (2010). An automatic fuzzy c-means algorithm for image segmentation. *Soft Comput.*, 14(2):123–128.
- [362] Libbrecht, M. W. and Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nat. Rev. Genet.*, 16(6):321.
- [363] Litjens, G., Debats, O., van de Ven, W., Karssemeijer, N., and Huisman, H. (2012). A pattern recognition approach to zonal segmentation of the prostate on MRI. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 7511 of *LNCS*, pages 413–420. Springer.
- [364] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.*, 42:60–88.
- [365] Litjens, G., Toth, R., van de Ven, W., Hoeks, C., Kerkstra, S., van Ginneken, B., et al. (2014). Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. *Med. Image Anal.*, 18(2):359–373.
- [366] Liu, M.-Y., Breuel, T., and Kautz, J. (2017). Unsupervised image-to-image translation networks. In *Proc. Advances in Neural Information Processing Systems (NIPS)*, pages 700–708.
- [367] Liu, X., Langer, D. L., Haider, M. A., Yang, Y., Wernick, M. N., and Yetik, I. S. (2009). Prostate cancer segmentation with simultaneous estimation of Markov random field parameters and class. *IEEE Trans. Med. Imaging*, 28(6):906–915.
- [368] Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Trans. Inform. Theory*, 28(2):129–137.
- [369] Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843):150.
- [370] Lopez-Paz, D. and Oquab, M. (2017). Revisiting classifier two-sample tests. In *Proc. International Conference on Learning Representations (ICLR)*, *arXiv preprint arXiv:1610.06545*.
- [371] Løvbjerg, M., Rasmussen, T. K., and Krink, T. (2001). Hybrid particle swarm optimiser with breeding and subpopulations. In *Proc. 3rd Annual Conference on Genetic and Evolutionary Computation (GECCO)*, pages 469–476. Morgan Kaufmann Publishers Inc.
- [372] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, 60(2):91–110.
- [373] Lu, J., Xue, S., Zhang, X., and Han, Y. (2015). A neural network-based interval pattern matcher. *Information*, 6(3):388–398.
- [374] Luan, S., Swanson, N., Chen, Z., and Ma, L. (2009). Dynamic Gamma Knife radiosurgery. *Phys. Med. Biol.*, 54(6):1579.

- [375] Lucignani, G., Paganelli, G., and Bombardieri, E. (2004). The use of standardized uptake values for assessing Fdg uptake with PET in oncology: a clinical perspective. *Nucl. Med. Commun.*, 25(7):651–656.
- [376] Lumley, M. A., Burgess, R., Billingham, L. J., McDonald, D. F., and Milligan, D. W. (1996). Colony counting is a major source of variation in CFU-GM results between centres. *Br. J. Haematol.*, 97(2):481–484.
- [377] Luxton, G., Petrovich, Z., Jozsef, G., Nedzi, L. A., and Apuzzo, M. L. (1993). Stereotactic radiosurgery: principles and comparison of treatment methods. *Neurosurgery*, 32(2):241–259.
- [378] Maas, A. L., Hannun, A. Y., and Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proc. 2013 ICML Workshop on Deep Learning for Audio, Speech and Language Processing*.
- [379] Machtinger, R., Inbar, Y., Cohen-Eylon, S., Admon, D., Alagem-Mizrachi, A., and Rabinovici, J. (2012). MR-guided focus ultrasound (MRgFUS) for symptomatic uterine fibroids: predictors of treatment success. *Hum. Reprod.*, 27(12):3425–3431.
- [380] Macovski, A. (1996). Noise in MRI. *Magn. Reson. Med.*, 36(3):494–497.
- [381] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 16(2):187–198.
- [382] Maes, F., Vandermeulen, D., and Suetens, P. (1999). Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Med. Image Anal.*, 3(4):373–386.
- [383] Mahapatra, D., Bozorgtabar, B., Hewavitharanage, S., and Garnavi, R. (2017). Image super resolution using generative adversarial networks and local saliency maps for retinal image analysis. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 382–390.
- [384] Maier, O., Menze, B. H., von der Gabelntz, J., Häni, L., Heinrich, M. P., Liebrand, M., Winzeck, S., Basit, A., Bentley, P., Chen, L., et al. (2017). ISLES 2015—a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Med. Image Anal.*, 35:250–269.
- [385] Makni, N., Iancu, A., Colot, O., Puech, P., Mordon, S., and Betrouni, N. (2011). Zonal segmentation of prostate using multispectral magnetic resonance images. *Med. Phys.*, 38(11):6093–6105.
- [386] Manousakas, I. N., Undrill, P. E., Cameron, G. G., and Redpath, T. W. (1998). Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions. *Comput. Biomed. Res.*, 31(6):393–412.
- [387] Marotz, J., Lübbert, C., and Eisenbeiss, W. (2001). Effective object recognition for automated counting of colonies in Petri dishes (automated colony counting). *Comput. Methods Programs Biomed.*, 66(2-3):183–198.



- [388] Martin, S., Daanen, V., and Troccaz, J. (2008). Atlas-based prostate segmentation using an hybrid registration. *Int. J. Comput. Assist. Radiol. Surg.*, 3(6):485–492.
- [389] Martin, S., Troccaz, J., and Daanen, V. (2010). Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model. *Med. Phys.*, 37(4):1579–1590.
- [390] Masciocchi, C., Arrigoni, F., Ferrari, F., Giordano, A. V., Iafrate, S., Capretti, I., Cannizzaro, E., Reginelli, A., Ierardi, A. M., Floridi, C., et al. (2017). Uterine fibroid therapy using interventional radiology mini-invasive treatments: current perspective. *Med. Oncol.*, 34(4):52.
- [391] Masson, M. H. and Denoeux, T. (2008). ECM: an evidential version of the fuzzy c-means algorithm. *Pattern Recognit.*, 41(4):1384–1397.
- [392] Mattes, D., Haynor, D. R., Vesselle, H., Lewellyn, T. K., and Eubank, W. (2001). Nonrigid multimodality image registration. In *Medical Imaging 2001: Image Processing*, volume 4322 of *Proc. SPIE*, pages 1609–1621. International Society for Optics and Photonics.
- [393] Matulewicz, L., Jansen, J. F., Bokacheva, L., Vargas, H. A., Akin, O., Fine, S. W., Shukla-Dave, A., Eastham, J. A., Hricak, H., Koutcher, J. A., et al. (2014). Anatomic segmentation improves prostate cancer detection with artificial neural networks analysis of 1h magnetic resonance spectroscopic imaging. *J. Magn. Reson. Imaging*, 40(6):1414–1421.
- [394] Maurer, C. R., Qi, R., and Raghavan, V. (2003). A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2):265–270.
- [395] Maycotte, P., Gearheart, C. M., Barnard, R., Aryal, S., Levy, J. M. M., Fosmire, S. P., Hansen, R. J., Morgan, M. J., Porter, C. C., Gustafson, D. L., et al. (2014). STAT3-mediated autophagy dependence identifies subtypes of breast cancer where autophagy inhibition can be efficacious. *Cancer Res.*, 74:2579–2590.
- [396] Mazzara, G. P., Velthuisen, R. P., Pearlman, J. L., Greenberg, H. M., and Wagner, H. (2004). Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. *Int. J. Radiat. Oncol. Biol. Phys.*, 59(1):300–312.
- [397] McAuliffe, M. J., Lalonde, F. M., McGarry, D., Gandler, W., Csaky, K., and Trus, B. L. (2001). Medical image processing, analysis and visualization in clinical research. In *Proc. Symposium on Computer-Based Medical Systems (CBMS)*, pages 381–386. IEEE.
- [398] McRobbie, D. W., Moore, E. A., Graves, M. J., and Prince, M. R. (2017). *MRI from Picture to Proton*. Cambridge University Press, Cambridge, UK.
- [399] Medina-Carnicer, R., Muñoz-Salinas, R., Carmona-Poyato, A., and Madrid-Cuevas, F. J. (2011). A novel histogram transformation to improve the performance of thresholding methods in edge detection. *Pattern Recognit. Lett.*, 32(5):676–693.

- [400] Meena Prakash, R. and Shantha Selva Kumari, R. (2016). Fuzzy c means integrated with spatial information and contrast enhancement for segmentation of MR brain images. *Int. J. Imaging Syst. Technol.*, 26(2):116–123.
- [401] Meier, R., Knecht, U., Loosli, T., Bauer, S., Slotboom, J., Wiest, R., and Reyes, M. (2016). Clinical evaluation of a fully-automatic segmentation method for longitudinal brain tumor volumetry. *Sci. Rep.*, 6.
- [402] Meijering, E. (2012). Cell segmentation: 50 years down the road [life sciences]. *IEEE Signal Process. Mag.*, 29(5):140–145.
- [403] Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A., and Olivo-Marin, J.-C. (2016). Imagining the future of bioimage analysis. *Nat. Biotechnol.*, 34(12):1250.
- [404] Menze, B. H., Jakab, A., Bauer, S., et al. (2015). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging*, 34(10):1993–2024.
- [405] Mescheder, L., Nowozin, S., and Geiger, A. (2017). Adversarial variational Bayes: unifying variational autoencoders and generative adversarial networks. *arXiv preprint arXiv:1701.04722*.
- [406] Metcalf, D., Milliard, S. T., Gomez, M., and Schwartz, M. (2016). Wearables and the Internet of Things for health: wearable, interconnected devices promise more efficient and comprehensive health care. *IEEE Pulse*, 7(5):35–39.
- [407] Metcalfe, P., Liney, G. P., Holloway, L., Walker, A., Barton, M., Delaney, G. P., Vinod, S., and Tome, W. (2013). The potential for an enhanced role for MRI in radiation-therapy treatment planning. *Technol. Cancer. Res. Treat.*, 12(5):429–446.
- [408] Mettler, F. A. and Guiberteau, M. J. (2012). *Essentials of Nuclear Medicine Imaging*. Elsevier Saunders, Philadelphia, PA, USA, 6 edition.
- [409] Meyer, F. (1994). Topographic distance and watershed lines. *Signal Process.*, 38(1):113–125.
- [410] Meystre, S. M., Savova, G. K., Kipper-Schuler, K. C., and Hurdle, J. F. (2008). Extracting information from textual documents in the electronic health record: a review of recent research. *Yearb. Med. Inform.*, 17(01):128–144.
- [411] Militello, C., Rundo, L., Conti, V., Minafra, L., Cammarata, F. P., Mauri, G., Gilardi, M. C., and Porcino, N. (2017). Area-based cell colony surviving fraction evaluation: a novel fully automatic approach using general-purpose acquisition hardware. *Comput. Biol. Med.*, 89:454–465.
- [412] Militello, C., Rundo, L., and Gilardi, M. C. (2014). Applications of imaging processing to MRgFUS treatment for fibroids: a review. *Transl. Cancer Res.*, 3(5):472–482.
- [413] Militello, C., Rundo, L., Vitabile, S., Russo, G., Pisciotta, P., Marletta, F., Ippolito, M., D’Arrigo, C., Midiri, M., and Gilardi, M. C. (2015a). Gamma Knife treatment planning: MR brain tumor segmentation and volume measurement based on unsupervised fuzzy c-means clustering. *Int. J. Imaging Syst. Technol.*, 25(3):213–225.

- [414] Militello, C., Vitabile, S., Rundo, L., Gagliardo, C., and Salerno, S. (2015b). An edge-driven 3D region-growing approach for upper airway morphology and volume evaluation in patients with pierre robin sequence. *Int. J. Adapt. Innov. Syst.*, 2(3):232–253.
- [415] Militello, C., Vitabile, S., Rundo, L., Russo, G., Midiri, M., and Gilardi, M. C. (2015c). A fully automatic 2D segmentation method for uterine fibroid in MRgFUS treatment evaluation. *Comput. Biol. Med.*, 62:277–292.
- [416] Militello, C., Vitabile, S., Russo, G., Candiano, G., Gagliardo, C., Midiri, M., and Gilardi, M. C. (2013). A semi-automatic multi-seed region-growing approach for uterine fibroids segmentation in MRgFUS treatment. In *Proc. International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS)*, pages 176–182. IEEE.
- [417] Miller, B. L. and Goldberg, D. E. (1995). Genetic algorithms, tournament selection, and the effects of noise. *Complex Syst.*, 9(3):193–212.
- [418] Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-Net: fully convolutional neural networks for volumetric medical image segmentation. In *Proc. International Conference on 3D Vision (3DV)*, pages 565–571. IEEE.
- [419] Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.
- [420] Mitchell, D. G. (1990). Color Doppler imaging: principles, limitations, and artifacts. *Radiology*, 177(1):1–10.
- [421] Mitra, S. and Shankar, B. U. (2015). Medical image analysis for cancer management in natural computing framework. *Inform. Sci.*, 306:111–131.
- [422] Miwa, K., Matsuo, M., Shinoda, J., Aki, T., Yonezawa, S., Ito, T., Asano, Y., Yamada, M., Yokoyama, K., Yamada, J., et al. (2012). Clinical value of [11C] Methionine PET for stereotactic radiation therapy with intensity modulated radiation therapy to metastatic brain tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 84(5):1139–1144.
- [423] Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal Proc. Mag.*, 13(6):47–60.
- [424] Morantz, R. and Wara, W. (1995). Gamma Knife radiosurgery in the treatment of brain tumors. *Cancer Control*, 2(4):300–308.
- [425] Moskvina, V., DesRosiers, C., Papiez, L., Timmerman, R., Randall, M., and DesRosiers, P. (2002). Monte carlo simulation of the Leksell Gamma Knife<sup>®</sup>: I. source modelling and calculations in homogeneous media. *Phys. Med. Biol.*, 47(12):1995.
- [426] Muangkote, N., Sunat, K., and Chiewchanwattana, S. (2017). R<sub>r-cr</sub>-IJADE: an efficient differential evolution algorithm for multilevel image thresholding. *Expert Syst. Appl.*, 90:272–289.
- [427] Mukherjee, D. P., Pal, A., Sarma, S. E., and Majumder, D. D. (1995). Bacterial colony counting using distance transform. *Int. J. Biomed. Comput.*, 38(2):131–140.

- [428] Muller, B. G., Shih, J. H., Sankineni, S., Marko, J., Rais-Bahrami, S., George, A. K., et al. (2015). Prostate cancer: interobserver agreement and accuracy with the revised prostate imaging reporting and data system at multiparametric MR imaging. *Radiology*, 277(3):741–750.
- [429] Munteanu, C. and Rosa, A. (2004). Gray-scale image enhancement as an automatic process driven by evolution. *IEEE Trans. Syst. Man Cybern.*, 34(2):1292–1298.
- [430] Nair, V. and Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proc. International Conference on Machine Learning (ICML)*, pages 807–814.
- [431] Najman, L., Couprie, M., and Bertrand, G. (2005). Watersheds, mosaics, and the emergence paradigm. *Discret. Appl. Math.*, 147(2-3):301–324.
- [432] NEMA (2011a). Digital Imaging and Communications in Medicine (DICOM) Standard, Part 1: Introduction and Overview. ISO NEMA PS3 / ISO 12052, National Electrical Manufacturers Association, Rosslyn, VA, USA.
- [433] NEMA (2011b). Digital Imaging and Communications in Medicine (DICOM) Standard, Part 3: Information Object Definitions. ISO NEMA PS3 / ISO 12052, National Electrical Manufacturers Association, Rosslyn, VA, USA.
- [434] Nguyen, T. M. and Wu, Q. J. (2013). Fast and robust spatially constrained Gaussian mixture model for image segmentation. *IEEE Trans. Circ. Syst. Video Technol.*, 23(4):621–635.
- [435] Niaf, E., Rouvière, O., Mège-Lechevallier, F., Bratan, F., and Lartizien, C. (2012). Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric MRI. *Phys. Med. Biol.*, 57(12):3833.
- [436] Nie, D., Trullo, R., Lian, J., Petitjean, C., Ruan, S., Wang, Q., and Shen, D. (2017). Medical image synthesis with context-aware generative adversarial networks. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 417–425.
- [437] Nieder, C., Grosu, A. L., and Gaspar, L. E. (2014). Stereotactic radiosurgery (SRS) for brain metastases: a systematic review. *Radiat. Oncol.*, 9(1):155.
- [438] Nobile, M. S., Cazzaniga, P., Besozzi, D., Colombo, R., Mauri, G., and Pasi, G. (2018a). Fuzzy self-tuning PSO: A settings-free algorithm for global optimization. *Swarm Evol. Comput.*, 39:70–85.
- [439] Nobile, M. S., Cazzaniga, P., Tangherloni, A., and Besozzi, D. (2016). Graphics processing units in bioinformatics, computational biology and systems biology. *Brief. Bioinform.*, 18(5):870–885.
- [440] Nobile, M. S., Pasi, G., Cazzaniga, P., Besozzi, D., Colombo, R., and Mauri, G. (2015). Proactive particles in swarm optimization: a self-tuning algorithm based on fuzzy logic. In *Proc. International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8. IEEE.

- [441] Nobile, M. S., Tangherloni, A., Rundo, L., Spolaor, S., Besozzi, D., Mauri, G., and Cazzaniga, P. (2018b). Computational intelligence for parameter estimation of biochemical systems. In *Proc. Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE.
- [442] Oh, C., Ham, B., and Sohn, K. (2017). Robust interactive image segmentation using structure-aware labeling. *Expert Syst. Appl.*, 79:90–100.
- [443] Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987.
- [444] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). Attention U-Net: learning where to look for the pancreas. In *Proc. Conference on Medical Imaging with Deep Learning (MIDL)*, pages 1–10. arXiv preprint arXiv:1804.03999.
- [445] Ortiz, A., Górriz, J., Ramírez, J., Salas-Gonzalez, D., and Llamas-Elvira, J. M. (2013). Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies. *Appl. Soft Comput.*, 13(5):2668–2682.
- [446] Osher, S. and Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79(1):12–49.
- [447] Osokin, A., Chessel, A., Carazo Salas, R. E., et al. (2017). GANs for biological image synthesis. In *Proc. International Conference on Computer Vision (ICCV)*, pages 2252–2261.
- [448] Otsu, N. (1975). A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.*, 11(285-296):23–27.
- [449] Ou, Y., Akbari, H., Bilello, M., Da, X., and Davatzikos, C. (2014). Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. *IEEE Trans. Med. Imaging*, 33(10):2039–2065.
- [450] Padhani, A. R., Lecouvet, F. E., Tunariu, N., Koh, D.-M., De Keyzer, F., Collins, D. J., et al. (2017). Rationale for modernising imaging in advanced prostate cancer. *Eur. Urol. Focus*, 3(2-3):223–239.
- [451] Pal, N. R. and Bezdek, J. C. (1995). On cluster validity for the fuzzy c-means model. *IEEE Trans. Fuzzy syst.*, 3(3):370–379.
- [452] Papa, G. (2013). Parameter-less algorithm for evolutionary-based optimization. *Comput. Optim. Appl.*, 56(1):209–229.
- [453] Papoulis, A. and Pillai, S. U. (2001). *Probability, Random Variables and Stochastic Processes*. McGraw Hill, New York, NY, USA, 2001, 4 edition.
- [454] Paranjape, R. B. (2009). Chapter 1 - Fundamental enhancement techniques. In Bankman, I. N., editor, *Handbook of Medical Image Processing and Analysis*, pages 3–18. Academic Press, Cambridge, MA, USA, 2 edition.

- [455] Pasero, E. and Castagneri, C. (2017). Application of an automatic ulcer segmentation algorithm. In *Proc. International Forum on Research and Technologies for Society and Industry (RTSI)*, pages 1–4. IEEE.
- [456] Pasquier, D., Lacornerie, T., Vermandel, M., Rousseau, J., Lartigau, E., and Betrouni, N. (2007). Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 68(2):592–600.
- [457] Patel, A. A., Gawlinski, E. T., Lemieux, S. K., and Gatenby, R. A. (2001). A cellular automaton model of early tumor growth and invasion: the effects of native tissue vascularity and increased anaerobic tumor metabolism. *J. Theor. Biol.*, 213(3):315–331.
- [458] Patton, J. A. and Turkington, T. G. (2008). SPECT/CT physical principles and attenuation correction. *J. Nucl. Med. Technol.*, 36(1):1–10.
- [459] Paulinas, M. and Ušinskas, A. (2007). A survey of genetic algorithms applications for image enhancement and segmentation. *Inf. Technol. Control*, 36(3).
- [460] Pelt, D. M. and Sethian, J. A. (2018). A mixed-scale dense convolutional neural network for image analysis. *Proc. Natl. Acad. Sci.*, 115(2):254–259.
- [461] Peng, H. (2008). Bioimage informatics: a new area of engineering biology. *Bioinformatics*, 24(17):1827–1836.
- [462] Peng, H., Bateman, A., Valencia, A., and Wren, J. D. (2012). Bioimage informatics: a new category in bioinformatics. *Bioinformatics*, 28(8):1057–1057.
- [463] Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, Cambridge, MA, USA, 1 edition.
- [464] Pereira, A. C., Reis, M. S., and Saraiva, P. M. (2008). Quality control of food products using image analysis and multivariate statistical tools. *Ind. Eng. Chem. Res.*, 48(2):988–998.
- [465] Peterson, A. M., Meltzer, C. C., Evanson, E. J., Flickinger, J. C., and Kondziolka, D. (1999). MR imaging response of brain metastases after Gamma Knife stereotactic radiosurgery. *Radiology*, 211(3):807–814.
- [466] Petit, S. F., Dekker, A. L., Seigneuric, R., Murrer, L., van Riel, N. A., Nordmark, M., Overgaard, J., Lambin, P., and Wouters, B. G. (2009). Intra-voxel heterogeneity influences the dose prescription for dose-painting with radiotherapy: a modelling study. *Phys. Med. Biol.*, 54(7):2179.
- [467] Pham, D. L. and Prince, J. L. (1999). Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Trans. Med. Imaging*, 18(9):737–752.
- [468] Pichler, B. J., Judenhofer, M. S., and Pfannenberger, C. (2008). Multimodal imaging approaches: PET/CT and PET/MRI. In *Molecular Imaging I*, pages 109–132. Springer.
- [469] Pinho, J., Sobral, J. L., and Rocha, M. (2013). Parallel evolutionary computation in bioinformatics applications. *Comput. Methods Programs Biomed.*, 110(2):183–191.

- [470] Pinker, K., Shitano, F., Sala, E., Do, R. K., Young, R. J., Wibmer, A. G., Hricak, H., Sutton, E. J., and Morris, E. A. (2018). Background, current role, and potential applications of radiogenomics. *J. Magn. Reson. Imaging*, 47(3):604–620.
- [471] Pivotal Software, Inc. (2018). RabbitMQ. <http://www.rabbitmq.com/>. Online; accessed 20 October 2018.
- [472] Pixmeo (2018). OsiriX, DICOM Viewer. <http://www.osirix-viewer.com/>. Online; Accessed on February 9, 2018.
- [473] Pizurica, A., Wink, A. M., Vansteenkiste, E., Philips, W., and Roerdink, B. J. (2006). A review of wavelet denoising in MRI and ultrasound brain imaging. *Curr. Med. Imaging Rev.*, 2(2):247–260.
- [474] Pluim, J. P., Maintz, J. A., and Viergever, M. A. (2003). Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imaging*, 22(8):986–1004.
- [475] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. (2000). Image registration by maximization of combined mutual information and gradient information. *IEEE Trans. Med. Imaging*, 19(8):809–814.
- [476] Poli, R. (2008). Analysis of the publications on the applications of particle swarm optimisation. *J. Artif. Evol. Appl.*, 2008:2–10.
- [477] Poli, R. and Cagnoni, S. (1997). Genetic programming with user-driven selection: experiments on the evolution of algorithms for image enhancement. In *Proc. Annual Conference on Genetic Programming*, pages 269–277.
- [478] Poli, R., Kennedy, J., and Blackwell, T. (2007). Particle swarm optimization: an overview. *Swarm Intell.*, 1(1):33–57.
- [479] Popovici, A. and Popovici, D. (2002). Cellular automata in image processing. In *Proc. 15th International Symposium on Mathematical Theory of Networks and Systems*, volume 1, pages 1–6.
- [480] Posse, S., Otazo, R., Dager, S. R., and Alger, J. (2013). MR spectroscopic imaging: principles and recent advances. *J. Magn. Reson. Imaging*, 37(6):1301–1325.
- [481] Potesil, V., Huang, X., and Zhou, X. S. (2007). Automated tumor delineation using joint PET/CT information. In *Medical Imaging 2007: Computer-Aided Diagnosis*, volume 6514 of *Proc. SPIE*, page 65142Y. International Society for Optics and Photonics.
- [482] Powell, J. and Buchan, I. (2005). Electronic health records should support clinical research. *J. Med. Internet Res.*, 7(1).
- [483] Powell, M. J. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer J.*, 7(2):155–162.
- [484] Prastawa, M., Bullitt, E., and Gerig, G. (2009). Simulation of brain tumors in MR images for evaluation of segmentation efficacy. *Med. Image Anal.*, 13(2):297–311.

- [485] Prior, F., Smith, K., Sharma, A., Kirby, J., Tarbox, L., Clark, K., et al. (2017). The public cancer radiology imaging collections of The Cancer Imaging Archive. *Sci. Data*, 4:170124.
- [486] Puck, T. T. and Marcus, P. I. (1956). Action of X-rays on mammalian cells. *J. Exp. Med.*, 103(5):653–666.
- [487] Qiu, W., Yuan, J., Ukwatta, E., Sun, Y., Rajchl, M., and Fenster, A. (2014). Dual optimization based prostate zonal segmentation in 3D MR images. *Med. Image Anal.*, 18(4):660–673.
- [488] Quinones-Hinojosa, A. (2012). *Schmidek and Sweet: Operative Neurosurgical Techniques 2-Volume Set: Indications, Methods and Results*. Elsevier Saunders, Philadelphia, PA, USA, 6 edition.
- [489] Radford, A., Metz, L., and Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. In *Proc. International Conference on Learning Representations (ICLR)*, arXiv preprint arXiv:1511.06434.
- [490] Raghavan, R. (1993). Cellular automata in pattern recognition. *Inform. Sci.*, 70(1-2):145–177.
- [491] Rajchl, M., Lee, M. C., Oktay, O., Kamnitsas, K., Passerat-Palmbach, J., Bai, W., et al. (2017). DeepCut: object segmentation from bounding box annotations using convolutional neural networks. *IEEE Trans. Med. Imaging*, 36(2):674–683.
- [492] Randazzo, V., Pasero, E., and Navaretti, S. (2018). VITAL-ECG: A portable wearable hospital. In *Proc. Sensors Applications Symposium (SAS)*, pages 1–6. IEEE.
- [493] Rangarajan, A., Chui, H., and Bookstein, F. L. (1997). The softassign Procrustes matching algorithm. In *Proc. International Conference on Information Processing in Medical Imaging (IPMI)*, volume 1230 of LNCS, pages 29–42. Springer.
- [494] Rangayyan, R. M. (2009). Part I - Enhancement. In Bankman, I. N., editor, *Handbook of Medical Image Processing and Analysis*, pages 1–2. Academic Press, Burlington, MA, USA, 2 edition.
- [495] Ratliff, F. (1984). Why Mach bands are not seen at the edges of a step. *Vision Res.*, 24(2):163–165.
- [496] Ravì, D., Wong, C., Deligianni, F., et al. (2017). Deep learning for health informatics. *IEEE J. Biomed. Health Inform.*, 21(1):4–21.
- [497] Ree, A. and Redalen, K. (2015). Personalized radiotherapy: concepts, biomarkers and trial design. *Br. J. Radiol.*, 88(1051):20150009.
- [498] Regge, D., Mazzetti, S., Giannini, V., Bracco, C., and Stasi, M. (2017). Big data in oncologic imaging. *Radiol. Med.*, 122(6):458–463.
- [499] Rezaei, M., Harmuth, K., Gierke, W., Kellermeier, T., Fischer, M., Yang, H., and Meinel, C. (2017). A conditional adversarial network for semantic segmentation of brain tumor. In *Proc. 2017 International MICCAI (BrainLes) Workshop*, volume 10670 of LNCS, pages 241–252. Springer.



- [500] Ridler, T. W. and Calvard, S. (1978). Picture thresholding using an iterative selection method. *IEEE Trans. Syst. Man Cybern.*, 8(8):630–632.
- [501] Roberts, A. (2008). Magnetic resonance-guided focused ultrasound for uterine fibroids. *Semin. Interv. Radiol.*, 25(4):394.
- [502] Roerdink, J. B. T. M. and Meijster, A. (2000). The watershed transform: definitions, algorithms and parallelization strategies. *Fundam. Inform.*, 41(1, 2):187–228.
- [503] Roethke, M., Anastasiadis, A., Lichy, M., Werner, M., Wagner, P., Kruck, S., Claussen, C. D., Stenzl, A., Schlemmer, H., and Schilling, D. (2012). MRI-guided prostate biopsy detects clinically significant cancer: analysis of a cohort of 100 patients after previous negative TRUS biopsy. *World J. Urol.*, 30(2):213–218.
- [504] Rogowska, J. (2009). Chapter 5 - Overview and fundamentals of medical image segmentation. In Bankman, I. N., editor, *Handbook of Medical Image Processing and Analysis*, pages 73–90. Academic Press, Burlington, MA, USA, 2 edition.
- [505] Rohr, K. (2000). Elastic registration of multimodal medical images: a survey. *Auszug aus: Künstliche Intelligenz*, 14(3):11–17.
- [506] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: convolutional networks for biomedical image segmentation. In *Proc. Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of LNCS, pages 234–241. Springer.
- [507] Rosenkrantz, A. B., Lim, R. P., Haghighi, M., Somberg, M. B., Babb, J. S., and Taneja, S. S. (2013). Comparison of interreader reproducibility of the prostate imaging reporting and data system and Likert scales for evaluation of multiparametric prostate MRI. *AJR Am. J. Roentgenol.*, 201(4):W612–W618.
- [508] Rosenstein, B. S., West, C. M., Bentzen, S. M., Alsner, J., Andreassen, C. N., Azria, D., Barnett, G. C., Baumann, M., Burnet, N., Chang-Claude, J., et al. (2014). Radiogenomics: radiobiology enters the era of big data and team science. *Int. J. Radiat. Oncol. Biol. Phys.*, 89(4):709–713.
- [509] Rosin, P. L. (2010). Image processing using 3-state cellular automata. *Comput. Vis. Image Underst.*, 114(7):790–802.
- [510] Rosset, A., Spadola, L., and Ratib, O. (2004). OsiriX: an open-source software for navigating in multidimensional dicom images. *J. Digit. Imaging*, 17(3):205–216.
- [511] Rother, C., Kolmogorov, V., and Blake, A. (2004). GrabCut: interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314.
- [512] Rouvière, O., Hartman, R. P., and Lyonnet, D. (2006). Prostate MR imaging at high-field strength: evolution or revolution? *Eur. Radiol.*, 16(2):276–284.
- [513] Rueckert, D., Glocker, B., and Kainz, B. (2016). Learning clinically useful information from images: past, present and future. *Med. Image Anal.*, 33:13–18.

- [514] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., and Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging*, 18(8):712–721.
- [515] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533.
- [516] Rundo, L., Han, C., Nagano, Y., Zhang, J., Hataya, R., Militello, C., et al. (2018a). USE-Net: Incorporating squeeze-and-excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets. *Neurocomputing*. (Under review).
- [517] Rundo, L., Han, C., Zhang, J., Hataya, R., Nagano, Y., Militello, C., et al. (2018b). CNN-based prostate zonal segmentation on T2-weighted MR images: a cross-dataset study. In *28th Italian Workshop on Neural Networks (WIRN) 2018, Vietri sul Mare (SA), Italy, June 13-15, 2018*, Smart Innovation, Systems and Technologies. Springer. (Accepted manuscript).
- [518] Rundo, L., Militello, C., Russo, G., D’Urso, D., Valastro, L. M., Garufi, A., et al. (2018c). Fully automatic multispectral MR image segmentation of prostate gland based on the fuzzy c-means clustering algorithm. In *Multidisciplinary Approaches to Neural Computing*, volume 69 of *Smart Innovation, Systems and Technologies*, pages 23–37. Springer.
- [519] Rundo, L., Militello, C., Russo, G., Garufi, A., Vitabile, S., Gilardi, M. C., and Mauri, G. (2017a). Automated prostate gland segmentation based on an unsupervised fuzzy c-means clustering technique using multispectral T1w and T2w MR imaging. *Information*, 8(2):49.
- [520] Rundo, L., Militello, C., Russo, G., Pisciotta, P., Valastro, L. M., Sabini, M. G., Vitabile, S., Gilardi, M. C., and Mauri, G. (2016a). Neuro-radiosurgery treatments: MRI brain tumor seeded image segmentation based on a cellular automata model. In *Proc. International Conference on Cellular Automata for Research and Industry (ACRI)*, volume 9863 of *LNCS*, pages 323–333. Springer.
- [521] Rundo, L., Militello, C., Russo, G., Vitabile, S., Gilardi, M. C., and Mauri, G. (2018d). GTVcut for neuro-radiosurgery treatment planning: an MRI brain cancer seeded image segmentation method based on a cellular automata model. *Nat. Comput.*, 17(3):521–536.
- [522] Rundo, L., Militello, C., Tangherloni, A., Russo, G., Lagalla, R., Mauri, G., Gilardi, M. C., and Vitabile, S. (2019a). Computer-assisted approaches for uterine fibroid segmentation in MRgFUS treatments: quantitative evaluation and clinical feasibility analysis. In *Quantifying and Processing Biomedical and Behavioral Signals*, volume 103 of *Smart Innovation, Systems and Technologies*, pages 229–241. Springer.
- [523] Rundo, L., Militello, C., Tangherloni, A., Russo, G., Vitabile, S., Gilardi, M. C., and Mauri, G. (2018e). NeXt for neuro-radiosurgery: a fully automatic approach for necrosis extraction in brain tumor MRI using an unsupervised machine learning technique. *Int. J. Imaging Syst. Technol.*, 28(1):21–37.

- [524] Rundo, L., Militello, C., Vitabile, S., Casarino, C., Russo, G., Midiri, M., and Gilardi, M. C. (2016b). Combining split-and-merge and multi-seed region growing algorithms for uterine fibroid segmentation in MRgFUS treatments. *Med. Biol. Eng. Comput.*, 54(7):1071–1084.
- [525] Rundo, L., Militello, C., Vitabile, S., Russo, G., Pisciotta, P., Marletta, F., Ippolito, M., D'Arrigo, C., Midiri, M., and Gilardi, M. C. (2016c). Semi-automatic brain lesion segmentation in Gamma Knife treatments using an unsupervised fuzzy c-means clustering technique. In *Advances in Neural Networks: Computational Intelligence for ICT*, volume 54 of *Smart Innovation, Systems and Technologies*, pages 15–26. Springer.
- [526] Rundo, L., Stefano, A., Militello, C., Russo, G., Sabini, M. G., D'Arrigo, C., Marletta, F., Ippolito, M., Mauri, G., Vitabile, S., and Gilardi, M. C. (2017b). A fully automatic approach for multimodal PET and MR image segmentation in Gamma Knife treatment planning. *Comput. Methods Programs Biomed.*, 144:77–96.
- [527] Rundo, L., Tangherloni, A., Cazzaniga, P., Nobile, M. S., Russo, G., Gilardi, M. C., Vitabile, S., Mauri, G., Besozzi, D., and Militello, C. (2018f). A genetic algorithm improves the segmentation of mr images with an underlying bimodal histogram. *Comput. Methods Programs Biomed.* (Submitted).
- [528] Rundo, L., Tangherloni, A., Militello, C., Gilardi, M. C., and Mauri, G. (2016d). Multimodal medical image registration using particle swarm optimization: a review. In *Proc. Symposium Series on Computational Intelligence (SSCI)*, pages 1–8. IEEE.
- [529] Rundo, L., Tangherloni, A., Nobile, M. S., Militello, C., Besozzi, D., Mauri, G., and Cazzaniga, P. (2019b). MedGA: a novel evolutionary method for image enhancement in medical imaging systems. *Expert Syst. Appl.*, 119:387–399.
- [530] Rundo, L., Tangherloni, A., Tyson, D. R., Betta, R., Nobile, M. S., Spolaor, S., Militello, C., Lubbock, A. L. R., Mauri, G., Quaranta, V., Besozzi, D., Lopez, C. F., and Cazzaniga, P. (2019c). ACDC: Automated cell detection and counting for time-lapse fluorescence microscopy. *Med. Biol. Eng. Comput.* (In preparation).
- [531] Rutman, A. M. and Kuo, M. D. (2009). Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. *Eur. J. Radiol.*, 70(2):232–241.
- [532] Ryan, G. L., Syrop, C. H., and Van Voorhis, B. J. (2005). Role, epidemiology, and natural history of benign uterine mass lesions. *Clin. Obstet. Gynecol.*, 48(2):312–324.
- [533] Saad, N. M., Abu-Bakar, S. A. R., Muda, S., and Mokji, M. (2010). Automated segmentation of brain lesion based on diffusion-weighted MRI using a split and merge approach. In *Proc. IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pages 475–480. IEEE.
- [534] Saad, N. M., Abu-Bakar, S. A. R., Muda, S., Mokji, M., and Abdullah, A. R. (2012). Automated region growing for segmentation of brain lesion in diffusion-weighted MRI. In *Lecture Notes in Engineering and Computer Science*, volume 1, pages 674–677.
- [535] Sadanandan, S. K., Ranefall, P., Le Guyader, S., and Wählby, C. (2017). Automated training of deep convolutional neural networks for cell segmentation. *Sci. Rep.*, 7(1):7860.

- [536] Sahoo, P. K., Soltani, S., and Wong, A. K. C., C. Y. C. (1988). A survey of thresholding techniques. *Comput. Vis. Graph. Image Process.*, 41(2):233–260.
- [537] Saitoh, F. (1999). Image contrast enhancement using genetic algorithm. In *Proc. International Conference on Systems, Man, and Cybernetics (SMC)*, volume 4, pages 899–904. IEEE.
- [538] Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., et al. (2008). Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, 132(3):487–498.
- [539] Sala, E., Mema, E., Himoto, Y., Veeraraghavan, H., Brenton, J., Snyder, A., Weigelt, B., and Vargas, H. (2017). Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin. Radiol.*, 72(1):3–10.
- [540] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. (2016). Improved techniques for training GANs. In *Proc. Advances in Neural Information Processing Systems (NIPS)*, pages 2234–2242.
- [541] Sandhu, R., Dambreville, S., and Tannenbaum, A. (2008). Particle filtering for registration of 2D and 3D point sets with stochastic dynamics. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE.
- [542] Sangwine, S. J. and Horne, R. E. (1998). *The Colour Image Processing Handbook*. Springer Science & Business Media, New York, NY, USA, 1 edition.
- [543] Sansone, M., Zeni, O., and Esposito, G. (2012). Automated segmentation of comet assay images using Gaussian filtering and fuzzy clustering. *Med. Biol. Eng. Comput.*, 50(5):523–532.
- [544] Sarvazyan, A. (1998). Mechanical imaging: a new technology for medical diagnostics. *Int. J. Med. Inf.*, 49(2):195–216.
- [545] Sarvazyan, A., J Hall, T., W Urban, M., Fatemi, M., R Aglyamov, S., and S Garra, B. (2011). An overview of elastography-an emerging branch of medical imaging. *Curr. Med. Imaging Rev.*, 7(4):255–282.
- [546] Sauvola, J. and Pietikäinen, M. (2000). Adaptive document image binarization. *Pattern Recognit.*, 33(2):225–236.
- [547] Scheenen, T. W., Rosenkrantz, A. B., Haider, M. A., and Fütterer, J. J. (2015). Multi-parametric magnetic resonance imaging in prostate cancer management: current status and future perspectives. *Invest. Radiol.*, 50(9):594–600.
- [548] Schettini, R., Gasparini, F., Corchs, S., Marini, F., Capra, A., and Castorina, A. (2010). Contrast image correction method. *J. Electr. Imaging*, 19(2):023005.
- [549] Scheuermann, T. and Hensley, J. (2007). Efficient histogram generation using scattering on GPUs. In *Proc. Symposium on Interactive 3D Graphics and Games*, pages 33–37. ACM.

- [550] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, 9(7):676.
- [551] Schlegl, T., Seeböck, P., Waldstein, S. M., et al. (2017). Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *Proc. International Conference on Information Processing in Medical Imaging (IPMI)*, pages 146–157. Springer.
- [552] Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7):671.
- [553] Schwab, L., Schmitt, M., and Wanka, R. (2015). Multimodal medical image registration using particle swarm optimization with influence of the data’s initial orientation. In *Proc. Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–8. IEEE.
- [554] Selman, S. H. (2011). The McNeal prostate: a review. *Urology*, 78(6):1224–1228.
- [555] Serbanati, L. D., Ricci, F. L., Mercurio, G., and Vasileanu, A. (2011). Steps towards a digital health ecosystem. *J. Biomed. Inform.*, 44(4):621–636.
- [556] Serra, A., Galdi, P., and Tagliaferri, R. (2018). Machine learning for bioinformatics and neuroimaging. *Wiley Interdisc. Rev. Data Min. Knowl. Discov.*, page e1248.
- [557] Shah, R., Vattoth, S., Jacob, R., Manzil, F. F. P., O’Malley, J. P., Borghei, P., Patel, B. N., and Curé, J. K. (2012). Radiation necrosis in the brain: imaging features and differentiation from tumor recurrence. *RadioGraphics*, 32(5):1343–1359.
- [558] Shamir, L., Delaney, J. D., Orlov, N., Eckley, D. M., and Goldberg, I. G. (2010). Pattern recognition software and techniques for biological image analysis. *PLoS Comput. Biol.*, 6(11):e1000974.
- [559] Shanmugavadivu, P. and Balasubramanian, K. (2014). Particle swarm optimized multi-objective histogram equalization for image enhancement. *Opt. Laser Technol.*, 57:243–251.
- [560] Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.*, 19:221–248.
- [561] Shi, B., Bai, X., and Yao, C. (2017). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(11):2298–2304.
- [562] Shrivastava, A., Pfister, T., Tuzel, O., et al. (2017). Learning from simulated and unsupervised images through adversarial training. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2107–2116. IEEE.
- [563] Siddique, N. and Adeli, H. (2013). *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*. John Wiley & Sons, Hoboken, NJ, USA.

- [564] Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J. Clin.*, 68(1):7–30.
- [565] Sijbers, J., Scheunders, P., Verhoye, M., Van der Linden, A., Van Dyck, D., and Raman, E. (1997). Watershed-based segmentation of 3D MR data for volume quantization. *Magn. Reson. Imaging*, 15(6):679–688.
- [566] Silva, S., Vanneschi, L., Cabral, A. I., and Vasconcelos, M. J. (2018). A semi-supervised genetic programming method for dealing with noisy labels and hidden overfitting. *Swarm Evol. Comput.*, 39:323–338.
- [567] Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *Proc. International Conference on Learning Representations (ICLR)*. arXiv preprint arXiv:1409.1556.
- [568] Singh, M., Verma, A., and Sharma, N. (2017). Bat optimization based neuron model of stochastic resonance for the enhancement of mr images. *Biocybern. Biomed. Eng.*, 37(1):124–134.
- [569] Sinop, A. K. and Grady, L. (2007). A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In *Proc. International Conference on Computer Vision (ICCV)*, pages 1–8. IEEE.
- [570] Smistad, E., Falch, T. L., Bozorgi, M., Elster, A. C., and Lindseth, F. (2015). Medical image segmentation on GPUs—a comprehensive review. *Med. Image Anal.*, 20(1):1–18.
- [571] Soille, P. (2004). *Morphological Image Analysis: Principles and Applications*. Springer Science & Business Media, Secaucus, NJ, USA, 2 edition.
- [572] Soille, P. and Talbot, H. (2001). Directional morphological filtering. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1313–1329.
- [573] Soille, P. J. and Ansoult, M. M. (1990). Automated basin delineation from digital elevation models using mathematical morphology. *Signal Process.*, 20(2):171–182.
- [574] Soliman, H., Das, S., Larson, D. A., and Sahgal, A. (2016). Stereotactic radio-surgery (SRS) in the modern management of patients with brain metastases. *Oncotarget*, 7(11):12318.
- [575] Sompong, C. and Wongthanavas, S. (2017). An efficient brain tumor segmentation based on cellular automata and improved tumor-cut algorithm. *Expert Syst. Appl.*, 72:231–244.
- [576] Song, Q., Bai, J., Han, D., Bhatia, S., Sun, W., Rockey, W., Bayouth, J. E., Buatti, J. M., and Wu, X. (2013). Optimal co-segmentation of tumor in PET-CT images with context information. *IEEE Trans. Med. imaging*, 32(9):1685–1697.
- [577] Sonka, M., Hlavac, V., and Boyle, R. (2014). *Image Processing, Analysis, and Machine Vision*. Cengage Learning, Boston, MA, USA, 4 edition.
- [578] Soret, M., Bacharach, S. L., and Buvat, I. (2007). Partial-volume effect in PET tumor imaging. *J. Nucl. Med.*, 48(6):932.

- [579] Sourbron, S. P. and Buckley, D. L. (2013). Classic models for dynamic contrast-enhanced MRI. *NMR Biomed.*, 26(8):1004–1027.
- [580] Spolaor, S., Tangherloni, A., Rundo, L., Nobile, M. S., and Cazzaniga, P. (2017). Reboot strategies in particle swarm optimization and their impact on parameter estimation of biochemical systems. In *Proc. Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–8. IEEE.
- [581] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958.
- [582] Starck, J. L., Murtagh, F., Candes, E. J., and Donoho, D. L. (2003). Gray and color image contrast enhancement by the curvelet transform. *IEEE Trans. Image Process.*, 12(6):706–717.
- [583] Stefano, A., Gallivanone, F., Messa, C., Gilardi, M., and Gastiglioni, I. (2014). Metabolic impact of partial volume correction of [18F] FDG PET-CT oncological studies on the assessment of tumor response to treatment. *Q. J. Nucl. Med. Mol. Imaging*, 58(4):413–423.
- [584] Stefano, A., Vitabile, S., Russo, G., Ippolito, M., Marletta, F., D’arrigo, C., D’Urso, D., Gambino, O., Pirrone, R., Ardizzone, E., et al. (2016). A fully automatic method for biological target volume segmentation of brain metastases. *Int. J. Imaging Syst. Technol.*, 26(1):29–37.
- [585] Stefano, A., Vitabile, S., Russo, G., Ippolito, M., Marletta, F., D’Arrigo, C., D’Urso, D., Sabini, M. G., Gambino, O., Pirrone, R., et al. (2015). An automatic method for metabolic evaluation of Gamma Knife treatments. In *Proc. International Conference on Image Analysis and Processing (ICIAP)*, volume 9279 of *LNCS*, pages 579–589. Springer.
- [586] Stefano, A., Vitabile, S., Russo, G., Ippolito, M., Sabini, M. G., Sardina, D., Gambino, O., Pirrone, R., Ardizzone, E., and Gilardi, M. C. (2017). An enhanced random walk algorithm for delineation of head and neck cancers in PET studies. *Med. Biol. Eng. Comput.*, 55(6):897–908.
- [587] Stewart, E. A. (2001). Uterine fibroids. *Lancet*, 357(9252):293–298.
- [588] Studholme, C., Drapaca, C., Iordanova, B., and Cardenas, V. (2006). Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. *IEEE Trans. Med. Imaging*, 25(5):626–639.
- [589] Studholme, C., Hill, D. L., and Hawkes, D. J. (1999). An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognit.*, 32(1):71–86.
- [590] Styner, M., Brechbuhler, C., Szckely, G., and Gerig, G. (2000). Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Trans. Med. Imaging*, 19(3):153–165.
- [591] Su, P., Yang, J., Li, H., Chi, L., Xue, Z., and Wong, S. T. (2013). Superpixel-based segmentation of glioblastoma multiforme from multimodal MR images. In *Proc. International Workshop on Multimodal Brain Image Analysis (MBIA)*, volume 8159 of *LNCS*, pages 74–83. Springer.

- [592] Suganthan, P. N. (1999). Particle swarm optimiser with neighbourhood operator. In *Proc. Congress on Evolutionary Computation (CEC)*, volume 3, pages 1958–1962. IEEE.
- [593] Summers, R. M., Yao, J., Pickhardt, P. J., Franaszek, M., Bitter, I., Brickman, D., Krishna, V., and Choi, J. R. (2005). Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. *Gastroenterology*, 129(6):1832–1844.
- [594] Sun, C. (2006). Moving average algorithms for diamond, hexagon, and general polygonal shaped window operations. *Pattern Recognit. Lett.*, 27(6):556–566.
- [595] Sun, J., Shi, Y., Gao, Y., and Shen, D. (2017). A point says a lot: an interactive segmentation method for MR prostate via one-point labeling. In *Proc. 2017 International Workshop on Machine Learning in Medical Imaging*, pages 220–228. Springer.
- [596] Suresh, S. and Lal, S. (2017). Modified differential evolution algorithm for contrast and brightness enhancement of satellite images. *Appl. Soft Comput.*, 61:622–641.
- [597] Suzuki, K., Abe, H., MacMahon, H., and Doi, K. (2006). Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Trans. Med. Imaging*, 25(4):406–416.
- [598] Suzuki, K., Horiba, I., and Sugie, N. (2003a). Linear-time connected-component labeling based on sequential local operations. *Comput. Vis. Image Underst.*, 89(1):1–23.
- [599] Suzuki, K., Horiba, I., and Sugie, N. (2003b). Neural edge enhancer for super-resolved edge enhancement from noisy images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(12):1582–1596.
- [600] Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A. (2017). Inception-v4, Inception-ResNet and the impact of residual connections on learning. In *Proc. 2017 AAAI Conference on Artificial Intelligence*.
- [601] Taha, A. A. and Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med. Imaging*, 15(1):29.
- [602] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., et al. (2016). Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans. Med. Imaging*, 35(5):1299–1312.
- [603] Talbi, H. and Batouche, M. (2004). Hybrid particle swarm with differential evolution for multimodal image registration. In *Proc. 2004 IEEE International Conference on Industrial Technology (ICIT)*, volume 3, pages 1567–1572. IEEE.
- [604] Tangherloni, A., Rundo, L., and Nobile, M. S. (2017). Proactive Particles in Swarm Optimization: a settings-free algorithm for real-parameter single objective optimization problems. In *Proc. Congress on Evolutionary Computation (CEC)*, pages 1940–1947. IEEE.
- [605] Tangherloni, A., Rundo, L., Spolaor, S., Nobile, M., Merelli, I., Besozzi, D., Mauri, G., Cazzaniga, P., and Liò, P. (2019). High performance computing for haplotyping: models and platforms. In Mencagli, G. et al., editors, *Proc. International European Conference on Parallel and Distributed Computing (Euro-Par), Workshop on Advances in*



- High-Performance Bioinformatics, Systems Biology (Med-HPC)*, volume 11339 of *LNCS*, pages 650–661. Springer.
- [606] Tangherloni, A., Spolaor, S., Cazzaniga, P., Besozzi, D., Rundo, L., Mauri, G., and Nobile, M. S. (2018a). Benchmark functions do not capture the complexity of biochemical parameter estimation. *Appl. Soft Comput.* (Under review).
- [607] Tangherloni, A., Spolaor, S., Rundo, L., Nobile, M. S., Cazzaniga, P., Mauri, G., Liò, P., Merelli, I., and Besozzi, D. (2018b). GenHap: a novel computational method based on genetic algorithms for haplotype assembly. *BMC Bioinform.* (Accepted), arXiv preprint arXiv:1812.07689.
- [608] Terakawa, Y., Tsuyuguchi, N., Iwai, Y., Yamanaka, K., Higashiyama, S., Takami, T., and Ohata, K. (2008). Diagnostic accuracy of  $^{11}\text{C}$ -methionine PET for differentiation of recurrent brain tumors from radiation necrosis after radiotherapy. *J. Nucl. Med.*, 49:694–699.
- [609] Thapaliya, K. and Kwon, G.-R. (2014). Identification and extraction of brain tumor from MRI using local statistics of Zernike moments. *Int. J. Imaging Syst. Technol.*, 24(4):284–292.
- [610] Thévenaz, P., Blu, T., and Unser, M. (2000). Interpolation revisited [medical images application]. *IEEE Trans. Med. Imaging*, 19(7):739–758.
- [611] Thielmann, H. W., Hagedorn, R., and Freber, W. (1985). Evaluation of colony-forming ability experiments using normal and DNA repair-deficient human fibroblast strains and an automatic colony counter. *Cytom. Part A*, 6(2):130–136.
- [612] Toennies, K. D. (2017). *Guide to Medical Image Analysis*. Springer-Verlag, London, UK, 3 edition.
- [613] Tofts, P. S., Brix, G., Buckley, D. L., Evelhoch, J. L., Henderson, E., Knopp, M. V., Larsson, H. B., Lee, T.-Y., Mayr, N. A., Parker, G. J., et al. (1999). Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J. Magn. Reson. Imaging*, 10(3):223–232.
- [614] Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Proc. International Conference on Computer Vision (ICCV)*, pages 839–846. IEEE.
- [615] Toth, R., Bloch, B. N., Genega, E. M., Rofsky, N. M., Lenkinski, R. E., Rosen, M. A., Kalyanpur, A., Pungavkar, S., and Madabhushi, A. (2011a). Accurate prostate volume estimation using multifeature active shape models on T2-weighted MRI. *Acad. Radiol.*, 18(6):745–754.
- [616] Toth, R., Ribault, J., Gentile, J., Sperling, D., and Madabhushi, A. (2013). Simultaneous segmentation of prostatic zones using active appearance models with multiple coupled level sets. *Comput. Vis. Image Underst.*, 117(9):1051–1060.
- [617] Toth, R., Tiwari, P., Rosen, M., Reed, G., Kurhanewicz, J., Kalyanpur, A., Pungavkar, S., and Madabhushi, A. (2011b). A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation. *Med. Image Anal.*, 15(2):214–225.

- [618] Trelea, I. C. (2003). The particle swarm optimization algorithm: convergence analysis and parameter selection. *Inform. Process. Lett.*, 85(5):317–325.
- [619] Trussell, H. J. (1979). Comments on picture thresholding using an iterative selection method. *IEEE Trans. Syst. Man Cybern.*, 9(5):311–311.
- [620] Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E., and Willsky, A. (2003). A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Med. Imaging*, 22(2):137–154.
- [621] Tsai, H.-Y., Zhang, H., Hung, C.-L., and Min, G. (2017). GPU-accelerated features extraction from magnetic resonance images. *IEEE Access*, 5:22634–22646.
- [622] Tyson, D. R., Garbett, S. P., Frick, P. L., and Quaranta, V. (2012). Fractional proliferation: a method to deconvolve cell population dynamics from single-cell data. *Nat. Methods*, 9(9):923.
- [623] Udupa, J. K. and Samarasekera, S. (1996). Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. *Graph. Models Image Process.*, 58(3):246–261.
- [624] Vallières, M., Freeman, C. R., Skamene, S. R., and El Naqa, I. (2015). A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys. Med. Biol.*, 60(14):5471.
- [625] van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.*, 9:2579–2605.
- [626] van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., and the scikit-image contributors (2014). scikit-image: image processing in Python. *PeerJ*, 2:e453.
- [627] van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G., Fillion-Robin, J.-C., Pieper, S., and Aerts, H. J. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Res.*, 77(21):e104–e107.
- [628] van Netten, J. J., Clark, D., Lazzarini, P. A., Janda, M., and Reed, L. F. (2017). The validity and reliability of remote diabetic foot ulcer assessment using mobile phone images. *Sci. Rep.*, 7(1):9480.
- [629] van Opbroek, A., Vernooij, M. W., Ikram, M. A., and de Bruijne, M. (2015). Weighting training images by maximizing distribution similarity for supervised segmentation across scanners. *Med. Image Anal.*, 24(1):245–254.
- [630] Vanneschi, L., Castelli, M., and Silva, S. (2014). A survey of semantic methods in genetic programming. *Genet. Program. Evolvable Mach.*, 15(2):195–214.
- [631] Venkatesh, M., Mohan, K., and Seelamantula, C. S. (2015). Directional bilateral filters for smoothing fluorescence microscopy images. *AIP Advances*, 5(8):084805.
- [632] Verkauf, B. S. (1993). Changing trends in treatment of leiomyomata uteri. *Curr. Opin. Obstet. Gynecol.*, 5(3):301–310.

- [633] Verma, H., Agrawal, R. K., and Kumar, N. (2014). Improved fuzzy entropy clustering algorithm for MRI brain image segmentation. *Int. J. Imaging Syst. Technol.*, 24(4):277–283.
- [634] Vezhnevets, V. and Konouchine, V. (2005). GrowCut: interactive multi-label nd image segmentation by cellular automata. In *Proc. Graphicon*, volume 1, pages 150–156.
- [635] Villeirs, G. M. and De Meerleer, G. O. (2007). Magnetic resonance imaging (MRI) anatomy of the prostate and application of MRI in radiotherapy planning. *Eur. J. Radiol.*, 63(3):361–368.
- [636] Vincent, G., Guillard, G., and Bowes, M. (2012). Fully automatic segmentation of the prostate using active appearance models. *Proc. 15th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2012 Grand Challenge: Prostate MR Image Segmentation*, pages 1–7.
- [637] Vincent, L. (1993). Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Trans. Image Process.*, 2(2):176–201.
- [638] Vincent, L. and Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(6):583–598.
- [639] Viola, P. and Wells III, W. M. (1997). Alignment by maximization of mutual information. *Int. J. Comput. Vis.*, 24(2):137–154.
- [640] Vitabile, S., Militello, C., Sorbello, F., and Barolli, L. (2011). A programmable networked processing node for 3D brain vessels reconstruction. In *Proc. International Conference on Network-Based Information Systems (NBIS)*, pages 294–301. IEEE.
- [641] Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans Inf. Theory*, 13(2):260–269.
- [642] Von Neumann, J. and Burks, A. W. (1966). *Theory of self-reproducing automata*. University of Illinois Press, Urbana, IL, USA.
- [643] Wachowiak, M. P., Smolíková, R., Zheng, Y., Zurada, J. M., Elmaghraby, A. S., et al. (2004). An approach to multimodal biomedical image registration utilizing particle swarm optimization. *IEEE Trans. Evol. Comput.*, 8(3):289–301.
- [644] Wahl, R. L., Jacene, H., Kasamon, Y., and Lodge, M. A. (2009). From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J. Nucl. Med.*, 50(Suppl 1):122S.
- [645] Wählby, C., Sintorn, I.-M., Erlandsson, F., Borgefors, G., and Bengtsson, E. (2004). Combining intensity, edge and shape information for 2d and 3d segmentation of cell nuclei in tissue sections. *J. Microsc.*, 215(1):67–76.
- [646] Walter, T., Shattuck, D. W., Baldock, R., Bastin, M. E., Carpenter, A. E., Duce, S., Ellenberg, J., Fraser, A., Hamilton, N., Pieper, S., et al. (2010). Visualization of image data from cells to organisms. *Nat. Methods*, 7:S26–S41.

- [647] Wang, G., Li, W., Zuluaga, M. A., Pratt, R., Patel, P. A., Aertsen, M., Doel, T., David, A. L., Deprest, J., Ourselin, S., et al. (2018a). Interactive medical image segmentation using deep learning with image-specific fine-tuning. *IEEE Trans. Med. Imaging*, 37(7):1562–1573.
- [648] Wang, S. and Summers, R. M. (2012). Machine learning and radiology. *Med. Image Anal.*, 16(5):933–951.
- [649] Wang, Y.-Y., Kumar, A., Liu, K.-C., Huang, S.-W., Huang, C.-C., Su, W.-C., Hsiao, F.-L., and Lie, W.-N. (2018b). Stereoscopic augmented reality for single camera endoscopy: a virtual study. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, 6(2):182–191.
- [650] Wang, Z. and Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Process. Lett.*, 9(3):81–84.
- [651] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Proc.*, 13(4):600–612.
- [652] Warfield, S. K., Kaus, M., Jolesz, F. A., and Kikinis, R. (2000). Adaptive, template moderated, spatially varying statistical classification. *Med. Image Anal.*, 4(1):43–55.
- [653] Warfield, S. K., Zou, K. H., and Wells, W. M. (2004). Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging*, 23(7):903–921.
- [654] Weinreb, J. C., Barentsz, J. O., Choyke, P. L., Cornud, F., Haider, M. A., Macura, K. J., et al. (2016). PI-RADS prostate imaging–reporting and data system: 2015, version 2. *Eur. Urol.*, 69(1):16–40.
- [655] Weissler, B., Gebhardt, P., Dueppenbecker, P. M., Wehner, J., Schug, D., Lerche, C. W., Goldschmidt, B., Salomon, A., Verel, I., Heijman, E., et al. (2015). A digital preclinical PET/MRI insert and initial results. *IEEE Trans. Med. Imaging*, 34(11):2258–2270.
- [656] Wernick, M. N. and Aarsvold, J. N. (2004). *Emission Tomography: The Fundamentals of PET and SPECT*. Academic Press, Cambridge, MA, USA, 1 edition.
- [657] Wernick, M. N., Yang, Y., Brankov, J. G., Yourganov, G., and Strother, S. C. (2010). Machine learning in medical imaging. *IEEE Signal Process. Mag.*, 27(4):25–38.
- [658] Win, K., Choomchuay, S., Hamamoto, K., and Raveesunthornkiat, M. (2018). Detection and classification of overlapping cell nuclei in cytology effusion images using a double-strategy random forest. *Appl. Sci.*, 8(9):1608.
- [659] Wojak, J., Angelini, E. D., and Bloch, I. (2010). Joint variational segmentation of ct-pet data for tumoral lesions. In *Proc. International Symposium on Biomedical Imaging (ISBI): From Nano to Macro*, pages 217–220. IEEE.
- [660] Wolfram, S. (1984). Cellular automata as models of complexity. *Nature*, 311(5985):419.

- [661] Wongthanavas, S. and Ponkaew, J. (2016). A cellular automata-based learning method for classification. *Expert Syst. Appl.*, 49:99–111.
- [662] Wu, A. (1992). Physics and dosimetry of the Gamma Knife. *Neurosurg. Clin. N. Am.*, 3(1):35–50.
- [663] Wu, K., Otoo, E., and Suzuki, K. (2009). Optimizing two-pass connected-component labeling algorithms. *Pattern Anal. Appl.*, 12(2):117–135.
- [664] Wu, Q. J. and Bourland, J. D. (1999). Morphology-guided radiosurgery treatment planning and optimization for multiple isocenters. *Med. Phys.*, 26(10):2151–2160.
- [665] Wu, T. and Felmlee, J. P. (2002). A quality control program for MR-guided focused ultrasound ablation therapy. *J. Appl. Clin. Med. Phys.*, 3(2):162–167.
- [666] Wu, W., Chen, A. Y., Zhao, L., and Corso, J. J. (2014). Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features. *Int. J. Comput. Assist. Radiol. Surg.*, 9(2):241–253.
- [667] Xia, Y., Eberl, S., Wen, L., Fulham, M., and Feng, D. D. (2012). Dual-modality brain PET-CT image segmentation based on adaptive use of functional and anatomical information. *Comput. Med. Imaging Graph.*, 36(1):47–53.
- [668] Xia, Y., Wen, L., Eberl, S., Fulham, M., and Feng, D. (2008). Segmentation of dual modality brain PET/CT images using the MAP-MRF model. In *Proc. Workshop on Multimedia Signal Processing*, pages 107–110. IEEE.
- [669] Xiao, B., Tang, H., Jiang, Y., Li, W., and Wang, G. (2018). Brightness and contrast controllable image enhancement based on histogram specification. *Neurocomputing*, 275:2798–2809.
- [670] Xiao, C.-y., Zhang, S., Cheng, S., and Chen, Y.-z. (2005). A novel method for speckle reduction and edge enhancement in ultrasonic images. In *Electronic Imaging and Multimedia Technology IV*, volume 5637 of *Proc. SPIE*, pages 469–477. International Society for Optics and Photonics.
- [671] Xie, K., Yang, J., Zhang, Z., and Zhu, Y. (2005). Semi-automated brain tumor and edema segmentation using MRI. *Eur. J. Radiol.*, 56(1):12–19.
- [672] Xu, C. and Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.*, 7(3):359–369.
- [673] Xu, S. and Rahmat-Samii, Y. (2007). Boundary conditions in particle swarm optimization revisited. *IEEE Trans. Antennas Propag.*, 55(3):760–765.
- [674] Xu, X., Xu, S., Jin, L., and Song, E. (2011). Characteristic analysis of Otsu threshold and its applications. *Pattern Recognit. Lett.*, 32(7):956–961.
- [675] Xu, Z., Bagci, U., Udupa, J. K., and Mollura, D. J. (2015). Fuzzy connectedness image co-segmentation for hybrid PET/MRI and PET/CT scans. In *Comput. Methods Mol. Imaging*, pages 15–24. Springer.

- [676] Xue, H., Srinivasan, L., Jiang, S., Rutherford, M., Edwards, A. D., Rueckert, D., and Hajnal, J. V. (2007). Automatic segmentation and reconstruction of the cortex from neonatal MRI. *Neuroimage*, 38(3):461–477.
- [677] Xue, J. H. and Zhang, Y. J. (2012). Ridler and Calvard’s, Kittler and Illingworth’s and Otsu’s methods for image thresholding. *Pattern Recognit. Lett.*, 33(6):793–797.
- [678] Xue, Y., Xu, T., Zhang, H., Long, R., and Huang, X. (2017). SegAN: adversarial network with multi-scale  $L_1$  loss for medical image segmentation. *arXiv preprint arXiv:1706.01805*.
- [679] Yang, X.-S. (2010). A new metaheuristic bat-inspired algorithm. In *Nature Inspired Cooperative Strategies for Optimization (NICSO)*, volume 284 of *Studies in Computational Intelligence*, pages 65–74. Springer.
- [680] Yankeelov, T. E., Mankoff, D. A., Schwartz, L. H., Lieberman, F. S., Buatti, J. M., Mountz, J. M., Erickson, B. J., Fennessy, F. M., Huang, W., Kalpathy-Cramer, J., et al. (2016). Quantitative imaging in cancer clinical trials. *Clin. Cancer Res.*, 22(2):284–290.
- [681] Yao, J., Chen, D., Lu, W., and Premkumar, A. (2006). Uterine fibroid segmentation and volume measurement on MRI. In *Medical Imaging 2006: Physiology, Function, and Structure from Medical Images*, volume 6143 of *Proc. SPIE*, page 614322. International Society for Optics and Photonics.
- [682] Yao, W., Zeng, Z., Lian, C., and Tang, H. (2018). Pixel-wise regression using U-Net and its application on pansharpening. *Neurocomputing*, 312:364–371.
- [683] Yao, Z. and Yi, W. (2016). Curvature aided Hough transform for circle detection. *Expert Syst. Appl.*, 51:26–33.
- [684] Ye, Q. Z. and Danielsson, P. E. (1988). On minimum error thresholding and its implementations. *Pattern Recognit. Lett.*, 7(4):201–206.
- [685] Yezzi, A., Zollei, L., and Kapur, T. (2001). A variational framework for joint segmentation and registration. In *Proc. Mathematical Methods in Biomedical Image Analysis (MMBIA)*, pages 44–51. IEEE.
- [686] Yin, Y., Fotin, S. V., Periaswamy, S., Kunz, J., Haldankar, H., Muradyan, N., et al. (2012). Fully automated 3D prostate central gland segmentation in MR images: a LOGISMOS based approach. In *Medical Imaging: Image Processing*, volume 8314 of *Proc. SPIE*, page 83143B. International Society for Optics and Photonics.
- [687] Yu, H., Caldwell, C., Mah, K., and Mozeg, D. (2009). Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans. Med. Imaging*, 28(3):374–383.
- [688] Yuan, J., Bae, E., and Tai, X. C. (2010). A study on continuous max-flow and min-cut approaches. In *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2217–2224. IEEE.

- [689] Yuan, Y., Failmezger, H., Rueda, O. M., Ali, H. R., Gräf, S., Chin, S.-F., Schwarz, R. F., Curtis, C., Dunning, M. J., Bardwell, H., et al. (2012). Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci. Transl. Med.*, 4(157):157ra143–157ra143.
- [690] Yuen, H. K., Princen, J., Illingworth, J., and Kittler, J. (1990). Comparative study of Hough transform methods for circle finding. *Image Vision Comput.*, 8(1):71–77.
- [691] Zadeh, L. A. (1965). Fuzzy sets. *Inf. Control*, 8(3):338–353.
- [692] Zadeh, L. A. (1996). Fuzzy logic = computing with words. *IEEE Trans. Fuzzy Syst.*, 4(2):103–111.
- [693] Zaidi, H. and Hasegawa, B. (2003). Determination of the attenuation map in emission tomography. *J. Nucl. Med.*, 44(2):291–315.
- [694] Zaitsev, D. A. (2017). A generalized neighborhood for cellular automata. *Theor. Comput. Sci.*, 666:21–35.
- [695] Zan, E., Yousem, D. M., Carone, M., and Lewin, J. S. (2010). Second-opinion consultations in neuroradiology. *Radiology*, 255(1):135–141.
- [696] Zasadny, K. R. and Wahl, R. L. (1993). Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-d-glucose: variations with body weight and a method for correction. *Radiology*, 189(3):847–850.
- [697] Zeiler, M. D. (2012). ADADELTA: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.
- [698] Zhang, D.-Q. and Chen, S.-C. (2004). A novel kernelized fuzzy c-means algorithm with application in medical image segmentation. *Artif. Intell. Med.*, 32(1):37–50.
- [699] Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., and Metaxas, D. (2017). StackGAN++: realistic image synthesis with stacked generative adversarial networks. *IEEE Trans. Pattern Anal. Mach. Intell.* (In Press).
- [700] Zhang, J. and Jiang, W. (2014). Segmentation for brain magnetic resonance images using dual-tree complex wavelet transform and spatial constrained self-organizing tree map. *Int. J. Imaging Syst. Technol.*, 24(3):208–214.
- [701] Zhang, Y. J. (2001). A review of recent evaluation methods for image segmentation. In *Proc. International Symposium on Signal Processing and its Applications (ISSPA)*, volume 1, pages 148–151. IEEE.
- [702] Zhou, W., Zhang, L., and Jiao, L. (2002). Linear programming support vector machines. *Pattern Recognit.*, 35(12):2927–2936.
- [703] Zhou, X., Li, Y., and Shen, L. (2016). A novel adaptive local thresholding approach for segmentation of HEp-2 cell images. In *Proc. International Conference on Signal and Image Processing (ICSIP)*, pages 174–178. IEEE.

- [704] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proc. International Conference on Computer Vision (ICCV)*, pages 2242–2251. IEEE.
- [705] Zijdenbos, A. P., Dawant, B. M., Margolin, R. A., and Palmer, A. C. (1994). Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Trans. Med. Imaging*, 13(4):716–724.
- [706] Zikic, D., Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O. M., Das, T., Jena, R., and Price, S. J. (2012). Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 7512 of *LNCS*, pages 369–376. Springer.
- [707] Zimmer, Y., Tepper, R., and Akselrod, S. (1997). An improved method to compute the convex hull of a shape in a binary image. *Pattern Recognit.*, 30(3):397–402.
- [708] Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M. C., Kaus, M. R., Haker, S. J., Wells, W. M., Jolesz, F. A., and Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index. *Acad. Radiol.*, 11(2):178–189.



# Appendix A

## Validation methodology

### A.1 Medical image segmentation evaluation metrics

We evaluate the segmentation methods by comparing the segmented images ( $\mathcal{S}$ ) with respect to the corresponding gold standard manual segmentation ( $\mathcal{G}$ ) [601, 182, 701]. In addition to evaluate the automated segmentation  $\mathcal{S}$  against a gold standard  $\mathcal{G}$ , these metrics can be extended to compare any pair of segmentation instances ( $\mathcal{A}_1, \mathcal{A}_2$ ).

#### Volume-based metrics

Accurate prostate volume measurement is crucial in diagnostic phases. Therefore, the volume  $V_S$  computed automatically by the proposed method, should be compared with the actual volume  $V_G$ , contoured manually by an expert radiologist.

First, to evaluate either over-estimation or under-estimation of the automated volume measurements with respect to the reference measure, the absolute average volume difference ( $AVD$ ) between the  $V_S$  and  $V_G$  values is calculated as:

$$AVD = \frac{\text{abs}(V_S - V_G)}{V_G}. \quad (\text{A.1})$$

However, the quantitative measure that best points out the similarity between the volumes  $V_S$  and  $V_G$  is the volumetric similarity ( $VS$ ), namely the absolute difference between volumes divided by the volume sum:

$$VS = \frac{\text{abs}(|V_S| - |V_G|)}{|V_S| + |V_G|} = 1 - VD, \quad (\text{A.2})$$

where  $VD$  is the volumetric distance and  $|\cdot|$  indicates the volume cardinality (in terms of the segmented voxels).

### Spatial overlap-based metrics

These metrics quantify the spatially-overlapping segmented ROI. Let true positives be  $TP = \mathcal{S} \cap \mathcal{G}$ , false negatives be  $FN = \mathcal{G} - \mathcal{S}$ , false positives be  $FP = \mathcal{S} - \mathcal{G}$ , and true negatives be  $TN = \mathcal{I}_\Omega - \mathcal{G} - \mathcal{S}$ . In what follows, we denote the cardinality of the pixels belonging to a region  $\mathcal{A}$  as  $|\mathcal{A}|$ .

- *Dice similarity coefficient* [152] is the most used measure in medical image segmentation to compare the overlap of two regions [708]:

$$DSC = \frac{2 \cdot |TP|}{|\mathcal{S}| + |\mathcal{G}|} \cdot 100. \quad (\text{A.3})$$

In the literature, it has been shown that a  $DSC$  above 70% are generally regarded as a satisfactory level of agreement between two segmentations (i.e., manual and automated delineations) in clinical applications [705, 676, 487].

- *Jaccard index* [290], also known as *Intersection over Union (IoU)*, is another similarity measure, which is the ratio between the cardinality of the intersection and the cardinality of the union calculated on the two segmentation results:

$$JI = \frac{|\mathcal{S} \cap \mathcal{G}|}{|\mathcal{S} \cup \mathcal{G}|} \cdot 100 = \frac{DSC}{2 - DSC}. \quad (\text{A.4})$$

As it can be seen in Eq. (A.4),  $JI$  is strongly related to  $DSC$ . Moreover  $DSC$  is always larger than  $JI$  except at extrema  $\{0\%, 100\%\}$ , where they take equal value.

- *Sensitivity* measures the portion of positive (foreground) pixels correctly detected by the proposed segmentation method with respect to the gold standard segmentation:

$$SEN = TPR = \frac{|TP|}{|TP| + |FN|} \cdot 100. \quad (\text{A.5})$$

- *Specificity* indicates the portion of negative (background) pixels correctly identified by the automatic segmentation against the gold standard segmentation:

$$TNR = \frac{|TN|}{|TN| + |FP|} \cdot 100. \quad (\text{A.6})$$

However, this formulation is ineffective when data are unbalanced (i.e., the ROI is much smaller than the whole image). Consequently, we use the following definition:

$$SPC = \left(1 - \frac{|FP|}{|S|}\right) \cdot 100. \quad (\text{A.7})$$

- *False positive ratio* denotes the presence of false positives compared to the reference region:

$$FPR = \frac{|FP|}{|FP + TN|} \cdot 100. \quad (\text{A.8})$$

- *False negative ratio* is dually defined as:

$$FNR = \frac{|FN|}{|FN + TP|} \cdot 100. \quad (\text{A.9})$$

### Spatial distance-based metrics

As precise boundary tracing plays an important role in clinical practice, overlap-based metrics have limitations in evaluating segmented images. In order to measure the distance between the two ROI boundaries, distance-based metrics can be considered. Let the manual contour  $\mathcal{G}$  consist in a set of vertices  $\{\mathbf{g}_a : a = 1, 2, \dots, A\}$  and the automatically-generated contour  $\mathcal{S}$  consist in a set of vertices  $\{\mathbf{s}_b : b = 1, 2, \dots, B\}$ . We calculate the absolute distance between an arbitrary element  $\mathbf{s}_b \in \mathcal{S}$  and all the vertices in  $\mathcal{G}$  as follows:

$$d(\mathbf{s}_b, \mathcal{G}) = \min_{a \in \{1, 2, \dots, A\}} \|\mathbf{s}_b - \mathbf{g}_a\|. \quad (\text{A.10})$$

- *Average absolute distance* measures the average difference between the ROI boundaries of  $\mathcal{S}$  and  $\mathcal{G}$ :

$$\text{AvgD} = \frac{1}{B} \sum_{b=1}^B d(\mathbf{s}_b, \mathcal{G}). \quad (\text{A.11})$$

- *Maximum absolute distance* represents the maximum difference between the ROI boundaries of  $\mathcal{S}$  and  $\mathcal{G}$ :

$$\text{MaxD} = \max_{b \in \{1, 2, \dots, B\}} d(\mathbf{s}_b, \mathcal{G}). \quad (\text{A.12})$$

- *Hausdorff distance* [82] measures the extent between the set of vertices  $\mathcal{S}$  and  $\mathcal{G}$ , corresponding to the boundaries of regions  $\mathcal{S}$  and  $\mathcal{G}$ , respectively. It is defined as:

$$HD = \max\{h(\mathcal{G}, \mathcal{S}), h(\mathcal{S}, \mathcal{G})\}, \quad (\text{A.13})$$

where  $h(\mathcal{G}, \mathcal{S}) = \max_{\mathbf{g} \in \mathcal{G}} \{\min_{\mathbf{s} \in \mathcal{S}} \{d(\mathbf{g}, \mathbf{s})\}\}$ ,  $h(\mathcal{S}, \mathcal{G}) = \max_{\mathbf{s} \in \mathcal{S}} \{\min_{\mathbf{g} \in \mathcal{G}} \{d(\mathbf{s}, \mathbf{g})\}\}$ , and  $d(\mathbf{g}, \mathbf{s})$  is the Euclidean distance in an  $n$ -dimensional Euclidean space  $\mathbb{R}^n$ .

- Lastly, in order to take into account the correlation of all samples belonging to two different points clouds, a variant of the Mahalanobis distance (*MHD*) is used. Again, let  $V_{\mathcal{S}_1}$  and  $V_{\mathcal{S}_2}$  be the sets of voxels segmented by two different procedures. The *MHD* between  $V_{\mathcal{S}_1}$  and  $V_{\mathcal{S}_2}$  is:

$$MHD = \sqrt{\left(\mu_{V_{\mathcal{S}_1}} - \mu_{V_{\mathcal{S}_2}}\right)^\top \Sigma^{-1} \left(\mu_{V_{\mathcal{S}_1}} - \mu_{V_{\mathcal{S}_2}}\right)}, \quad (\text{A.14})$$

where  $\mu_{V_{\mathcal{S}_1}}$  and  $\mu_{V_{\mathcal{S}_2}}$  are the means of the two 3D segmented volumes. Furthermore,  $\Sigma$  is the common covariance of the two sets and is given by  $\Sigma = \frac{|V_{\mathcal{S}_1}| \Sigma_{V_{\mathcal{S}_1}} + |V_{\mathcal{S}_2}| \Sigma_{V_{\mathcal{S}_2}}}{|V_{\mathcal{S}_1}| + |V_{\mathcal{S}_2}|}$ .

## A.2 Medical image enhancement evaluation metrics

In this section, the definition of the metrics typically used to evaluate image enhancement approaches are recalled, which will be exploited to assess the performance of MedGA. These metrics are essential to quantitatively evaluate the effects of image enhancement techniques, since measuring the “quality” of an image might be strongly subjective. In particular, we benefit here of the metrics considered in [252] to assess the capability of the image enhancement approaches in improving contrast, details and human visual perception.

Let  $\mathcal{I}_{\text{orig}}$  and  $\mathcal{I}_{\text{enh}}$  be the original input image and the enhanced image, respectively, consisting in  $M$  rows and  $N$  columns. Considering that the range of gray levels of the output image  $\mathcal{L}_{\text{out}} = [l_{\text{out}}^{(\min)}, \dots, l_{\text{out}}^{(\max)}]$  could be different from the original range  $\mathcal{L}_{\text{in}} = [l_{\text{in}}^{(\min)}, \dots, l_{\text{in}}^{(\max)}]$ , as a first step before computing the metrics we re-map the output pixel intensities (i.e., the gray levels) onto the original range, as follows:

$$\tilde{\mathcal{I}}_{\text{enh}}(a, b) = \frac{\left(\mathcal{I}_{\text{enh}}(a, b) - l_{\text{out}}^{(\min)}\right) \cdot \left(l_{\text{in}}^{(\max)} - l_{\text{in}}^{(\min)}\right)}{\left(l_{\text{out}}^{(\max)} - l_{\text{out}}^{(\min)}\right)} + l_{\text{in}}^{(\min)}, \quad (\text{A.15})$$

where  $a = 1, \dots, M$  and  $b = 1, \dots, N$ . Note that we actually consider the extended range  $\mathcal{L}'_{in}$  for  $I_{orig}$ .

The Peak Signal-to-Noise Ratio (*PSNR*) denotes the ratio between the maximum possible intensity value of a signal and the distortion between the input and output images:

$$\begin{aligned} PSNR &= 10 \cdot \log_{10} \left( \frac{(I_{in}^{(\max)})^2}{MSE} \right) \\ &= 20 \cdot \log_{10} \left( \frac{I_{in}^{(\max)}}{\sqrt{MSE}} \right), \end{aligned} \quad (\text{A.16})$$

where  $MSE = \frac{1}{M \times N} \sum_{a=1}^M \sum_{b=1}^N \|\mathcal{I}_{orig}(a, b) - \tilde{\mathcal{I}}_{enh}(a, b)\|^2$  is the Mean Squared Error, which allows to compare the pixel values of  $\mathcal{I}_{orig}$  to those of  $\tilde{\mathcal{I}}_{enh}$ .

Furthermore, the *PSNR* is usually expressed in terms of the logarithmic Decibel scale. With regard to our application, we employ only a limited portion of the full dynamic range of 16-bit images, such as in the case of DICOM medical images; we thus use as the largest possible value the maximum intensity value present in the original image (i.e.,  $I_{in}^{(\max)} = \max\{\mathcal{L}_{in}\} = \max\{\mathcal{L}'_{in}\}$ ) instead of the maximum representable value in a 16-bit image (i.e.,  $2^{16} - 1 = 65,535$ ).

[429] stated that good contrast and enhanced images are characterized by high numbers of *edgels* (i.e., pixels belonging to an edge), and that an enhanced image should have a higher intensity of the edges, compared to its non-enhanced counterpart [537]. Therefore, a good enhancement technique should yield satisfactory results in the case of standard vision processing tasks, such as segmentation or edge detection [582]. Here, to evaluate image enhancement of MRI data, we employ the [79]'s method, which is a highly reliable and mathematically well-defined edge detector. This approach deals with weak edges and accurately determines edgels, by applying a double threshold (to identify potential edges) and a hysteresis-based edge tracking. Let  $\mathcal{M}_{Canny}$  be the edge map yielded by the Canny's edge detector, which is a binary image wherein only edgels are set to 1. The number of detected edges (#DE) in  $\mathcal{M}_{Canny}$  is computed as:

$$\#DE = \sum_{a=1}^M \sum_{b=1}^N \mathcal{M}_{Canny}(a, b). \quad (\text{A.17})$$

An additional metrics, called Absolute Mean Brightness Error (*AMBE*) [101, 26], can be employed to measure the brightness preservation of the enhanced image:

$$AMBE = \frac{|\mathbb{E}[\mathcal{I}_{orig}] - \mathbb{E}[\tilde{\mathcal{I}}_{enh}]|}{L}, \quad (\text{A.18})$$

where  $\mathbb{E}[\cdot]$  denotes the expected (mean) value of a gray level distribution. *AMBE* is normalized in  $[0, 1]$ , dividing by  $L = l_{in}^{(\max)} - l_{in}^{(\min)}$ , which is the dynamic range of the input gray-scale (in our case,  $\mathcal{L}'_{in}$ ). Note that low values of *AMBE* denote that the mean brightness of the original image is preserved.

Finally, we consider an alternative quality metrics called Structural Similarity Index (*SSIM*) [651], used to assess the image degradation perceived as variations in structural information [55]. The structural information defines the attributes that represent the structure of objects in the image, independently of the average luminance and contrast. In particular, local luminance and contrast are taken into account since overall values of luminance and contrast can remarkably vary across the whole image. *SSIM* is based on the degradation of structural information—assuming that human visual perception is highly adapted for extracting structural information from a scene—and compares local patterns of pixel intensities. As a matter of fact, natural image signals are highly structured, since pixels are strongly dependent on each other, especially those close by. These dependencies convey important information about the structure of the objects in the viewing field. Let  $\mathbf{X}$  and  $\mathbf{Y}$  be the  $\mathcal{I}_{\text{orig}}$  and  $\mathcal{I}_{\text{enh}}$  image signals, respectively; *SSIM* combines three relatively independent terms:

- the luminance comparison  $l(\mathbf{X}, \mathbf{Y}) = \frac{2\mu_{\mathbf{X}}\mu_{\mathbf{Y}} + \kappa_1}{\mu_{\mathbf{X}}^2 + \mu_{\mathbf{Y}}^2 + \kappa_1}$ ;
- the contrast comparison  $c(\mathbf{X}, \mathbf{Y}) = \frac{2\sigma_{\mathbf{X}}\sigma_{\mathbf{Y}} + \kappa_2}{\sigma_{\mathbf{X}}^2 + \sigma_{\mathbf{Y}}^2 + \kappa_2}$ ;
- the structural comparison  $s(\mathbf{X}, \mathbf{Y}) = \frac{\sigma_{\mathbf{X}\mathbf{Y}} + \kappa_3}{\sigma_{\mathbf{X}}\sigma_{\mathbf{Y}} + \kappa_3}$ ;

where  $\mu_{\mathbf{X}}$ ,  $\mu_{\mathbf{Y}}$ ,  $\sigma_{\mathbf{X}}$ ,  $\sigma_{\mathbf{Y}}$ , and  $\sigma_{\mathbf{X}\mathbf{Y}}$  are the local means, standard deviations, and cross-covariance for the images  $\mathbf{X}$  and  $\mathbf{Y}$ , while  $\kappa_1, \kappa_2, \kappa_3 \in \mathbb{R}^+$  are regularization constants for luminance, contrast, and structural terms, respectively, exploited to avoid instability in the case of image regions characterized by local mean or standard deviation close to zero. Typically, small non-zero values are employed for these constants; according to [651], an appropriate setting is  $\kappa_1 = (0.01 \cdot L)^2$ ,  $\kappa_2 = (0.03 \cdot L)^2$ ,  $\kappa_3 = \kappa_2/2$ , where  $L$  is the dynamic range of the pixel values in  $\mathcal{I}_{\text{orig}}$  (represented in  $\mathcal{L}'_{in}$ ). *SSIM* is then computed by combining the components described above:

$$SSIM = l(\mathbf{X}, \mathbf{Y})^\alpha \cdot c(\mathbf{X}, \mathbf{Y})^\beta \cdot s(\mathbf{X}, \mathbf{Y})^\gamma, \quad (\text{A.19})$$

where  $\alpha, \beta, \gamma > 0$  are weighting exponents. As reported in [651], if  $\alpha = \beta = \gamma = 1$  and  $\kappa_3 = \kappa_2/2$ , the *SSIM* becomes:

$$SSIM = \frac{(2\mu_{\mathbf{X}}\mu_{\mathbf{Y}} + \kappa_1)(2\sigma_{\mathbf{X}\mathbf{Y}} + \kappa_2)}{(\mu_{\mathbf{X}}^2 + \mu_{\mathbf{Y}}^2 + \kappa_1)(\sigma_{\mathbf{X}}^2 + \sigma_{\mathbf{Y}}^2 + \kappa_2)}. \quad (\text{A.20})$$

*SSIM* generalizes the Universal Quality Index (*UQI*), defined in [650], which corresponds to the specific settings having  $\kappa_1 = \kappa_2 = 0$  and yields unstable results when either  $(\mu_X^2 + \mu_Y^2)$  or  $(\sigma_X^2 + \sigma_Y^2)$  tends to zero. Notice that the higher the *SSIM* value, the higher the structural similarity, implying that the enhanced image  $\mathcal{I}_{\text{enh}}$  and the original image  $\mathcal{I}_{\text{orig}}$  are quantitatively similar.

## A.3 Quantitative validation in clonogenic assays

This Section describes the procedure to validate our method based on ACC alone [411] against the conventional cell colony counting.

Generally, for higher robustness, several wells bearing the same experimental configuration are considered. Therefore, a total of different  $r$  repetitions (i.e., number of wells with the number of cells and treatment conditions) is performed.

Moreover, the obtained values are averaged over three counting measurements for each well, performed by a staff composed of three different expert biologists, represent our ground-truth used as a reference for the fully automated approach.

### A.3.1 Conventional evaluation

According to [191], the conventional cell colony counting takes into account the Plating Efficiency (PE) and the Surviving Fraction (SF) calculated using Eqs. (A.21) and (A.22), respectively:

$$\text{PE} = \frac{\# \text{ grown colonies}}{\# \text{ plated cells}} \times 100, \quad (\text{A.21})$$

$$\text{SF} = \frac{\# \text{ colonies grown after treatment}}{\# \text{ plated cells}} \times \text{PE}. \quad (\text{A.22})$$

The PE represents the fraction of colonies formed from cells that were not exposed to the treatment, considering the number of grown colonies and the number of plated cells, and must be determined for each experiment. The SF represents the surviving fraction of cells after any treatment (i.e., irradiation or cytotoxic agent), calculated taking into account the PE of untreated cells (i.e., control well), the number of colonies grown after treatment and the number of plated cells.

### A.3.2 Area-based evaluation

As for the conventional method in laboratory practice, in our approach the SF is measured considering that the ACC of untreated cells is 100, expressed as a percentage, by normalizing as in Eq. (A.23):

$$SF_{\text{Treated cells}} = \frac{ACC_{\text{Treated cells}}}{ACC_{\text{Untreated cells}}} \times 100. \quad (\text{A.23})$$

## A.4 Clinical evaluation based on a Likert scale

The multimodal segmentation method proposed in [526] computes the tumor volumes defined in MRI as well as in MET-PET, simultaneously. Two boundaries are separately generated: one on MRI (i.e., GTV) and one on PET (i.e., BTV or BTV<sub>MRI</sub>). Ideally, both GTV and BTV should be validated against a manual delineation by the experts, considered as our “ground-truth”, using the same PET/MRI datasets. However, due to the current clinical protocol in Leksell Gamma Knife treatment planning, the manual reference contours are available either on MR or PET images only, but not on the fused PET/MR images. Currently, it is not possible to define a real “gold standard” CTV according to both morphologic MRI and metabolic PET images without treatment response evaluation, since the treatment is planned on MRI datasets alone in the clinical practice. As a matter of fact, ground truth refers to having exact knowledge of the tumor size. In PET imaging, and consequently in multimodal PET/MR imaging, the metabolic ground truth is impossible to be obtained for images concerning living humans. Thus, the validation of PET methods with real patient images lacks of an actual gold standard. In addition, manual delineation by expert physicians is subject to intra- and inter-operator variability, especially in tumor segmentation on PET images. Histological images provide the only valid ground truth for quantitative segmentation evaluation in human studies. Obviously, the histology of brain metastases treated using neuro-radiosurgery is unavailable. For this reason, the proposed monomodal PET image segmentation method was assessed using phantom experiments [584]. The same algorithm has been already applied in a real clinical environment to assess the clinical feasibility of the proposed automated approach [584].

So, to assess the clinical utility and feasibility of the proposed multimodal method and to evaluate retrospectively the clinical impact of BTV integration in Gamma Knife treatment planning, a qualitative evaluation, using a five-point Likert scale ranging from 1 (worst) to 5 (best), was carried out by a team of experienced physicians. The clinical staff, composed of a neurosurgeon, a nuclear medicine physician, and a radiation oncologist, jointly analyzed the brain tumors without any information of the used segmentation method



(BTV or  $BTV_{MRI}$ ). By comparing their perspectives by different clinical backgrounds, the physicians were able to gain insights about the tumors and provide a more careful evaluation for each case. As a matter of fact, a Likert scale, rather than providing fixed criteria, allows clinicians to give a score based on overall impression, focusing mainly on the clinical value of BTV in an accurate treatment planning. Although possibly prone to greater inter-observer variability, this approach may be more familiar to some physicians and may perhaps be more straightforward to apply for observers with previous experience in brain PET imaging.

The used Likert scale, according to retrospective clinical value for treatment planning, was defined as:

1. strong worsening in CTV definition;
2. moderate worsening in CTV definition;
3. indifferent, neither enhancement nor worsening in CTV definition;
4. moderate enhancement in CTV definition;
5. strong enhancement in CTV definition.

## A.5 Visual Turing test

To quantitatively evaluate how realistic the synthetic images are, a visual Turing test performed by an expert could be useful [207]. The so-called Visual Turing Test [540] uses binary questions to probe a human ability to identify attributes and relationships in images [207]. For these motivations, it is commonly used to evaluate GAN-generated images, such as for SimGAN [562]. This applies also to medical images in clinical environments [108], wherein physicians' expertise is critical.

