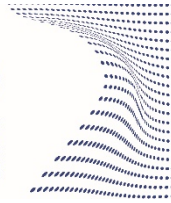




ScuDo
Scuola di Dottorato - Doctoral School
WHAT YOU ARE, TAKES YOU FAR



**UNIVERSITÀ
DEGLI STUDI
DI TORINO**

Doctoral Dissertation
Doctoral Program in Bioengineering and Medical-Surgical Sciences (31th Cycle)

Modeling and Interpretation of Physiopathological Processes using Multimodal and Multiscale Medical Imaging

Massimo Salvi

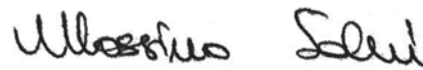
* * * * *

Supervisor
Prof. F. Molinari

Politecnico di Torino
February 2019

This thesis is licensed under a Creative Commons License, Attribution - Noncommercial - NoDerivative Works 4.0 International: see www.creativecommons.org. The text may be reproduced for non-commercial purposes, provided that credit is given to the original author.

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.



.....
Massimo Salvi
Turin, February 2019

Abstract

Nowadays, medical imaging plays an essential role in the detection and diagnosis of several diseases, ranging from the extraction of anatomical and functional information to cellular and molecular expressions. However, visual inspection of biomedical images is often time-consuming, highly subjective and requires experienced operators. Computer aided diagnosis systems (CADx) have been extensively used in clinical practise to support the interpretation of medical images; however, most of the current CADx approaches still entail substantial user-dependency. On the other hand, the development of fully automated solutions is still challenging due to the heterogeneity of the analysed medical images. In fact, even if images are acquired with the same imaging device and standardized protocol, the shape, appearance, and size of internal body structures may vary for different patients (inter-subject variability) and for the same patient at different times (intra-subject variability).

The aim of this thesis work is to develop and validate a series of automated solutions applied to different medical imaging modalities and scales for the modeling and interpretation of physiopathological processes. The proposed approach aims to overcome the limitation of traditional CADx systems and become the bridge technology that enables the effective extraction of quantitative data from biomedical images.

This thesis work can be divided into two macro-sections: firstly, a novel and fully automated strategy for the detection of biological structures is presented. Secondly, six automated algorithms for the reliable quantification and characterization of biomedical images are described.

In the first part, an adaptive algorithm, named ARCO (Adaptive Rapid Curve Optimization), is proposed for the detection of relevant objects in medical images. The ARCO algorithm is the first fully automated method for a fast and accurate segmentation of biological structures in several imaging modalities, where objects

exhibit different shapes, dimensions, and color appearance. The proposed technique has been used as starting point to develop more complex algorithms for the extraction of functional and anatomical information in medical imaging.

In the second part, six algorithms are proposed for the automatic and objective analysis of medical images, ranging from the microscale (optical microscopy) to in-vivo imaging (ultrasound). The first algorithm, named CARE (CARDiosphere Evaluation), is designed for the architectural modeling of cardiac stromal cells in fluorescence images. Then, three fully automated methods are proposed for the analysis of histopathological images. The MANA (Multiscale Adaptive Nuclei Analysis) algorithm is a multi-tissue strategy for nuclei detection, while two other presented methods are for the detection of cancer tissue in prostate and breast histopathological images. Finally, two algorithms are proposed for the extraction of architectural muscle parameters in ultrasound images. The Muscle UltraSound Analysis (MUSA) algorithm is designed to measure the muscle thickness on longitudinal images while the TRAnsverse Muscle Analysis (TRAMA) algorithm is able to measure the muscle cross-sectional area (CSA) in transverse scans.

In conclusion, the proposed techniques achieve high quality results in the architectural and functional modelling of healthy and pathological structures. These algorithms can be extended in the investigation of other organs, diseases and embedded in CADx systems for obtaining a reliable and user-independent diagnosis.

*“Not everything that
can be counted counts,
and not everything that
counts can be counted”*

W. B. Cameron

Contents

1 Introduction	1
1.1 Quantitative Imaging	1
1.2 Optical Microscopy	2
1.2.1 Bright-field Microscopy	3
1.2.2 Confocal Microscopy	4
1.3 Ultrasound Imaging	5
1.4 Automated Algorithms for Quantitative Analysis in Medical Imaging	7
1.5 Objective of the Thesis	8
References	11
2 Multimodal and Multiscale Strategy for Automatic Object Detection in Medical Images	13
2.1 Introduction	14
2.2 Materials and Methods	15
2.2.1 ARCO Algorithm	15
2.2.2 Workflow Description	19
2.3 Results	21
2.3.1 Segmentation results in Confocal Imaging	21
2.3.2 Segmentation results in Bright-field Imaging	23
2.3.3 Segmentation results in other Imaging Modalities	25
2.4 Discussion	27
2.5 Conclusion	28
References	29
3 Architectural and Functional Modeling of Human Cardiospheres	32
3.1 Introduction	33
3.2 Materials and Methods	34
3.2.1 Automated Cell Detection in Human-Derived Cardiospheres: CARE Algorithm	35
3.2.2 Quantitative Spatial Assessment of Markers in 3D Cardiospheres	42
3.3 Results	44
3.3.1 Segmentation results of CARE Algorithm	44
3.3.2 Functional Characterization of 3D Cardiospheres	48
3.4 Discussion	49
3.5 Conclusion	50
References	51

4 Digital Histopathology Image Analysis	53
4.1 Introduction	54
4.2 Materials and Methods	57
4.2.1 Multi-tissue and Multiscale Nuclei Segmentation: MANA Algorithm	57
4.2.2 Automated Gleason Score in Prostate Histological Images	61
4.2.3 Automated Detection of Neoplastic Epithelium in Breast Carcinoma ..	66
4.3 Results	71
4.3.1 Segmentation results of MANA Algorithm	71
4.3.2 Tissue Characterization in Prostate Cancer	74
4.3.3 Tissue Characterization in Breast Carcinoma	76
4.4 Discussion	78
4.5 Conclusion	80
References	82
5 Architectural Analysis in Skeletal Muscle Ultrasound	87
5.1 Introduction	88
5.2 Materials and Methods	89
5.2.1 Longitudinal Muscle Ultrasound Analysis: MUSA Algorithm	89
5.2.2 Transverse Muscle Ultrasound Analysis: TRAMA Algorithm	96
5.3 Results	102
5.3.1 Segmentation results of MUSA Algorithm	102
5.3.2 Segmentation results of TRAMA Algorithm	104
5.4 Discussion	106
5.5 Conclusion	107
References	108
Conclusions and Final Remarks	111
Appendix A	112
Appendix B	116
Appendix C	120
Appendix D	121
List of Contributions	122

Chapter 1

Introduction

1.1 Quantitative Imaging

The recent advance of medical science and the creation of several medicines have benefited humankind and the whole civilization [1]. In this context, medical images play a crucial role in clinical diagnosis and therapy prescription. The term “medical imaging” concerns different technologies which give different information about anatomic structures or physiologic functions to diagnose, monitor or treat medical conditions [2].

The extraordinary technological advancement of imaging methods has led to a significant increase and enrichment of the information that can be extracted from the images [3]. Medical research continues to explore new quantitative imaging biomarkers with the aim to improve the patients’ health [4], [5]. For this reason, quantitative imaging is becoming a reference tool in medicine [6], [7]. Following the QIBA (Quantitative Imaging Biomarkers Alliance) definition [8]:

“Quantitative imaging is the extraction of quantifiable features from medical images for the assessment of normal conditions or the severity, degree of change, or status of a disease, injury, or chronic condition relative to normal. Quantitative imaging includes the development, standardization, and optimization of anatomical, functional, and molecular imaging acquisition protocols, data analyses, display methods, and reporting structures. These features permit the validation of accurately and precisely obtained image-derived metrics with anatomically and physiologically relevant parameters, including treatment response and outcome, and the use of such metrics in research and patient care.”

Nowadays, quantitative imaging is applied through several modalities, including optical and fluorescence microscopy, nuclear medicine, magnetic resonance imaging and ultrasound. Obviously, quantitative imaging is enhanced by wide datasets, which facilitate the evaluation of morphological and functional quantitative features [9].

1.2 Optical Microscopy

For more than two centuries until now, the optical microscope has been a standard tool in life and material science [9], [10]. Often referred to as light microscope, it is a kind of microscope that produces a magnified image of small samples using a visible light and a system of lenses. From an optical microscope, image can be captured by light sensitive cameras to generate digital images.

Basic optical microscopes can be very simple, although more complex models can improve resolution and sample contrast. The most widely used optical microscope configuration is the compound light microscope. Modern compound microscopes have a two-stage magnifying design built around separate lens systems, the *objective* and the *eyepiece* [11]. These two convex lenses converge the light rays that pass through them. The objective lens is positioned close to the object to be viewed and it forms a magnified and upside-down image (*real image*). This image is viewed by the eyepiece lens, acting simply as a magnifying glass. This lens is positioned so that it does not form a second real image, but it makes the light rays spread more so that as they enter the observer's eye, they appear to come from larger inverted image beyond the object lens (*virtual image*). The scheme of a simple compound microscope is shown in Figure 1.1.

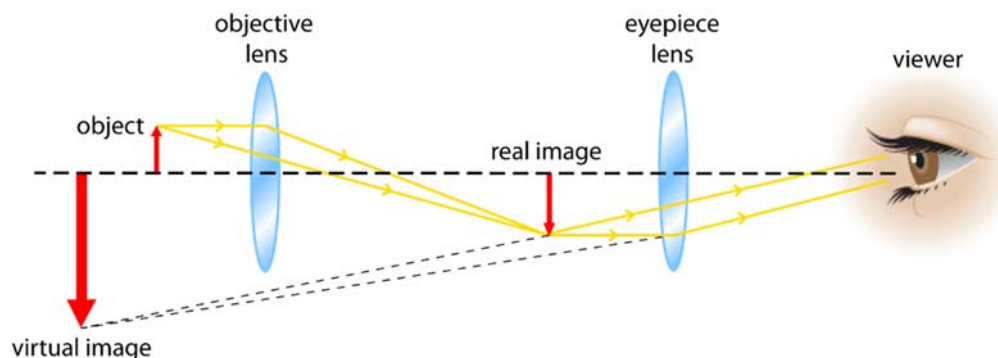


Fig. 1.1 Diagram of a simple compound microscope. The magnification of the object is obtained using two lenses: the objective and the eyepiece lens.

The total magnification of a microscope is calculated by multiplying the individual magnifications of the objective and the eyepiece. For example, an objective lens of 5x combined with an eyepiece lens of 20x will give a total magnification of 100x.

The resolving power of a microscope determines the degree of detail that is visible. Resolution is defined as the minimum distance between two points that can still be distinguished as two separate structures to an unaided eye. The microscope resolution is generally calculated using the Abbe's equation of diffraction limit [12], as shown in Eq. 1.1:

$$r = \frac{\lambda}{n \sin \alpha} \quad (1.1)$$

where r is the minimum resolution, λ is the wavelength of illumination light, n is the refraction index of the medium between lens and point source, and $\sin \alpha$ is half

of the angle of the cone of light from specimen plane accepted by the objective lens. The Abbe's equation shows that the resolution of an optical microscope depends directly on the wavelength of light used to illuminate the specimen and inversely on the numerical aperture ($n \sin \alpha$) of the objective lens.

1.2.1 Bright-field Microscopy

Several types of illumination light can be adopted by a compound light microscope to view magnified objects. The most common and simplest microscope configuration is called bright-field microscopy [13].

In brightfield microscopy, the sample is placed on the microscope translating table and an incandescent light is pointed at a lens, called *condenser*, located beneath the specimen. The condenser usually contains an aperture diaphragm to focus light on the sample, light passes through it and then is collected by an objective lens above the translating table. The objective lens magnifies the light and transmits it to the eyepiece and finally into the user's eyes. Some of the light is absorbed by pigmentation, dense areas or stains of the sample and this contrast allows the user to see the specimen. Figure 1.2 shows a sketch of a bright-field microscope.

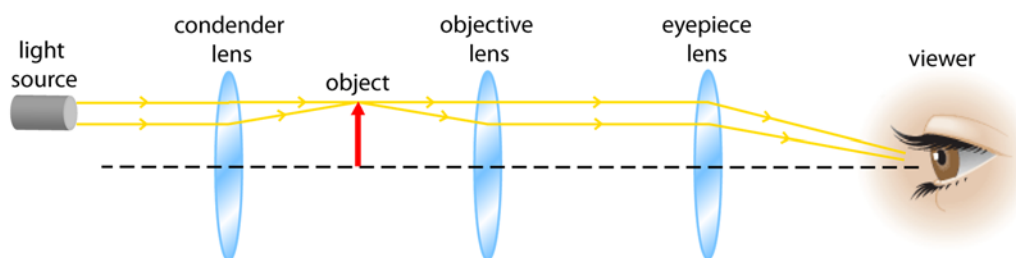


Fig. 1.2 Diagram of a bright-field microscope. The illumination light is focused on the object thanks to the condenser lens, then object magnification is achieved through the objective and eyepiece lens.

The main advantage of this technique is the simplicity of sample preparation. Brightfield microscopy is very easy to use because few adjustments are necessary to view specimens. In addition, the optics used in the brightfield technique don't alter the colour of the specimen.

Bright-field microscopy is intensively used to view live cells or fixed specimens. Most of biomedical samples contain very little intrinsic contrast when viewed directly with transmitted light, many of them are opaque or transparent, so staining is required to increase the contrast between object and background light [14], [15]. An example of images acquired with and without staining is presented in Figure 1.3. Illumination of the specimen is the most important variable to obtain high-quality images in microscopy or digital imaging. Optimal specimen illumination should be glare-free, bright and spread in all the field of view. In addition, the magnification of bright-field microscopy is limited by the resolving power possible with the wavelength of visible light.

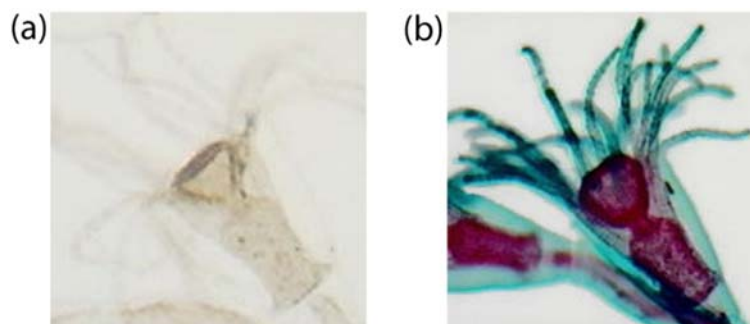


Fig. 1.3 Example of images acquired using bright-field illumination. (a) unstained specimen; (b) stained specimen. Images available from <https://olympus.magnet.fsu.edu/>

1.2.2 Confocal Microscopy

A second type of optical technique is the confocal microscopy. In a conventional bright-field microscope, the specimen is fully illuminated, and the image can be viewed directly by eye. In contrast, the image formation technique in a confocal microscope is quite different. Illumination comes from the scanning of one or more focused beams of light, usually from an arc-discharge source or a laser, across the specimen [16]. A beam splitter (dichroic mirror) is placed between the illumination light and the specimen to allow light of a certain wavelength to pass through, while light of other wavelengths is reflected. Then, the illumination is focused in the specimen by the objective lens and collected by a scanning device under computer control. The sequences of points of light from the specimen are detected by a photomultiplier tube (PMT) through a pinhole, and the output from the PMT is built into an image and displayed by the computer [17]. The confocal microscope configuration is illustrated in Figure 1.4.

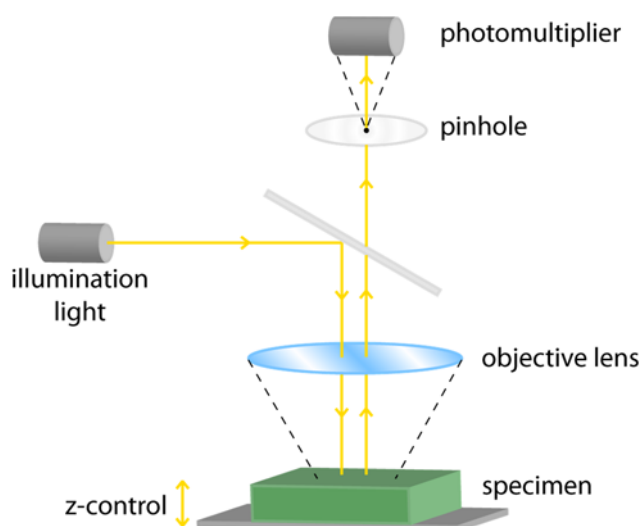


Fig. 1.4 Principle of confocal imaging. Illumination light is focused to the sample by the objective lens, then reflected light is detected by the photomultiplier to create the image.

Confocal microscopy offers several advantages over conventional optical microscopy, including the elimination of image degrading out-of-focus

information, controllable depth of field, and the ability to collect serial optical sections from thick specimens. As only light produced by fluorescence very close to the focal plane can be detected, the image's optical resolution, particularly in the sample depth direction, is much better than bright-field microscopes. However, long exposure times are often required to form the image as much of the light from sample fluorescence is blocked at the pinhole.

In recent years, confocal microscopy has become very popular, mostly for the high-quality images obtained, and for the increasing number of applications in the biological field, based on imaging of tissues and living and fixed cells [17], [18]. In fact, confocal technology is proving to be one of the most important advances ever achieved in optical microscopy. An example of images acquired using a confocal microscope is shown in Figure 1.5.

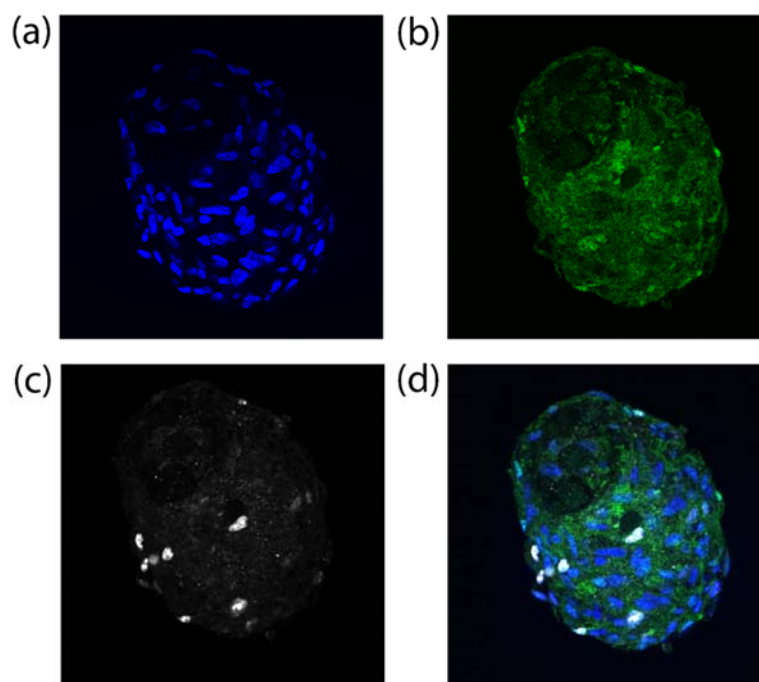


Fig. 1.5 Example of cell images acquired with a confocal microscope. (a) Cell nuclei are highlighted in blue using a fluorescent protein that binds to DNA. (b) Cytoskeletal tension of cells is assessed in green through a fluorescent marker. (c) Proliferative cell activity is shown in white using a fluorescent marker. (d) Combined image obtained by merging the three previous ones.

1.3 Ultrasound Imaging

Ultrasound waves are mechanical vibrations with a frequency greater than 20kHz, higher than the upper limit of the human hearing range [19]. Ultrasound imaging uses high-frequency sound waves (2-20 MHz) to view inside the body. In an ultrasound exam, a transducer (probe) is placed directly on the skin and a thin layer of gel is applied in order to transmit the ultrasound waves from the probe into the body. The interaction between the ultrasound beam and the medium (in this case tissue) involves the formation of the image [20]. This interaction depends on wave characteristics and medium characteristics. The most important

medium property is the *acoustical impedance* which corresponds to the resistance that the tissue opposes to mechanical vibrations. Ultrasonography is based on the fact that ultrasound waves are partially reflected when they meet an interface between two tissues with different acoustic impedance (Figure 1.6) [21].

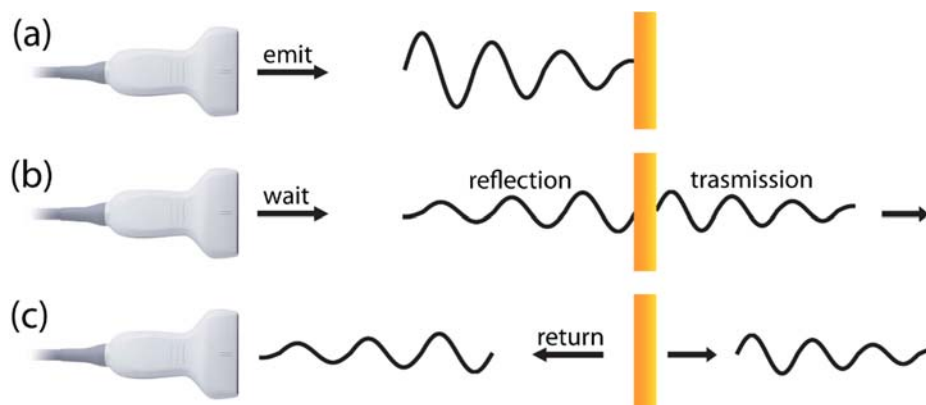


Fig. 1.6 Example of the emission, reflection and transmission of the ultrasound waves. The US probe emits the ultrasound pulse, then waits for returning echoes to reconstruct the image.

In medical imaging the most common mode of representation of the ultrasound signal is the Brightness Mode (B-mode): the result is a two-dimensional grayscale image composed by pixels at different brightness. The image formation from all returning echoes is based on a computer analysis of the acoustic and temporal proprieties of the echoes [22]. The intensity of each pixel is proportional to the amplitude of the return echo while pixel depth depends on the depth at which the echo is produced (d), defined as:

$$d = \frac{1}{2} c \Delta t \quad (1.2)$$

where Δt is the echo travelling time (time between the wave emission and its reception) and c is the mean propagation speed of sound waves in biological tissues (1540 m/s).

As ultrasound images are captured in real-time, they can show the movement of internal organs. In addition, there is no ionizing exposure associated with ultrasound unlike other imaging techniques as X-rays or CT (computed tomography). On the other hand, the ultrasound beam is attenuated through tissues in direct proportion to its frequency: absorption of ultrasound rises with the increasing of beam frequency. For this reason, high frequencies (7.5-15 MHz) are suited to study superficial structures, while low frequencies (2.5-5 MHz) are used to study deeper structures. A higher frequency allows also to achieve a better resolution but a less depth of penetration. For this reason, it is not possible to have at the same time the best resolution and the deeper scan.

In the last decades, medical ultrasonography was used in clinical practise for the evaluation and visualization of abdominal tissues [23], neck and organs like breasts [24], liver [25], thyroid [26] and skeletal muscles [27]. Figure 1.7 shows some application of ultrasound images for different organs.

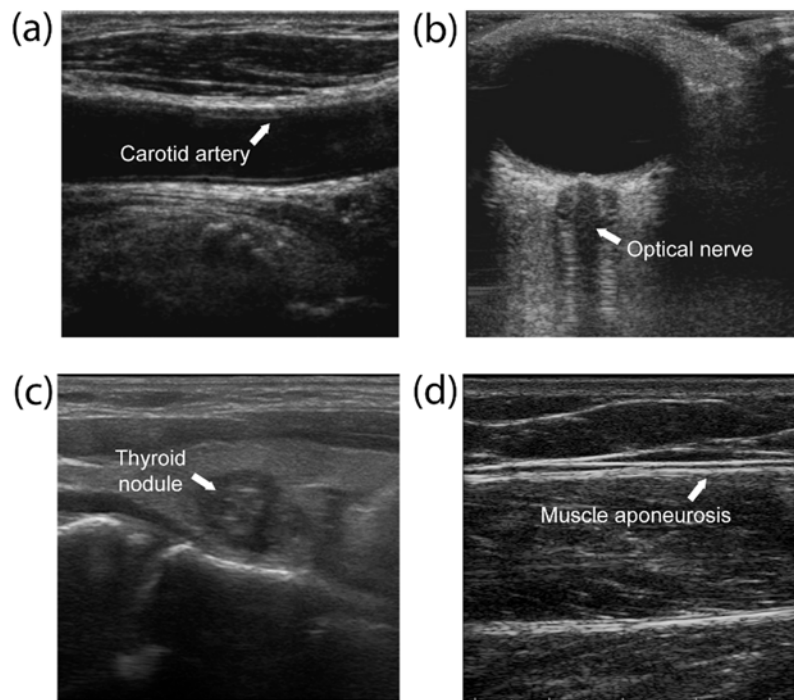


Fig. 1.7 Ultrasound images of different anatomical structures. (a) US image of a carotid artery. (b) Optical nerve evaluation through ultrasound imaging. (c) Thyroid nodule detection in US. (d) Muscle aponeurosis in musculoskeletal ultrasound image.

1.4 Automated Algorithms for Quantitative Analysis in Medical Imaging

Nowadays, medical imaging plays a fundamental role in the diagnosis of several diseases, ranging from anatomical and functional information to cellular and molecular expressions. However, visual inspection of biomedical images is often time-consuming, highly subjective and requires experienced operators. In fact, due to various subjective factors as well as limited analysis time and tools, it is quite common that different medical doctors may come up with diverse interpretations, leading to different diagnoses.

The first idea of developing a computerized scheme for quantitative analysis in medical imaging was reported and discussed by Lodwick in 1960s [28] while the concept of computer-aided detection (CAD) and diagnosis (CADx) was introduced between 1980 and 1990. Nevertheless, computers were not enough powerful at that time and digital images were not easily accessible. In addition, advanced biomedical image processing methods were not available.

At the beginning of 21st century, CAD and CADx systems have spread in many diagnostic field thanks to the progress in medical imaging techniques. In the last years, several approaches have been developed to assist the specialist during the interpretation of biomedical images. CADx systems are used in various medical applications as lung [29], [30], liver [31], prostate [32] and breast cancer [33], cardiovascular diseases [34] and endocrine disorders.

Despite being a powerful tool to obtain more accurate and reliable diagnosis, most of the current CADx approaches still entail substantial user dependency.

Semi-automated tools do not provide repeatable results as their output is closely related to the user interaction (i.e. two specialists may make a different diagnosis on the same set of images). Recently, particular attention was turned on fully automated solutions that allow overcoming the limitations of semi-automatic methods with the aim to provide automatic and objective analysis of medical images. The development of an automatic algorithm is still challenging due to the heterogeneity of the medical images. In fact, even if the images are acquired with the same standardized protocol, the shape, appearance, and size of the body internal structures may vary for different patients (inter-subjected variation) and for the same patient at different time (intra-subject variation). To be useful within a biological context, a fully automated method should be:

- *Fast*: the computational time may not exceed the time of a manual operator analysis. Even if very efficient, an automated method cannot take hours to perform an analysis that is usually done by a manual operator in a few minutes.
- *Robust*: automatic analysis of images in medicine may not provide wrong measurements. The algorithm should be able to identify images that cannot be processed correctly, reject and withdrawn them from further processing.

Modern automated approaches for medical imaging are usually bound to a single application. In addition, the computational cost is often huge and requires high-performance computers to perform the analysis. Finally, most of the current automated solutions are not able to adapt themselves to the inherent variability of biomedical images.

1.5 Objective of the Thesis

The aim of the work presented here is to develop and validate a set of multimodal and multiscale automated solutions for the modeling and interpretation of physiopathological processes in the field of medical imaging. The proposed approaches aim to overcome the limitation of traditional CADx systems and become the bridge technology that enables the effective extraction of quantitative data from biomedical images. The described strategies can be applied to a variety of imaging modalities, such as optical and fluorescence imaging, Magnetic Resonance Imaging (MRI) and ultrasound (US) imaging. The multiscale imaging modalities that directly benefited from the developed solutions in this work are summarized in Figure 1.8.

This thesis work can be divided into two macro-sections: in the first part, a novel and fully automated strategy for the detection of biological structures is presented. In the second part, six automated algorithms for the reliable quantification and characterization of biomedical images are described. Each chapter of this thesis consists of one or more published articles or manuscripts

submitted for peer-reviewed journals. Additional validation and supplemental material complete each study.

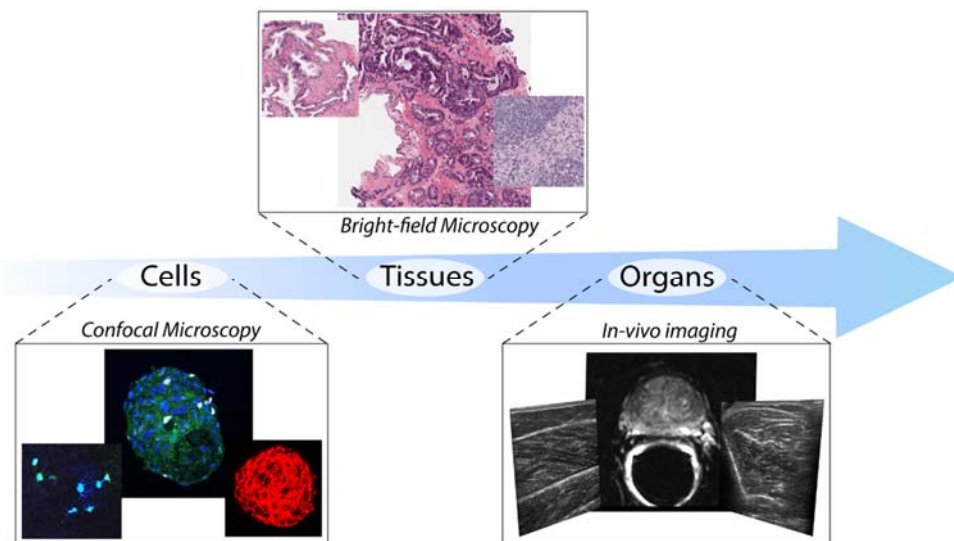


Fig. 1.8 Multimodal and multiscale approach of this thesis work.

In Chapter 2, a multiscale and multimodal algorithm, named ARCO (Adaptive Rapid Curve Optimization), is proposed for the detection of relevant objects in medical images. The ARCO algorithm is the first fully automated method for a fast and accurate segmentation of biological structures in several imaging modalities, where objects exhibit different shapes, dimensions, and color appearance. The proposed technique has been used as starting point to develop more complex algorithms (Chapter 3-4-5) for the extraction of functional and anatomical information in medical imaging.

In Chapter 3, a robust and innovative algorithm, named CARE (CARDiosphere Evaluation), is proposed for the characterization of cardiosphere geometry and architecture. The cardiosphere is a multicellular 3D spheroid used as culture model to reproduce the cell interactions of the cardiac micro-domain and the molecular mechanisms underlying myocardial diseases. The 3D-image processing provided by CARE enabled the quantitative evaluation of the geometrically arranged markers distribution within this model of cardiac niche.

In Chapter 4, three fully automated methods are presented for the analysis and extraction of quantitative data from histopathological images. The first algorithm, named MANA (Multiscale Adaptive Nuclei Analysis), is a multi-tissue method for nuclei segmentation in histological images. The second and third methods are designed for cancer detection in prostate and breast histopathological images respectively. These three algorithms are used to provide an objective, fast and accurate characterization of the histological specimen.

In Chapter 5, two algorithms named MUSA (Muscle UltraSound Analysis) and TRAMA (TRAnverse Muscle Analysis) are proposed for the architectural analysis of four skeletal muscles (medial gastrocnemius, vastus lateralis, rectus femoris and tibialis anterior) in ultrasound imaging. The MUSA algorithm is developed to measure the muscle thickness on longitudinal images while the

TRAMA algorithm is designed for the segmentation of the muscle cross-sectional area (CSA) in transverse scans. The architectural muscle parameters provided by MUSA and TRAMA can be used for the investigation of muscle size and for the detection of muscle atrophy, neuromuscular disorders and other pathological conditions.

The conclusions and final remarks of this work are reported in the last section of the thesis.

References

- [1] D. Ganguly, S. Chakraborty, M. Balitanas, and T. Kim, “Medical imaging: A review,” in *Security-Enriched Urban Computing and Smart Grid*, Springer, 2010, pp. 504–516.
- [2] J. T. Bushberg and J. M. Boone, *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011.
- [3] P. Lambin *et al.*, “Radiomics: extracting more information from medical images using advanced feature analysis,” *Eur. J. Cancer*, vol. 48, no. 4, pp. 441–446, 2012.
- [4] A. J. Buckler, L. Bresolin, N. R. Dunnick, D. C. Sullivan, and Group, “Quantitative imaging test approval and biomarker qualification: interrelated but distinct activities,” *Radiology*, vol. 259, no. 3, pp. 875–884, 2011.
- [5] M. Swan, “Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking,” *Int. J. Environ. Res. Public Health*, vol. 6, no. 2, pp. 492–525, 2009.
- [6] L. G. Kessler *et al.*, “The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions,” *Stat. Methods Med. Res.*, vol. 24, no. 1, pp. 9–26, 2015.
- [7] D. L. Raunig *et al.*, “Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment,” *Stat. Methods Med. Res.*, vol. 24, no. 1, pp. 27–67, 2015.
- [8] S. G. Armato, G. V. Goldmacher, and L. H. Schwartz, “Quantitative imaging biomarkers alliance (QIBA),” in *Radiological Society of North America*, 2014.
- [9] A. B. Rosenkrantz *et al.*, “Clinical utility of quantitative imaging,” *Acad. Radiol.*, vol. 22, no. 1, pp. 33–49, 2015.
- [10] T. Wilson and C. Sheppard, *Theory and practice of scanning optical microscopy*, vol. 180. Academic Press London, 1984.
- [11] A. Boyde, “The Principles and Practice of Electron Microscopy,” *J. Anat.*, vol. 148, p. 268, 1986.
- [12] H. Köhler, “On Abbe’s theory of image formation in the microscope,” *Opt. Acta Int. J. Opt.*, vol. 28, no. 12, pp. 1691–1701, 1981.
- [13] A. G. Valdecasas, D. Marshall, J. M. Becerra, and J. J. Terrero, “On the extended depth of focus algorithms for bright field microscopy,” *Micron*, vol. 32, no. 6, pp. 559–569, 2001.
- [14] D. Ploton, M. Menager, P. Jeannesson, G. Himber, F. Pigeon, and J. J. Adnet, “Improvement in the staining and in the visualization of the argyrophilic proteins of the nucleolar organizer region at the optical level,” *Histochem. J.*, vol. 18, no. 1, pp. 5–14, 1986.
- [15] K. Bamba, “Evaluation of acrosomal integrity of boar spermatozoa by bright field microscopy using an eosin-nigrosin stain,” *Theriogenology*, vol. 29, no. 6, pp. 1245–1251, 1988.
- [16] G. McMahon, *Analytical instrumentation: a guide to laboratory, portable and miniaturized instruments*. John Wiley & Sons, 2008.
- [17] J. Pawley, *Handbook of biological confocal microscopy*. Springer Science & Business Media, 2010.
- [18] C. J. R. Sheppard, D. M. Shotton, and C. Sheppard, *Confocal Laser Scanning Microscopy. Microscopy Handbook*. New York: BIOS Scientific

- Publishers Ltd, 1997.
- [19] J. E. Aldrich, “Basic physics of ultrasound imaging,” *Crit. Care Med.*, vol. 35, no. 5, pp. S131–S137, 2007.
 - [20] P. R. Hoskins, K. Martin, and A. Thrush, *Diagnostic ultrasound: physics and equipment*. Cambridge University Press, 2010.
 - [21] W. R. Hendee, *Medical imaging physics*. John Wiley & Sons, 2003.
 - [22] S. Pillen, I. M. P. Arts, and M. J. Zwarts, “Muscle ultrasound in neuromuscular disorders,” *Muscle Nerve Off. J. Am. Assoc. Electrodiagn. Med.*, vol. 37, no. 6, pp. 679–693, 2008.
 - [23] J. A. Hides, T. Miokovic, D. L. Belavý, W. R. Stanton, and C. A. Richardson, “Ultrasound imaging assessment of abdominal muscle function during drawing-in of the abdominal wall: an intrarater reliability study,” *J. Orthop. Sport. Phys. Ther.*, vol. 37, no. 8, pp. 480–486, 2007.
 - [24] D. D. Adler, P. L. Carson, J. M. Rubin, and D. Quinn-Reid, “Doppler ultrasound color flow imaging in the study of breast cancer: preliminary findings,” *Ultrasound Med. Biol.*, vol. 16, no. 6, pp. 553–559, 1990.
 - [25] A. Berzigotti and L. Castera, “Update on ultrasound imaging of liver fibrosis,” *J. Hepatol.*, vol. 59, no. 1, pp. 180–182, 2013.
 - [26] R. Smith-Bindman *et al.*, “Risk of thyroid cancer based on thyroid ultrasound imaging characteristics: results of a population-based study,” *JAMA Intern. Med.*, vol. 173, no. 19, pp. 1788–1795, 2013.
 - [27] S. Pillen *et al.*, “Skeletal muscle ultrasound: correlation between fibrous tissue and echo intensity,” *Ultrasound Med. Biol.*, vol. 35, no. 3, pp. 443–446, 2009.
 - [28] G. S. Lodwick, C. L. Haun, W. E. Smith, R. F. Keller, and E. D. Robertson, “Computer diagnosis of primary bone tumors: A preliminary report,” *Radiology*, vol. 80, no. 2, pp. 273–275, 1963.
 - [29] R. Agarwal, A. Shankhadhar, and R. K. Sagar, “Detection of lung cancer using content based medical image retrieval,” in *Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on*, 2015, pp. 48–52.
 - [30] H. Yoshimura, M. L. Giger, K. Doi, H. MacMAHON, and S. M. Montner, “Computerized scheme for the detection of pulmonary nodules. A nonlinear filtering technique.,” *Invest. Radiol.*, vol. 27, no. 2, pp. 124–129, 1992.
 - [31] S. Bansal, G. Chhabra, B. S. Chandra, and J. Virmani, “A hybrid CAD system design for liver diseases using clinical and radiological data,” in *U-Healthcare Monitoring Systems*, Elsevier, 2019, pp. 289–314.
 - [32] I. Reda *et al.*, “A new NMF-autoencoder based CAD system for early diagnosis of prostate cancer,” in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, 2016, pp. 1237–1240.
 - [33] M. L. Giger, “Update on the potential of computer-aided diagnosis for breast cancer,” *Futur. Oncol.*, vol. 6, no. 1, pp. 1–4, 2010.
 - [34] H. Fujita, K. Doi, L. E. Fencil, and K. G. Chua, “Image feature analysis and computer-aided diagnosis in digital radiography. 2. Computerized determination of vessel sizes in digital subtraction angiography,” *Med. Phys.*, vol. 14, no. 4, pp. 549–556, 1987.

Chapter 2

Multimodal and Multiscale Strategy for Automatic Object Detection in Medical Images

Part of this chapter has been patented as:

F. Molinari and M. Salvi, **Method for an automatic segmentation and classification of relevant objects in a digital tissue sample image**, Patent application on 30/05/2018, patent number: 102018000005848.

2.1 Introduction

As discussed in the Chapter 1, medical image analysis is an important tool for the detection of several diseases. An automatic processing of a medical image has several advantages including the decrease of the overall time to provide a professional diagnosis and of errors due to a human factor when the image is analyzed.

During medical image analysis, the first step of an automatic algorithm is always the identification of the objects of interest (cells, tissues, structures, etc.). Then, different methods can be applied to the detected objects (morphological or texture analysis) with the aim to provide a diagnosis or to extract quantitative data from the image. For this reason, the correct identification of the relevant objects plays a crucial role during image processing.

Over the years, several automated strategies have been proposed for object recognition in medical images [1], [2]. These segmentation techniques can be divided into macro-classes based on the approach used to recognize the objects of interest. Segmentation algorithms can be grouped into five categories: *pixel-based* models, *edge-based* methods, *region-based* methods, *statistical* models and *deep learning* algorithms.

In *pixel-based* methods, each pixel is segmented according to its intensity value: this operation is called thresholding [3], [4]. Typically, these models use the grayscale image histogram to perform the object detection. The segmentation result strongly depends on the characteristics of the histogram and the image itself. Although they are very simple to implement, these algorithms often fail to achieve satisfactory performance in high-noise images where the gray-level intensity distribution of the object is similar to that of the background.

Edge-based models take advantage of image gradients to perform object recognition [5], [6]. Over the years, complex edge-based methods have been implemented for structures segmentation within medical images (parametric deformable models). These models consist of a parametric and deformable curve guided and influenced by image forces (external constraint forces) that pull it toward features such as edges or lines. However, these parametric models do not adapt itself to the topological variations of the objects (i.e. structures fused) and do not generate high-quality outputs in low contrast images.

Region-based work under the assumption that different objects are separated by perceptual boundaries like neighbourhood or texture features. Through these models, good segmentation results can be achieved even in images where objects exhibit different intensities [7]. Moreover, thanks to the implementation of geometrical deformable models, topological variations of the object can be detected [8]. Nevertheless, region-based methods are often slow and have a high computational cost.

The model-based segmentation (*statistical models*) is performed by matching a model that contains information about the expected shape and appearance of the

structure of interest to new images [9]. Due to the inherent a-priori information, this approach is more stable against local image perturbations and artifacts respect to other techniques. Information about common variations of the object of interest must be included in the model: in fact, a wide and representative training set of all the possible variations of the anatomical structure is needed to obtain reliable performance.

Deep learning approaches exploit deep neural networks (DNNs) for the recognition and classification of objects inside medical images [10], [11]. These types of networks are able to automatically extract features from the image. DNNs are often constructed with a layer-by-layer method and their architecture is much more complex and ‘deep’ than traditional networks. Recently, DNNs achieved state-of-art performance in many segmentation tasks of medical imaging [12]. However, DNNs need a large annotated training set to obtain satisfactory performance and they are not suitable for multimodal or multiscale approaches.

Besides their accuracy and efficiency, all the previously cited methods for object detection present several weak points, such as the low adaptability, the need of wide datasets, the manual initialization and the design oriented for a single application, which make them unsuitable for clinical use. In addition, a completely multimodal and fully automated algorithm for the extraction of relevant objects in medical images has never been proposed in literature.

It is desirable that the automatic processing of a medical image be scalable to a wide variety of object classes, e.g. from sub-cellular objects such as nuclei, to bigger and more complex structures like tissues and organs. It is further desirable that an automatic algorithm be adaptable to the natural variability (size, morphology and color texture) of biological objects. An optimal method should also be robust, fast and with a low computational cost.

In this Chapter, a fully automated strategy for the detection of relevant objects in medical images is described. This method is fast, accurate and can be applied to a variety of imaging modalities such as fluorescence and optical imaging, CT, MRI and so on. To the best of our knowledge, the proposed method is the first multiscale, multimodal and fully automated algorithm for the segmentation of the objects of interest within medical images.

2.2 Materials and Methods

2.2.1 ARCO Algorithm

The ‘core’ of the proposed method for the identification of the relevant objects is the ARCO (Adaptive Rapid Curve Optimization) algorithm. The ARCO algorithm is a multiscale method designed for object segmentation in medical images with different light conditions. The proposed strategy is composed by two steps: *object-based thresholding* and *post-processing refining*. In the following paragraphs, an exhaustive description of the ARCO algorithm is provided.

Object-based detection

This step represents the technical innovation to achieve a first object-based detection. Firstly, the original RGB image is converted into grayscale by eliminating the hue and saturation information while retaining the luminance. Then, the grayscale histogram of the image is computed and the Progressive Weighted Mean (PWM_{CURVE}) of the histogram is calculated.

Considering a grayscale image with pixel intensities expressed by integer numbers between 0 and N. The histogram is defined as a distribution with N+1 classes where the frequency of each gray level within the image is displayed. For the generic class K of the histogram ($0 \leq K \leq N$), the value of PWM_{CURVE} for that class is defined as follows:

$$PWM_{CURVE}(K) = \frac{\sum_{i=0}^K w_i x_i}{\sum_{i=0}^K w_i} \quad (\text{Eq. 2.1})$$

where w_i represents the histogram count of the i^{th} class and x_i is the respective gray level value. For each class of the histogram, the value of the PWM_{CURVE} is defined as the weighted average of all the grayscale histogram values up to that class. The trend and the shape of PWM_{CURVE} strongly depend on the image histogram so significant characteristics of the grayscale distribution can be extracted using this function. Notably, if the histogram shows a significant color variation from a certain point with respect to the distribution that precedes it, here we can expect to see a change of concavity in the PWM_{CURVE} . Inflection points of the PWM_{CURVE} may suggest potential threshold values for objects detection as they are the local stability points of the grayscale histogram.

Conceptually, PWM_{CURVE} is an alternative representation of the color distribution that makes it easier to apply object-based thresholds. Therefore, PWM_{CURVE} can be used to automatically spot objects and structures within the image.

First of all, the PWM_{CURVE} is calculated as described in Equation 2.1 and it is fitted using a polynomial function to estimate its inflection points (*candidate thresholds*). Then, the grayscale image is segmented using all the candidate thresholds and, for each of them, the median area of objects found is evaluated. Among all the candidate thresholds, the algorithm defines as the *optimal threshold* the one that has the segmented objects with the highest median area. Finally, the ARCO algorithm applies the optimal threshold to the grayscale image to obtain a first raw mask of relevant objects inside the image.

The processing for obtaining the optimal threshold is shown in Figure 2.1, where an image acquired using a confocal microscope is used as example. In this case, the objects of interest are the colored and bright spots (brain cells) within the image.

In conclusion, this method implements a novel object-based thresholding and it is robust to different image resolutions, staining, and tissue types. Being a method that can be applied to different types of imaging and objects, some input parameters can be modified by the user to achieve the best performance. The list of the input parameters of the object-based detection is listed in Table 2.1.

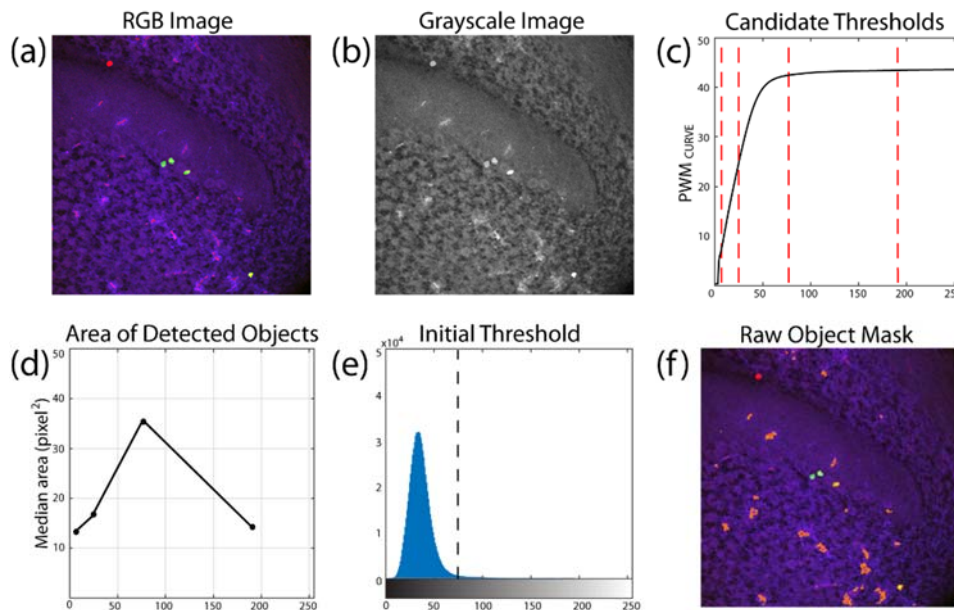


Fig. 2.1 Processing for obtaining the optimal threshold for a confocal microscopy image. (a) Original RGB image. (b) Grayscale image. (c) The PWM_{CURVE} is evaluated from the grayscale histogram and the candidate thresholds (red dotted lines) are estimated as inflection points of the PWM_{CURVE} . (d) Median area of segmented objects for each candidate threshold. (e) The optimal threshold is chosen as the one with the objects with the highest median area. (f) Application of the initial threshold on the RGB image.

Table 2.1 Input parameters of the object-based detection provided by the ARCO algorithm.

Parameters	Description
$image_{LAYER}$	Layer of the image to be processed
$object_{TYPE}$	Specify if the object of interest is bright on a dark background or vice versa
$polynomial_{ORDER}$	Order of the polynomial function that fit the PWM_{CURVE}
$object_{PROPERTY}$	Condition to be imposed on the objects to find the optimal threshold among the candidate ones

$Image_{LAYER}$ defines the layer of the image that should be processed by the ARCO algorithm (i.e. grayscale, red layer, ...). $Object_{TYPE}$ indicates the intensity of relevant objects with respect to the background. This parameter affects the sign applied by the object-based thresholding ('>' if the object is lighter than background, '<' if the object is darker than background). $Polynomial_{ORDER}$ specifies the polynomial order of the function that fit the PWM_{CURVE} . Low values of this parameter may not provide a good 'fit' of the PWM_{CURVE} function, while too high values of $polynomial_{ORDER}$ can generate 'ghost' inflection points due to mathematical interpolation. Finally, $object_{PROPERTY}$ defines which criteria should be imposed on the objects segmented with the candidate thresholds to find the optimal one. For example, if the algorithm should perform a cell nuclei detection,

the optimal threshold is the one that has the objects with the highest median area (Figure 2.1). On the other hand, if the algorithm should identify homogeneous structures, the optimal threshold is defined as the one has the objects' intensity with the lowest standard deviation.

All the ARCO input parameters strongly depend on the type of objects that should be segmented (see next Chapters for more applications). Given a specific application, the ARCO parameters are optimized following the procedure described in Appendix A. Briefly, different combinations of *polynomialORDER* and *objectPROPERTY* are defined (*creation of all the configurations*). Then, for each configuration generated in the previous step, the ARCO algorithm is applied to the entire image dataset. This step is repeated until all the configurations are tested. The optimal combination of the ARCO input parameters is chosen as the one that maximize the object-level $F1_{SCORE}$, which is defined as:

$$F1_{SCORE} = \frac{2TP}{2TP + FN + FP} \quad (\text{Eq. 2.2})$$

where TP (true positive) represents the number of objects identified by ARCO with an overlap higher than 80% with the manual annotations. FN (false negative) denotes all the objects not found by the proposed method (or with an overlap lower than 20% with the manual objects). FP (false positive) represents all objects obtained by ARCO without a corresponding manual drawing.

In general, the value of *polynomialORDER* is correlated with the imaging modality while *objectPROPERTY* depends on the number of expected objects within the image (some sample images are reported in Appendix A).

Post-processing refining

This step is required to correct over-segmentation from the object-based detection as the thresholding can lead to small or too large structures. Too small structures may be under-segmented or wrong objects, whereas too large areas may consist of a fusion of different objects. For this reason, a posteriori check of the output of the object-based detection is needed to increase the robustness of the proposed method. The ARCO algorithm integrates several post-processing refining with the aim to optimize the segmentation result. These post-processing algorithms are grouped into two categories:

- *structure separation*: methods based on the morphological and/or chromatic characteristics of the objects; they perform separation between structures by implementing techniques such as watershed transform [13], [14] or morphological operators [15];
- *structure removal*: selective removal of segmented structures using morphological [16], [17] and/or chromatic characteristics [18]. Geometric constraints can also be applied on the segmented object to correctly recognize the structures of interest.

For example, since the goal of the segmentation shown in Figure 2.1 is the identification of cell nuclei, the ARCO algorithm applies a *structure removal* strategy by deleting all the segmented objects below a certain solidity. Solidity of a structure is defined as the ratio between its actual area and its convex area. Since it is expected that brain cells are convex objects, a segmented region containing a nucleus should have a solidity near to one. For this reason, objects solidity can be used for this specific application as a feature for the removal of unwanted objects. The segmentation result after the post-processing refining is shown in Figure 2.2.

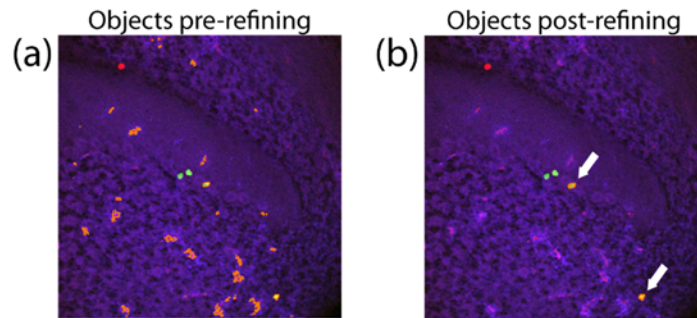


Fig. 2.2 Post-processing refining of the ARCO algorithm. (a) Segmentation pre-refining. (b) Segmentation post-refining. White arrows indicate the objects found at the end of the post-processing refining.

The ARCO algorithm is designed as a high PPV (Positive Predictive Value) method for object segmentation. PPV [19] is defined as:

$$PPV = \frac{TP}{TP + FP} \quad (\text{Eq. 2.3})$$

where TP (True Positive) denotes the number of objects correctly identified by the ARCO algorithm and FP (False Positive) represents all the structures obtained by ARCO without a corresponding manual object. This means that ARCO may miss some objects (False Negative) but it is unlikely that the algorithm finds incorrect objects (False Positives). The ARCO algorithm is integrated into a more complex workflow in order to make the algorithm more reliable and complete. A detailed description of the workflow is provided in the next section.

2.2.2 Workflow Description

To improve the results provided by the ARCO algorithm, an iterative workflow is designed for the segmentation of objects of interest. The workflow adopted by the proposed method is schematically described in Figure 2.3.

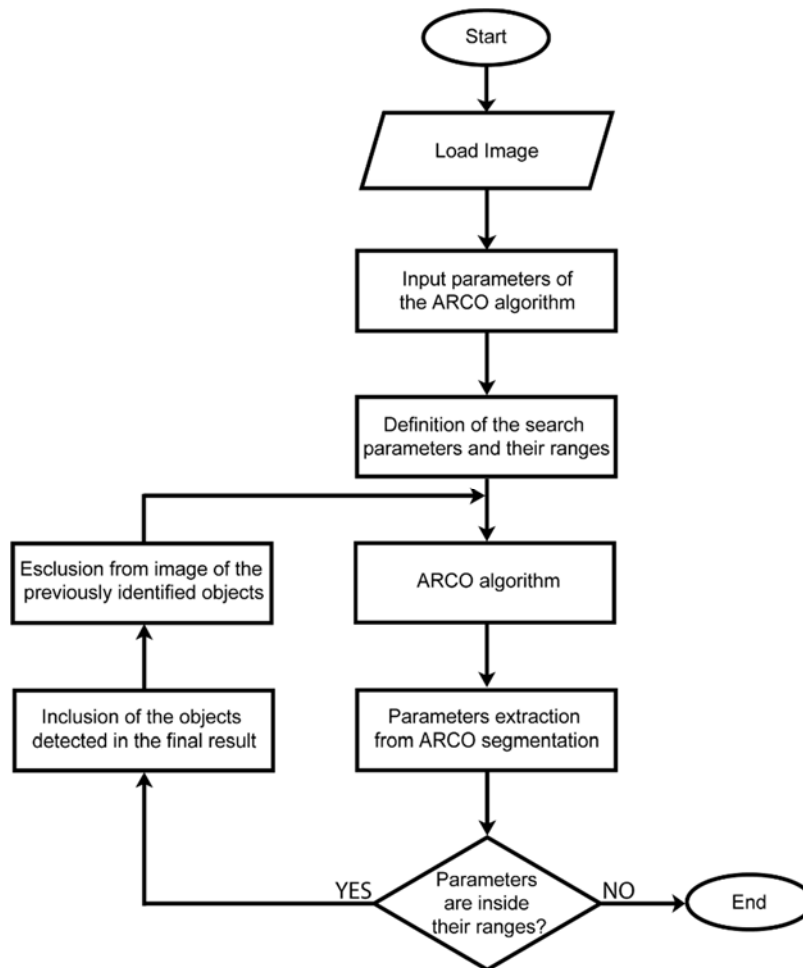


Fig. 2.3 Schematic representation of the proposed workflow for relevant objects detection.

The input data of the workflow are defined in the first three steps. Firstly, the original image is read and stored in memory (*load image*). Then, the *input parameters of the ARCO algorithm* (Table 2.1) are set. Finally, the *definition of the search parameters and their ranges* is performed. The search parameters are defined by the user and they represent the morphological and chromatic characteristics of the relevant objects. On the other hand, the ranges of the search parameters are used as a stopping condition of the workflow: if one or more segmented objects have at least one search parameter outside its range, the processing stops.

Once the input parameters of the workflow have been defined, the ARCO algorithm is applied to the image. Being a high-PPV algorithm, ARCO does not recognize false-objects, but it may not find all the relevant objects at once. Therefore, additional steps are needed to identify all the objects of interest present within the image. First of all, the chromatic and morphological characteristics of the objects segmented with ARCO are extracted (*parameters extraction from ARCO segmentation*). If the extracted parameters are within the ranges of the search parameters, the newly segmented objects are added to the final binary mask (*inclusion of the objects detected in the final result*) and, at the same time,

they are deleted from the original image (*exclusion from image of the previously identified objects*); otherwise the segmentation ends.

All the previous steps are repeated until the objects segmented with ARCO have the characteristics within the ranges of the search parameters. Using this strategy, multiple segmentations can be applied on the same image in order to identify all the objects of interest. The workflow previously described is applied to the image in Figure 2.2 to trace all the cells boundaries. Figure 2.4 shows the iterations performed by the algorithm and the result of the automatic segmentation. Figure 2.4 also provides the comparison between the automatic result (final iteration) and the manual annotation (ground truth).

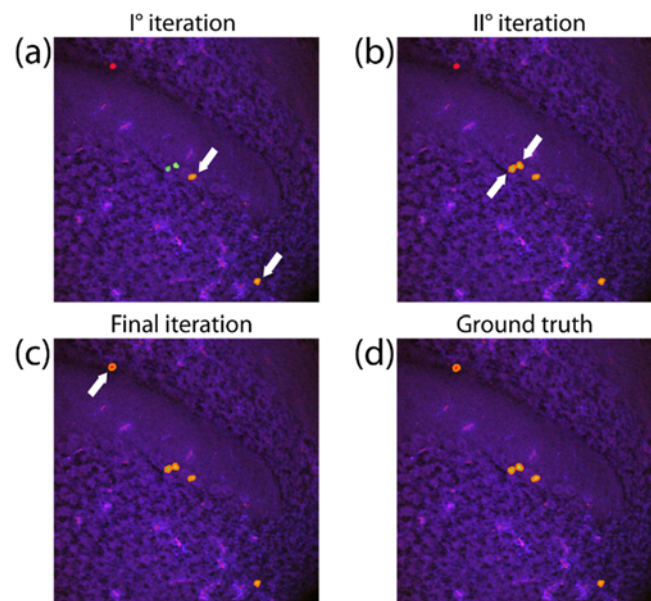


Fig. 2.4 Iterations of the proposed workflow for object segmentation. (a) Result after the first iteration. (b) Result after the second iteration. (c) Final automatic result. (d) Manual annotation. White arrows indicate the objects found at the end of the i -th iteration.

2.3 Results

The ARCO algorithm and the iterative workflow have been applied to different imaging modalities to find all the objects of interest within the image. In the next subsections, some results obtained with the proposed method are shown.

2.3.1 Segmentation results in Confocal Imaging

Brain cells segmentation in fluorescence images

The understanding of how cell diversity within and across distinct brain regions is ontogenetically achieved is a pivotal topic in neuroscience [20], [21]. Clonal analyses based on multicolor cell labeling represent a powerful tool to tackle this issue and disclose lineage relationships [22], but produce enormous sets of fluorescence images, leading to time consuming analyses that may be biased by the operator's subjectivity. For this reason, the ARCO algorithm and the

entire workflow are applied to automatically spot all the brain cells. The segmentation result is illustrated in Figure 2.5. The same figure also shows the robustness of the proposed method, where the cells are found regardless objects' appearance or background intensity.

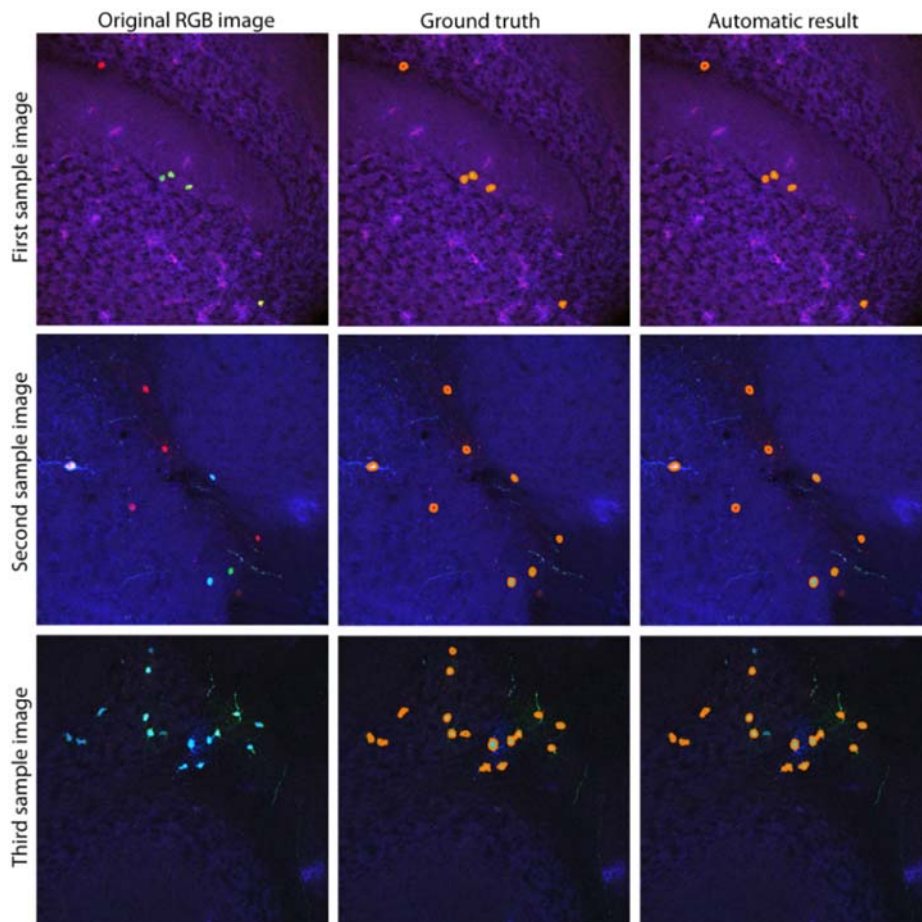


Fig. 2.5 Comparison between manual annotation and automatic result for three samples (rows) of brain cells imaging, showing a high variation of laser intensities and cells appearance. The original RGB image, manual annotation (ground truth) and automatic segmentation are shown in columns.

Membrane segmentation of stem cells in fluorescence images

The ‘cardiosphere’ is a 3D cluster of cardiac progenitor cells recapitulating a stem cell niche-like microenvironment with a potential for disease and regeneration modelling of the failing human myocardium [23], [24]. In this multicellular 3D context, it is extremely important to decrypt the spatial distribution of cells for dissecting the evolution of cellular phenotypes by direct quantification of fluorescent signals in confocal microscopy [25]. An automated method is developed using ARCO to perform cardiosphere membrane segmentation. Figure 2.6 shows the results obtained by applying the proposed algorithm on this type of images. A detailed description of this application is provided in Chapter 3.

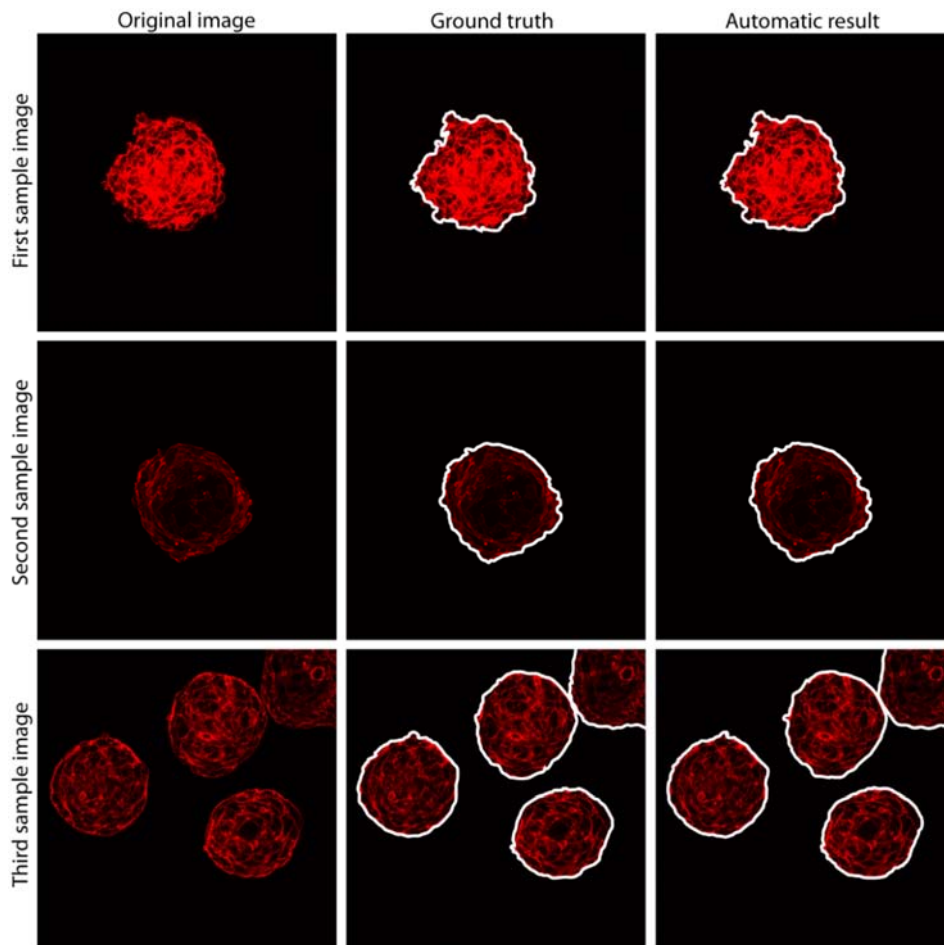


Fig. 2.6 Comparison between manual and automatic segmentation of stem cells for three samples (rows), showing images with a high variation of laser intensities and membranes appearance. The original RGB image, manual annotation (ground truth) and automatic segmentation are shown in columns.

2.3.2 Segmentation results in Bright-field Imaging

Cell nuclei segmentation in digital pathology

Accurate nuclei detection in histopathological images is essential for many clinical purposes like cancer detection and reporting [26], [27]. While manual annotations are operator-dependent and time-consuming, fully automatic solutions still remains challenging for the high variability of cells' appearance [27]. An ARCO-based method is developed for the recognition and segmentation of nuclei in histological images [28]. Figure 2.7 shows the segmentation results obtained on three sub-images taken from different tissues. A detailed description of the algorithm and the application is provided in Chapter 4.

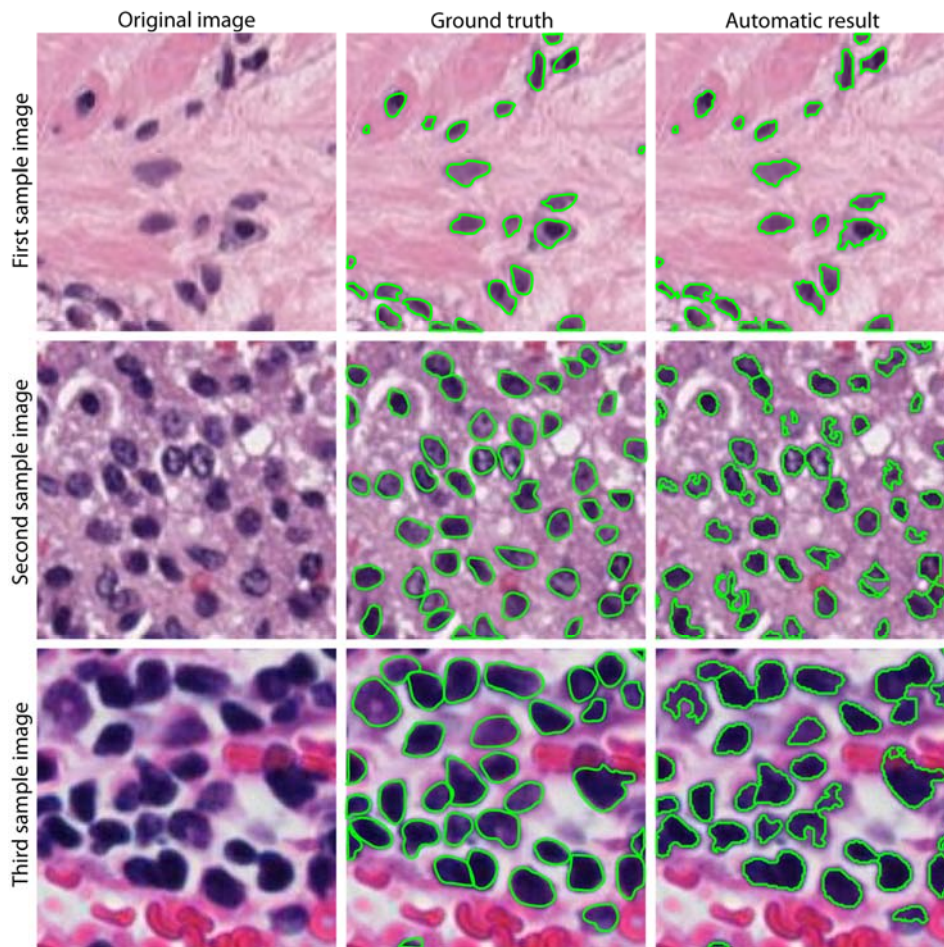


Fig. 2.7 Comparison between manual and automatic segmentation of nuclei cells for three tissues (rows), showing images with a high variation of cells number, size and color. The original RGB image, manual annotation (ground truth) and automatic segmentation are shown in columns.

Lipid vesicles segmentation in bright-field microscopy

The ability to provide a certain dose of medicine inside the body in a specific time and place through artificially designed drug carriers is one of the core challenges in biomedical research [29], [30]. Recent developments in the design and interaction of nanoparticles have led to applications in drug delivery with liposome-based vehicles (lipid vesicles). For this application, an automated algorithm is designed using ARCO for segmentation and tracking of liposomes [31]. Figure 2.8 shows some examples of automatic recognition of liposomes.

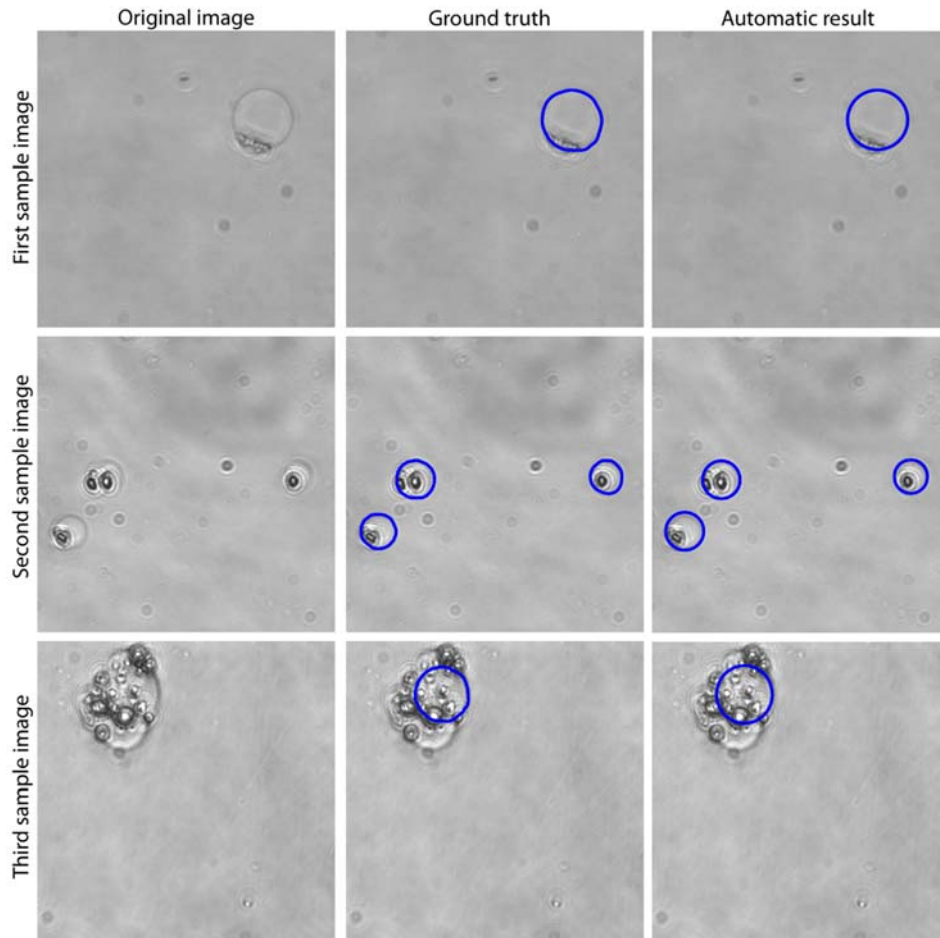


Fig. 2.8 Comparison between manual and automatic segmentation of lipid vesicles for three images (rows), showing samples with a high variation of vesicles size and appearance. The original image, manual annotation (ground truth) and automatic segmentation are shown in columns.

2.3.3 Segmentation results in other Imaging Modalities

Muscle aponeuroses detection in ultrasound imaging

Musculoskeletal ultrasound imaging allows a real-time and non-invasive measurement of the muscle thickness and cross-sectional which are clinically relevant descriptors of muscle size [32]. The extraction of quantitative parameters from muscle ultrasound images has proven to be useful to assess muscle force and its changes during musculoskeletal rehabilitation and training [33] as well as to diagnose neuromuscular disorders [34]. An automated strategy is developed using the proposed workflow for the detection of muscular aponeuroses in longitudinal and transverse ultrasound images [35]. Some examples of segmentation are shown in Figure 2.9. An exhaustive description of this application is provided in Chapter 5.

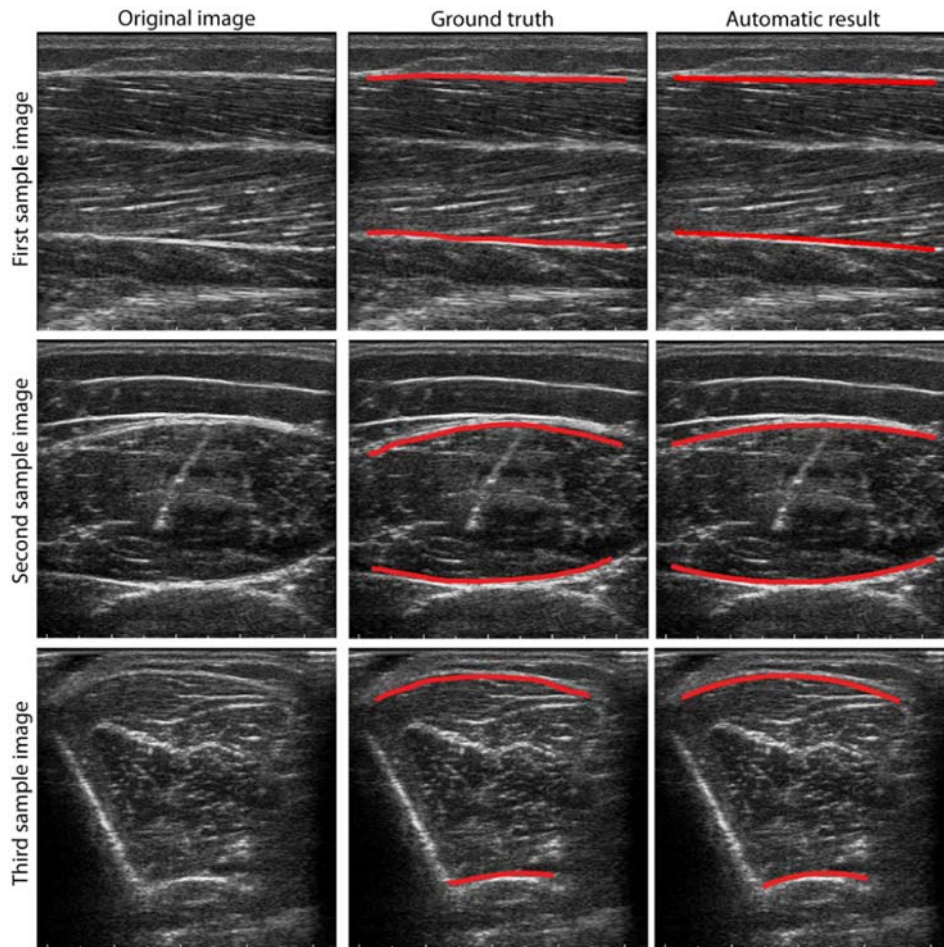


Fig. 2.9 Comparison between manual and automatic segmentation of muscle aponeuroses for three images (rows), showing muscles with a high variation of muscle shape and size. The original image, manual annotation (ground truth) and automatic segmentation are shown in columns.

Prostate segmentation in MR imaging

Magnetic resonance imaging (MRI) has an increasing role in the clinical workup of prostate cancer [36]. Recently, MRI functional imaging (DWI – diffusion weighted imaging) is emerging as a promising tool for the quantification of the prostate cancer aggressiveness [37]. For this reason, prostate gland boundary delineation in DWI represents an important role for MR-guided biopsy, radiation therapy planning and development of computer-aided diagnosis system for the early detection [38], [39]. Starting from ARCO, an automated and unsupervised methodology for the prostate segmentation in DWI sequences is developed. The comparison between manual and automatic tracing is shown in Figure 2.10.

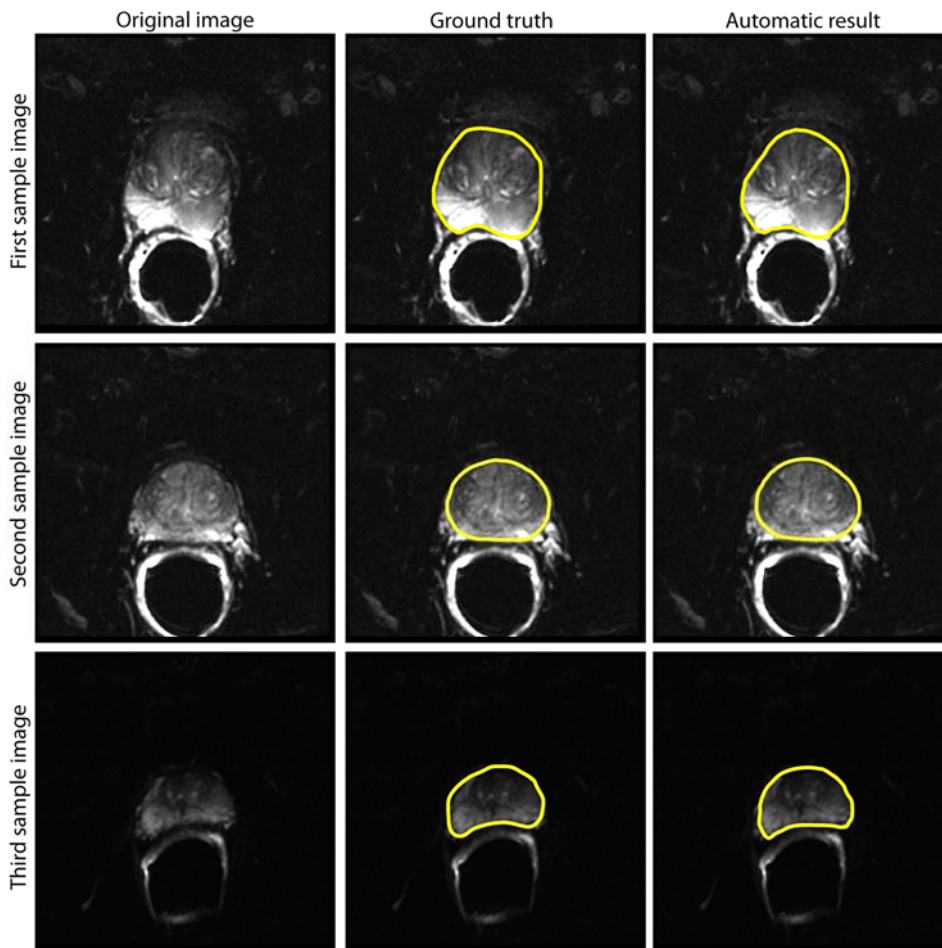


Fig. 2.10 Comparison between manual and automatic segmentation of prostate gland for three images (rows), showing scans with a high variation of intensities and glands appearance. The original image, manual annotation (ground truth) and automatic segmentation are shown in columns.

2.4 Discussion

The segmentation of a bio-image represents a fundamental step in the context of evidence-based medicine [40], radiomics and decision support systems [41]. Object segmentation in the medical field is notoriously a difficult task due to the high variability of the biological structures (in both physiological and pathological conditions) and for the different characteristics of the data acquired with different imaging systems (ultrasound, radiography, magnetic resonance, computed tomography).

In this Chapter, a multiscale and multimodal technique for object segmentation in medical images is presented. The developed method did not require any user interaction and automatically detected different relevant objects in several imaging modalities. The proposed workflow integrates several image processing and segmentation techniques to identify the object contours (post-processing refining). The observed robustness provided by the proposed method is mainly due to the use of a novel object-based detection (ARCO algorithm).

Thanks to the ARCO algorithm, it is possible to recognize and segment objects coming from different types of images, brightness and magnification. Once the user defined the layer to analyze ($image_{LAYER}$), the algorithm can process any type of image, format and color space (RGB, CMYK, YCbCr, ...).

The proposed method analyzes the intensity histogram of the image to perform the segmentation, therefore the computational cost is low, and the computational time is independent of the image complexity. Furthermore, the proposed algorithm does not take as input the image resolution, so it can be used to identify any objects, ranging from cellular and sub-cellular structures to tissues or entire organs. The entire workflow can be applied to the analysis of two-dimensional images of 3D volumes (“plane-wise” analysis) or to the study of entire volumes (volumetric analysis). Moreover, the ARCO algorithm can also be used for the analysis of video frames (dynamic analysis).

Being a fast method, the workflow can also be implemented in a cloud computing architecture. In addition, the proposed algorithm does not need a training set like deep learning algorithms to work properly. The simple and adaptive structure of the algorithm makes it easily integrated into commercial medical devices as a ready-to-use tool for quantitative analysis. This method can also be used for research/discovery in all those applications where the manual operator takes a long time to annotate the data.

Further studies are required to test the reliability and the robustness of the ARCO algorithm in other imaging modalities and applications. In the future, thanks to the speed and robustness provided by this method, fast and high-throughput analyses on medical images could be performed.

2.5 Conclusion

To the best of our knowledge, ARCO is the first automated multimodal and multiscale method for the segmentation of relevant objects in medical images. The developed method is capable of processing images acquired from different acquisition systems where objects present different geometries, dimension and texture.

In the future, thanks to the reliable object detection provided by ARCO, the proposed workflow could be the starting point for the development of automatic methods for quantitative analyses in the medical field. Since it is a completely automated technique and does not require any user-interaction, ARCO could also be used in future works to process large amounts of data and extract 'knowledge' in the field of big data and radiomics.

In the next Chapters, a detailed description of different ARCO-based algorithms designed for morphological and textural characterization of biological structures is provided.

References

- [1] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993.
- [2] P. Sivakumar and S. Meenakshi, "A review on image segmentation techniques," *Int. J. Adv. Res. Comput. Eng. Technol*, vol. 5, no. 3, pp. 641–647, 2016.
- [3] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.
- [4] M. Cheriet, J. N. Said, and C. Y. Suen, "A recursive thresholding technique for image segmentation," *IEEE Trans. image Process.*, vol. 7, no. 6, pp. 918–921, 1998.
- [5] R. Maini and H. Aggarwal, "Study and comparison of various image edge detection techniques," *Int. J. image Process.*, vol. 3, no. 1, pp. 1–11, 2009.
- [6] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, 1990.
- [7] A. Morar, F. Moldoveanu, and E. Gröller, "Image segmentation based on active contours without edges," in *2012 IEEE 8th International Conference on Intelligent Computer Communication and Processing*, 2012, pp. 213–220.
- [8] T. F. Chan, B. Y. Sandberg, and L. A. Vese, "Active contours without edges for vector-valued images," *J. Vis. Commun. Image Represent.*, vol. 11, no. 2, pp. 130–141, 2000.
- [9] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: a review," *Med. Image Anal.*, vol. 13, no. 4, pp. 543–563, 2009.
- [10] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1. MIT press Cambridge, 2016.
- [11] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, 2018.
- [12] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
- [13] V. Grau, A. U. J. Mewes, M. Alcaniz, R. Kikinis, and S. K. Warfield, "Improved watershed transform for medical image segmentation using prior information," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 447–458, 2004.
- [14] P. R. Hill, C. N. Canagarajah, and D. R. Bull, "Image segmentation using a texture gradient based watershed transform," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1618–1633, 2003.
- [15] P. Soille, *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.
- [16] S. Chen and R. M. Haralick, "Recursive erosion, dilation, opening, and closing transforms," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 335–345, 1995.
- [17] E. R. Dougherty and R. A. Lotufo, *Hands-on morphological image processing*, vol. 59. SPIE press, 2003.
- [18] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 7–27,

- 2001.
- [19] R. H. Fletcher, S. W. Fletcher, and G. S. Fletcher, *Clinical epidemiology: the essentials*. Lippincott Williams & Wilkins, 2012.
- [20] J. García-Marqués and L. López-Mascaraque, “Clonal identity determines astrocyte cortical heterogeneity,” *Cereb. cortex*, vol. 23, no. 6, pp. 1463–1472, 2012.
- [21] K. Loulier *et al.*, “Multiplex cell and lineage tracking with combinatorial labels,” *Neuron*, vol. 81, no. 3, pp. 505–520, 2014.
- [22] M. Figueres-Oñate, J. García-Marqués, and L. López-Mascaraque, “UbC-StarTrack, a clonal method to target the entire progeny of individual progenitors,” *Sci. Rep.*, vol. 6, p. 33896, 2016.
- [23] M. T. Woon and T. J. Kamp, “Stem Cells: Put to the test,” *Elife*, vol. 6, p. e24276, 2017.
- [24] D. L. Jones and A. J. Wagers, “No place like home: anatomy and function of the stem cell niche,” *Nat. Rev. Mol. cell Biol.*, vol. 9, no. 1, p. 11, 2008.
- [25] M. Pesce *et al.*, “P4225 YAP-based position sensing in myofibroblast evolution of cardiac stromal cells in human cardiospheres,” *Eur. Heart J.*, vol. 39, no. suppl_1, pp. ehy563-P4225, 2018.
- [26] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, “Histopathological Image Analysis: A Review,” *Biomed. Eng. IEEE Rev.*, vol. 2, pp. 147–171, 2009.
- [27] S. Doyle, M. D. Feldman, N. Shih, J. Tomaszewski, and A. Madabhushi, “Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer,” *BMC Bioinformatics*, vol. 13, no. 1, p. 282, Oct. 2012.
- [28] M. Salvi and F. Molinari, “Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images,” *Biomed. Eng. Online*, vol. 17, no. 1, 2018.
- [29] D. Kagan *et al.*, “Rapid delivery of drug carriers propelled and navigated by catalytic nanoshuttles,” *Small*, vol. 6, no. 23, pp. 2741–2747, 2010.
- [30] T. M. Allen and P. R. Cullis, “Drug delivery systems: entering the mainstream,” *Science (80-.)*, vol. 303, no. 5665, pp. 1818–1822, 2004.
- [31] M. Albini *et al.*, “Movement of giant lipid vesicles induced by millimeter wave radiation change when they contain magnetic nanoparticles,” *Drug Deliv. Transl. Res.*, Sep. 2018.
- [32] O. R. Seynnes *et al.*, “Effect of androgenic-anabolic steroids and heavy strength training on patellar tendon morphological and mechanical properties,” *J. Appl. Physiol.*, 2013.
- [33] F. Damas *et al.*, “Early resistance training-induced increases in muscle cross-sectional area are concomitant with edema-induced muscle swelling,” *Eur. J. Appl. Physiol.*, vol. 116, no. 1, pp. 49–56, Jan. 2016.
- [34] S. Pillen, I. M. P. Arts, and M. J. Zwarts, “Muscle ultrasound in neuromuscular disorders,” *Muscle Nerve*, vol. 37, no. 6, pp. 679–693, Jun. 2008.
- [35] C. Caresio, M. Salvi, F. Molinari, K. M. Meiburger, and M. A. Minetto, “Fully Automated Muscle Ultrasound Analysis (MUSA): Robust and Accurate Muscle Thickness Measurement,” *Ultrasound Med. Biol.*, vol. 43, no. 1, 2017.
- [36] S. Ghai and M. A. Haider, “Multiparametric-MRI in diagnosis of prostate cancer,” *Indian J. Urol. IJU J. Urol. Soc. India*, vol. 31, no. 3, p. 194, 2015.

- [37] T. Hambrock *et al.*, “Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort,” *Eur. Urol.*, vol. 61, no. 1, pp. 177–184, 2012.
- [38] G. Litjens *et al.*, “Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge,” *Med. Image Anal.*, vol. 18, no. 2, pp. 359–373, 2014.
- [39] T. Clark, J. Zhang, S. Baig, A. Wong, M. A. Haider, and F. Khalvati, “Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks,” *J. Med. Imaging*, vol. 4, no. 4, p. 041307, 2017.
- [40] A. Lopez-Perez, R. Sebastian, and J. M. Ferrero, “Three-dimensional cardiac computational modelling: methods, features and applications,” *Biomed. Eng. Online*, vol. 14, no. 1, p. 35, 2015.
- [41] R. J. Gillies, P. E. Kinahan, and H. Hricak, “Radiomics: images are more than pictures, they are data,” *Radiology*, vol. 278, no. 2, pp. 563–577, 2015.

Chapter 3

Architectural and Functional Modeling of Human Cardiospheres

Part of this chapter has been published as:

M. Pesce, M. Salvi, F. Amedeo, F. Angelini, I. Chimenti, M. Agrifoglio, F. Molinari, R. Santoro, E. Messina, and U. Morbiducci, **YAP-based position sensing in myofibroblast evolution of cardiac stromal cells in human cardiospheres**, *European Heart Journal*, issue supplement, 2018; 39.

3.1 Introduction

Heart failure is a clinical condition in continuous growth worldwide, characterized by inflammation and a progressive fibrosis that negatively impacts on the contractility of the heart. This is particularly relevant in elderly people with long course consequences of acute or chronic ischemia and increased physical frailty. Recent studies show cardiac stem cells have a sensing apparatus that makes them specifically reactive to environmental mechanics [1], [2]. Experimental evidences establish the role of heart resident (stem) cells as active “mechanosensors”, capable of showing alterations of the myocardial compliance. Previous studies also demonstrate the presence of a strong mechanical and biological connection in chronic diseases like heart failure [3].

Monitoring the differentiation process of stem/progenitor cells is important either to devise new regenerative medicine approaches, or to understand the molecular basis of chronic diseases involving modifications in tissue structure and property. Until now, this issue has remained unaddressed, also given the lack of systematic tools enabling quantitative investigation (even in real time) of cells dynamics [4] inside vivo micro-domains (the niches) or in disease models [5], [6]. The need for quantitative, cost-effective methods for analyzing, e.g., cell-matrix or cell-cell dynamic interactions, in the last decade has become more and more compelling, stimulated by the plethora of methods recently proposed to engineer tissue-specific 3D microenvironments mimicking the native architecture, i.e. the so-called ‘organoid’ approach [7]. This approach is expected in the future to support ‘synthetic’ tissue/niche modelling [8] for enhanced regenerative medicine applications [9], pathology decryption [10] or fundamental cell differentiation programs in developmental processes [11], [12].

In this context, the cardiosphere is a representative model of cardiac niche, which may be suitable for myocardial regeneration/engineering approaches [13], [14], [15], as well for decryption and modelling of molecular mechanisms underlying myocardial diseases, such as, for example, cardiac fibrosis [16], [3]. The cardiosphere is a multicellular 3D spheroid used as culture model to reproduce the cell interactions of the cardiac micro-domain. Cardiospheres are cultured taking advantage of the natural ability of stromal progenitors to (1) outgrow from explanted human adult-derived atrial appendage tissue, and (2) aggregate onto cell-repulsive culture substrates [17], thereby maintaining cell to cell contacts.

Previous studies have already demonstrated that the outer and the inner cardiosphere environments present remarkable differences, as far as differentiation potency, paracrine signaling and metabolism are concerned [3]. Up to now, only approaches with limited quantitative throughput have been applied to assess cell dynamics inside these complex structures, mostly based on fluorescence-based imaging technology. The quantitative and spatial analysis of fluorescent markers within the cardiosphere has shown promising results for the functional characterization of these 3D cell aggregates [18]. However, in order to quantify the spatial arrangement of fluorescent markers (e.g. marker inside/outside cell nuclei), robust and automated segmentation algorithms are needed.

Cell segmentation within fluorescence microscopy images is a challenging task for an automated algorithm. The main reason is the uneven background intensity caused by the fluorescence of the out-of-focus structures. For an automated method, this unevenness makes the separation between background and foreground a non-trivial task. In addition, intensity variations inside cell nuclei may lead to an over-segmentation as one nucleus can be split into more than one object. Most current nuclei detection approaches in fluorescence microscopy images are based on intensity thresholding [19] and gradients [20]. However, all of these methods have been developed to analyze 2D images and none of these has been applied in a multicellular 3D context.

To bridge the gap of knowledge derived by a paucity of automatic solutions for the specific characterization of cells inside in vitro 3D aggregates, including cardiospheres, here an adaptive algorithm is presented, CARE ('CARdiosphere Evaluation'), for automatic cardiosphere segmentation in fluorescence microscopy images. The proposed technique takes a 3D stacked image from confocal microscopy as input and performs the segmentation and 3D rendering of cardiosphere membranes and nuclei. Starting from the 3D rendering, the spatial distribution of three markers involved in the evolution of cardiac stromal cells (YAP, GATA4 and Ki67) is also evaluated to provide a functional characterization of the cardiosphere.

The aim of this study is to modeling the outer and the inner cardiosphere environments to provide quantitative data on the functional properties of the cardiosphere.

In this Chapter, a robust and innovative automatic algorithm, named CARE is proposed to characterize, in few minutes of processing, the cardiosphere geometry and architecture. To the best of our knowledge, the proposed method is the first fully automated algorithm for the characterization and segmentation of in vitro 3D cell spheroids, including cardiospheres.

3.2 Materials and Methods

Cardiosphere culture

Primary cardiospheres (CSs) were isolated as described in [3] from right atrial appendage biopsies obtained from three donor patients undergoing elective cardiac surgery. All patients signed an informed consent and the entire study was approved by the institutional review board of "Umberto I" Hospital, "La Sapienza" University of Rome.

Briefly, explant cultures were obtained after mechanical fragmentation and enzymatic digestion (trypsin/EDTA 0.05% for 15 minute at room temperature) of myocardial tissue, and plated on fibronectin-coated petri dishes in the following media recipe: Iscove's modified Dulbecco's medium (IMDM) (Sigma-Aldrich) supplemented with 20% FBS (Sigma-Aldrich), 1% penicillin-streptomycin (Sigma-Aldrich), 1% L-glutamine (Lonza, Basel, Switzerland), and 0.1mM2-

mercaptoethanol (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). After 4 weeks, explant cells spontaneously migrated from tissue fragments were harvested with EDTA wash and mild trypsinization (trypsin/EDTA 0.05% for 2-3 minute at room temperature).

Cells were then plated on poly-D-lysine (BD-Biosciences) coated wells (9000 cells/cm²) in the following media: 35% IMDM / 65% DMEM/F-12 Mix (Gibco and Lonza), 3.5% FBS, 1% penicillin-streptomycin, 1% L-glutamine, 0.1mM 2-mercaptoethanol, 1 unit/ml thrombin (Sigma-Aldrich), 1:50 B-27 (Invitrogen), 80 ng/ml bFGF, 25 ng/ml EGF, and 4 ng/ml cardiotrophin-1 (all Peptotech).

Primary cardiospheres (CSs) were harvested by pipetting and centrifugation at 50rcf after 1 week and plated in fibronectin-coated 8-well chamber-slides (Eppendorf) for 3-4 hours to allow attachment. Cardiospheres were then fixed with 4% paraformaldehyde for 10 minutes at room temperature, and then subjected to immunofluorescence staining protocols.

3.2.1 Automated Cell Detection in Human-Derived Cardiospheres: CARE Algorithm

Image database

Twenty 3D-stacks of human-derived cardiospheres obtained from different patients, for a total number of 27 cardiospheres and 1160 slides, are analyzed. Each 3D-stack is acquired using two different lasers to highlight cell membranes (PHAL) and nuclei (DAPI). PHAL (phalloidin) is a peptide used to stain actin filaments that compose the cardiac stem cell membrane (fluorescent color: red). DAPI is a fluorescent protein that bind to DNA and highlights cell nuclei (fluorescent color: blue). Two expert biologists, with more than 10 years of experience, manually annotated membranes and nuclei boundaries.

CARE algorithm architecture

The algorithm is designed to automatically segment cardiosphere-derived cells in fluorescence microscopy imaging. The CARE algorithm is developed in MATLAB environment (MathWorks, Natick, MA, USA). Image processing and analysis is carried out on a workstation with a 2.8 GHz exa-core CPU and 64-GB of RAM. Three main steps compose the processing: PHAL processing, DAPI processing and 3D rendering. The procedure of the CARE algorithm is schematically described in Figure 3.1. In the next paragraphs, an exhaustive description of the proposed method is provided.

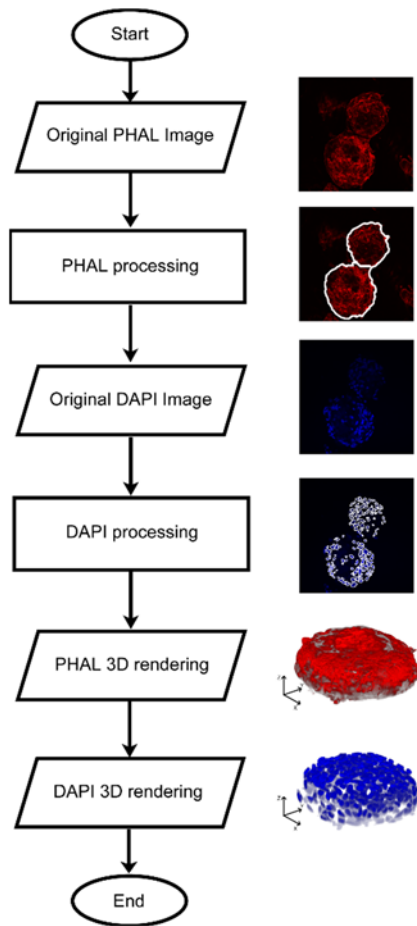


Fig. 3.1 Schematic representation of the CARE algorithm. The proposed technique takes as input the 3D stack of PHAL images to perform membranes segmentation. Then, DAPI images are used to detect all cell nuclei. Finally, CARE performs a 3D rendering of the membranes and nuclei within the cardiosphere.

PHAL processing

The first step of the CARE algorithm is the identification of the cardiosphere membranes by analyzing the 3D stack of the PHAL layer. Then, the identification of the external borders of the cardiospheres is performed by applying an object-based detection scheme to each image of the stack. The core of this step is the ARCO algorithm that we previously developed and adapted to these images (Chapter 2). The input parameters of the ARCO algorithm are listed in Table 3.1.

Table 3.1 Input parameters of ARCO for membrane segmentation in the CARE algorithm.

Parameters	Value
<i>image LAYER</i>	Red layer of the PHAL image
<i>object TYPE</i>	Object of interest is bright on a dark background
<i>polynomial ORDER</i>	PWM_{CURVE} is fitted with a 7 th order polynomial function
<i>object PROPERTY</i>	Minimum standard deviation of the segmented objects is imposed as condition to find the optimal threshold among candidate ones.

Briefly, the red layer of the PHAL image is extracted. From the grayscale histogram of the image, the Progressive Weighted Mean (PWM_{CURVE}) is computed as described in Chapter 2. Then, the PWM_{CURVE} is fitted with a 7th order polynomial function in order to estimate its inflection points (candidate thresholds). Among candidate thresholds, the optimal threshold is chosen as the one that identifies objects with the lowest standard deviation. This condition on the standard deviation is imposed to obtain homogeneous objects. Finally, cardiosphere membranes are found as the regions with an intensity higher than the optimal threshold (bright objects on a dark background).

The processing steps for obtaining the optimal threshold are shown in Figure 3.2, where images with three different laser intensities are presented as explanatory examples. From the results presented in Figure 3.2, it can be appreciated the robustness of the proposed method for cardiospheres border identification, where an optimal threshold image intensity value is selected, regardless of the shape both of the image histogram and of the cardiosphere.

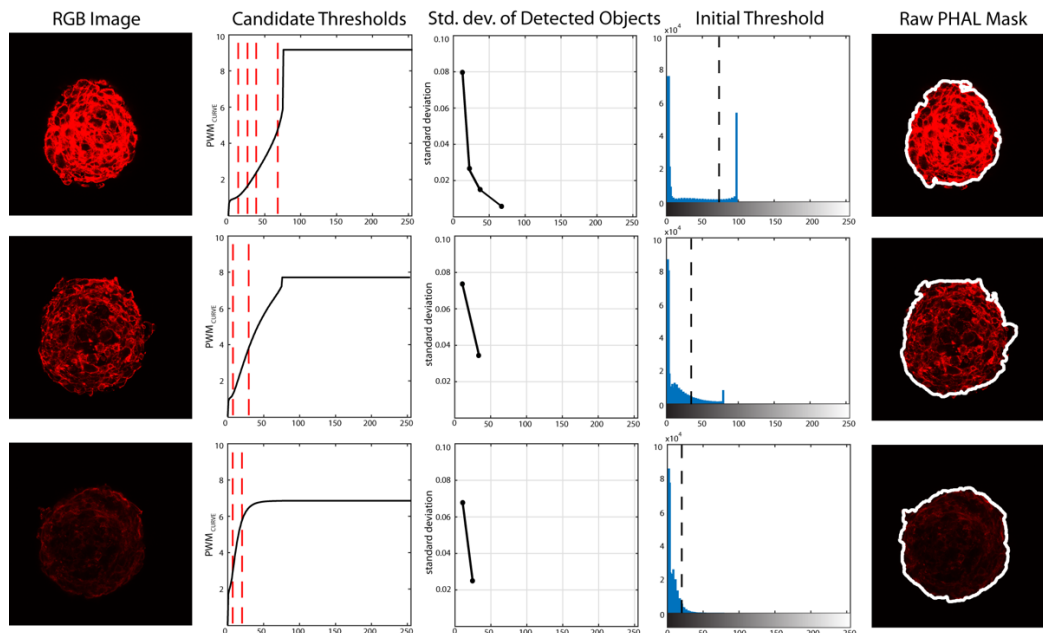


Fig. 3.2 Steps for obtaining the optimal threshold for three images of the stack with a different variation of laser intensity. Starting from the red layer of the RGB image, the PWM_{CURVE} is calculated from the grayscale histogram. Then, the inflection points of the PWM_{CURVE} (red dotted lines) are estimated (candidate thresholds). For each candidate threshold, the standard deviation of detected objects intensity is evaluated. The optimal threshold (black-dotted line) is chosen as the one that identifies objects with the lowest standard deviation. In the rightmost column, the application of the optimal threshold on the PHAL image is illustrated.

The herein develop method also includes an automatic strategy for the refinement of the shapes of the objects detected. Preliminarily, detected objects with area less than $1200 \mu\text{m}^2$ are deleted because they are too small to be considered as

cardiospheres. Then, starting from the first frame of the stack, the CARE algorithm also performs an iterative four-steps processing procedure to further clean the obtained masks:

1. definition of *reference frame* as the current frame;
2. definition of *realign frame* as the next frame after the reference frame;
3. deletion of all the objects inside the *realign frame* with an overlapping lower than 75% with *reference frame* objects;
4. move on to next frame (with the current *realign frame* becoming the next current frame).

The procedure described above is extended to all the images of the stack. With this operation, all previously segmented objects that do not belong to the cardiosphere are deleted. The refining process for a sample image is presented in Figure 3.3a-c.

The object-based detection provided by ARCO has a high sensitivity, but sometimes it may lead to suboptimal profiles, given by two or more cardiospheres that are very close to each other. In such a case, the automatic algorithm may depict them as a single object. For this reason, the ARCO algorithm incorporates a post-processing refining to overcome this issue.

The aim of this step is to separate cardiospheres still fused together. In previous works, the watershed transform was successfully used to isolate touching structures [21]. Using this transform, the image is treated as a topographic map where the intensity of each point represents its height. In order to separate touching objects, the watershed transform finds the lines that run along the tops of ridges. However, objects of interest may have uneven color, causing low performance of the classical watershed. For this reason, seed-based methods were proposed [22]. The first step in seed-based techniques is to identify seeds close to the object center and use them as starting points for watershed transform. Because of the advantages of simplicity and speed, several marker-based watershed methods have been developed for image segmentation [22], [23], [24].

A marked-based watershed [25] is used to separate “fused” cardiospheres. Firstly, the distance transform (DT) of the membrane binary mask is calculated. To identify marker positions, local maxima of the DT are identified using the extended-maxima transform [24]. Technically, the extended-maxima transform estimates regional maxima by searching in N-connected neighborhoods. For this application, a neighborhood size of $N=20$ pixels (equal to $6.91 \mu\text{m}$) is empirically set, based on the observation that it guarantees an effective and affordable output in terms of cardiospheres separation. The separation process of “fused” cardiospheres is illustrated in Figure 3.3.

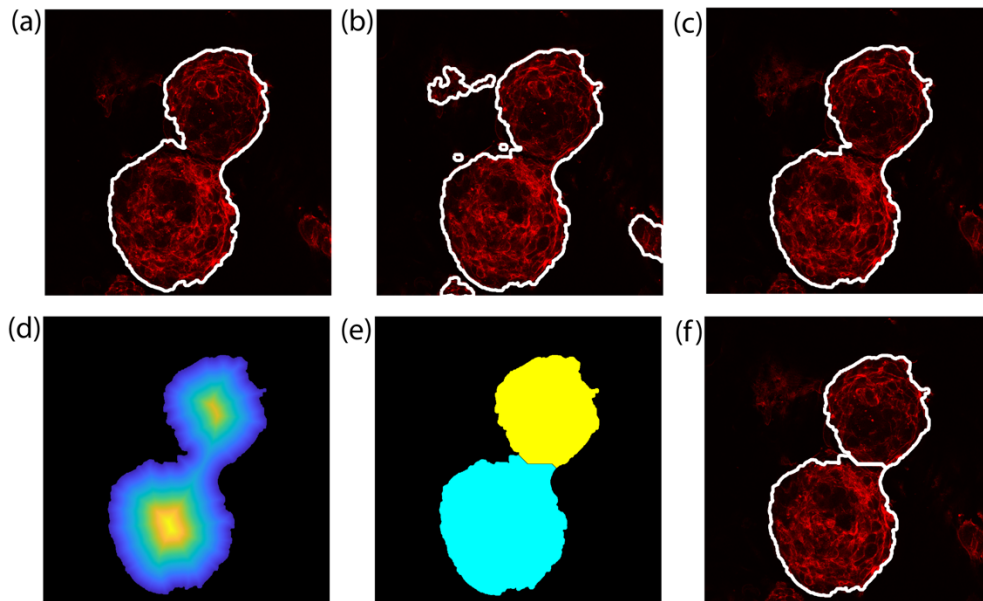


Fig. 3.3 Processing of cardiospheres membranes. First row shows the refining operation of cardiospheres external edges while second row illustrates the separation process of touching cardiospheres. (a) Current frame (reference frame). (b) Realign frame (next frame). (c) Realign frame after refining. (d) Distance transform of the membrane mask. (e) Application of the marker-based watershed for cardiosphere separation. (f) Final membrane mask on RGB image.

DAPI processing

After cardiospheres border segmentation, the proposed method analyzes the 3D stack of the DAPI layer. Starting from the original RGB image, the membranes segmentation mask is applied to each frame (Figure 3.4a). All objects outside the mask are excluded from the analysis, as they do not belong to the cardiosphere. The same object-based detection used for the PHAL processing is applied for cell nuclei segmentation to obtain a raw mask of cells inside each cardiosphere (Figure 3.4b). The input parameters of the ARCO algorithm for cell detection in fluorescent images are listed in Table 3.2.

Table 3.2 Input parameters of ARCO for cell detection in the CARE algorithm.

Parameters	Value
<i>image LAYER</i>	Blue layer of the DAPI image
<i>object TYPE</i>	Object of interest is bright on a dark background
<i>polynomial ORDER</i>	PWM _{CURVE} is fitted with a 7 th order polynomial function
<i>object PROPERTY</i>	Maximum median area of the segmented objects is imposed as condition to find the optimal threshold among candidate ones.

In the acquired images, cell nuclei are very often close to each other and the algorithm connects them as a single structure. For this reason, also at this stage of the investigation a marker-based watershed is applied in order to separate fused nuclei (Figure 3.4c).

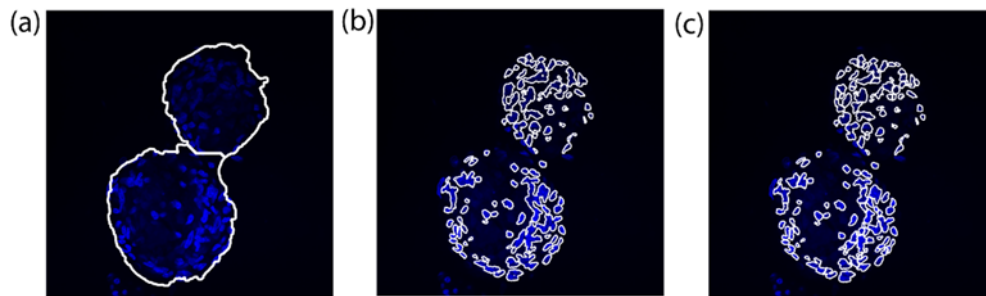


Fig. 3.4 Processing of DAPI layer. (a) Membrane mask applied to the original image. (b) Raw nuclei detection. (c) Cell nuclei separation using a marker-based watershed.

3D rendering

The 3D rendering of cardiospheres is obtained combining the segmentation masks obtained as mentioned above with the corresponding RGB image (Figure 3.5). Unfortunately, this operation is not sufficient to ensure a proper 3D reconstruction of the volume of cardiospheres, because it is affected by the border effect in the region where the cardiosphere is in contact with the surface on which it grows. To overcome this limitation, the method proposed here identifies the frame where the first contact between the cardiosphere and the support surface occurs (*cut-off frame*). To do that, starting from the first image of the stack, three conditions are checked on the border segmentation mask. If at least one of these conditions is satisfied, the slide is labeled as cut-off frame and the remaining images are not used for 3D rendering:

1. grayscale intensity - if in the image i^{th} of the stack the grayscale average intensity inside the segmented border mask is lower than 0.20, then the image is too dark to be considered for 3D rendering;
2. shape difference - if the area difference between the segmented border in frame $i-1$ and frame i is greater than 30%, then the cardiosphere is starting to spread on the surface;
3. shape solidity - if the segmented border mask solidity is less than 0.60, then the shape is so irregular that it cannot belong to a single cardiosphere. Solidity of a structure is defined as the ratio between its actual area and its convex area. Since it is expected that cardiospheres are convex objects, the solidity is used as feature for the identification of the *cut-off frame*.

Figure 3.5 shows a 3D rendering before and after the estimation of the cut-off frame. Through the process described above, the CARE algorithm produces two renderings: (i) the 3D volume of the external border of the cardiosphere and (ii) the 3D volume of all the cell nuclei inside the cardiosphere.

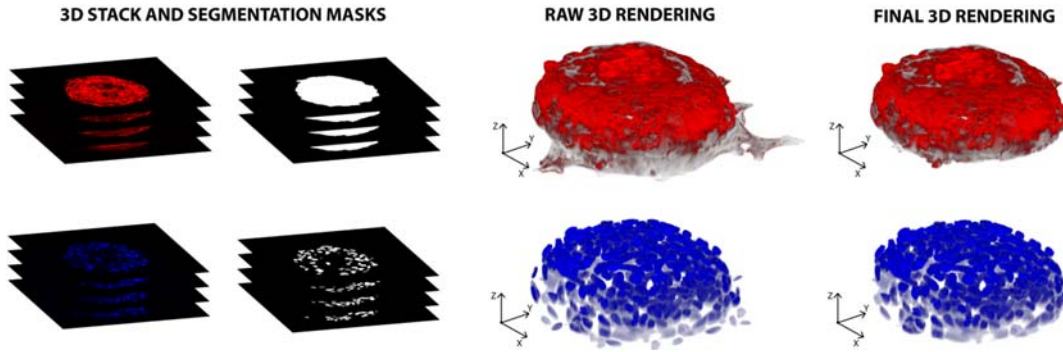


Fig. 3.5 Cardiosphere 3D rendering before and after the cut-off frame estimation. PHAL and DAPI mask are combined to obtain the raw volumes. Then, the proposed algorithm identifies the frame in which there is the first contact between the cardiosphere and the surface. The final 3D rendering is achieved by excluding all the slides after the cut-off frame.

Performance indicators

The parameters tuning of the CARE algorithm is reported in Appendix B. All the 1160 images of the dataset are used to validate the performance of CARE in segmenting cardiospheres borders respect to two manual operators (OP1, OP2). Given the presence of a high number of nuclei in each image, only part of them are used to validate the DAPI layer. In particular, five random images are extracted from each stack, for a total of 100 images. The same two operators manually draw each cell in order to assess inter-operator variability in the cell nuclei detection.

A comparison between masks drawn by a manual operator ($MASK_{MANUAL}$) and those provided by CARE ($MASK_{AUTOMATIC}$) is also carried out to assess the algorithm performance in the segmentation of cardiosphere borders and cell nuclei. True positive (TP) denotes the number of pixels in common between manual and automatic masks, false negative (FN) represents all pixels not identified by CARE and false positive (FP) are all the pixels identified by CARE but not by the manual operator. The segmentation performance is assessed by calculating the *recall*, *precision*, *F1SCORE* and *jaccardINDEX*, defined as follows:

$$recall = \frac{TP}{TP + FN} \quad (\text{Eq. 3.1})$$

$$precision = \frac{TP}{TP + FP} \quad (\text{Eq. 3.2})$$

$$F1_{SCORE} = \frac{2 \times precision \times recall}{precision + recall} \quad (\text{Eq. 3.3})$$

$$jaccard_{INDEX} = \frac{|MASK_{MANUAL} \cap MASK_{AUTOMATIC}|}{|MASK_{MANUAL} \cup MASK_{AUTOMATIC}|} \quad (\text{Eq. 3.4})$$

In detail, *recall* measures the missed detection of ground truth shapes, *precision* indicates the false detection of ghost objects and *F1SCORE* is defined as the harmonic mean between precision and recall [26]. The *jaccardINDEX* measures the similarity

between two different shapes and it is defined as the ratio between the size of the intersection and the size of the union of the segmented objects [27].

3.2.2 Quantitative Spatial Assessment of Markers in 3D Cardiospheres

Image database

Fifty-four 3D-stacks of human-derived cardiospheres obtained from different patients are analyzed to quantitatively characterize the fluorescence distribution within the cardiosphere of three markers involved in the evolution of cardiac stromal cell: YAP, GATA4 and Ki67. YAP is a transcription factor whose nuclear translocation dynamics is regulated by the generation of cellular cytoskeletal tension [28]. GATA4 is a transcription factor involved in cardiac differentiation while Ki67 is use to marker the proliferative cell activity. Firstly, each 3D-stack is acquired using two different lasers to highlight cell membranes (PHAL) and nuclei (DAPI); then, immunofluorescence is performed for YAP, GATA4 and Ki67, followed by high-resolution confocal imaging. All the cardiospheres completely within the field of view (FOV) of the image are analyzed for the spatial evaluation of the markers. The overall dataset composition is reported in Table 3.3.

Table 3.3 Database composition for quantitative assessment of markers inside cardiospheres.

Marker	Fluorescent color	# 3D-stacks	# Cardiospheres
<i>YAP</i>	Green	24	43
<i>GATA4</i>	Green	6	18
<i>Ki67</i>	White	24	50

Markers evaluation

First of all, the CARE algorithm is applied to all the 3D-stacks to identify membranes and cell nuclei within the cardiosphere. Then, the binary masks obtained with CARE are used to evaluate the concentration of the three markers (YAP, GATA4 and Ki67) inside and outside the cell nuclei. Figure 3.6 shows an example of the starting images and the respective masks used for marker evaluation.

In order to assess the spatial distribution of markers, the volume of the cardiosphere is divided into two sub-volumes: *internal* and *external* volumes. Firstly, each cardiosphere is automatically cut into two non-overlapped areas: ‘cap’ (spherical cap of the sample) and ‘center’ (central body of the sample). Then, the ‘center’ body is further divided into two sub-volumes: ‘center external’ and ‘center internal’.

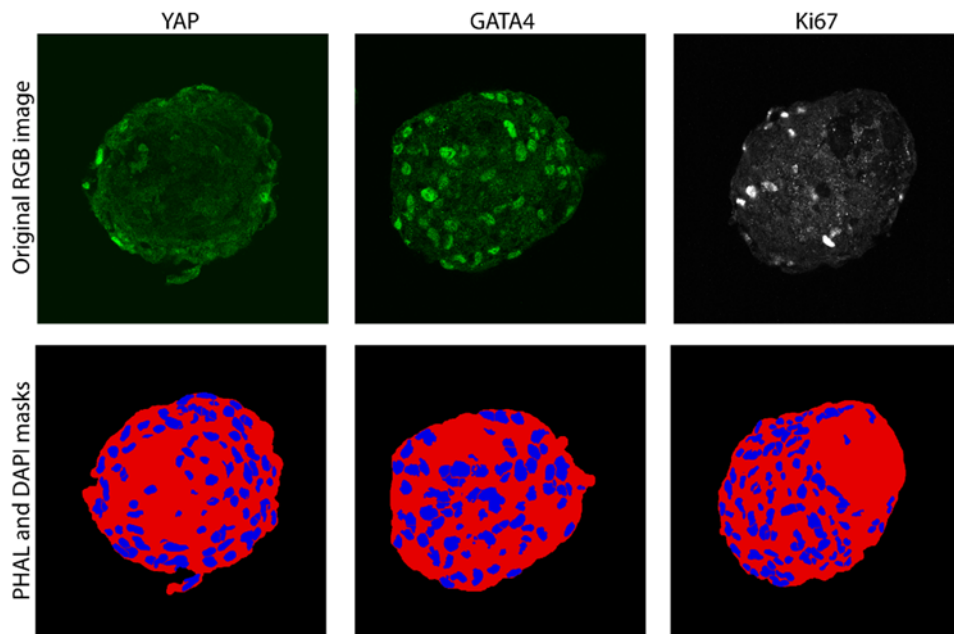


Fig. 3.6 Starting images and masks for protein markers evaluation. The first column shows the original YAP image and the corresponding automatic PHAL (red) and DAPI (blue) masks. The second and columns illustrate the GATA4 and the Ki67 images respectively.

The *external* volume represents the actual external volume of the cardiosphere and it is given by the sum of the 'cap' and 'center external' volumes. The *internal* volume represents the internal core of the cardiosphere and it is composed of the 'center internal' volume. Figure 3.7 shows the processing for obtaining the *internal* and *external* volumes starting from the 3D reconstruction of the PHAL layer provided by CARE.

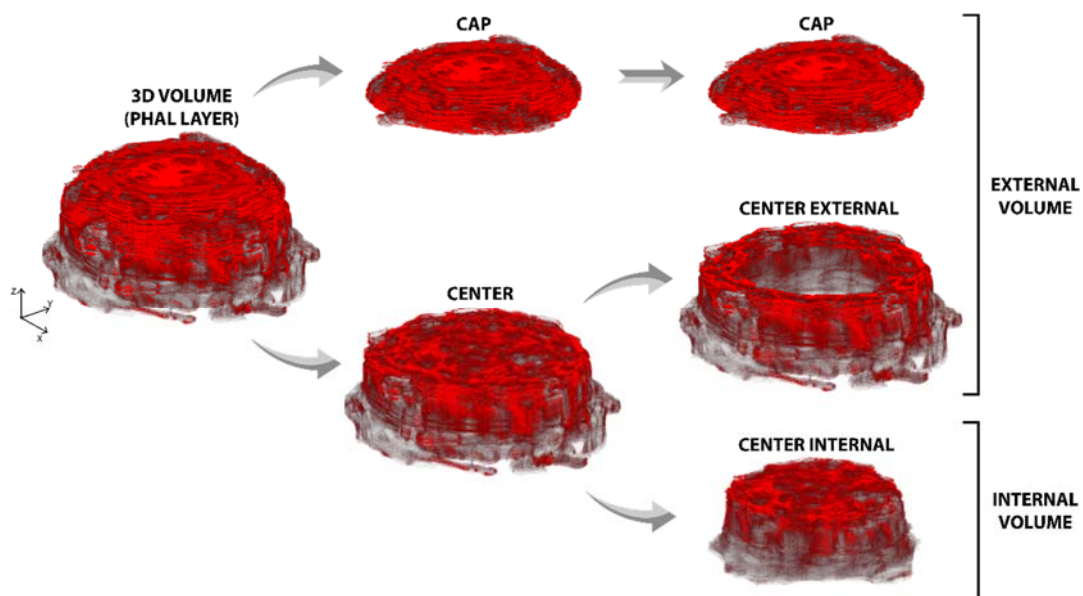


Fig. 3.7 Steps for obtaining *external* and *internal* volumes of the cardiosphere. The *external* volume consists of the spherical 'cap' and the external border of the central body ('center external'). The *internal* volume represents the core of the cardiosphere ('center internal').

For both *internal* and *external* volumes, the ratio between the fluorescence concentration expressed inside the cell nuclei and the overall fluorescence concentration is evaluated. This ratio indicates how much fluorescence has been internalized by the cell nuclei of the cardiosphere respect to the entire volume considered. For each of the two volumes (*internal* and *external*), the following parameters have been evaluated:

$$YAP_{INT/TOT} = \frac{\sum_{volume} YAP_{INT}}{\sum_{volume} YAP_{TOT}} \quad (\text{Eq. 3.5})$$

$$GATA4_{INT/TOT} = \frac{\sum_{volume} GATA4_{INT}}{\sum_{volume} GATA4_{TOT}} \quad (\text{Eq. 3.6})$$

where $(\cdot)_{INT}$ indicates the amount of fluorescence expressed within the nuclei and $(\cdot)_{TOT}$ represents the total fluorescence expressed in the volume considered. Since Ki67 is a nuclear marker, $Ki67_{INT/TOT}$ has been computed as the ratio between the fluorescence within nuclei and the total fluorescence of the entire cardiosphere:

$$Ki67_{INT/TOT} = \frac{\sum_{volume} Ki67_{INT}}{\sum_{cardiosphere} Ki67_{TOT}} \quad (\text{Eq. 3.7})$$

$YAP_{INT/TOT}$, $GATA4_{INT/TOT}$ and $Ki67_{INT/TOT}$ are compared in *internal* and *external* volumes using two-tailed paired sample. All statistical tests are carried out with a significance level (p) of 0.05.

3.3 Results

Cardiospheres are stained with DAPI, and TRITC-labelled phalloidin to highlight, respectively, cells and nuclei distribution and shapes [17]. Conventional confocal microscopy is employed to obtain images of the cardiospheres by 3D-stack acquisition with a relatively high definition. By visual inspection of Figure 3.8, cardiospheres exhibit a complex structure emerging above the culture plate as hemispheres, made of cells distributed with apparent multiple orientation and cells/nuclei shapes. The presence of internal cavities with a non-uniform dimension can be also appreciated.

3.3.1 Segmentation results of CARE Algorithm

CARE vs manual operator image segmentation

The results of the comparison between automatic and manual segmentation are summarized in Table 3.4. CARE demonstrates excellent performances in segmenting cardiospheres borders (PHAL), with very high average values of *precision*, *recall*, $F1_{SCORE}$ and $jaccard_{INDEX}$ respect to two expert operators (OP1, OP2) thus demonstrating the accuracy of the method (Table 3.4).

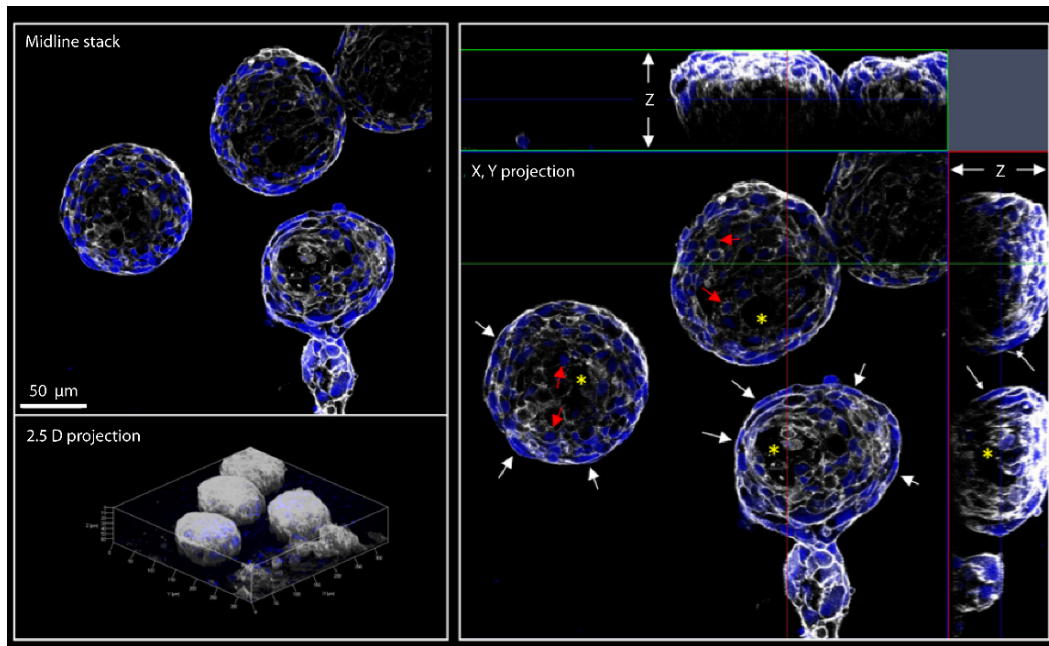


Fig. 3.8 Structure of the cardiospheres as observed by confocal microscopy. White color represents the cytoskeleton as evidenced by F-actin staining with Phalloidin. Blue color represents the nuclei as revealed by DAPI, an intercalant of the DNA. The three panels represent the midline stack (upper left) the 2.5 projection of the cardiospheres with the X and Y dimensions (lower left) and the projection of the whole stack along the indicated X and Y axes (right) respectively.

Table 3.4 Performance of the CARE algorithm in the cardiosphere border (PHAL) and nuclei segmentation (DAPI). Data are reported as mean \pm standard deviation.

Layer	#Images	Validation	Recall	Precision	F1 _{SCORE}	jaccard _{INDEX}
PHAL	1160	OP1 vs OP2	0.9410 \pm 0.0285	0.9789 \pm 0.343	0.9588 \pm 0.0158	0.9170 \pm 0.0281
		OP1 vs CARE	0.9339 \pm 0.0241	0.9728 \pm 0.0345	0.9497 \pm 0.0157	0.9079 \pm 0.0255
		OP2 vs CARE	0.9508 \pm 0.0264	0.9717 \pm 0.0351	0.9606 \pm 0.0201	0.9271 \pm 0.0255
DAPI	100	OP1 vs OP2	0.7602 \pm 0.0486	0.8215 \pm 0.0507	0.7872 \pm 0.0234	0.6497 \pm 0.0318
		OP1 vs CARE	0.9001 \pm 0.0329	0.6713 \pm 0.0506	0.7679 \pm 0.0370	0.6157 \pm 0.0431
		OP2 vs CARE	0.9210 \pm 0.0316	0.6517 \pm 0.0599	0.7615 \pm 0.0439	0.6174 \pm 0.0586

As for nuclei segmentation (DAPI), the average $F1_{SCORE}$ calculated between the two operators (0.7872) is comparable with the one obtained between CARE and each of them (0.7679 and 0.7615). The algorithm exhibits an excellent performance in the recognition of cell nuclei, compared to manual operators (OP1: 0.9001 and OP2: 0.9210). Being very sensitive, CARE tends to slightly overestimate the nuclei surface, and this leads to a lower *precision*, compared to manual operators (OP1: 0.6713 and OP2: 0.6517). Finally, the values of $jaccard_{INDEX}$, between OP1 and

CARE (0.6157), and OP2 and CARE (0.6174) are comparable to the value between OP1 and OP2 (0.6497).

An explanatory example comparing the output of the segmentation obtained by applying CARE and by manual operators is presented in Figure 3.9.

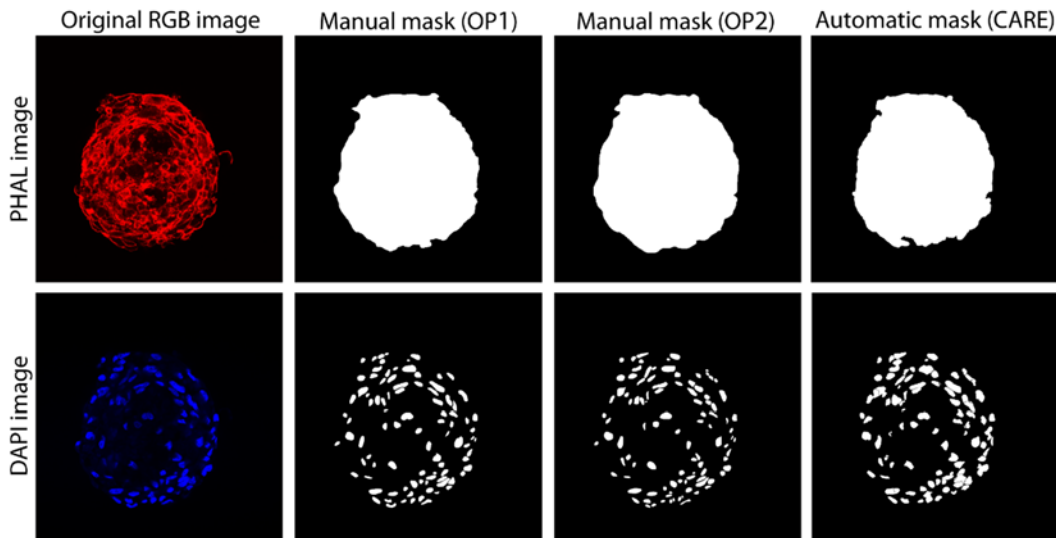


Fig. 3.9 Comparison between manual and automatic segmentation (rows). First column shows the original RGB image while manual masks performed by two expert operators (OP1, OP2) are reported in the second and third columns. The result provided by the proposed method is shown in the rightmost column.

Comparison with open-source software

Automatic results provided by CARE are also compared with two open-source software (CellProfiler and Fiji) widely applied to the analysis of fluorescence microscopy images. CellProfiler [29] is a software with a modular structure that integrates several image-processing techniques to perform automatic analyses on biological images. Fiji [30] is a Java-based software with several plugins which facilitate scientific image analysis based on a semi-automatic pipeline consisting of: (i) conversion of RGB image into grayscale, (ii) manual intensity thresholding, (iii) hole filling and (iv) small particles removal. For the nuclei segmentation, here an additional step is included in the analysis: (v) automatic cell separation. A visual inspection of Figure 3.10 allows to compare the performances of CellProfiler, Fiji and CARE in cardiospheres segmentation.

A quantitative comparison of the performances offered by the two open-source software with CARE is reported in Table 3.5 and Table 3.6. As can be seen from Table 3.5 and Table 3.6, the Cell Profiler segmentation is characterized by a low recall (PHAL: 0.7997, DAPI: 0.8053) and this lead to a lowering of the average $F1_{SCORE}$ (PHAL: 0.7997, DAPI: 0.8053). Moreover, the mean $jaccard_{INDEX}$ (PHAL: 0.7939, DAPI: 0.5031) is lower than the proposed one for more than 10%.

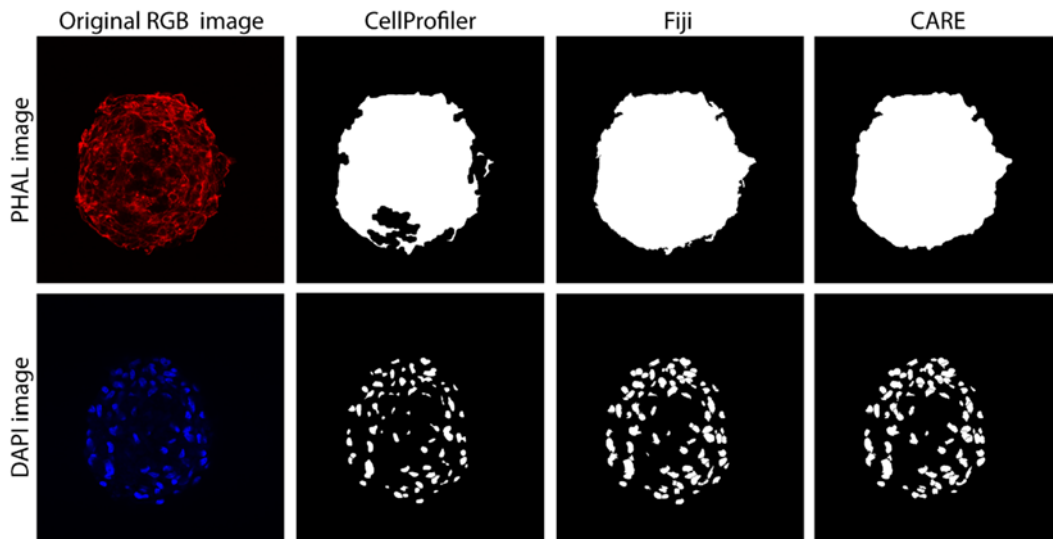


Fig. 3.10 Comparison between CellProfiler, Fiji and CARE in the cardiosphere segmentation (rows). First column shows the original RGB image while CellProfiler and Fiji result are provided in the second and third columns respectively. The automatic mask obtained with CARE is shown in the last column.

Fiji segmentation performance is quite similar to CARE results. The average $F1_{SCORE}$ achieved with Fiji is slightly lower than those obtained with CARE (PHAL: 0.9249, DAPI: 0.7504). This software is semi-automatic and requires user intervention to function properly. For this reason, the average computational time is about 10 times higher than CARE algorithm.

Table 3.5 Performance of CellProfiler and Fiji software in the segmentation of the external cardiosphere membrane (PHAL layer). Data are reported as mean \pm standard deviation.

Method	Computational Time (sec)	PHAL layer			
		Recall	Precision	$F1_{SCORE}$	$jaccard_{INDEX}$
CellProfiler (automatic)	13.18 \pm 4.51	0.7997 \pm 0.0655	0.9912 \pm 0.0028	0.8724 \pm 0.0482	0.7939 \pm 0.0671
Fiji (semi-automatic)	117.21 \pm 13.91	0.8865 \pm 0.0559	0.9902 \pm 0.0005	0.9249 \pm 0.0387	0.8775 \pm 0.0548
CARE (proposed)	8.27 \pm 1.31	0.9339 \pm 0.0241	0.9728 \pm 0.0345	0.9497 \pm 0.0157	0.9079 \pm 0.0255

Table 3.6 Performance of CellProfiler and Fiji software in the segmentation of the cardiosphere cell nuclei (DAPI layer). Data are reported as mean \pm standard deviation.

Method	Computational Time (sec)	DAPI layer			
		Recall	Precision	$F1_{SCORE}$	$jaccard_{INDEX}$
CellProfiler (automatic)	14.26 \pm 3.02	0.8053 \pm 0.1198	0.5818 \pm 0.1704	0.6536 \pm 0.1341	0.5031 \pm 0.1365
Fiji (semi-automatic)	123.81 \pm 19.38	0.8457 \pm 0.0944	0.6845 \pm 0.0453	0.7504 \pm 0.0522	0.6045 \pm 0.0633
CARE (proposed)	12.42 \pm 2.36	0.9001 \pm 0.0329	0.6713 \pm 0.0506	0.7679 \pm 0.0370	0.6157 \pm 0.0431

3.3.2 Functional Characterization of 3D Cardiospheres

For each 3D-stack, the three parameters $YAP_{INT/TOT}$, $GATA4_{INT/TOT}$ and $Ki67_{INT/TOT}$ are calculated. Then, a comparison of those parameters in *external* and *internal* volumes is carried out. Figure 3.11 shows the variation of $YAP_{INT/TOT}$, $GATA4_{INT/TOT}$ and $Ki67_{INT/TOT}$ in *internal* and *external* volumes.

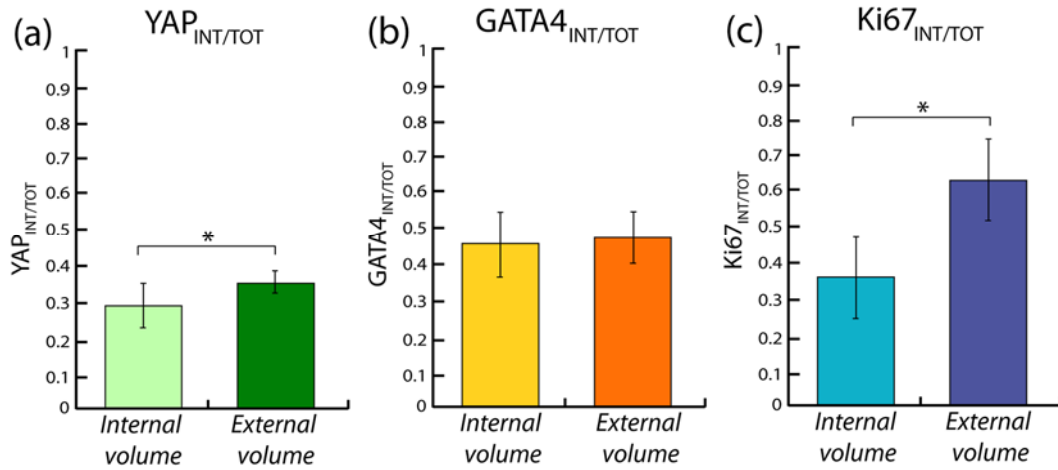


Fig. 3.11 Comparison of average $YAP_{INT/TOT}$, $GATA4_{INT/TOT}$ and $Ki67_{INT/TOT}$ in internal vs external volume. Asterisk denotes statistically significant difference ($p < 0.05$).

As shown in Figure 3.11, the values of $YAP_{INT/TOT}$ and $Ki67_{INT/TOT}$ increase in the external volume. On the contrary, $GATA4_{INT/TOT}$ shows nearly the same values. The paired t-test reveals a significant difference in $YAP_{INT/TOT}$ ($p = 1.74 \cdot 10^{-6}$) and $Ki67_{INT/TOT}$ ($p = 1.05 \cdot 10^{-10}$) between *internal* and *external* volumes, whereas no significant difference is found for $GATA4_{INT/TOT}$ ($p = 0.55$). On average, $YAP_{INT/TOT}$ increases from 0.2899 to 0.3604, whereas $Ki67_{INT/TOT}$ elevates from 0.3706 to 0.6294. In contrast, the values of $GATA4_{INT/TOT}$ on the *internal* volume remain almost equal to the ones on the *external* volume (0.4687 vs 0.4873). The results of the statistical analysis are shown in Table 3.7.

Table 3.7 Paired t-test of $YAP_{INT/TOT}$, $GATA4_{INT/TOT}$ and $Ki67_{INT/TOT}$ for *external* and *internal* volume. Data are reported as mean \pm standard deviation.

Parameter	Internal volume	External volume	p value
$YAP_{INT/TOT}$	0.2899 ± 0.0704	0.3604 ± 0.033	$1.74 \cdot 10^{-6} *$
$GATA4_{INT/TOT}$	0.4687 ± 0.098	0.4873 ± 0.074	0.5583
$Ki67_{INT/TOT}$	0.3706 ± 0.1109	0.6294 ± 0.1109	$1.05 \cdot 10^{-10} *$

* denotes a statistically significant difference compared with border volume

3.4 Discussion

In the present study, a fully automatic algorithm is proposed for human-derived cardiosphere segmentation and characterization in fluorescence microscopy images. A cardiosphere is a 3D cluster of cardiac stromal cells that show a promising phenotype for regeneration of the failing human myocardium [13], [15], [3] but also as a model of cardiac pathologies such as heart failure [16].

The proposed technique is able to detect cardiosphere membrane and cells inside fluorescence images. The CARE algorithm does not require any user interaction and it is capable of automatically detecting membranes and cell nuclei in a 3D context. The algorithm is tested on twenty 3D-stacks of human-derived cardiospheres, for a total number of 27 cardiospheres and 1160 slides. Two expert biologists manually annotate cardiosphere membranes for all the images of our dataset. To assess the inter-operator variability in nuclei segmentation, the same manual operators also draw each cell boundary on 100 random images.

The comparison between automatic results and manual annotations show high performance of the proposed method. The CARE algorithm showed excellent performance in the membranes segmentation, with an average $F1_{SCORE}$ of 0.9497 ± 0.0157 and $jaccard_{INDEX}$ of 0.9079 ± 0.0255 . In the cell segmentation, the proposed algorithm obtains a mean $F1_{SCORE}$ and $jaccard_{INDEX}$ comparable with respect to two expert operators (Table 3.4). In addition, the CARE algorithm obtains the highest $F1_{SCORE}$ compared to other software (Fiji and CellProfiler) designed for cell detection in fluorescence microscopy. Respect to other automatic and semi-automatic methods, the proposed algorithm has also the lowest computational time and the best $jaccard_{INDEX}$.

When CARE is employed to assess markers distribution in the cardiosphere, a prevalent distribution of cells with high YAP nuclear concentration is observed in the external shell. This suggests dependence of YAP-nuclear translocation process on cytoskeleton tensioning as already shown in literature for 2D systems [31]. The results obtained with CARE demonstrate a tight correlation between cell mechanical tension and YAP nuclear translocation in the 3D cardiosphere. Localization of the Ki67 proliferation marker exhibits a similar trend, confirming the correlation between control of cell proliferation and YAP nuclear signaling in the 3D model. These results suggest that cells that are located in the external border of the cardiosphere are exposed to a surface tension and are more actively proliferating [3].

This study demonstrated, for the first time, the existence of spatially regulated mechanosensing mechanism linking cytoskeleton tensioning to activity of cardiac stromal cells evolution in pro-fibrotic phenotype [18].

Thanks to the fast and robust cell detection provided by CARE, fully automatic systems for morphological/antigenic characterization of cells inside 3D aggregates can be easily developed. Our research group is currently working on a CARE-based algorithm for the automated quantitative analysis of other markers distribution within the cardiosphere (alphaSMA, Collagen I), with the aim to discover predictive associations between cell mechanical cues and dynamic phenotypic changes. The

assessment of cardiosphere proprieties could help to understand the mechanisms that regulate cardiac biology and disease [17], [32], and it could also open new technological and therapeutic perspectives [33], [34].

In the future, a 3D object separation will be integrated within the CARE algorithm in order to increase the cell nuclei segmentation performance. Future studies are also required to test the accuracy and reliability of the proposed technique in cell nuclei detection in other 3D aggregates grown in vitro.

3.5 Conclusion

In this Chapter, an adaptive algorithm for segmentation and characterization of 3D cardiosphere in fluorescence microscopy images is presented. To the best of our knowledge, CARE is the first fully automated solution for the detection of cells inside in vitro 3D aggregates.

The CARE algorithm is tested on twenty 3D-stacks of human-derived cardiospheres, in which nuclei have different morphologies and laser intensities. For each image of the dataset, high segmentation performances are achieved. The observed accuracy of CARE in cell segmentation is mainly due to the use of the adaptive thresholding provided by the ARCO algorithm. CARE takes around 25 seconds to perform membrane and cells detection in images with hundreds of nuclei, thus indicating the efficiency of the proposed technique.

The 3D-image processing provided by CARE enabled the automatic quantitative evaluation of the geometrically arranged markers distribution in the cardiospheres. The proposed algorithm allows to establish a criterion to dissect the effect of cell straining and the relative amount of YAP, GATA4 and Ki67 nuclear concentrations. After the 3D analysis, a recurrent positioning of cells with nuclear localized YAP and Ki67 in the external volume of the cardiosphere is observed. Cells with nuclear-localized YAP and exhibiting proliferation marker Ki67 are abundant at the cardiosphere periphery, but undetectable in their core [18]. On the other hand, this asymmetric distribution is not observed in GATA4, a transcription factor expressed in cardiac cells.

Being totally automated, CARE could be used in the future as starting point to realize rapid and accurate systems for functional characterization of cell aggregates in 3D contexts. In the future, the combination of innovative cell engineering, 3D biology tools, and 3D image processing [16], [3] will provide advanced methods to obtain an integrated representation of cell forces sensing in the context of the cardiac stiffing process and myocardial inflammation.

References

- [1] S. van Putten, Y. Shafieyan, and B. Hinz, “Mechanical control of cardiac myofibroblasts,” *J. Mol. Cell. Cardiol.*, vol. 93, pp. 133–142, 2016.
- [2] A. K. Schroer and W. D. Merryman, “Mechanobiology of myofibroblast adhesion in fibrotic cardiac disease,” *J Cell Sci*, p. jcs-162891, 2015.
- [3] I. Chimenti, D. Massai, U. Morbiducci, A. P. Beltrami, M. Pesce, and E. Messina, “Stem cell spheroids and ex vivo niche modeling: rationalization and scaling-up,” *J. Cardiovasc. Transl. Res.*, vol. 10, no. 2, pp. 150–166, 2017.
- [4] D. L. Jones and A. J. Wagers, “No place like home: anatomy and function of the stem cell niche,” *Nat. Rev. Mol. cell Biol.*, vol. 9, no. 1, p. 11, 2008.
- [5] I. Roeder, M. Loeffler, and I. Glauche, “Towards a quantitative understanding of stem cell–niche interaction: Experiments, models, and technologies,” *Blood Cells, Mol. Dis.*, vol. 46, no. 4, pp. 308–317, 2011.
- [6] I. Roeder and R. Lorenz, “Asymmetry of stem cell fate and the potential impact of the niche,” *Stem Cell Rev.*, vol. 2, no. 3, pp. 171–180, 2006.
- [7] L. M. Murrow, R. J. Weber, and Z. J. Gartner, “Dissecting the stem cell niche with organoid models: an engineering-based approach,” *Development*, vol. 144, no. 6, pp. 998–1007, 2017.
- [8] D. C. Kirouac and P. W. Zandstra, “The systematic production of cells for cell therapies,” *Cell Stem Cell*, vol. 3, no. 4, pp. 369–381, 2008.
- [9] N. Shakiba and P. W. Zandstra, “Engineering cell fitness: lessons for regenerative medicine,” *Curr. Opin. Biotechnol.*, vol. 47, pp. 7–15, 2017.
- [10] K. B. Blagoev, “Organ aging and susceptibility to cancer may be related to the geometry of the stem cell niche,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 48, pp. 19216–19221, 2011.
- [11] A. Yachie-Kinoshita *et al.*, “Modeling signaling-dependent pluripotency with Boolean logic to predict cell fate transitions,” *Mol. Syst. Biol.*, vol. 14, no. 1, p. e7952, 2018.
- [12] M. Tewary *et al.*, “A stepwise model of Reaction-Diffusion and Positional-Information governs self-organized human peri-gastrulation-like patterning,” *Development*, p. dev-149658, 2017.
- [13] R. Gaetani *et al.*, “Different types of cultured human adult cardiac progenitor cells have a high degree of transcriptome similarity,” *J. Cell. Mol. Med.*, vol. 18, no. 11, pp. 2147–2151, 2014.
- [14] R. Gaetani *et al.*, “Cardiac tissue engineering using tissue printing technology and human cardiac progenitor cells,” *Biomaterials*, vol. 33, no. 6, pp. 1782–1790, 2012.
- [15] L. Barile *et al.*, “Cardiac stem cells: isolation, expansion and experimental use for myocardial regeneration,” *Nat. Rev. Cardiol.*, vol. 4, no. S1, p. S9, 2007.
- [16] M. Pesce, E. Messina, I. Chimenti, and A. P. Beltrami, “Cardiac mechanoperception: a life-long story from early beats to aging and failure,” *Stem Cells Dev.*, vol. 26, no. 2, pp. 77–90, 2017.
- [17] E. Messina *et al.*, “Isolation and expansion of adult cardiac stem cells from human and murine heart,” *Circ. Res.*, vol. 95, no. 9, pp. 911–921, 2004.
- [18] M. Pesce *et al.*, “P4225 YAP-based position sensing in myofibroblast evolution of cardiac stromal cells in human cardiospheres,” *Eur. Heart J.*, vol. 39, no. suppl_1, pp. ehy563-P4225, 2018.
- [19] L. P. Coelho, A. Shariff, and R. F. Murphy, “Nuclear segmentation in

- microscope cell images: a hand-segmented dataset and comparison of algorithms,” in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, 2009, pp. 518–521.
- [20] G. Li *et al.*, “Segmentation of touching cell nuclei using gradient flow tracking,” *J. Microsc.*, vol. 231, no. 1, pp. 47–58, 2008.
- [21] S. Beucher and F. Meyer, “The morphological approach to segmentation: the watershed transformation,” *Opt. Eng. York-Marcel Dekker Inc.*, vol. 34, p. 433, 1992.
- [22] Y. Sun and G. He, “Segmentation of high-resolution remote sensing image based on marker-based watershed algorithm,” in *Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on*, 2008, vol. 4, pp. 271–276.
- [23] H. Gao, P. Xue, and W. Lin, “A new marker-based watershed algorithm,” in *Circuits and Systems, 2004. ISCAS'04. Proceedings of the 2004 International Symposium on*, 2004, vol. 2, p. II-81.
- [24] H. Xu, C. Lu, and M. Mandal, “An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm,” *IEEE J. Biomed. Heal. informatics*, vol. 18, no. 5, pp. 1729–1741, 2014.
- [25] M. Salvi and F. Molinari, “Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images,” *Biomed. Eng. Online*, vol. 17, no. 1, p. 89, 2018.
- [26] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.
- [27] R. Real and J. M. Vargas, “The probabilistic basis of Jaccard’s index of similarity,” *Syst. Biol.*, vol. 45, no. 3, pp. 380–385, 1996.
- [28] S. Dupont *et al.*, “Role of YAP/TAZ in mechanotransduction,” *Nature*, vol. 474, no. 7350, p. 179, 2011.
- [29] A. E. Carpenter *et al.*, “CellProfiler: image analysis software for identifying and quantifying cell phenotypes,” *Genome Biol.*, vol. 7, no. 10, p. R100, 2006.
- [30] J. Schindelin *et al.*, “Fiji: an open-source platform for biological-image analysis,” *Nat. Methods*, vol. 9, no. 7, p. 676, 2012.
- [31] A. Elosegui-Artola *et al.*, “Force triggers YAP nuclear entry by regulating transport across nuclear pores,” *Cell*, vol. 171, no. 6, pp. 1397–1410, 2017.
- [32] I. Chimenti *et al.*, “Serum and supplement optimization for EU GMP-compliance in cardiospheres cell culture,” *J. Cell. Mol. Med.*, vol. 18, no. 4, pp. 624–634, 2014.
- [33] M. H. Yacoub and J. Terrovitis, “CADUCEUS, SCIPIO, ALCADIA: cell therapy trials using cardiac-derived cells for patients with post myocardial infarction LV dysfunction, still evolving,” *Glob. Cardiol. Sci. Pract.*, vol. 2013, no. 1, p. 3, 2013.
- [34] M. Peruzzi *et al.*, “State of the art on the evidence base in cardiac regenerative therapy: overview of 41 systematic reviews,” *Biomed Res. Int.*, vol. 2015, 2015.

Chapter 4

Digital Histopathology Image Analysis

Part of this chapter has been published as:

M. Salvi and F. Molinari, **Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images**, *Biomedical Engineering Online*, 2018, 17: 89-101.

and

M. Salvi, K. M. Meiburger, B. De Santi, N. Michielli, L. Molinaro, and F. Molinari, **Automated gleason grading in prostate cancer histopathology images**, *Sixth National Congress of Bioengineering*, 2018.

4.1 Introduction

Digital pathology is the process where the histological specimen is digitized in order to generate a high-resolution image. Nowadays, the digital pathology is spreading due to the growing availability of whole-slide digital scanners [1]. These digitized slides afford the possibility of applying image analysis techniques for applications in detection, segmentation, and classification of cellular structures. In the last years, automated algorithms have shown to be beneficial in many contexts as they have the capacity to not only significantly reduce the laborious and tedious nature of providing accurate quantifications, but to act as a second reader helping to reduce inter-reader variability among pathologists [2]. The variety of image analysis tasks in the context of digital pathology includes segmentation (e.g., nuclei) and tissue classification (e.g., cancerous vs. non-cancerous).

The evaluation of cell nuclei plays a fundamental role in the analysis of histological images. In particular, parameters like nuclei spatial distribution, cells shape and size are generally used by pathologists for tumor detection and characterization [3]. In routine histology, the most widely used staining method is the hematoxylin and eosin (H&E). This coloration allows the specialist to discriminate between cell nuclei (hematoxylin: bluish color) from cytoplasm (eosin: pinkish color) [4]. Cell nuclei counting is a time-consuming operation and it is prone to intra- and inter-observer variability, which results in a limited reliability.

Manual annotation of cell nuclei boundaries is a cumbersome operation, which is never performed in routine, but which would be required to precisely assess nuclei size and morphology. The spatial and architectural dispersion of nuclear structures in histopathological images is highly relevant in the context of tumor grading [5]. Cancer grade is a key-feature used to predict the patient prognosis and to prescribe a therapy [3].

Recently, many efforts have been devoted to developing fully automated strategies to perform cell nuclei segmentation with the aim to improve the accuracy and efficiency of histological analysis. Most of the current nuclei detection approaches on H&E stained images are based on color information [6], [7]. Using these techniques, an segmentation accuracy of 80% can be achieved [8]. Since these approaches are dependent on either color and intensity-related attributes, none of these works have been tested on multi-tissue data or in pathological conditions, where nuclei may exhibit irregular shapes and different intensities.

In the last decade, several methods have been developed to perform nuclei detection using image gradients [9] or morphological operators [10]. Nevertheless, automated methods using of a prior knowledge of nuclei shape are prone to fail due to the high variation of tissue preparation procedures (staining and sectioning). Moreover, the existence of touching cells makes the nuclei separation quite challenging for an automatic algorithm [1].

In recent times, deep neural networks (DNNs) drove advances in computer vision tasks and achieved the same performance of many segmentation algorithms in medical imaging [11], [12]. Above all, convolutional neural networks (CNNs) have shown a great potential in the detection of cell nuclei for different tissues [13]. These networks estimate a probability map of the nuclear regions based on the learned cellular appearances. In this way, convolutional networks can generalize across various nuclear color variations [13]. However, CNNs need a large annotated training set to obtain adequate performance and the network architecture must be changed in case of variation in the magnification. This is because CNNs fail to generalize if the nuclei, in addition to changing color, also change size. For this reason, deep neural networks are not suitable for multiscale approaches.

In recent times, several automated strategies have shown promising results in tumor detection and classification within histopathological images [14], [15]. Consequently, researchers have begun to develop computer-aided diagnosis strategies using computer vision and image processing methods in order to identify the location and spatial extent of diseases such as prostate [16], [17] and breast cancer [18], [19].

Prostate cancer is among the most diagnosed forms of tumor in men worldwide [20]. Recently, a new, clinically validated, more accurate Gleason scoring system was proposed [21] in order to reduce overtreatments of low-grade cases. Briefly, the Gleason scoring method classifies the architecture of prostate glands into 5 patterns, ranging from 1 to 5. Patterns 1+2 are considered obsolete and no more assigned in prostate biopsies, pattern 3 is low-grade, and patterns 4 and 5 are high-grade cancers. The final Gleason Score (GS) is obtained by the most predominant pattern plus the secondary pattern (if only two patterns present) or the most aggressive pattern (if more than two patterns are present). So, five Gleason Grade Groups (GG) are defined: GG 1 (GS ≤ 6), GG 2 (GS 3+4=7), GG 3 (GS 4+3=7), GG 4 (GS 4+4=8, 3+5=8, 5+3=8), and GG 5 (GS = 9-10).

The manual visual assessment of prostate histopathological images is time-consuming and prone to errors due to intra- and inter-operator variability [22]. An increasing interest in developing robust algorithms for the rapid and reproducible Gleason scoring of digital images has therefore developed. Several methods are reported in literature, many that focus on training a classifier and then classifying a gland image into the correct pattern [23], [24]. The downfall to these techniques is that they are based on analyzing images containing only one gland and therefore only one pattern, making it necessary to first isolate the glands within an entire biopsy image. Other techniques are nuclei distribution-based [5] or lumen-based [25], and then employ a subsequent classification, requiring manual identification of the pixels of interest to train a classifier. However, most of the available techniques either distinguish between cancerous and non-cancerous glands or give a primary score of a specific pattern, and rarely focus on the final Gleason GG and the discrimination between Gleason GG 2 and 3, which is crucial for the prognosis of the patient [21].

Besides prostate cancer, one of the common cancer types is breast cancer [20]. A generally underrated problem related to image analysis of breast cancer on histopathological slides is the correct identification of neoplastic epithelium and its distinction from stromal reaction. In most cases, the focus is on the neoplastic epithelium where morphological measures (nuclear pleomorphism, glands formation) and immuno-phenotypical quantifications (proteins expression) are performed. In other cases, the stromal component is analyzed (quantification of tumor infiltrating lymphocytes, characterization of stromal reaction type). In both cases, morphological and immuno-phenotypic characterization of the tumor can predict the breast carcinoma behavior and its prognosis [26], [27], [28] and, of course, should only be evaluated on the correct tumor component (epithelial vs stromal). In addition, the direct estimation of tumor/stroma ratio (the ratio of tumor volume occupied by neoplastic epithelium and stroma) has relevance in the determination of the pathological response to neoadjuvant therapy [29] and represents a novel prognostic parameter [30].

Since most of the current diagnosis processes are based on the subjective opinion of pathologists, automated and quantitative solutions for the assessment of histopathological images would have scope of application. In particular, an automatic system for the analysis of histological tissue could assist pathologists by providing objective results, quantitative measurements or even a second opinion. In the last decade, several algorithms have been proposed for the automatic analysis of breast carcinoma [14]. These methods are based on multi-resolution algorithms [31], nuclear features [32], [33], and deep neural networks [34], [35] to perform cancer detection and grading. Given a digital histological image, these techniques are trained to recognize the presence or the absence of tumor tissue within the image, but none of them is capable of discriminating between neoplastic epithelium and stromal response within the same image. These methods do not perform a segmentation of the tumor areas; therefore, it is not possible to obtain accurate contours of the cancer tissue. In fact, the extreme inter-tumoral and intra-tumoral variability of neoplastic epithelium morphology is the main challenge for an automatic algorithm aiming at the segmentation of the tumoral area.

In this Chapter, three robust and novel algorithms are proposed for the quantitative analysis of histopathological images. The first method, named MANA (Multiscale Adaptive Nuclei Analysis), is a multi-tissue and multiscale method for cell detection and segmentation in histological images. The second method is able to perform an automated gleason grading in prostate cancer histopathological images. Finally, the third method consists of a full-automated solution for the segmentation of neoplastic epithelium in H&E stained images of breast tissue.

4.2 Materials and Methods

4.2.1 Multi-tissue and Multiscale Nuclei Segmentation: MANA Algorithm

Image database

The histological dataset consists of 30 H&E stained images taken from six different organ tissues (bone, liver, thyroid, colon, adrenal gland and prostate). Moreover, images were digitalized using three magnifications (10x, 20x, 40x) in order to test the multiscale approach of MANA algorithm. In each image, one expert pathologist (with 15 years of experience in histopathology) manually marked the nuclei centers, for a total of more than 59000 cells. The images were acquired at “Città della Salute” Hospital – Molinette (Torino, Italy). The overall dataset composition is shown in Table 4.1.

Table 4.1 Dataset composition of the MANA algorithm.

Tissue	Magnification	# Images	# Nuclei
Bone	40x	5	6889
Liver	10x, 20x	5	5051
Thyroid	10x, 20x	5	19050
Colon	20x	5	9166
Adrenal gland	10x, 20x	5	12972
Prostate	20x	5	5995
<i>Total</i>	<i>10x, 20x, 40x</i>	<i>30</i>	<i>59123</i>

MANA algorithm architecture

The proposed algorithm is designed to automatically segment nuclei in H&E staining images. The MANA algorithm is developed using MATLAB environment and the image processing is carried out on a workstation with a 2.8 GHz exa-core CPU and 64 GB of RAM. The processing is composed by three main steps: raw nuclei detection, area-based correction and nuclei separation. In the next paragraphs, an exhaustive description of the proposed method is provided.

Raw nuclei detection

The first step of the MANA algorithm consists of a preliminary cell segmentation by analyzing the grayscale image of the histological sample. Then, the identification of nuclei boundaries is performed by applying the object-based detection provided by the ARCO algorithm (Chapter 2). The input parameters of the ARCO algorithm are listed in Table 4.2.

Table 4.2 Input parameters of the ARCO algorithm for the MANA algorithm.

Parameters	Value
<i>image</i> <small>LAYER</small>	Grayscale image
<i>object</i> <small>TYPE</small>	Object of interest is dark on a bright background
<i>polynomial</i> <small>ORDER</small>	PWM _{CURVE} is fitted with a 5 th order polynomial function
<i>object</i> <small>PROPERTY</small>	Maximum median area of the segmented objects is imposed as condition to find the optimal threshold among candidate ones.

Briefly, the RGB image of the histological specimen is converted into grayscale. From the grayscale histogram of the image, the Progressive Weighted Mean (PWM_{CURVE}) is computed as described in Chapter 2. Then, the PWM_{CURVE} is fitted with a 5th order polynomial function in order to estimate its inflection points (candidate thresholds). Among candidate thresholds, the optimal threshold is chosen as the one that identifies objects with the highest median area. Finally, cell nuclei are found as the regions with an intensity lower than the optimal threshold (dark objects on a bright background). The processing steps for obtaining the optimal threshold are shown in Figure 4.1, where images taken from three different tissues are presented as explanatory examples. From the results presented in Figure 4.1, it can be appreciated the robustness of the proposed algorithm for cell nuclei detection, where an optimal threshold value is selected, regardless of the cells' appearance or image histogram shape.

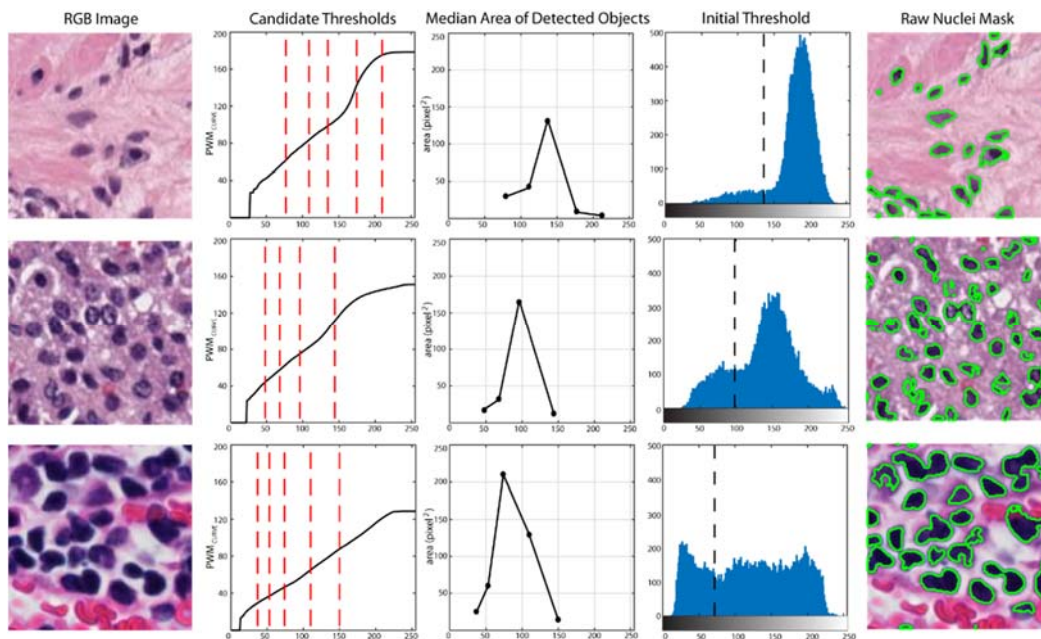


Fig. 4.1 Steps for obtaining the optimal threshold for three images of different tissues with a high variation of cells color and size. Starting from the RGB image, the PWM_{CURVE} is calculated from the grayscale histogram. Then, the inflection points of the PWM_{CURVE} (red dotted lines) are estimated (candidate thresholds). For each candidate threshold, the median area of detected objects is evaluated. The optimal threshold (black-dotted line) is chosen as the one that identifies objects with the highest median area. In the rightmost column, the application of the optimal threshold on the RGB image is illustrated.

Area-based correction

This step is devoted to correcting the over-segmentation of the previous step as the ARCO algorithm may lead to small or too large objects. Small regions could be wrong or over-segmented objects while too large areas may consist of a fusion of different nuclei. To lessen the over-segmentation and to optimize the nuclei detection, the mean area of segmented objects ($\text{mean}_{\text{AREA}}$) is calculated. Then, all the objects are labelled as ‘small’, ‘normal’, or ‘big’. ‘Small’ objects are regions smaller than 25% of $\text{mean}_{\text{AREA}}$ while all the objects greater than 5 times the $\text{mean}_{\text{AREA}}$ are classified as ‘big’. The remaining structures are labeled as ‘normal’.

‘Small’ objects are deleted because they are too little to be potentially considered as nuclei, whereas ‘big’ structures should be split, in case they are nuclei agglomerates. A first nuclei separation is obtained by iteratively decreasing the optimal threshold for all the ‘big’ objects until they are labeled as ‘normal’ (area less than 5 times the $\text{mean}_{\text{AREA}}$). The effect of this procedure is sketched in Figure 4.2a. Using these criteria, the optimal threshold found by the ARCO algorithm is locally modified to identify the highest number of nuclei within the histopathological image.

Nuclei separation

This step is needed to further separate remaining fused nuclei. In previous works, the watershed transform was successfully used to isolate touching structures [36]. The proposed method implements a marked-based watershed to separate “fused” nuclei [37]. In marker-based techniques, seeds close to nuclear centers are used as starting points for watershed transform. Firstly, the distance transform (DT) of the nuclei binary mask is calculated. To identify marker positions, local maxima of the DT are identified using the extended-maxima transform [38]. This transform estimates the regional maxima by searching in N-connected neighborhoods. The size of neighborhood (N) affects the sensitivity of the maxima-extended transform in the detection of nuclear seeds.

Additionally, the solidity of all objects is also evaluated. Solidity of a structure is defined as the ratio between its actual area and its convex area. Since it is expected that nuclei are convex objects, a detected shape containing an actual nucleus should have a solidity approximately equal to one. Hence, the solidity of a region is used as a discriminant feature for varying the sensitivity of the watershed transform by changing the neighborhood size of the maxima-extended. The MANA algorithm applies a high-sensitive watershed for low solidity shapes (object to split) while sensitivity is decreased for high solidity regions (single nuclei). The application of a marker-based watershed sensitive to shapes solidity is shown in Figure 4.2c. As a final operation, the mean area of the structures obtained after watershed is calculated and all the regions smaller than 25% of mean area are erased by MANA.

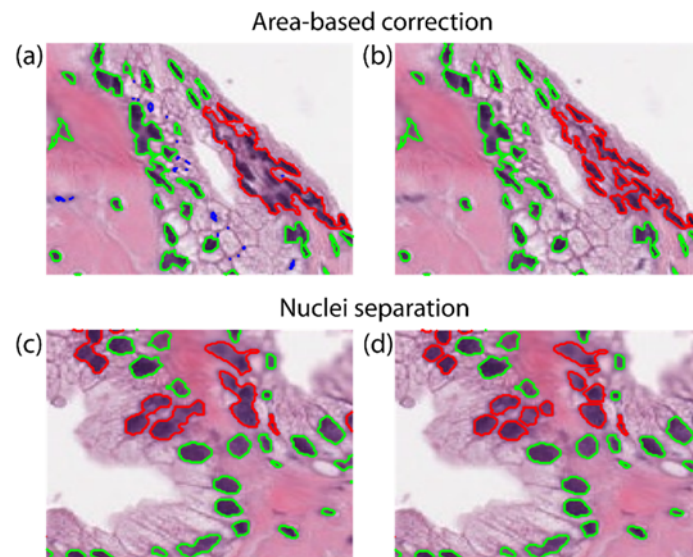


Fig. 4.2 Area correction and nuclei separation of the MANA algorithm. The input image is shown on the left column while, on the right one, the corresponding output image is represented. The first row (a,b) is relative to the area-based correction, where 'big' regions (red) are partially split and 'small' structures (blue) are deleted. The second row (c,d) shows the nuclei separation in which a marker-based watershed is applied on structures with a small (red) and high solidity (green).

Performance indicators

The parameters tuning of the MANA algorithm is reported in Appendix B. A comparison between manual annotations and automatic segmentation is carried out to assess the MANA algorithm performance in the detection and segmentation of nuclei. The segmentation performance is assessed by calculating the *recall*, *precision* and *F1_{SCORE}*, defined as follows:

$$recall = \frac{TP}{TP + FN} \quad (\text{Eq. 4.1})$$

$$precision = \frac{TP}{TP + FP} \quad (\text{Eq. 4.2})$$

$$F1_{SCORE} = \frac{2 \times precision \times recall}{precision + recall} \quad (\text{Eq. 4.3})$$

In detail, *recall* measures the missed detection of ground truth shapes and *precision* indicates the false detection of ghost objects [39]. *F1_{SCORE}* is a common used object detection metric [40], but it penalizes only object-level errors [13]. In fact, *F1_{SCORE}* does not take into account all pixel-level errors (e.g. under-segmentation of correctly detected objects). Let N_{CS} , N_{US} and N_{SE} represent the numbers of correct-segmentation (*CS*), under-segmentation (*US*) and segmentation-error (*SE*). The pixel-level performance of the MANA algorithm is assessed by calculating the *CS*, *US* and *SE* rates [41], which are defined as follows:

$$CS = \frac{N_{CS}}{N_{GT}} \times 100 \quad (\text{Eq. 4.4})$$

$$US = \frac{N_{US}}{N_{GT}} \times 100 \quad (\text{Eq. 4.5})$$

$$SE = \frac{N_{SE}}{N_{GT}} \times 100 \quad (\text{Eq. 4.6})$$

where N_{GT} (GT: ground truth) represents the number of nuclei identified by the pathologist. The US rate reveals the failure to split nuclear regions in the correct number of nuclei while the SE rate indicates the missed detection of cells. An example of correct-segmentation (CS), under-segmentation (US) and segmentation-error (SE) is provided in Figure 4.3.

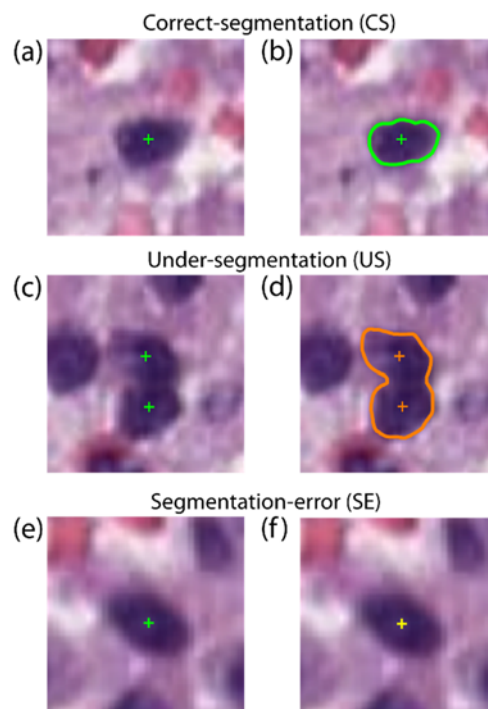


Fig. 4.3 Evaluation of the pixel-level performance. Manual annotations are illustrated in the left column while, in the right one, the corresponding automatic result is shown. (a,b) Correct-segmentation (color: green). (c,d) Under-segmentation (color: orange). (e,f) Segmentation-error (color: yellow).

4.2.2 Automated Gleason Score in Prostate Histological Images

Image database

The histological dataset consists of 25 biopsy images collected from 11 patients. The samples were H&E stained and digitalized at a 10x magnification (Aperio Scanscope XT) at “Città della Salute” Hospital – Molinette (Torino, Italy). Four images from each biopsy were extracted for a total of 100 images. An expert pathologist with over 15 years of experience assigned a Gleason Score to each image and was blinded from the automatic algorithm results. The overall dataset composition is shown in Table 4.3.

Table 4.3 Dataset composition of the algorithm for automatic gleason scoring.

Gleason Group	Gleason Score	# Images
1	≤ 6	30
2	3+4	20
3	4+3	25
4	4+4	25

Gleason grading algorithm architecture

The algorithm is designed to automatically assign the Gleason score in prostate histopathological images. The proposed method is developed using MATLAB environment and the image processing is carried out on a workstation with a 2.8 GHz exa-core CPU and 64 GB of RAM. The procedure of the proposed method is schematically described in Figure 4.4. Three main steps compose the processing: nuclei and lumen detection, glands segmentation and glands classification. In the next paragraphs, an exhaustive description of the proposed method is provided.

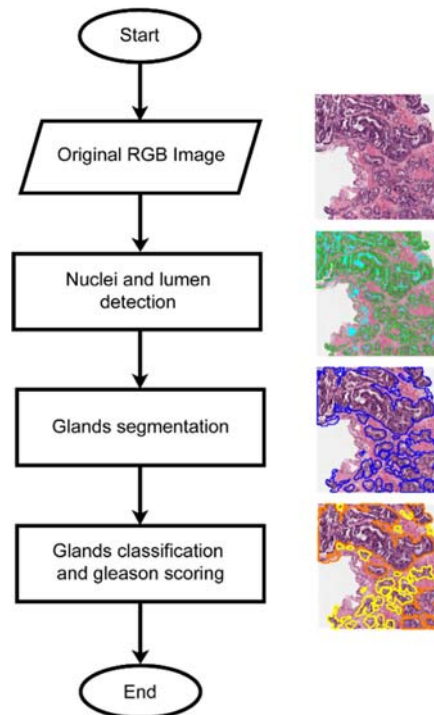


Fig. 4.4 Schematic representation of the algorithm for gleason scoring. The proposed technique takes as input the RGB image of the histological specimen to perform nuclei and lumen detection. Then, prostate glands are detected and classified to obtain the gleason score.

Nuclei and lumen detection

Starting from the original RGB image of the specimen (Figure 4.5a), the proposed algorithm performs the discrimination between the histological tissue and the background. The aim of this step is to process only the cellular structures within the image. The tissue recognition occurs thanks to the application of an

RGB high-pass filter [42] where the RGB color of each pixel is treated as a 3D vector, and the strength of the edge is the magnitude of the maximum gradient. Then, Otsu thresholding [43] is applied to extract a raw binary mask of the tissue and a morphological opening with a disk of $50\mu\text{m}$ radius is carried out to obtain smoother tissue contours (Figure 4.5b).

Once obtained the tissue boundary, the algorithm performs the nuclei detection (Figure 4.5c) using the MANA algorithm [44]. Then, the proposed method performs the lumen detection. Simple thresholding may be ineffective to segment gland lumina, since they often have different intensity and uniformity depending on their size, shape, and the presence of artefacts. Therefore, in order to properly segment the lumen, a series of Gabor kernels are applied to the original grayscale image, which are defined as follows:

$$g(x, y, \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(\frac{2\pi x'}{\lambda} + \psi\right) \quad (\text{Eq. 4.7})$$

where $x' = x\cos(\theta) + y\sin(\theta)$, $y' = -x\sin(\theta) + y\cos(\theta)$, γ is the spatial aspect ratio that specify the ellipticity of the Gabor function, ψ represents the phase offset, λ is the wavelength of the sinusoidal factor, θ denotes the orientation of the normal to the parallel stripes of a Gabor function and σ is the standard deviation of the Gaussian envelope. For this application, we imposed $\gamma=1.2$, $\psi=0$, $\lambda=10$, $\sigma=1$ and eight directions (θ) are considered to make the algorithm faster and to still reduce the noise level. The obtained eight filtered images are summed and normalized; then a threshold equal to 90% of the image maximum is applied. Finally, all small elements are removed to obtain the final lumen mask (Figure 4.5d).

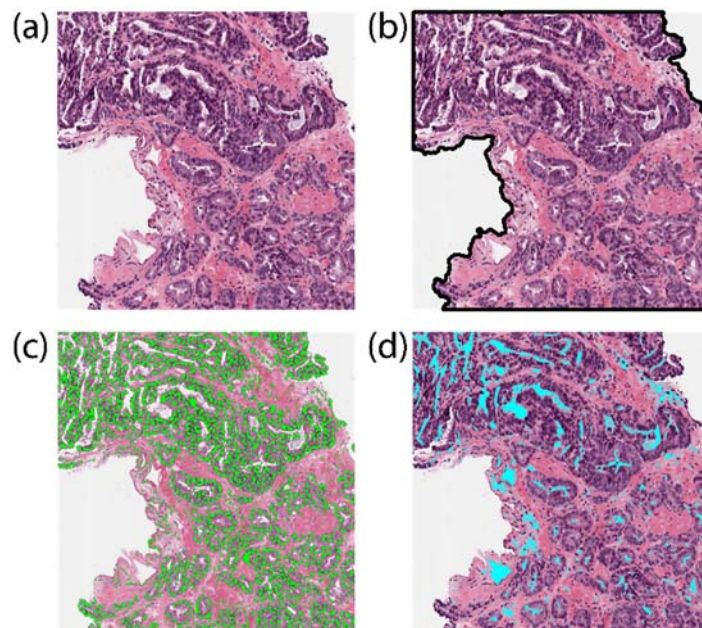


Fig. 4.5 Processing steps for nuclei and lumen detection. (a) Original RGB image. (b) Tissue identification. (c) Nuclei detection (green). (d) Lumina segmentation (cyan).

Glands segmentation

A first initial glands segmentation is obtained by means of a k-means algorithm [45]. The k-means algorithm is an iterative and unsupervised method that assigns n observations to exactly one of k clusters defined by centroids, where k is fixed a priori. The main idea is to define k centroids, one for each cluster. Then, the algorithm assigns each observation to the cluster with the closest centroid. At this point the average of the observations in each cluster is computed to obtain k new centroid locations. The algorithm repeats the previous step until cluster assignments do not change, or the maximum number of iterations ($iter_{MAX}$) is reached. The k-means algorithm is applied to the image $I_1 = I_R - I_B$, where I_R and I_B are the red and blue channels of the considered image I , respectively. This channel subtraction aims to brighten the stroma while simultaneously darkening everything belonging to the glands (Figure 4.6a). In order to obtain a correct classification, the background, nuclei and lumen mask are subtracted from image I_1 . The initial centroids for clustering are respectively the pixel with the highest and lowest intensity within image I_1 . $iter_{MAX}$ is set to 100 and the k-means is applied with a number of clusters (k) equal to 2: stroma and glands (Figure 4.6b). Gland pixels are defined as the pixels belonging to the cluster with the lower mean intensity.

Since the k-means result is noisy, a refinement process is necessary before obtaining the final gland segmentation. First of all, the k-means glands mask is merged with the lumen and nuclei masks to obtain a raw glands mask (Figure 4.6c). Then, small objects are removed (isolated nuclei), and a morphological closing is applied to smoothen the gland contours (Figure 4.6d). The processing for obtaining the prostate glands contours is shown in Figure 4.6.

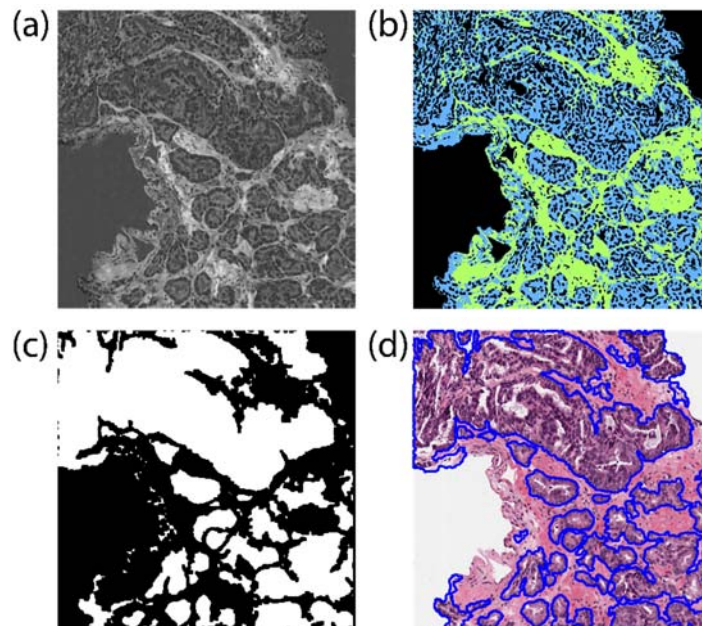


Fig. 4.6 Processing steps for prostate glands segmentation. (a) starting image for k-means algorithm. (b) K-means result for two classes: glands (blue) and stroma (green), (c) raw glands mask. (d) final glands contours (blue).

Glands classification

In order to obtain the Gleason Score, each gland needs to be classified according to a certain pattern (normal, pattern 3, pattern 4, and pattern 5).

A feed-forward neural network (NN) is used to classify each gland. To train the network, three glands are extracted from each image and the pathologist assigns a pattern for each of them. The total number of glands (300) is equally divided among the 3 classes (100 'normal', 100 'pattern3' and 100 'pattern4'). Then, the 300 extracted glands are randomly divided into 3 sets: the training set (70% = 210 glands), development set (15% = 45 glands), and test set (15% = 45 glands). Seven features are extracted from each gland: two morphological descriptors and five first-order texture descriptors [46], [47]. The detailed features description is provided in Table 4.4.

Table 4.4 Input features of the neural network to perform glands pattern classification.

Feature	Mathematical definition
Area (A)	$ \{(x_i, y_i)\} , \quad \forall (x_i, y_i) \in Gland$
Solidity (S)	$S = \frac{A}{A_{CONVEX}}$
Mean (m)	$m = \frac{1}{N} \sum_{i=1}^N I(x_i, y_i)$
Variance (σ^2)	$\sigma^2 = \frac{1}{N} \sum_{i=1}^N [I(x_i, y_i) - m]^2$
Skewness (S_k)	$S_k = \frac{1}{N} \frac{\sum_{i=1}^N [I(x_i, y_i) - m]^3}{(\sigma^2)^{\frac{3}{2}}}$
Kurtosis (K_t)	$K_t = \frac{1}{N} \frac{\sum_{i=1}^N [I(x_i, y_i) - m]^4}{(\sigma^2)^2}$
Entropy (E)	$E = - \sum_{j=1}^{256} p_j \log_2 p_j$

where A_{CONVEX} indicates the area of the smallest convex polygon that can contain the gland, $I(x,y)$ denotes the grayscale image of the gland, N indicates the glands area and p represents its normalized histogram.

In order to understand the feature's relation to the response variable (gland pattern), the Pearson correlation coefficient is also computed [48]. All seven input features are used to train the neural network because, for each of them, the Pearson coefficient is always higher than 0.3. Feature standardization is applied to obtain variables with zero-mean and unit-variance. During network training, the Levenberg-Marquardt backpropagation [49] and the mean squared normalized error (mse) are used as optimization algorithm and cost function respectively. Finally, the number of epochs is set to 100.

Several network configurations are tested, varying the network parameters (#units, #layers, activation functions and learning rate) and, for each configuration, the errors on the train and dev sets are evaluated. The optimal neural network is chosen in two steps. First of all, the NNs with a maximum

difference between the train and development error equal to 3% are selected to avoid overfitting. Among these, the final configuration is chosen as the one with the minimum train/dev mean error to ensure best performance.

The optimal network is then used to classify all the glands detected by the proposed algorithm (Figure 4.7). After glands classification, the Gleason Score is assigned according to two rules: i) the main pattern is defined as the most dominant pattern in the image, ii) the secondary pattern is set as the second most dominant pattern as long as it covers at least 5% of the image; otherwise the secondary pattern is imposed as the primary one.

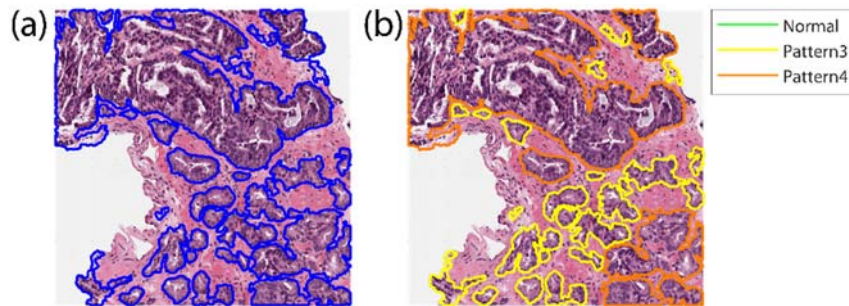


Fig. 4.7 Glands classification for gleason scoring. (a) Glands mask. (b) Glands classification (yellow: pattern 3, orange: pattern 4).

Performance indicators

The parameters tuning of the proposed algorithm is reported in Appendix B. The classification *accuracy* is evaluated to measure the algorithm performance in glands classification and gleason scoring:

$$accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (\text{Eq. 4.8})$$

Accuracy represents the proportion of correct classifications among all classifications, eventually multiplied by 100 to turn it into a percentage.

4.2.3 Automated Detection of Neoplastic Epithelium in Breast Carcinoma

Image database

The histological dataset consists of 100 H&E stained images from 10 different cases of invasive carcinoma, non-special type (according to WHO 2012 classification), each with a dimension of 600x1200 pixel and extracted from a whole-slide digital scan. The samples were digitalized at a 20x magnification at the “ASL CN2 Alba-Bra” (Alba, Italy). 50 images include both neoplastic epithelium and stromal reaction, 25 has only neoplastic epithelium and 25 images are extracted without cancer tissue in the field-of-view. One expert pathologist manually annotated the neoplastic epithelium contours in each image. The overall dataset composition is shown in Table 4.5.

Table 4.5 Dataset composition of the algorithm for neoplastic epithelium detection in breast carcinoma.

Tissue condition	# Images
Neoplastic epithelium and stromal reaction	50
All neoplastic epithelium	25
No neoplastic epithelium	25
<i>Total</i>	<i>100</i>

Neoplastic detection algorithm architecture

The proposed algorithm is designed to automatically detect neoplastic epithelium contours in H&E stained images of breast tissue. The algorithm is developed using MATLAB and runs on a workstation with a 2.8 GHz exa-core CPU and 64 GB of RAM. The procedure of the proposed method is schematically described in Figure 4.8. Two main steps compose the processing: i) nuclei segmentation and classification, ii) neoplastic epithelium detection. In the next paragraphs, an exhaustive description of the method is provided.

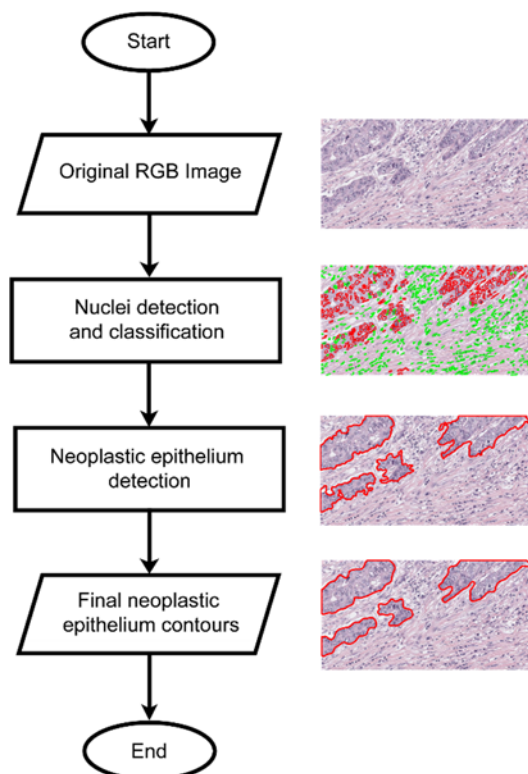


Fig. 4.8 Schematic representation of the algorithm for neoplastic epithelium detection in breast tissue. The proposed technique takes as input the RGB image of the histological specimen to perform nuclei detection and classification. Then, neoplastic epithelium is detected as the area containing tumoral nuclei.

Nuclei segmentation and classification

The first step of the proposed algorithm is the cell nuclei segmentation inside the image. This task is performed using the nuclei detection provided by the MANA algorithm [44]. Briefly, the RGB image of the histological specimen (Figure 4.9a) is converted into grayscale and the object-based detection provided by ARCO is applied to obtain a raw mask of nuclei. Then, area-based corrections and nuclei separation are performed to identify the cells boundaries (Figure 4.9b).

In order to find tumor areas, each nucleus needs to be classified according to a certain pattern (*tumor* or *no-tumor*). A feed-forward neural network (NN) is used to classify each cell. Five images are randomly selected to train the network. For each image, the pathologist manually draws all nuclei and assigns a pattern for each of them. Then, the total number of nuclei (9316) is randomly divided into 3 sets: the training set (70% = 6522 cells), development set (15% = 1397 cells), and test set (15% = 1397 cells). The same five first-order texture descriptors used for gleason scoring (Table 4.4) are extracted from each cell and used as input features for the NN: *mean*, *variance*, *skewness*, *kurtosis* and *entropy*. Nuclear texture is used to classify each cell as cancer nuclei have a distinct morphology characterized by coarse chromatin texture [50].

The optimal network configuration is chosen using the same criteria adopted for gleason scoring (section 4.2.2). The resulting network configuration has three hidden layers, with respectively 8, 4 and 3 units. All layers have a tan-sigmoid transfer function while the output layer has a linear transfer function. The network learning rate is set to 0.01. The optimal network is then used to classify all the cell nuclei detected by the MANA algorithm (Figure 4.9c). Since some nuclei can be misclassified, an iterative three-steps refining process is applied to better define the tumor areas:

1. for each cell classified as *no-tumor*, a Region of Interest (ROI) of 40x40 μm surrounding that cell is defined. If inside that ROI at least 95% (*conversion percentage*) of the nuclei has been classified as *tumor*, then the cell is now labeled as *tumor*;
2. for each cell classified as *tumor*, a Region of Interest (ROI) of 40x40 μm is also defined. If inside that ROI at least 95% of the nuclei has been classified as *no-tumor*, then the cell is now labeled as *no-tumor*;
3. lowering of the *conversion percentage* by 5%. These steps are repeated until the conversion percentage reaches 75%.

The procedure described above allowed to decrease the number of misclassified nuclei. An example of the refining process is shown in Figure 4.9d.

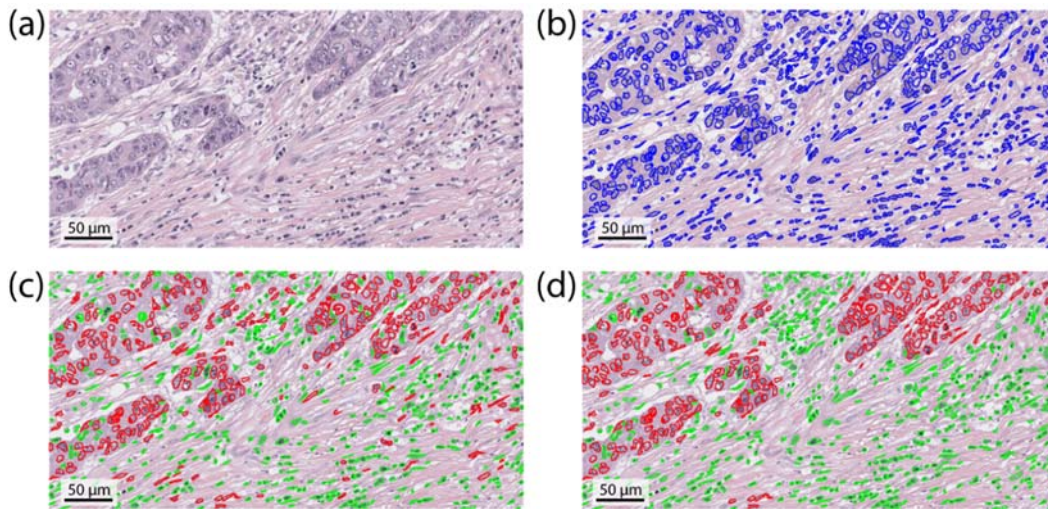


Fig. 4.9 Nuclei detection and classification. (a) Original RGB image. (b) Nuclei detection (blue). (c) Raw nuclei classification (red: tumor nuclei, green: no-tumor nuclei). (d) Nuclei classification after refining process.

Neoplastic epithelium detection

Once performed nuclei detection and classification, only nuclei labelled as *tumor* are considered for the next steps of the processing (Figure 4.10a). Centroids of tumor cells are extracted, and a density-based clustering is performed to segment tumor areas. In particular, a spatial clustering named DBSCAN (Density-Based Spatial Clustering of Application with Noise) is applied to each nuclei centroid [51]. This algorithm finds all the neighbors of data points, within a circle of radius ε , and adds them into the same cluster. For any neighbor point, if its ε -neighborhood contains at least a predefined number of points (*minPOINTS*), the cluster is expanded to contain its neighbors, as well. However, if the number of points in the neighborhood is less than *minPOINTS*, the point is considered to be noise and it is deleted. An iterative two-steps DBSCAN is performed by the proposed method, starting from $\varepsilon = 200 \mu\text{m}$ and *minPOINTS* = 5:

1. application of the DBSCAN on nuclei centroids using the current values of ε and *minPOINTS*;
2. comparison between the number of cells before and after the clustering. If the difference is less than 5%, ε is decreased by $20 \mu\text{m}$ and the algorithm returns to step 1.

These two steps are repeated until more than 5% of nuclei are deleted by DBSCAN or ε reaches $50 \mu\text{m}$. In this way, the clustering sensitivity is adapted for each image by varying the neighborhood size (ε). The application of the iterative DBSCAN is shown in Figure 4.10b.

For each cluster, the proposed method extracts the circumscribed polygon to all the nuclei of that cluster (Figure 4.10c). However, the result obtained is still sub-optimal, so further steps are performed to get an accurate contour of the neoplastic epithelium.

The proposed method applies an active contour model to improve the detection of tumor borders. In particular, the Chan-Vese region-based energy model is implemented as described in [52]. This model could detect objects whose boundaries are not necessarily defined by gradient. The Chan-Vese active contours is based on techniques of curve evolution [53], Mumford–Shah functional for segmentation [54] and level sets [55]. For each image of the dataset, this model allows to obtain accurate contours of tumor areas, even if the gradient between neoplastic epithelium and stromal response is very low. The Chan-Vese active contours is applied to the image $I_1 = I_R - I_B$, where I_R and I_B are the red and blue channels of the considered image I , respectively. This channel subtraction aims to brighten the stromal response while simultaneously darkening everything belonging to the neoplastic epithelium (Figure 4.10d). The tumor polygons obtained in the previous step are used as the initial contour at which the evolution of the segmentation begins, and the number of iterations is set to 100. The result obtained after the application of Chan-Vese model is illustrated in Figure 4.10e.

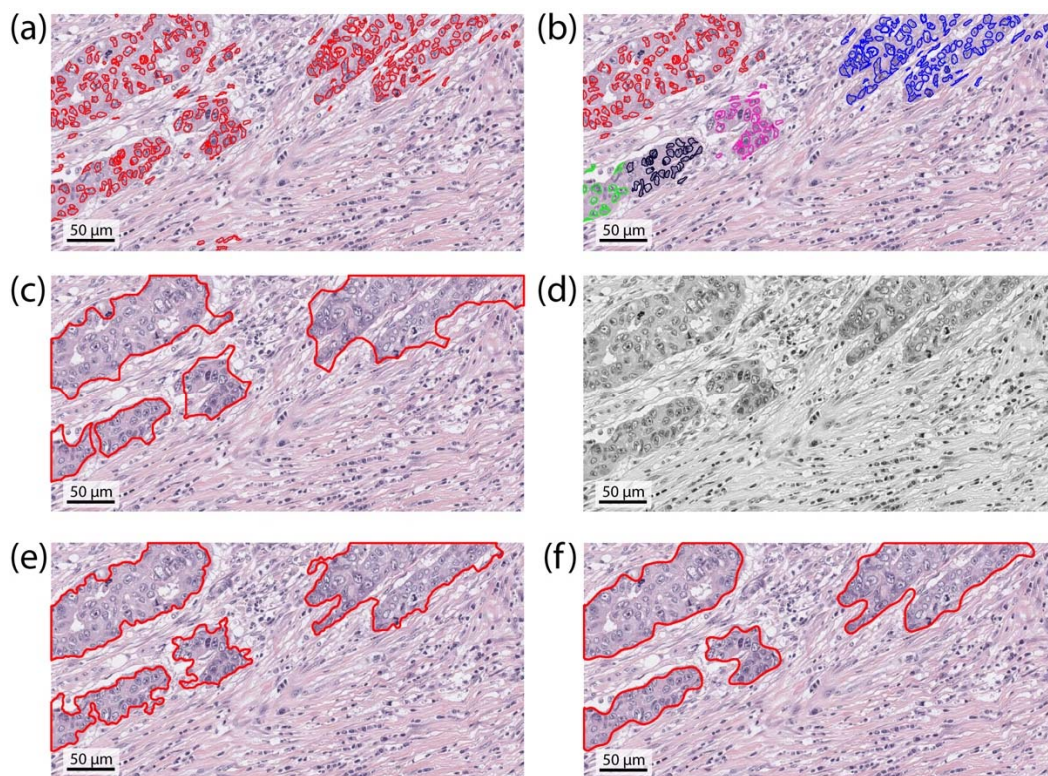


Fig. 4.10 Neoplastic epithelium detection. (a) Nuclei classified as tumor, (b) Density-based nuclei clustering (DBSCAN) where each color represents one cluster. (c) Circumscribed polygon for each tumor zone. (d) Starting image for active contours. (e) Tumor boundaries after active contours. (f) Final neoplastic epithelium contours.

Morphological operators are then applied to tumor areas to obtain smoother contours. This process is composed of three steps: i) morphological erosion using a disk with 5 μm radius, ii) removal of areas smaller than 100 μm² and iii) morphological dilation using the same structural element of the previous erosion. Finally, boundaries between neoplastic epithelial and stromal response are

interpolated using the Savitzky-Golay filter [56]. The Savitzky-Golay filter is designed to smoothing data points using a polynomial function fitted to N neighboring points. For this application, the polynomial order is set to 11 with a neighborhood size of 201. The final result provided by the proposed algorithm is illustrated in Figure 4.10f.

Finally, one of the three tissue conditions (Table 4.5) is associated to each image following these rules:

1. if the whole image is recognized as a tumor area, then the tissue condition is *all neoplastic epithelium*;
2. if there is no tumor area, then the tissue condition is *no neoplastic epithelium*;
3. in all other cases the tissue condition is *neoplastic epithelium and stromal reaction*.

Performance indicators

The parameters tuning of the proposed algorithm is reported in Appendix B. The *accuracy* (Eq.4.8) of the algorithm in the association of the tissue condition (*all neoplastic epithelium*, *no neoplastic epithelium*, *neoplastic epithelium and stroma reaction*) is evaluated. Then, a comparison between masks drawn by a manual operator ($MASK_{MANUAL}$) and those provided by the proposed method ($MASK_{AUTOMATIC}$) is carried out by calculating the *recall* (Eq. 4.1), *precision* (Eq. 4.2), *F1SCORE* (Eq. 4.3), *specificity* and *jaccardINDEX*:

$$specificity = \frac{TN}{TN + FP} \quad (\text{Eq. 4.9})$$

$$jaccard_{INDEX} = \frac{|MASK_{MANUAL} \cap MASK_{AUTOMATIC}|}{|MASK_{MANUAL} \cup MASK_{AUTOMATIC}|} \quad (\text{Eq. 4.10})$$

In detail, *specificity* assesses the true negative fraction and the *jaccardINDEX* measures similarity between two different shapes, defined as the size of the intersection divided by the size of the union of the segmented object [57].

4.3 Results

4.3.1 Segmentation results of MANA Algorithm

MANA vs manual operator image segmentation

An example of the validation process for each of the six organs analyzed (colon, liver, bone, prostate, adrenal gland and thyroid) is shown in Figure 4.11. Table 4.6 summarizes both the object-level performance (*recall*, *precision* and *F1SCORE*) and pixel-level performance (*CS*, *US*, *SE* rates) of the proposed technique. As can be seen from Table 4.6, the MANA algorithm achieves an average *F1SCORE* of 0.9305 on 30 images. The accuracy and the robustness of the

proposed algorithm are also demonstrated by the similar values of recall and precision obtained for the entire dataset.

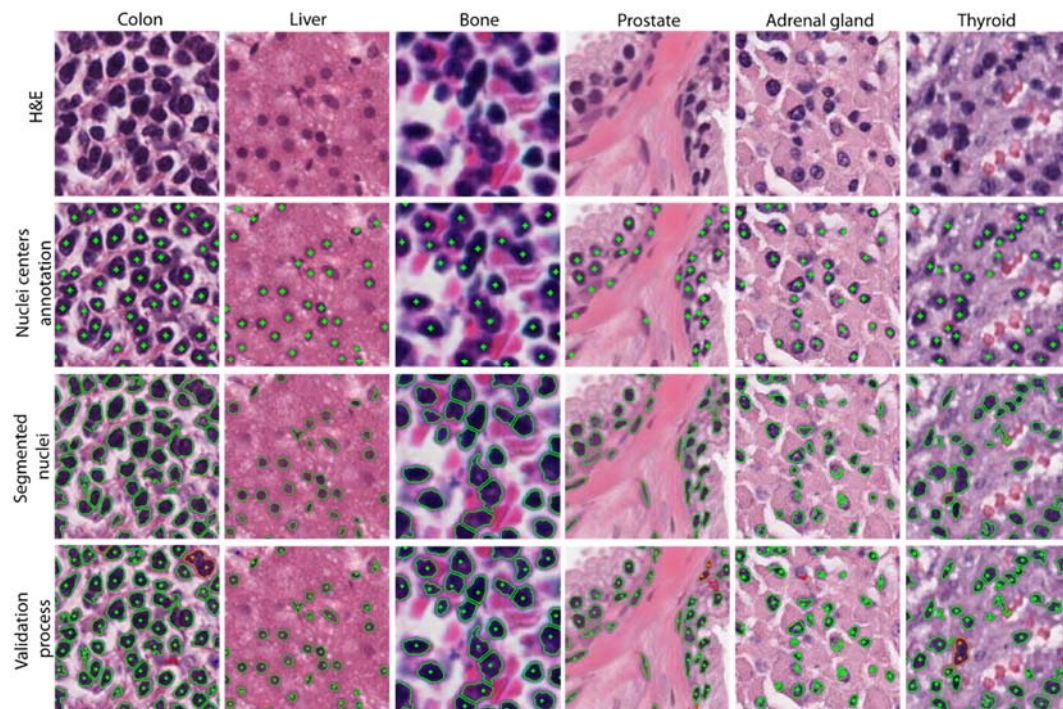


Fig. 4.11 Examples of sub-images taken from each tissue analyzed (columns), showing challenging cases based on variation in nuclear size, color appearance and crowding. In the first and second row are shown the original image and the manual annotations respectively. The automatic result of the MANA algorithm is presented in the third rows. In the last row, the validation process is illustrated (correct-segmentation: green, under-segmentation: orange, segmentation-error: blue, false-negative: red).

Table 4.6 Performance of the MANA algorithm in the nuclei segmentation (mean \pm standard deviation).

Organ	Computational Time (sec)	Object-level performance			Pixel-level performance		
		Recall	Precision	F1 _{SCORE}	CS (%)	US (%)	SE (%)
Colon	22.89 \pm 2.15	0.9505 \pm 0.0121	0.9048 \pm 0.0114	0.9270 \pm 0.0086	86.78 \pm 1.97	8.69 \pm 2.00	4.53 \pm 1.14
Liver	11.32 \pm 1.25	0.9249 \pm 0.0267	0.9547 \pm 0.0114	0.9392 \pm 0.0101	87.17 \pm 4.55	5.81 \pm 2.42	7.02 \pm 2.37
Bone	13.10 \pm 1.13	0.9486 \pm 0.0290	0.9362 \pm 0.0203	0.9417 \pm 0.0077	74.15 \pm 4.17	21.35 \pm 4.30	4.07 \pm 2.49
Prostate	12.28 \pm 1.31	0.9533 \pm 0.0127	0.9404 \pm 0.0147	0.9467 \pm 0.0106	77.47 \pm 7.72	18.79 \pm 7.31	3.74 \pm 0.88
Adrenal Gland	18.02 \pm 1.26	0.9126 \pm 0.0300	0.9239 \pm 0.0312	0.9174 \pm 0.0129	84.60 \pm 4.44	7.33 \pm 2.55	8.06 \pm 2.62
Thyroid	23.71 \pm 5.94	0.9335 \pm 0.0296	0.8914 \pm 0.0221	0.9112 \pm 0.0038	81.62 \pm 6.18	12.61 \pm 5.39	5.77 \pm 2.60
<i>Overall</i>	<i>16.89 \pm 5.72</i>	<i>0.9372 \pm 0.0288</i>	<i>0.9253 \pm 0.0293</i>	<i>0.9305 \pm 0.0161</i>	<i>81.97 \pm 7.05</i>	<i>12.43 \pm 7.32</i>	<i>5.53 \pm 2.66</i>

An average correct-segmentation (*CS*) rate of 81.97% coupled to a segmentation-error (*SE*) rate of 5.53% is also obtained. Furthermore, the under-segmentation (*US*) rate is around 7% in tissues with well-defined nuclei (e.g. liver and adrenal gland) while the percentage increases up to 20% in images containing a lot of touching cells (e.g. bone and prostate).

Finally, the computational time is slightly dependent on the number of detected nuclei, ranging between 11 and 23 seconds (average computational time: 16.89 seconds).

Comparison with open-source software

Automatic results provided by MANA are also compared with three open-source software (CellProfiler, QuPath and Fiji) widely applied to the analysis of histological images [58]. CellProfiler [59] is a software with a modular structure that integrates several image-processing techniques to perform automatic analyses on biological images. QuPath [60] is a new bio-image software designed to provide an open-source solution for whole slide images in digital pathology. This software allows performing different automatic analyses of histopathological images, including cell nuclei segmentation. Fiji [61] is a Java-based software with several plugins which facilitate scientific image analysis based on a semi-automatic pipeline consisting of: (i) conversion of RGB image into grayscale, (ii) manual intensity thresholding, (iii) automatic nuclei separation using watershed transform and (iv) small particles removal.

A quantitative comparison of the performances offered by the three open-source software with MANA is reported in Figure 4.12 and Table 4.7.

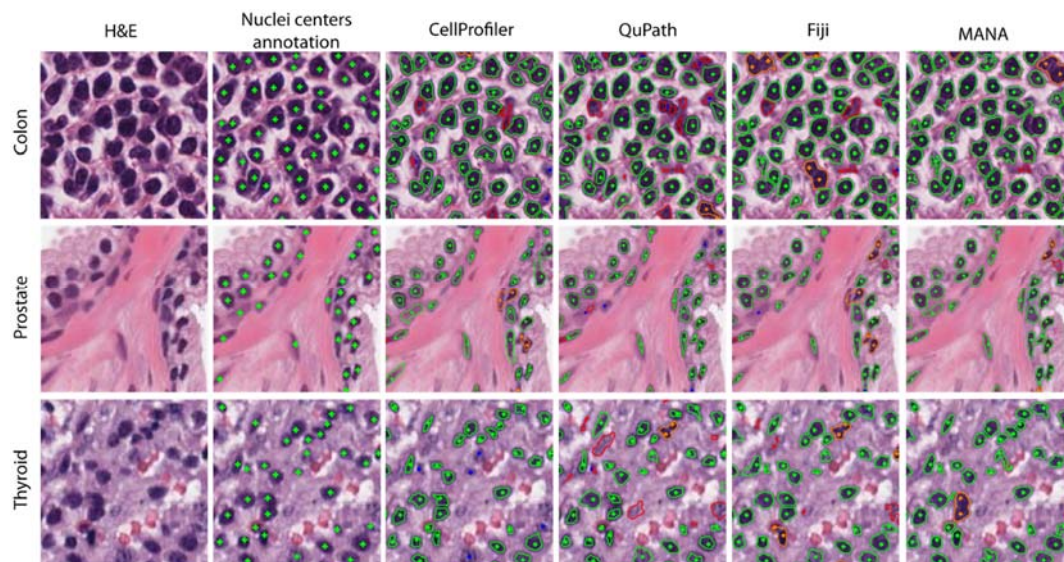


Fig. 4.12 Comparison between CellProfiler, QuPath, Fiji and MANA in the nuclei segmentation (columns). The first and second column show the original image and the corresponding manual annotation. Sub-images taken from three different organs are illustrated in rows.

Table 4.7 Performance of CellProfiler, QuPath and Fiji software in the segmentation of the cell nuclei (mean \pm standard deviation).

Method	Computational Time (sec)	Object-level performance			Pixel-level performance		
		Recall	Precision	F1 _{SCORE}	CS (%)	US (%)	SE (%)
CellProfiler (automatic)	21.13 \pm 3.78	0.6866 \pm 0.2421	0.8274 \pm 0.0820	0.7154 \pm 0.2030	65.47 \pm 23.23	4.79 \pm 3.21	29.73 \pm 23.44
QuPath (automatic)	11.37 \pm 1.96	0.9248 \pm 0.0552	0.7120 \pm 0.0916	0.8004 \pm 0.0638	74.24 \pm 10.41	9.29 \pm 7.70	6.47 \pm 3.23
Fiji (semi-automatic)	252.73 \pm 76.11	0.9462 \pm 0.0386	0.8658 \pm 0.0424	0.9030 \pm 0.0248	82.59 \pm 7.05	12.69 \pm 5.88	4.65 \pm 3.30
MANA (proposed)	16.89 \pm 5.72	0.9372 \pm 0.0288	0.9253 \pm 0.0293	0.9305 \pm 0.0161	81.97 \pm 7.05	12.43 \pm 7.32	5.53 \pm 2.66

As can be seen from Table 4.7, the nuclei segmentation provided by CellProfiler has a low *recall*. In fact, several nuclei are not detected by the software, and this generates a high number of false negative (FN) cells. In addition, the average *F1_{SCORE}* of CellProfiler is lower than the proposed one for more than 20 % (0.7154 vs 0.9305). This software has also the poorest pixel-level performance, with a low *CS* rate (65.45%) and a large mean *SE* rate (29.73%).

The QuPath software proves to be an efficient tool for nuclei segmentation, with an average *recall* of 0.9248 and a fast-computational time (11.37 seconds). On the other hand, this software detects a lot of false-positive (FP) nuclei, causing a very low *precision* (0.7120). This low *precision* leads to a lowering of the average *F1_{SCORE}* (0.8004).

The average *F1_{SCORE}* obtained with Fiji is slightly lower than those achieved with the MANA algorithm (0.9030 vs 0.9305). In fact, the Fiji processing is based on a manual single threshold whereas MANA can locally modify the optimal threshold to detect the highest number of actual nuclei. Being a semi-automated pipeline, the Fiji processing is also 15 times higher than MANA algorithm (252.73 vs 16.89 seconds).

4.3.2 Tissue Characterization in Prostate Cancer

The optimal network configuration used for glands classification has four hidden layers, with respectively 9, 5, 7 and 8 units (Figure 4.13). The first two and the last layers has a tan-sigmoid transfer function, the third layer has a log sigmoid transfer function and the output layer has a linear transfer function. The network learning rate is 0.01. This network achieves an average accuracy of 82.85% (training set) and 84.54% (development set). In the test set, the network obtains a mean accuracy of 81.98% in glands classification (76.92% for normal pattern, 82.35% for pattern 3 and 86.67% for pattern 4).

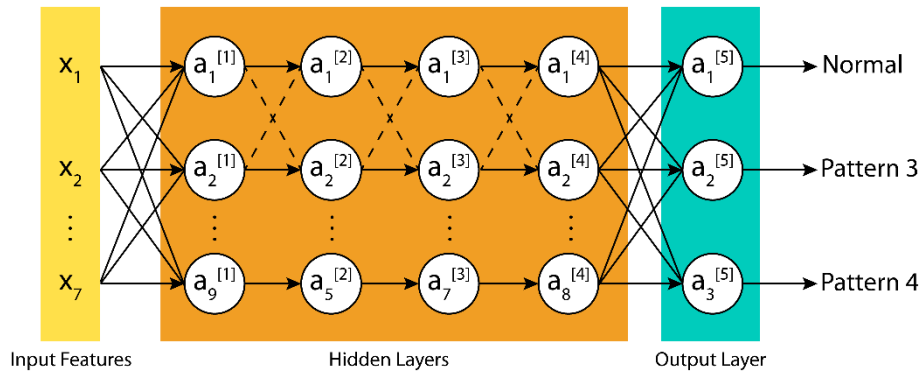


Fig. 4.13 Neural network architecture for glands classification.

To evaluate the performance of the algorithm, a comparison between the automatic Gleason Group and the one assigned by the pathologist is carried out. The overall accuracy of the proposed method for gleason scoring is reported in a confusion matrix in Figure 4.14.

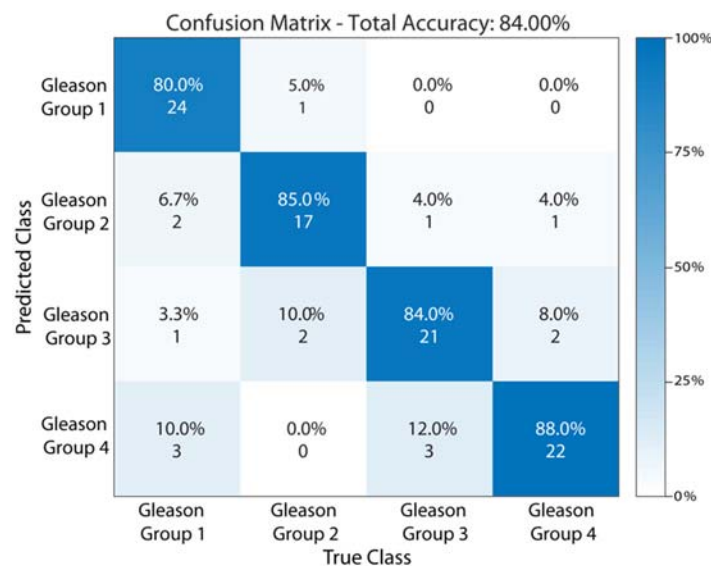


Fig. 4.14 Confusion matrix of the classification accuracy for all the Gleason Group (GG).

It can be observed that the algorithm correctly assigns the Grade Group in 84% of the images (84 out of 100 images). Importantly, the difference between Gleason Group 2 (3+4) and 3 (4+3) is correctly distinguished with a rate of 84.4% (38/45 images). In five cases the algorithm underestimates the Gleason Group, and a GG 2 is misclassified in GG 1 in only one image. Results also show how the proposed algorithm never classify a high Gleason Group (GG3, GG4) as a low one (GG1), showing excellent sensitivity. Figure 4.15 shows the comparison between manual and automatic gleason scoring (GG_{MANUAL} , $GG_{\text{AUTOMATIC}}$) for four sample images.

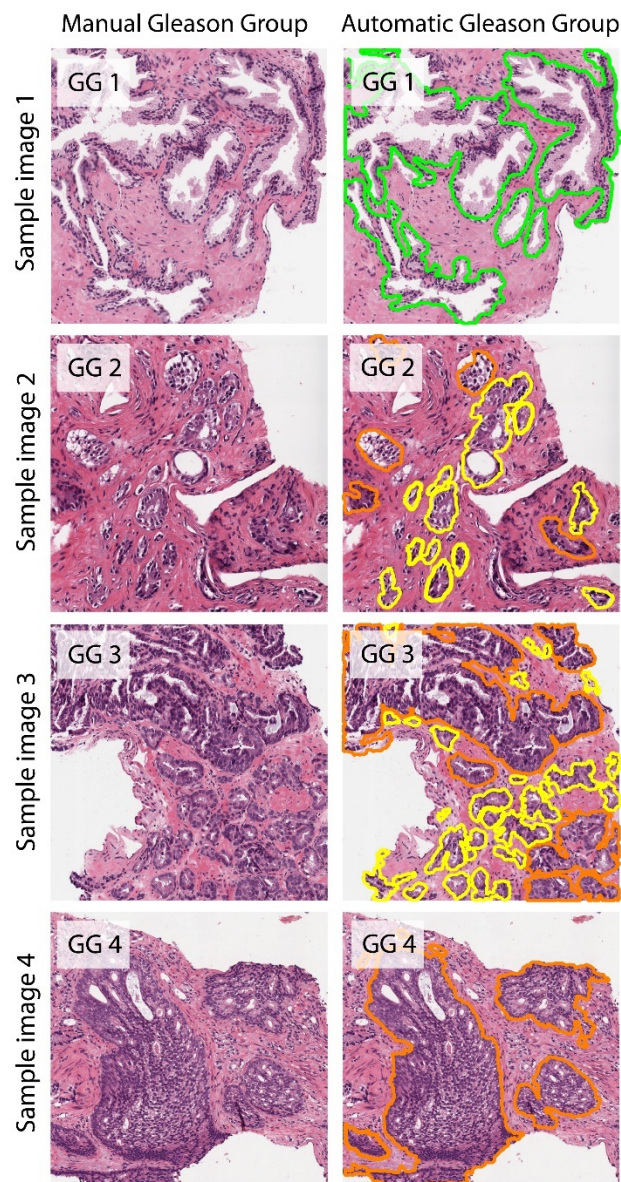


Fig. 4.15 Comparison between manual and automatic gleason scoring. First row illustrates the manual Gleason Group while second row shows the glands classification (green: normal, yellow: pattern 3, orange: pattern 4) and the corresponding automatic Gleason Group.

4.3.3 Tissue Characterization in Breast Carcinoma

The network for nuclei classification achieves an average accuracy of 94.64% (training set) and 94.29% (development set). In the test set, the network obtains a mean accuracy of 95.22% (93.81% for *no-tumor* nuclei and 96.62% for *tumor* nuclei).

The algorithm also obtains a 100% of *accuracy* in the identification of the tissue condition for all the 100 images of the dataset. An example of manual and automatic segmentation in the three tissue conditions is shown in Figure 4.16.

The performance of the proposed method in the discrimination between neoplastic epithelium and stromal response is also assessed. In order to perform

this kind of evaluation, images should contain both neoplastic epithelium and stromal response. For this reason, only images labeled as *neoplastic epithelium and stromal response* are used. The results of the comparison between manual and automatic segmentation are summarized in Table 4.9.

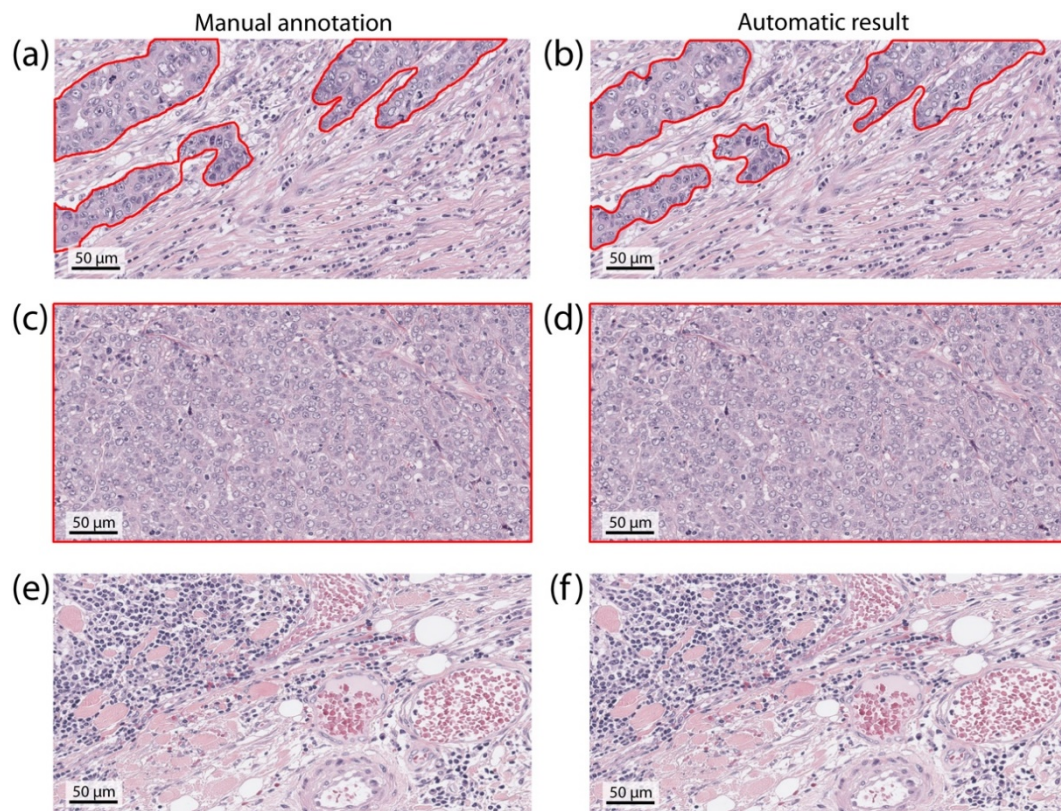


Fig. 4.16 Comparison between manual and automatic segmentation in the three tissue conditions. First column illustrates the manual annotations while second column shows the corresponding automatic results. (a,b) image labeled as neoplastic epithelium and stromal response, (c,d) image with all neoplastic epithelium, (e,f) image without neoplastic epithelium.

Table 4.9 Performance of the proposed algorithm in the discrimination between neoplastic epithelium and stromal response (mean \pm standard deviation).

Tissue condition	Computational Time (sec)	Recall	Precision	F1 _{SCORE}	Specificity	jaccard INDEX
Neoplastic epithelium and stromal reaction	18.11 \pm 3.86	0.8680 \pm 0.0418	0.9326 \pm 0.0660	0.8894 \pm 0.0736	0.9407 \pm 0.0683	0.8481 \pm 0.1114

The proposed method can be considered very performing in the detection and segmentation of neoplastic epithelium, with very high average values of *recall*, *precision*, *F1_{SCORE}*, *specificity* and *jaccard_{INDEX}* thus demonstrating the accuracy of the method (Table 4.9). For all images, *precision* and *specificity* are always higher than 0.90. In addition to being accurate, the proposed method is also fast, with an average computational time of 18.21 seconds.

Figure 4.17 shows some examples of manual and automatic segmentation of neoplastic epithelium in different patients.

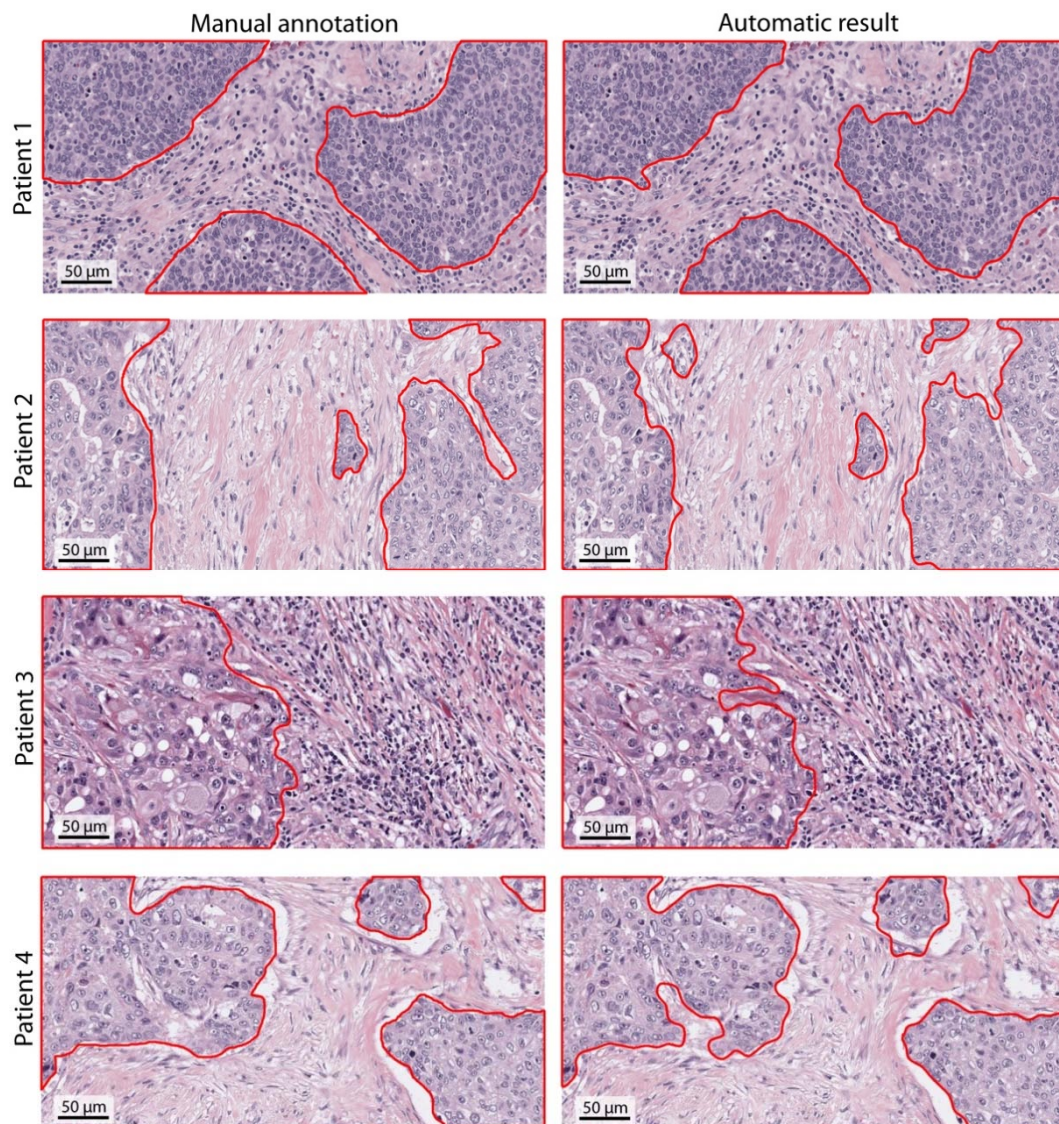


Fig. 4.17 Comparison between manual and automatic segmentation of the neoplastic epithelium for four different patients, showing challenging cases with high variation of staining intensities, cells morphology and tumor-stromal architecture.

4.4 Discussion

In this Chapter, three fully automated methods are proposed for the quantitative analysis of histopathological images. The analysis of histopathological images is crucially important and has a wide range of applications, like cancer grading [62], cancer diagnosis [63] and molecular markers quantification in healthy and pathological samples specimens [10].

The first algorithm, named MANA, is able to detect cell nuclei boundaries within H&E stained images. Nuclei segmentation in histopathological images is a challenging task due to the high variability of nuclei size, shape and color intensity [64], [65]. The MANA algorithm does not need any user interaction and

it is able to automatically detect cell nuclei in several tissues and magnifications. Nuclei centers were manually marked by an expert pathologist, for a total of 59,123 cells. It was not necessary to segment nuclei boundaries since the MANA algorithm does not require a training set as deep learning-based methods. For this reason, having a faster manual segmentation, the number of annotated nuclei was increased, creating a dataset that contains more than twice the number of marked nuclei compared to previous works [2], [13], [63].

The MANA algorithm is validated using metrics that penalizes both detection (object-level) and segmentation (pixel-level) errors. The comparison between manual annotations and automatic results shows high performances of the proposed algorithm. For each of the six tissues analyzed, the MANA algorithm obtains always an $F1_{SCORE}$ higher than 0.90, with an average $F1_{SCORE}$ of 0.9305 (Table 4.6). Compared with the only multi-tissue nuclei segmentation system proposed in literature [13], the proposed method achieves a large margin, with more than 10% improvement in nuclei detection ($F1_{SCORE}$: 0.8267 vs 0.9305). Object-level and pixel-level performance of the MANA algorithm are also comparable to previous works on cell nuclei detection in H&E stained images [15], [64], [66]. Overall, the accuracy and robustness of MANA allows achieving, on different magnifications and tissues, performances in line or better than those of state-of-art algorithms designed for single tissues [59], [60]. The proposed technique also obtains the highest $F1_{SCORE}$ compared to other open-source software (CellProfiler, QuPath and Fiji) designed for nuclei segmentation. Compared to other automatic methods, MANA has the best pixel-level performance and one of the lowest computational time.

The second method presented is a fully automated strategy for gleason scoring in prostate histopathological images. The gleason grading system is used to help evaluate the prognosis of prostate cancer using samples from a prostate biopsy. The proposed method can operate at the commonly used magnification (10x) and discriminate between normal glands, Gleason patterns 3, and Gleason patterns 4, and finally provide the correct Gleason Group to each image. The algorithm is tested on 100 H&E stained images of prostate tissue and automatic results are compared with manual scoring of an expert pathologist. To the best of our knowledge, there is no published work reporting about an algorithm able to distinguish between the ambiguous Gleason GG 2 (3+4) and 3 (4+3). Compared to other works found in literature [25], [67], the proposed approach does not require a very high magnification and provides good results of glands segmentation even without well-defined lumina.

As can be seen from the classification performance (Figure 4.14), only 16 images out of 100 are incorrectly classified (total *accuracy* of 84%). In 11 cases the algorithm overestimates the GG while in only 5 images the GG is underestimated. The algorithm is very performing in the distinction between GG 2 and GG3, with an accuracy of 84.4% (38/45 images). An integration of our glands dataset is required in order to improve the neural network performance. Currently, our research team is collecting new prostate images of GG 4 (5 + 3 and 3 + 5) and GG 5 (GS = 9-10) to include patter 5 glands in the dataset.

The third proposed technique is a fully automatic method for the discrimination between neoplastic epithelium and stromal reaction in H&E stained images of breast tissue. The distinction between neoplastic epithelium and stromal reaction is essential in the correct extraction of predictive and prognostic biomarkers for breast carcinoma [26], [28]. Starting from cell nuclei, our technique is able to detect neoplastic epithelium boundaries without any user interaction. The proposed method is tested on 100 H&E stained images of breast tissue and automatic results are compared with manual annotations of an expert pathologist.

The comparison between manual and automatic segmentation shows high performances of the proposed technique (Table 4.9). In particular, an average *F1SCORE* of 0.8894 coupled to a mean *jaccardINDEX* of 0.8481 is obtained. These high performances are mainly due to the combination of an accurate cell nuclei detection/classification and adaptive techniques (spatial clustering and active contours). The proposed method needs an image acquired with at least 20x magnification otherwise the neural network fails the nuclear texture classification due to low resolution. As a consequence, the algorithm is not able to correctly discriminate between neoplastic epithelium and stromal response for images acquired with a magnification lower than 20x. The proposed technique shows excellent performance in images with large extension of tumor areas (Figure 4.17, patient 1-3) while the algorithm accuracy slightly decreases with a lobular-structure tumor (Figure 4.17, patient 2-4) due to the imposed parameters (i.e. settings of the spatial clustering and active contours). The proposed method exhibits excellent performance in images of invasive non-special type carcinoma, but future studies are required to test the accuracy of the algorithm for tumor segmentation in other types of cancer like tubular and lobular breast carcinoma.

Thanks to the accurate and fast nuclei segmentation of the MANA algorithm, fully-automated systems for markers quantification [10], histological lesions evaluation [65] and tumor patterns recognition [68] can be easily developed. In addition, automated systems for tumor characterization in whole-slide tissue can be easily developed thanks to the speed and robustness of the proposed methods.

4.5 Conclusion

The visual inspection of histological images is often subjective and time-consuming, so reliable automatic algorithms could help the pathologists in their everyday work.

In this Chapter, three methods are presented as fully-automated solutions for the quantitative analysis of histopathological images. The three developed algorithms are designed to: i) perform cell nuclei detection on different organs and magnification (MANA algorithm); ii) perform the Gleason Group of a histological prostate specimen; iii) perform the neoplastic epithelium segmentation in H&E stained images of breast tissue. The proposed algorithms are compared with manual operators and high performances are obtained for each image of the dataset. The observed reliability and robustness of these techniques are mainly

due to the use of the object-based detection of the ARCO algorithm and an optimized and adaptive workflow.

Starting from the proposed algorithms, our research group is currently working on a system for the discrimination of cancerous areas from non-cancerous areas on entire biopsies of prostate and breast tissue. In addition, we are also working on a MANA-based algorithm for the automatic detection of cellular structures in other organs (e.g. steatosis and necrosis in liver tissue, glomeruli and tubules in kidney tissue).

References

- [1] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, and P. A. Heng, "DCAN: Deep contour-aware networks for object instance segmentation from histology images," *Med. Image Anal.*, vol. 36, pp. 135–146, 2017.
- [2] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *J. Pathol. Inform.*, vol. 7, 2016.
- [3] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological Image Analysis: A Review," *Biomed. Eng. IEEE Rev.*, vol. 2, pp. 147–171, 2009.
- [4] A. H Fischer, K. A Jacobson, J. Rose, and R. Zeller, "Hematoxylin and Eosin Staining of Tissue and Cell Sections," *CSH Protoc.*, vol. 2008, p. pdb.prot4986, 2008.
- [5] S. Doyle, M. D. Feldman, N. Shih, J. Tomaszewski, and A. Madabhushi, "Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer," *BMC Bioinformatics*, vol. 13, no. 1, p. 282, 2012.
- [6] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," *Anal. Quant. Cytol. Histol.*, vol. 23, no. 4, pp. 291–299, 2001.
- [7] M. Macenko *et al.*, "A method for normalizing histology slides for quantitative analysis," *Proc. - 2009 IEEE Int. Symp. Biomed. Imaging From Nano to Macro, ISBI 2009*, pp. 1107–1110, 2009.
- [8] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved Automatic Detection and Segmentation of Cell Nuclei in Histopathology Images," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 841–852, Apr. 2010.
- [9] S. Ali and A. Madabhushi, "An Integrated Region-, Boundary-, Shape-Based Active Contour for Multiple Object Overlap Resolution in Histological Imagery," *IEEE Trans. Med. Imaging*, vol. 31, no. 7, pp. 1448–1460, 2012.
- [10] S. Ram and J. J. Rodriguez, "Size-Invariant Detection of Cell Nuclei in Microscopy Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 7, pp. 1753–1764, Jul. 2016.
- [11] Y. Zheng, D. Liu, B. Georgescu, H. Nguyen, and D. Comaniciu, "3D Deep Learning for Efficient and Robust Landmark Detection in Volumetric Data," in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I*, N. Navab, J. Hornegger, W. M. Wells, and A. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 565–572.
- [12] H. Chen *et al.*, "Automatic Localization and Identification of Vertebrae in Spine CT via a Joint Learning Model with Deep Neural Networks," in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I*, N. Navab, J. Hornegger, W. M. Wells, and A. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 515–522.
- [13] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, and A. Sethi, "A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology," *IEEE Trans. Med. Imaging*, vol. 36, no. 7, pp.

- 1550–1560, 2017.
- [14] M. Veta, J. P. W. Pluim, P. J. Van Diest, and M. A. Viergever, “Breast cancer histopathology image analysis: A review,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1400–1411, 2014.
- [15] J. Xu *et al.*, “Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 119–130, 2016.
- [16] A. Madabhushi, “Digital pathology image analysis: opportunities and challenges,” 2009.
- [17] M. J. Donovan *et al.*, “Development and validation of a novel automated Gleason grade and molecular profile that define a highly predictive prostate cancer progression algorithm-based test,” *Prostate Cancer Prostatic Dis.*, p. 1, 2018.
- [18] A. Basavanthally, S. Agner, G. Alexe, G. Bhanot, S. Ganesan, and A. Madabhushi, “Manifold learning with graph-based features for identifying extent of lymphocytic infiltration from high grade, her2+ breast cancer histology,” *Image Anal. Appl. Biol.(in Conjunction MICCAI)*, New York [Online]. Available <http://www.miaab.org/miaab-2008-papers/27-miaab-2008-paper-21.pdf>, 2008.
- [19] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, and N. Navab, “Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1313–1321, 2016.
- [20] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2017,” *CA. Cancer J. Clin.*, vol. 67, no. 1, pp. 7–30, 2017.
- [21] J. I. Epstein, L. Egevad, M. B. Amin, B. Delahunt, J. R. Srigley, and P. A. Humphrey, “The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma,” *Am. J. Surg. Pathol.*, p. 1, Oct. 2015.
- [22] S. G. Veloso, M. F. Lima, P. G. Salles, C. K. Berenstein, J. D. Scalon, and E. A. Bambirra, “Interobserver agreement of gleason score and modified gleason score in needle biopsy and in surgical specimen of prostate cancer,” *Int. braz j urol*, vol. 33, no. 5, pp. 639–651, Oct. 2007.
- [23] K. Jafari-Khouzani and H. Soltanian-Zadeh, “Multiwavelet grading of pathological images of prostate,” *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 697–704, Jun. 2003.
- [24] H.-J. Yoon, C.-C. Li, C. Christudass, R. Veltri, J. I. Epstein, and Z. Zhang, “Cardinal Multiridgelet-based Prostate Cancer Histological Image Classification for Gleason Grading,” in *2011 IEEE International Conference on Bioinformatics and Biomedicine*, 2011, pp. 315–320.
- [25] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi, “Gland Segmentation and Computerized Gleason Grading of Prostate Histology by Integrating Low-, High-level and Domain Specific Information,” in *MIAAB Workshop*, 2007, pp. 1–8.
- [26] E. A. Rakha, M. E. El-Sayed, A. R. Green, A. H. S. Lee, J. F. Robertson, and I. O. Ellis, “Prognostic markers in triple-negative breast cancer,” *Cancer*, vol. 109, no. 1, pp. 25–32, 2007.
- [27] C. Denkert *et al.*, “Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer,” *J. Clin. Oncol.*, vol. 28, no. 1, pp. 105–113, 2009.
- [28] S. Adams *et al.*, “Prognostic value of tumor-infiltrating lymphocytes in

- triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199,” *J. Clin. Oncol.*, vol. 32, no. 27, p. 2959, 2014.
- [29] S. E. Pinder, E. Provenzano, H. Earl, and I. O. Ellis, “Laboratory handling and histology reporting of breast specimens from patients who have received neoadjuvant chemotherapy,” *Histopathology*, vol. 50, no. 4, pp. 409–417, 2007.
- [30] J. Wu, C. Liang, M. Chen, and W. Su, “Association between tumor-stroma ratio and prognosis in solid tumor patients: a systematic review and meta-analysis,” *Oncotarget*, vol. 7, no. 42, p. 68954, 2016.
- [31] J.-R. Dalle, W. K. Leow, D. Racoceanu, A. E. Tutac, and T. C. Putti, “Automatic breast cancer grading of histopathological images,” in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008, pp. 3052–3055.
- [32] S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman, and J. Tomaszewski, “Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology,” in *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, 2008, pp. 284–287.
- [33] A. Phinyomark, S. Jitreee, P. Phukpattaranont, and P. Boonyapiphat, “Texture analysis of breast cancer cells in microscopic images using critical exponent analysis method,” *Procedia Eng.*, vol. 32, pp. 232–238, 2012.
- [34] T. Wan, J. Cao, J. Chen, and Z. Qin, “Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features,” *Neurocomputing*, vol. 229, pp. 34–44, 2017.
- [35] M. Karabatak and M. C. Ince, “An expert system for detection of breast cancer based on association rules and neural network,” *Expert Syst. Appl.*, vol. 36, no. 2, pp. 3465–3469, 2009.
- [36] J. Cheng and J. C. Rajapakse, “Segmentation of Clustered Nuclei With Shape Markers and Marking Function,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 741–748, 2009.
- [37] H. Xu, C. Lu, and M. Mandal, “An Efficient Technique for Nuclei Segmentation Based on Ellipse Descriptor Analysis and Improved Seed Detection Algorithm,” *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 5, pp. 1729–1741, Sep. 2014.
- [38] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2003.
- [39] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, Jul. 2009.
- [40] C. D. Malon and E. Cosatto, “Classification of mitotic figures with convolutional neural networks and seeded blob features,” *J. Pathol. Inform.*, vol. 4, p. 9, 2013.
- [41] H. Hui Kong, M. Gurcan, and K. Belkacem-Boussaid, “Partitioning Histopathological Images: An Integrated Framework for Supervised Color-Texture Segmentation and Cell Splitting,” *IEEE Trans. Med. Imaging*, vol. 30, no. 9, pp. 1661–1677, Sep. 2011.
- [42] R. Bala and R. Eschbach, “Spatial color-to-grayscale transform preserving chrominance edge information,” in *Color and Imaging Conference, 2004*, vol. 2004, no. 1, pp. 82–86.
- [43] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE*

- Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.
- [44] M. Salvi and F. Molinari, “Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images,” *Biomed. Eng. Online*, vol. 17, no. 1, 2018.
- [45] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 2007, pp. 1027–1035.
- [46] J. Portilla and E. P. Simoncelli, “A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients,” *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 49–70, 2000.
- [47] F. Molinari, C. Caresio, U. R. Acharya, M. R. K. Mookiah, and M. A. Minetto, “Advances in quantitative muscle ultrasonography using texture analysis of ultrasound images,” *Ultrasound Med. Biol.*, vol. 41, no. 9, pp. 2520–2532, 2015.
- [48] S. Wright, “Correlation and causation,” *J. Agric. Res.*, vol. 20, no. 7, pp. 557–585, 1921.
- [49] J. J. Moré, “The Levenberg-Marquardt algorithm: implementation and theory,” in *Numerical analysis*, Springer, 1978, pp. 105–116.
- [50] A. Chekkoury *et al.*, “Automated malignancy detection in breast histopathological images,” in *Medical Imaging 2012: Computer-Aided Diagnosis*, 2012, vol. 8315, p. 831515.
- [51] K. M. Kumar and R. M. Reddy, “A fast DBSCAN clustering algorithm by accelerating neighbor searching using Groups method,” *Pattern Recognit.*, vol. 58, pp. 39–48, Oct. 2016.
- [52] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, 2001.
- [53] D. Peng, B. Merriman, S. Osher, H. Zhao, and M. Kang, “Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulation,” *J. Comput. Phys.*, vol. 155, pp. 410–438, 1999.
- [54] D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Commun. pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.
- [55] J. A. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, vol. 3. Cambridge university press, 1999.
- [56] S. J. Orfanidis, “Introduction to Signal Processing, Prentice Hall,” *Englewood Cliffs, NJ*, 1996.
- [57] R. Real and J. M. Vargas, “The Probabilistic Basis of Jaccard’s Index of Similarity,” *Syst. Biol.*, vol. 45, no. 3, p. 380, Sep. 1996.
- [58] V. Wiesmann, D. Franz, C. Held, C. Münzenmayer, R. Palmisano, and T. Wittenberg, “Review of free software tools for image analysis of fluorescence cell micrographs,” *J. Microsc.*, vol. 257, no. 1, pp. 39–53, Jan. 2015.
- [59] A. E. Carpenter *et al.*, “CellProfiler: image analysis software for identifying and quantifying cell phenotypes,” *Genome Biol.*, vol. 7, no. 10, p. R100, Oct. 2006.
- [60] P. Bankhead *et al.*, “QuPath: Open source software for digital pathology image analysis,” *Sci. Rep.*, vol. 7, no. 1, p. 16878, Dec. 2017.
- [61] J. Schindelin *et al.*, “Fiji: an open-source platform for biological-image analysis,” *Nat. Methods*, vol. 9, no. 7, pp. 676–682, Jun. 2012.

- [62] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, “Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential,” *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 97–114, 2014.
- [63] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. J. Snead, I. A. Cree, and N. M. Rajpoot, “Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1196–1206, 2016.
- [64] S. Wienert *et al.*, “Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach,” *Sci. Rep.*, vol. 2, p. 503, 2012.
- [65] W. Wang, J. A. Ozolek, and G. K. Rohde, “Detection and classification of thyroid follicular lesions based on nuclear structure from histopathology images,” *Cytom. Part A J. Int. Soc. Adv. Cytom.*, vol. 77, no. 5, pp. 485–494, 2010.
- [66] C. Lu, M. Mahmood, N. Jha, and M. Mandal, “A robust automatic nuclei segmentation technique for quantitative histopathological image analysis,” *Anal. Quant. Cytol. Histol.*, vol. 34, pp. 296–308, 2012.
- [67] Y. Peng, Y. Jiang, L. Eisengart, M. A. Healy, F. H. Straus, and X. J. Yang, “Segmentation of prostatic glands in histology images,” in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, 2011, pp. 2091–2094.
- [68] K. Nguyen, A. K. Jain, and B. Sabata, “Prostate cancer detection: Fusion of cytological and textural features,” *J. Pathol. Inform.*, vol. 2, 2011.

Chapter 5

Architectural Analysis in Skeletal Muscle Ultrasound

Part of this chapter has been published as:

C. Caresio, M. Salvi, F. Molinari, K.M. Meiburger and M.A. Minetto, **Fully Automated Muscle Ultrasound Analysis (MUSA): Robust and Accurate Muscle Thickness Measurement.**, *Ultrasound Med. Biol.*, 2017; 43: 195–205.

5.1 Introduction

Thanks to ultrasound (US) imaging, a non-invasive and reliable investigation of skeletal muscles in both healthy and pathological conditions can be easily performed. In the last decades, musculoskeletal ultrasonography has been extensively used to characterize muscles both morphologically [1], [2], and architecturally [3], [4], [5].

The appearance of skeletal muscles in US is clearly distinctive and can be easily discriminated through visual inspection [6]. Skeletal muscles are identified by the superficial and deep aponeuroses, two boundary highly reflective connective tissue. Muscle ultrasound scans are often collected in the transverse and longitudinal plane through the placement of the US probe in correspondence of the maximum muscle diameter or the muscle region of interest.

In particular, the measurement of muscle thickness has been widely used in clinical studies to investigate several pathological conditions like disuse atrophy [7], ageing (e.g. primary sarcopenia) [8], [9], [10] and muscle hypertrophy [11], [12]. Furthermore, previous studies adopted the muscle thickness as a measure to predict the total body fat-free mass [13], the leg skeletal muscle mass [14], [15], as well as to indirectly evaluate the cross-sectional area [16], [17] and volume [18], [19], [20] of superficial muscles.

In transverse plane, musculoskeletal ultrasound imaging enables to quantify muscle characteristics (e.g. size and echogenicity) relative to the maximal cross-sectional area (CSA). The CSA represents a fundamental architectural parameter directly correlated with to the maximum force generated by the muscle [21], [22]. Previous studies have shown the feasibility of the quantitative assessment of cross-sectional area for different skeletal muscles [2], [23], [24], [25], [26], in both static and dynamic conditions. These studies demonstrated that the CSA quantification is a reliable tool for the assessment of muscle size during musculoskeletal rehabilitation [27], [28] and training [29] as well as to neuromuscular disorders [30].

In most of the previously cited works, the muscle parameters extraction (CSA and thickness) from ultrasound images is manually performed, which is a time-consuming task and is also prone to errors. Several semi-automatic and fully-automated approaches have been proposed in the last decays to extract quantitative data on muscle ultrasound images. In 2010, a muscle boundary tracking algorithm to measure the thickness of the pectoralis major muscle was presented [31]. Few years later, a sequential programming approach was proposed to measure the abdominal muscles thickness [32]. In another work, a semi-automatic approach was implemented to detect the aponeuroses and estimate the gastrocnemius muscle thickness as the distance between the deep and superficial aponeurosis [33].

Although the previous cited studies were applicable in the study of muscle ultrasound architecture, they were all designed and optimized for a specific muscle, making them of difficult application in a clinical scenario. In addition, all these methods were developed for the extraction of the parameters from

longitudinal scans and most of them are manually driven and require the user input to work properly. Furthermore, an automatic strategy for the quantitative evaluation of the muscle in transverse scans has never been proposed in literature. To the best of our knowledge, no previous work presented a fully automated algorithm for muscle thickness measurement and CSA segmentation for different superficial skeletal muscles.

In this Chapter, two robust and novel algorithms are presented to characterize the skeletal muscle geometry and architecture in both longitudinal and transverse planes. The first method, named MUSA (Muscle UltraSound Analysis), is an automatic technique designed for the muscle thickness measurement of different skeletal muscles acquired in longitudinal plane. The second method, named TRAMA (TRAnverse Muscle ultrasound Analysis), is a completely automatic algorithm for the CSA segmentation in transverse ultrasound images taken from four different skeletal muscles of the lower limb.

5.2 Materials and Methods

5.2.1 Longitudinal Muscle Ultrasound Analysis: MUSA algorithm

Image database

A total of 200 images were acquired from 50 subjects (25 males and 25 females). All the subjects were free from neuromuscular or skeletal impairments. For each subject, four skeletal muscles were investigated during the ultrasound session: medial gastrocnemius, tibialis anterior, vastus lateralis and rectus femoris. The dominant side was acquired in the longitudinal plan for each of the four muscles and a total of 200 images (4 muscles x 50 subjects) were collected and analyzed. All the subjects signed an informed consent. The entire study was approved by the local ethics committee and was conducted at "Città della Salute" Hospital – Molinette (Torino, Italy).

A detailed explanation of the ultrasound device settings and acquisition protocol is described in Appendix C and D.

MUSA algorithm architecture

The algorithm is designed to automatically identify the aponeuroses in muscle ultrasound images acquired in the longitudinal plane. Once the aponeuroses are correctly located, the proposed method estimates the muscle thickness by measuring the distance between the superficial and deep aponeurosis. The MUSA algorithm is developed in MATLAB environment and image processing is carried out on a workstation with a 2.8 GHz exa-core CPU and 64 GB of RAM. The procedure of the MUSA algorithm is schematically described in Figure 5.1. Three main steps compose the processing: search for candidate aponeuroses, search for muscle fascicles and selection of the actual aponeuroses. In the next paragraphs, an exhaustive description of the proposed method is provided.

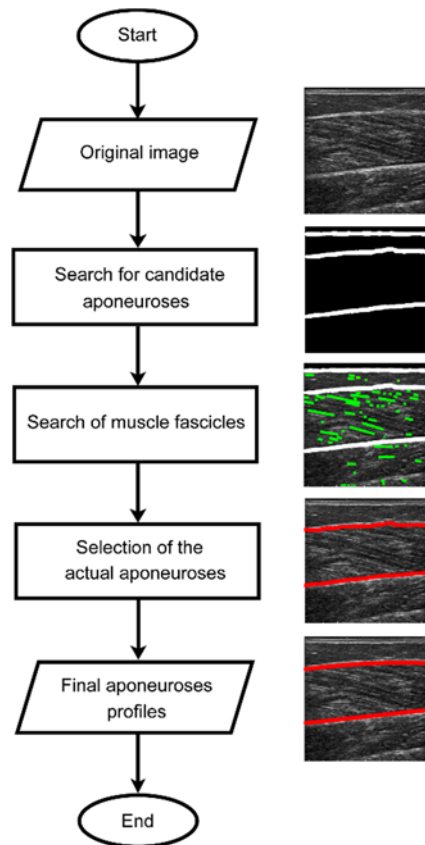


Fig. 5.1 Schematic representation of the MUSA algorithm. Starting from the original US image, the proposed method finds all the candidate muscle aponeuroses. Then, after the fascicle detection, the MUSA algorithm detects the two actual aponeuroses and perform the muscle thickness measurement.

Search for candidate aponeuroses

The first step of the MUSA algorithm is the automatic search of all the possible positions of the muscle aponeuroses within the image. Starting from the original image (Figure 5.2a), the vertical Sobel gradient is computed to enhance all the aponeuroses-like structures (Figure 5.2b). Then, a vertical FODG filter (First-Order Derivative Gaussian) is applied to the processed image. This filter is obtained by the convolution between a simple derivative filter and a Gaussian kernel [34]. The dimension of the structures enhanced by the FODG is directly proportional to its kernel size. For this reason, a kernel dimension of 1 mm is selected to enhance all the aponeurosis-like structures with a size equal to or bigger than 1 mm. The output of the FODG filter (Figure 5.2c) is thresholded using the ARCO algorithm hence obtaining the FODG binary mask shown in Figure 5.2d. The input parameters of the ARCO algorithm are listed in Table 5.1.

As can be seen in Figure 5.2d, the FODG binary mask contains the two muscle aponeuroses as well as other structures with a thickness comparable to an aponeurosis. A heuristic cleaning step is then performed to delete all the objects that cannot be qualified as candidate aponeuroses. In order to discard non-aponeurosis structures, each object within the FODG binary mask is approximated to an ellipsis and its eccentricity and its major axis length is evaluated. The

eccentricity is defined as the ratio of the distance between the foci of the ellipse and its major axis length (an ellipse whose eccentricity is 1 is a line segment, while an ellipse whose eccentricity is 0 is actually a circle). All the regions with a major axis length lower than 60% of the image column width or with an eccentricity lower than 0.995 are deleted (Figure 5.2e).

Table 5.1 Input parameters of ARCO for candidate aponeuroses detection in the MUSA algorithm.

Parameters	Value
<i>image LAYER</i>	Grayscale image
<i>object TYPE</i>	Object of interest is bright on a dark background
<i>polynomial ORDER</i>	PWM _{CURVE} is fitted with a 7 th order polynomial function
<i>object PROPERTY</i>	Minimum standard deviation of the segmented objects is imposed as condition to find the optimal threshold among candidate ones.

Finally, branch removal is carried out to smooth the morphology of the remaining regions. In Figure 5.2f, the white arrow illustrates the correction made by the branch removal with respect to the irregularity of the superficial aponeurosis of Figure 5.2e.

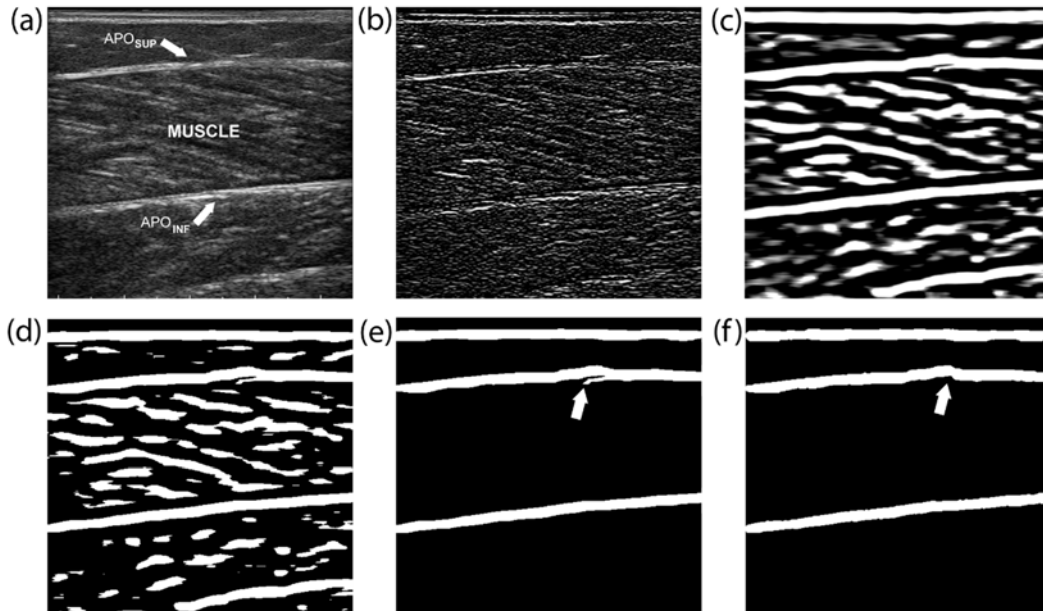


Fig. 5.2 First steps of the MUSA algorithm. (a) Original muscle ultrasound image of a medial gastrocnemius. (b) Vertical Sobel gradient applied to the original image. (c) FODG output image. (d) FODG binary mask obtained with the ARCO algorithm. (e) FODG mask after the removal of no-aponeuroses structures. (f) Binary mask after the branch removal step. In panels (e) and (f), the arrow illustrates the effect of branch removal on the FODG binary mask. APO_{INF} = deep aponeurosis, APO_{SUP} = superficial aponeurosis.

Search for muscle fascicles

After the branch removal of the previous section (Figure 5.2f), the FODG binary mask often contains more than two candidate aponeuroses. For this reason, the presence of muscle fascicles is checked between each pair of candidate aponeuroses to detect the two actual muscle aponeuroses among all the candidate ones. In fact, a muscle region always contains fascicles in longitudinal US scans. Firstly, the original image (Figure 5.2a) is equalized and then Otsu thresholding [35] is applied to obtain a fascicles binary mask (Figure 5.3a). This mask contains the muscle fascicles and aponeuroses within the image. All the candidate aponeuroses are removed from the binary fascicles mask (Figure 5.3b), together with all the structures with an area lower than 1 mm^2 (which are likely to represent over-segmentation due to speckles).

In the longitudinal plane, muscle fascicles are represented by a line with a specific orientation. For this reason, a line detection based on Hough transform [36] is computed on the fascicles binary mask. The variable θ is the angle that the line forms with the origin, measured in degrees anticlockwise from the positive x-axis ($0^\circ \leq \theta \leq 180^\circ$). For this application, all the detected lines with an orientation outside the range $90^\circ \leq \theta \leq 180^\circ$ are deleted. Figure 5.3c illustrates the detected fascicles (green line) overlaid with the original ultrasound image.

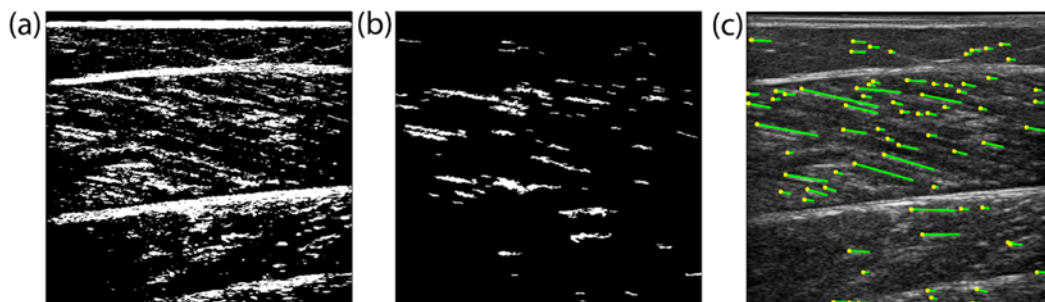


Fig. 5.3 Processing steps for fascicle detection. (a) Fascicles binary mask obtained with histogram equalization and Otsu thresholding. (b) Fascicles mask without aponeurosis candidates. (c) Line detection overlaid on the original image: fascicles endpoints are portrayed in yellow while muscle fascicles are shown in green.

Selection of the actual aponeuroses

This step is devoted to the detection of the actual muscle aponeuroses in longitudinal plane. First of all, the final fascicles mask is merged with FODG mask (Figure 5.4a). At this stage, we do not need an accurate profile of the muscle aponeuroses, but we just need to select the correct aponeuroses among the candidate ones. A heuristic search is carried out to identify the two muscle aponeuroses. This heuristic search is applied in the central region of the image (between 20% and 80% of the image width) since, in this region, the muscle geometry is reproduced without distortions (the ultrasound beams are perfectly

perpendicular to the muscle aponeuroses). Seven columns spaced by 10% are defined to search the two muscle aponeuroses. The number of search columns is set experimentally to guarantee the best performance coupled with the lowest computational cost. The heuristic search follows three steps:

1. the proposed method detects all the intersection points between the aponeurosis candidates and each search column. Columns with only one intersection point are discarded from the search. Each intersection point is labelled as $Y_{m,n}$ (n identifies the search column while m denotes the corresponding structure intersected). The MUSA algorithm also calculates the distance between consecutive intersection points on the same search column. If two points are closer than 15 mm, then those points are deleted. The minimal vertical distance is set to 15 mm based on the minimum thickness of the four investigated muscles [37]. If no couple of points have a distance higher than 15 mm, the entire search column is discarded;
2. if only one pair of intersection points is detected for a single search column, then the uppermost point represents the superficial aponeurosis (APO_{SUP}) while the lowest point indicates the deep aponeurosis (APO_{INF});
3. if more than one pair of intersection points are detected for a column, then the proposed technique selects the ROI (region of interest) with the highest number of muscle fascicles: for each couple of points, a ROI with width equal to the 10% of the entire image width and height equal to the distance between the two points is defined (Figure 5.4b). The Hough transform emphasize muscle fascicles within the image; therefore, in order to define the APO_{SUP} and APO_{INF} , the number of fascicles within each ROI of the analyzed column is calculated. The APO_{SUP} and APO_{INF} coordinates are defined as the pair of points of the ROI with the highest number of muscle fascicles.

The result of this heuristic search is a sequence of points that delineate the position of the superficial aponeurosis APO_{SUP} and the deep aponeurosis APO_{INF} . The output of the heuristic search is illustrated in Figure 5.4c. The region within the FODG binary mask mostly recognized as APO_{INF} is labelled as the final deep aponeurosis; similarly, for the superficial aponeurosis (Figure 5.4d). After the heuristic search, the two profiles are roughly placed on the aponeuroses (Figure 5.4d), therefore a refinement is required to obtain more accurate muscle profiles. A DoG (Difference of Gaussians) filter is then applied to detect the actual aponeurosis edges. For an input image I , the DoG filtered image is defined as:

$$DoG_{IMAGE} = (I \cdot G_1) - (I \cdot G_2) \quad (\text{Eq. 5.1})$$

where G_1 and G_2 are the kernes of two low-pass Gaussian filters. For this application, the size of each kernel (N) and the corresponding standard deviation σ is set as follow: $N_1=9$, $\sigma_1=101$, $N_2=83$ and $\sigma_2=21$. The DoG_{IMAGE} is shown in Figure 5.4e. In the DoG_{IMAGE} , the transitions from dark to bright (i.e. from muscle, which is dark, to APO_{INF} , which is bright) are negative, whereas the transitions

from bright to dark (i.e. from APO_{SUP} to muscle) are positive. In this way, it is possible to locate the exact position of the interface between aponeurosis and muscle for each column of the image.

Finally, the aponeuroses profiles are interpolated by a bicubic spline as illustrated in Figure 5.4f.

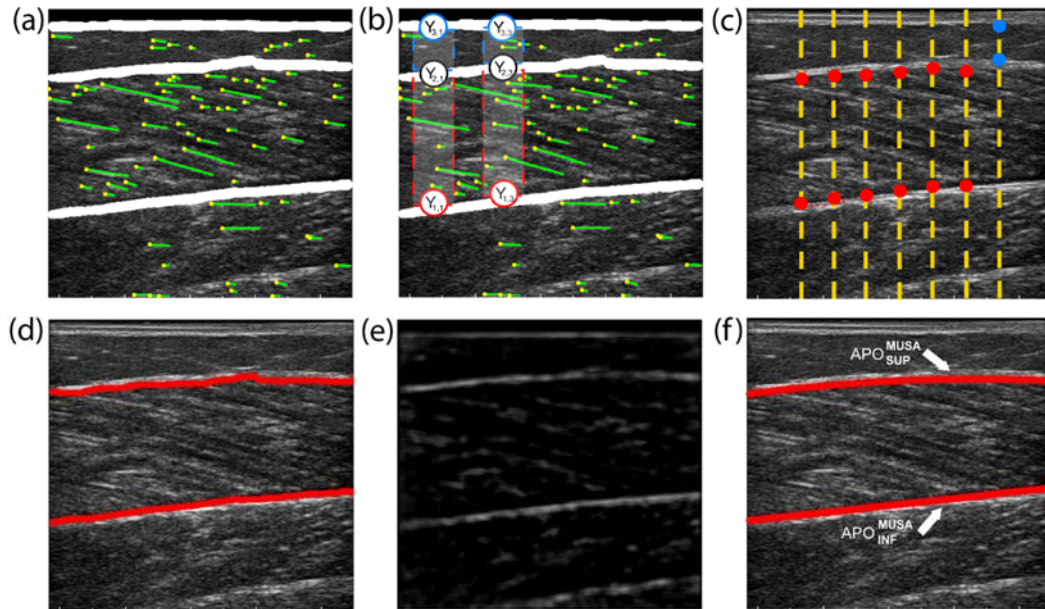


Fig. 5.4 MUSA heuristic process for the selection of actual aponeuroses. (a) Fascicles and FODG masks overlaid on the original image. (b) Sketch of the heuristic search where two ROI (regions of interest) are illustrated in dashed rectangles. (c) APO_{INF} and APO_{SUP} identified for each search column. (d) Profiles obtained at the end of the heuristic process. (e) Application of the DoG filter on the original image. (f) Final interpolated profiles of the deep and superficial aponeuroses.

Tibialis anterior images

The tibialis anterior is a bipennate muscle and it is characterized by a unique architecture. The fibers of this muscle branch out at a specific angle (pennation angle), from the central intramuscular fascia to the deep and superficial aponeuroses [38]. As can be seen in Figure 5.5a, the muscle fascicles of the two compartments exhibit opposite pennation angles. The MUSA processing can be applied and repeated for each muscle compartment. In this way, the proposed method is able to process ultrasound images containing multi-compartmental muscles. For the tibialis anterior muscle, the proposed method performs the following three steps:

1. starting from the FODG binary mask (Figure 5.5b), MUSA finds the lower compartment (between central fascia and deep aponeurosis). Among all candidates, the ones that correspond to the central fascia and the deepest aponeurosis are traced (Figure 5.5c) by applying the same strategy as described for mono-compartmental muscles;

2. the central fascia also defines the bottom limit of muscle upper compartment. As shown in Figure 5.5d, the MUSA algorithm is then applied from the central fascia upwards in order to identify the upper compartment (between central fascia and superficial aponeurosis);
3. once the superficial and deep aponeuroses are correctly located (Figure 5.5e), the final muscle profiles are obtained by applying the DoG filter (Figure 5.5f).

Since pennation angles are apposite in the two compartments of the tibialis anterior, the range of θ for the Hough transform is set to $0^\circ \leq \theta \leq 90^\circ$ for the lower compartment and $90^\circ \leq \theta \leq 180^\circ$ for the upper compartment.

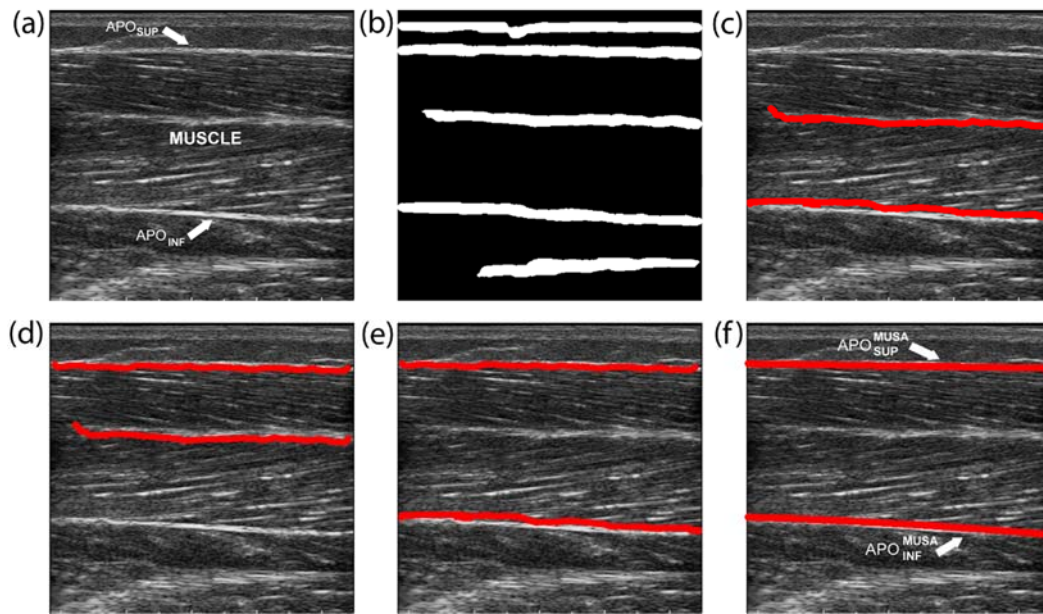


Fig. 5.5 Processing steps for a tibialis anterior muscle. (a) Original image of a representative tibialis anterior. (b) Binary FODG mask of the image. (c) Automatic profiles of the lower compartment (central fascia and deep aponeurosis). (d) Automatic profiles of the upper compartment (superficial aponeurosis and central fascia). (e) APO_{SUP} and APO_{INF} profiles obtained at the end of the heuristic process. (f) Final profiles of the deep (APO_{INF}^{MUSA}) and superficial (APO_{SUP}^{MUSA}) deep aponeuroses.

Performance indicators

The parameters tuning of the MUSA algorithm is reported in Appendix B. To estimate the muscle thickness, the MUSA algorithm implements the centerline distance metric. This metric has been widely used for thickness measurements in ultrasound images [39], [40]. The first step of the centerline distance metric is the determination of the centerline between the two profiles. Then, a chord perpendicular to each point of the centerline is plotted (the chord length represents the measure of muscle thickness at that point). The final thickness measurement is computed as the mean distance of all the chords along the centerline.

The measure of muscle thickness provided by MUSA is compared with the manual measures performed by three experienced operators. The manual operators drew five-line segments that connects the two muscle aponeuroses at

around 10%, 30%, 50%, 70% and 90% of the entire length of the muscle profile in the image. The Euclidean distance is calculated between each segment's endpoints and the final manual measurement is obtained as the average of these five distances. The absolute error between the mean of the three manual measurements and the automatic muscle thickness is evaluated in terms of mean and standard deviation.

Finally, automatic results and the mean of the three manual measurements are compared using the Kruskal-Wallis analysis of variance and the Bland-Altman plots [41]. A correlation analysis between the differences and averages of the two measurement methods (MUSA vs mean of manual operators) is also performed using the Spearman test to evaluate if there is any statistical dependence between the two datasets (manual vs automatic). All statistical tests are carried out with a significance level (p) of 0.05.

5.2.2 Transverse Muscle Ultrasound Analysis: TRAMA Algorithm

Image database

A total of 200 images were acquired from 50 subjects (25 males and 25 females). Twenty subjects were healthy, while the remaining 30 subjects were affected by different hormonal disorders (growth hormone deficiency: $n=5$; type 2 diabetes: $n=5$; transgender individuals under hormonal treatment: $n=5$; acromegaly: $n=5$; obesity: $n=5$) possibly affecting muscle size and/or structure. For each subject and patients, the same four skeletal muscles of the MUSA algorithm were investigated (medial gastrocnemius, tibialis anterior, vastus lateralis and rectus femoris). The dominant side was acquired in the longitudinal plan for each of the four muscles and a total of 200 images (4 muscles x 50 subjects) were collected and analyzed. All the subjects signed an informed consent. The entire study was approved by the local ethics committee and was conducted at "Città della Salute" Hospital – Molinette (Torino, Italy).

A detailed explanation of the ultrasound device settings and acquisition protocol is described in Appendix C and D.

TRAMA algorithm architecture

The TRAMA algorithm is designed to detect and segment the muscle cross-sectional area (CSA). This method is developed in MATLAB environment and image processing is carried out on a workstation with a 2.8 GHz exa-core CPU and 64-GB of RAM. The processing of TRAMA is sketched in Figure 5.6. The proposed technique has been built following a similar architecture already proposed in the MUSA algorithm [42]. In particular, a similar approach is adopted during the search of aponeurosis candidates and the final muscle border refinement. Three main steps compose the processing: search for candidate aponeuroses, search of muscle connective tissue and selection of the actual aponeuroses. In the next paragraphs, an exhaustive description of the proposed method is provided.

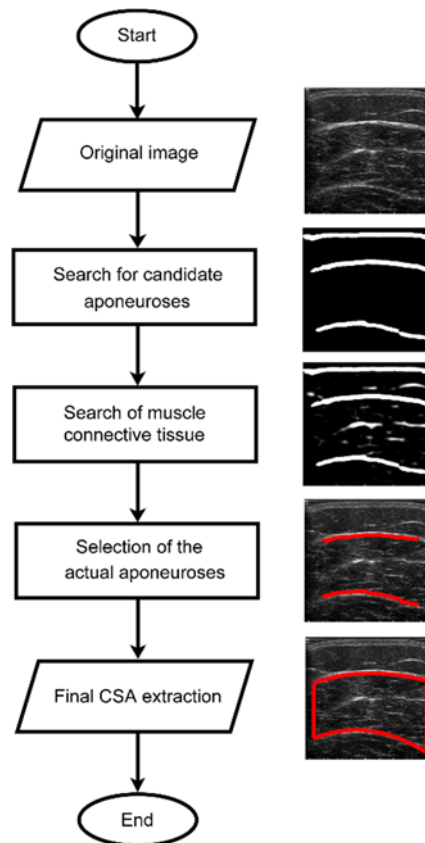


Fig. 5.6 Schematic representation of the TRAMA algorithm. Starting from the original US image, the proposed method finds all the candidate muscle aponeuroses. Then, after the connective tissue detection, the TRAMA algorithm detect the two actual aponeuroses and extract the cross-sectional area (CSA).

Search for candidate aponeuroses

In order to find the aponeurosis candidates, the TRAMA algorithm follows the same steps of the MUSA algorithm. Briefly, the vertical Sobel gradient is applied to the original image (Figure 5.7a) to enhance all the aponeuroses-like structures (Figure 5.7b). The resulting image is filtered with a FODG filter (Figure 5.7c) and then thresholded using the ARCO algorithm (Figure 5.7d). In order to discard no-aponeurosis structures, each object with a major axis length shorter than 60% of the image column width or with an eccentricity lower than 0.995 is deleted (Figure 5.7e). Finally, branch removal is carried out to smooth and fix potential irregularities on the muscle aponeuroses. Figure 5.7f shows the result obtained where only aponeurosis-like objects are preserved.

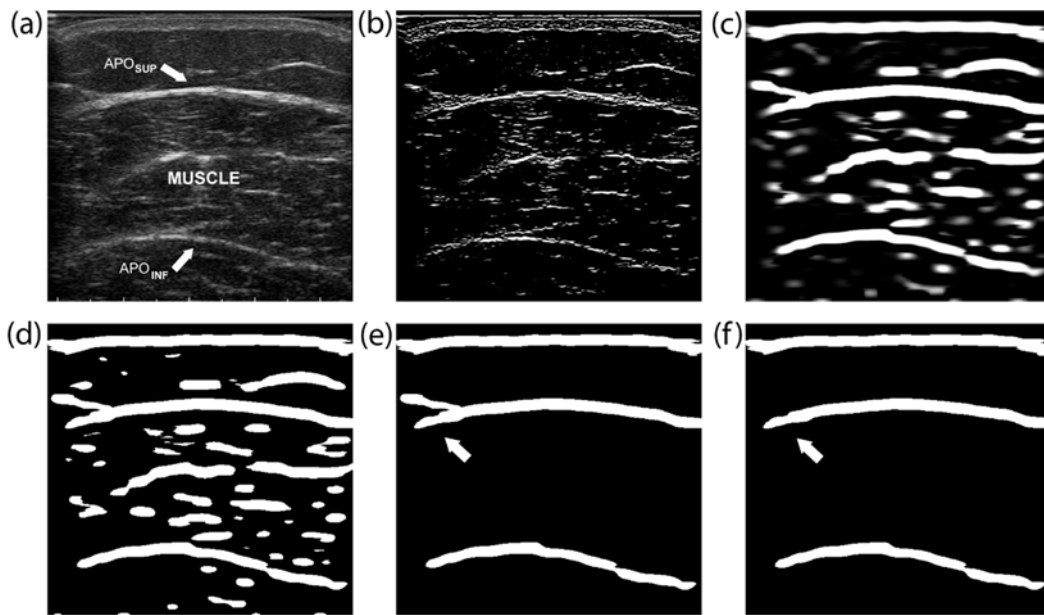


Fig. 5.7 First steps of the TRAMA algorithm. (a) Original muscle ultrasound image of a vastus lateralis. (b) Vertical Sobel gradient applied to the original image. (c) FODG output image. (d) Binary FODG mask obtained with the ARCO algorithm. (e) FODG mask after the removal of no-aponeuroses structures. (f) Binary mask after the branch removal step. In panels (e) and (f), the arrow illustrates the effect of branch removal on the FODG binary mask. APO_{INF} = deep aponeurosis, APO_{SUP} = superficial aponeurosis.

Search of muscle connective tissue

The presence of connective tissue is checked between each pair of candidate aponeuroses to detect the two actual muscle aponeuroses among all the candidate ones. A Frangi filter [43] is applied to the original image (Figure 5.8a) for its intrinsic versatility in detecting objects with different shapes and sizes within an image. This multiscale filter is based on the local second-order structure of the image: second-order partial derivatives are calculated by convolution of the image with the second-order derivative of Gaussian kernels of different sizes. For this application, three values of sigma are defined (6, 8 and 10 pixels) since they represent the sizes of the muscle connective tissue in transverse scans. Then, the second derivative Gaussian kernels are iteratively computed according to the chosen sigma (σ) values as a $3\sigma \times 3\sigma$ grid. The resulting image after the application of the Frangi filter is defined as follows:

$$V_{\sigma} = \begin{cases} 0 & \text{if } \lambda_2 > 0 \\ \left(1 - e\left(-\frac{S^2}{2c^2}\right)\right) e\left(\frac{R_b^2}{2\beta^2}\right) & \text{in all the other cases} \end{cases} \quad (\text{Eq. 5.2})$$

where β and c are sensitivity parameters, λ_1 and λ_2 are the eigenvalues of the Hessian matrix and S and R_b measure the second-order structureness and the blob-like structure respectively. In our case, β is set to 0.5 and c is equal to 15. The pixel values of the final filtered image (Figure 5.8b) are the maximum pixel values among all the output images generated using the three sigma values.

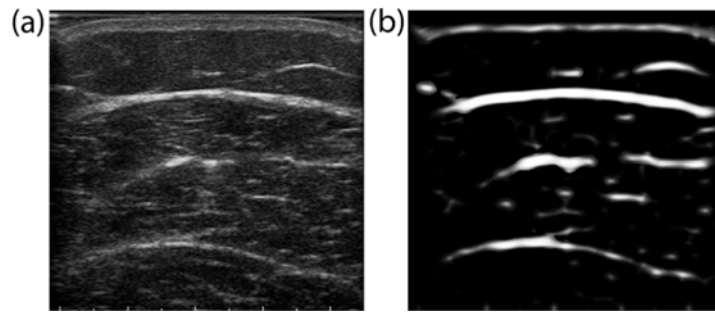


Fig. 5.8 Detection of muscle connective tissue. (a) Original image. (b) Connective tissue structures highlight using Frangi filter.

Selection of the actual aponeuroses

This step is devoted to the detection of the actual muscle aponeuroses in transverse plane. First of all, the FODG binary mask is merged with the output of the Frangi filter (Figure 5.9a) and seven columns spaced by 10% are defined to search the two muscle aponeuroses. A heuristic search is carried out to reduce the number of candidate aponeuroses. This heuristic search is applied in the central region of the image (between 20% and 80% of the image width). The heuristic search follows three steps:

1. the proposed method detects all the intersection points between the aponeurosis candidates and each search column. Columns with only one intersection point are discarded from the search. Each intersection point is labelled as $Y_{m,n}$ (n identifies the search column while m denotes the corresponding structure intersected). The TRAMA algorithm also calculates the distance between consecutive intersections points on the same search column. If two points are closer than 15 mm, then those points are deleted (15 mm is the minimum thickness for the four investigated muscles);
2. if only one pair of intersection points is detected for a single search column, then the uppermost point represents the superficial aponeurosis (APO_{SUP}) while the lowest point indicates the deep aponeurosis (APO_{INF});
3. if more than one pair of intersection points are detected for a column, then the proposed technique selects the ROI (region of interest) with more presence of connective tissue: for each couple of points, a ROI with width equal to the 10% of the entire image width and height equal to the distance between the two points is defined (Figure 5.9b). This operation is possible thanks to the Frangi filter that enhance the connective tissue regions within the image. Therefore, in order to define the APO_{SUP} and APO_{INF} , the mean gray-scale intensity of the output of Frangi filter within each ROI is calculated. The APO_{SUP} and APO_{INF} coordinates are defined as the pair of points of the ROI with the highest grayscale value.

The results of this heuristic search is shown in Figure 5.9c. The region within the FODG binary mask mostly recognized as APO_{INF} is labelled as the final deep aponeurosis; similarly, for the superficial aponeurosis (Figure 5.9d).

Figure 5.9e shows the interpolation of the aponeurosis profiles. The interpolation is performed only on the central 90% of the image width to exclude all the region with a poorer skin-probe contact due to the curvature of the lower limbs. Finally, the endpoints of the deep and superficial aponeurosis are connected together through two vertical lines (Figure 5.9f).

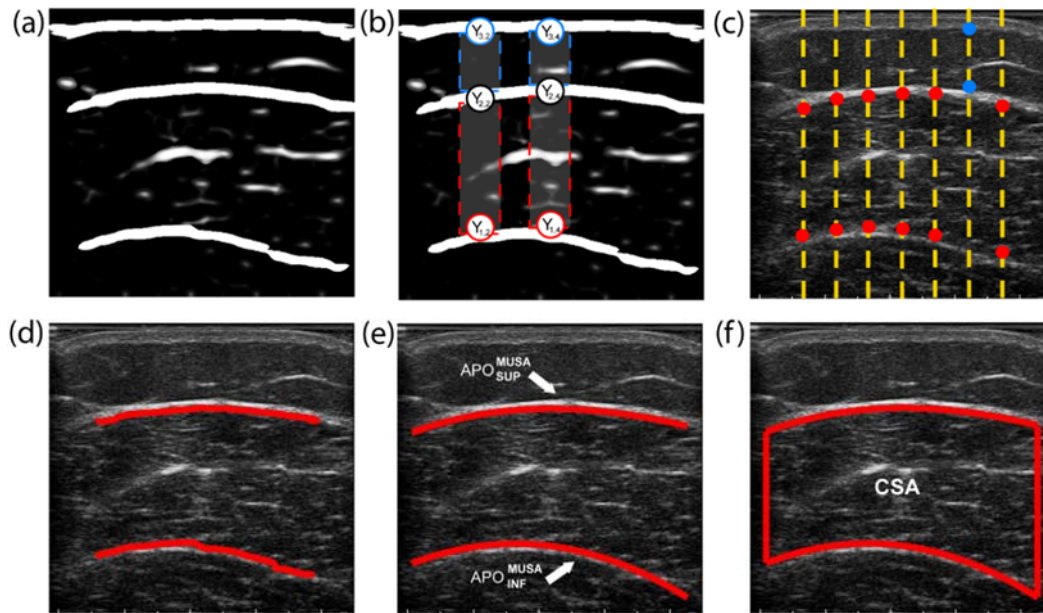


Fig. 5.9 TRAMA heuristic process for the selection of actual aponeuroses. (a) FODG masks overlaid on the multiscale output image. (b) Sketch of the heuristic search where two ROI (regions of interest) are illustrated in dashed rectangles. (c) APO_{INF} and APO_{SUP} identified for each search column. (d) Profiles obtained at the end of the heuristic process. (e) Interpolation of the deep and superficial aponeurosis profiles. (f) Muscle cross-sectional area (CSA).

Tibialis anterior images

Since the tibialis anterior muscle is a bipennate muscle, the cross-sectional area detection is slightly different due to its specific circular crown shape (Figure 5.10a). The detection of deep and superficial aponeuroses follows the same approach adopted for the other muscles (Figure 5.10b). Then, Frangi filter is applied to enhance the tibia-bone interface (Figure 5.10c). The filtered image is thresholded using Otsu thresholding and the same cleaning structures step is performed. This new image presents a L-shaped structure where the horizontal profile is the deep aponeurosis while the vertical profile represents the tibia-bone interface (Figure 5.10d). These two profiles are interpolated and prolonged until their intersection (Figure 5.10e, point H). The automatic cross-sectional area is obtained connecting the superficial aponeurosis endpoints with the deep aponeurosis and the lateral interface profile, as shown in Figure 5.10f.

Performance indicators

The parameters tuning of the TRAMA algorithm is reported in Appendix B. The measure of CSA provided by TRAMA is compared with the manual measures performed by two experienced operators. The absolute error between the mean of the two manual measurements and the automatic CSA measurement is evaluated in terms of mean and standard deviation. A Kruskal-Wallis analysis of variance is used to perform a statistical comparison between automatic and manual measurements. To evaluate the accuracy of the automatic segmentation, SDC (Sorensen-Dice Coefficient) and HD (Hausdorff Distance) are also evaluated:

$$SDC = \frac{2 |A \cap B|}{|A| + |B|} \quad (\text{Eq. 5.3})$$

$$HD_{(a \in A, b \in B)} = \max\{d(a, b)\} \quad (\text{Eq. 5.4})$$

where A is the manual CSA shape drawn by the operator, B is the CSA identified by the algorithm and $d(a, b)$ is the distance between point a (on border of shape A) and point b (on border of shape B). SDC measures similarity between two different shapes, defined as twice the size of their intersection divided by the sum of the two shapes areas [44]. HD measures how far two borders are from each other. Two contours are close in the Hausdorff distance if every point of either set is close to some point of the other set [45]. All statistical tests were carried out with a significance level (p) of 0.05.

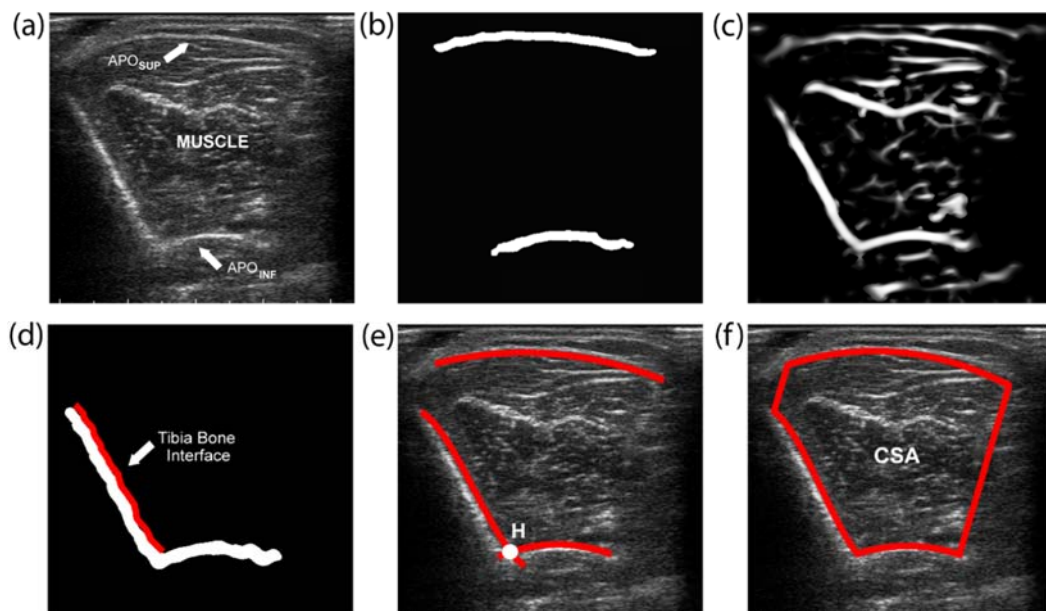


Fig. 5.10 Processing steps for a tibialis anterior muscle. (a) Original image of a representative tibialis anterior. (b) Superficial and deep aponeurosis profiles obtained at the end of the heuristic process. (c) Frangi filter output image. (d) Identification of the Tibia Bone Interface (red). (e) Intersection between deep aponeurosis and Tibia Bone Interface (point H). (f) Final CSA (cross-sectional area) segmentation.

5.3 Results

5.3.1 Segmentation results of MUSA Algorithm

Segmentation results

For each of the four muscles investigated, Figure 5.11 illustrates a visual comparison between manual thickness measurement (right panel) and the corresponding result provided by the proposed method (left panel). The MUSA algorithm achieved a 100% segmentation success rate as it is able to compute the muscle thickness in the entire image dataset.

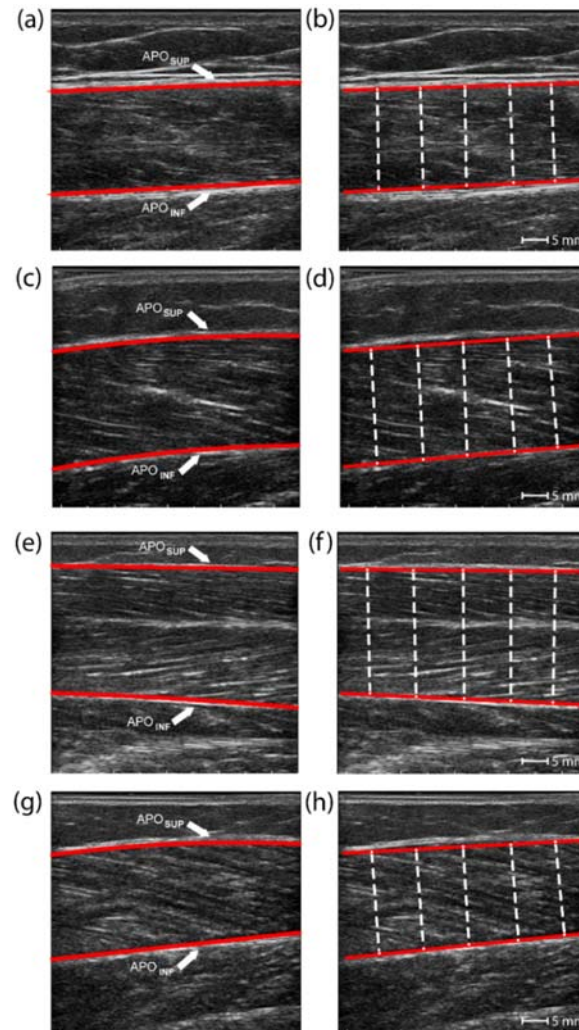


Fig. 5.11 Comparison between automatic (left column) and manual (right column) thickness measurement for rectus femoris (a,b), vastus lateralis (c,d), tibialis anterior (e,f) and medial gastrocnemius (g,h).

Performance evaluation

Table 5.2 reports mean and standard deviation of the errors between muscle thickness provided by MUSA algorithm and the mean of the three manual measurements. The error is computed as the difference between the automatic thickness and the mean of the three manual measurements while absolute error is defined as the absolute value of the error. The difference between the mean of

manual measurements and the automatic value is in the range of 0.06–0.45 mm. Analysis of the group data (Table 5.2) shows comparable values obtained by automatic and manual measurements. In fact, no significant differences are observed between the mean of the three operators and MUSA for any muscle ($p = 0.63$ for rectus femoris, $p = 0.70$ for vastus lateralis, $p = 0.85$ for tibialis anterior, and $p = 0.23$ for medial gastrocnemius).

Table 5.2 Manual and automatic muscle thickness measurements for each muscle (mean \pm standard deviation). RF = rectus femoris, VL = vastus lateralis, TA = tibialis anterior, MG = medial gastrocnemius.

Muscle	Operator 1 (mm)	Operator 2 (mm)	Operator 3 (mm)	Mean of Operators (mm)	MUSA (mm)	Error (mm)	Absolute Error (mm)
RF	22.3 \pm 3.8	22.4 \pm 3.7	21.7 \pm 3.7	22.1 \pm 3.7	21.8 \pm 3.8	0.43 \pm 0.78	0.49 \pm 0.80
VL	21.9 \pm 4.0	22.0 \pm 4.0	21.4 \pm 4.0	21.7 \pm 4.0	21.3 \pm 4.0	0.45 \pm 0.56	0.51 \pm 0.58
TA	28.2 \pm 3.6	28.0 \pm 3.6	27.6 \pm 3.5	27.9 \pm 3.5	27.9 \pm 3.7	0.06 \pm 0.70	0.47 \pm 0.56
MG	19.9 \pm 3.1	20.0 \pm 3.0	19.3 \pm 3.1	19.7 \pm 3.1	19.4 \pm 3.1	0.30 \pm 0.45	0.48 \pm 0.35

The Bland-Altman plots in Figure 5.12 show that the mean difference between the manual and automatic measurements is in the range of 0.06–0.45 mm. Moreover, most of the differences are within the 95% limits of agreement, which indicates that the two measurement methods (manual and MUSA) are interchangeable. For the four muscles investigated, the Spearman test shows no significant correlation between the means and the differences of the two measurements, thus suggesting that the accuracy of MUSA is not correlated with the muscle thickness value (medial gastrocnemius: $R = -0.07$, $p = 0.64$; vastus lateralis: $R = -0.01$, $p = 0.95$; tibialis anterior: $R = -0.14$, $p = 0.33$ and rectus femoris: $R = -0.09$, $p = 0.61$).

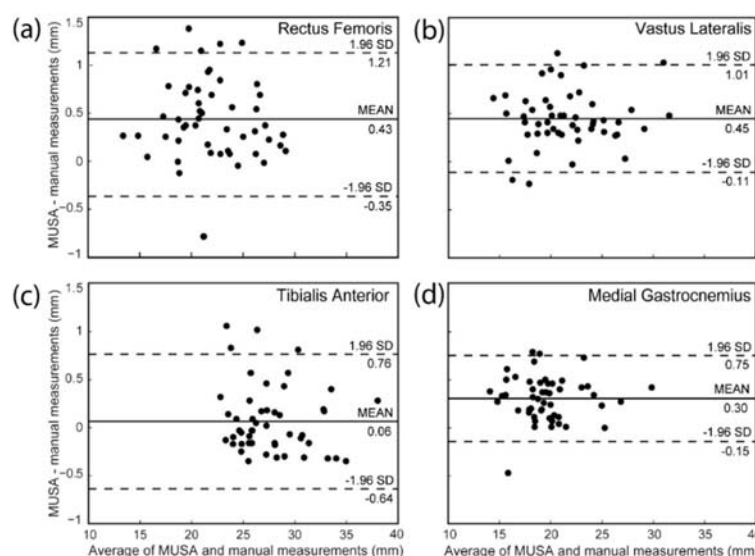


Fig. 5.12 Bland-Altman plots for between the mean of the three manual operators and MUSA algorithm. (a) Rectus femoris. (b) Vastus lateralis. (c) Tibialis anterior. (d) Medial gastrocnemius.

5.3.2 Segmentation results of TRAMA Algorithm

Segmentation results

For each of the four muscles investigated, Figure 5.13 illustrates a visual comparison between manual CSA tracing (right panel) and the one obtained with TRAMA (left panel). The TRAMA algorithm achieved a 100% segmentation success rate as it is able to detect the muscle CSA in the entire image dataset.

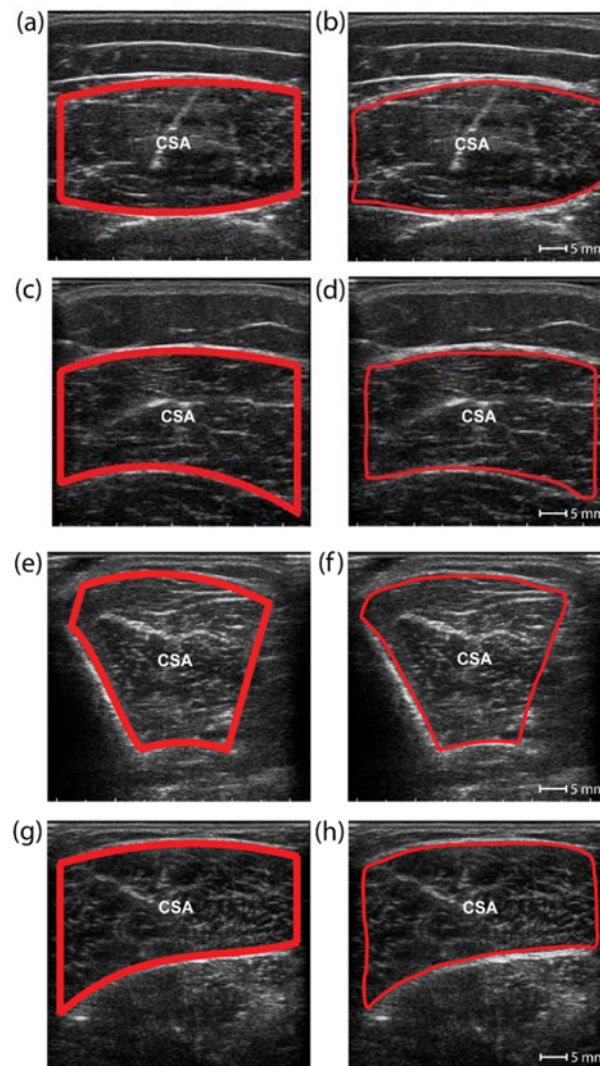


Fig. 5.13 Comparison between automatic (left column) and manual (right column) cross-sectional area (CSA) identification. (a,b) rectus femoris. (c,d) vastus lateralis. (e,f) tibialis anterior. (g,h) medial gastrocnemius.

Performance evaluation

Table 5.3 reports mean and standard deviation of the errors between manual CSA measurements obtained by the two operators and automatic CSA detection. The error is computed as the difference between the mean of the two manual CSAs and the automatic CSA measurements while absolute error is the absolute value of the error. Analysis of the group data (Table 5.3) shows comparable

values (no statistical differences) between automatic and manual measurements ($p = 0.32$ for rectus femoris, $p = 0.34$ for vastus lateralis, $p = 0.14$ for tibialis anterior, and $p = 0.10$ for medial gastrocnemius).

Table 5.3 CSA measurements for each muscle (mean \pm standard deviation) obtained by each operator and TRAMA. RF = rectus femoris, VL = vastus lateralis, TA = tibialis anterior, MG = medial gastrocnemius.

Muscle	Operator 1 (mm ²)	Operator 2 (mm ²)	Mean of Operators (mm ²)	TRAMA (mm ²)	Error (mm ²)	Absolute Error (mm ²)
RF	833.3 \pm 160.0	807.6 \pm 152.0	820.5 \pm 155.7	851.1 \pm 159.6	31.1 \pm 17.5	37.1 \pm 17.4
VL	876.2 \pm 149.1	883.3 \pm 137.1	879.8 \pm 142.1	906.7 \pm 146.7	26.9 \pm 19.4	27.8 \pm 18.1
TA	797.0 \pm 145.2	792.1 \pm 153.0	794.6 \pm 148.1	892.3 \pm 152.7	34.8 \pm 23.6	35.4 \pm 22.6
MG	782.4 \pm 128.5	747.1 \pm 125.4	764.8 \pm 126.0	806.4 \pm 128.6	41.7 \pm 17.4	41.7 \pm 17.4

The mean and standard deviation values of *SDC* and *HD* are reported in Table 5.4. Comparison between the two operators and TRAMA shows average *SDC* in the range of 94-97% and mean *HD* values from 1.7 to 3.5 mm, which are comparable to the average values of the inter-operator variability (*HD* from 1.7 to 2.6 mm and *SDC* in the range of 95-96% and). These results indicates that the two measurement methods (manual and TRAMA) are interchangeable. In order to provide a more clinical context of the TRAMA performances, we divided our dataset into three different sets (normal-BMI vs increased-BMI subjects, shallow vs. deep images and normal vs. patients). No significant difference in the *SDC* values is found between the groups in all three cases, further underlying the accuracy and versatility of the developed technique.

Table 5.4 Sorensen-dice coefficient (SDC) and Hausdorff distance (HD) (mean \pm standard deviation) computed to compare the two operators (Operator 1 and Operator 2) with TRAMA for each of the four muscles. RF = rectus femoris, VL = vastus lateralis, TA = tibialis anterior, MG = medial gastrocnemius.

Muscle	Operator 1 vs Operator 2		TRAMA vs Operator 1		TRAMA vs Operator 2	
	SDC (%)	HD (mm)	SDC (%)	HD (mm)	SDC (%)	HD (mm)
RF	96.7 \pm 0.9	1.7 \pm 0.7	96.1 \pm 1.2	2.3 \pm 0.7	97.2 \pm 0.8	1.8 \pm 0.8
VL	95.8 \pm 1.4	2.3 \pm 1.1	95.6 \pm 1.4	2.8 \pm 1.3	96.9 \pm 1.0	1.9 \pm 0.8
TA	95.8 \pm 1.8	2.4 \pm 1.2	94.7 \pm 1.8	3.5 \pm 1.3	96.1 \pm 1.6	2.4 \pm 1.0
MG	95.3 \pm 1.9	2.6 \pm 1.5	94.6 \pm 1.7	3.3 \pm 1.3	96.9 \pm 0.7	2.3 \pm 0.9

5.4 Discussion

The extraction of the muscle architectural parameters is used to investigate the muscle characteristics (size, echogenicity and texture) in both physiological [12] and pathological conditions [30]. The variability in depth, direction, and shape of muscle aponeuroses makes the automatic segmentation of muscle contours a challenging task. This variability is caused by several physiological and anatomical factors, like the thickness of the subcutaneous adipose layer, the health status and the age of the subject, and the contraction status and the anatomy of the analyzed muscle. In this Chapter, two fully automated algorithms are proposed for quantitative analysis of skeletal muscle ultrasound images.

The first proposed technique, named MUSA, is able to recognize the muscle aponeuroses and measure the muscle thickness in longitudinal scans. The comparison between automatic and manual measurements shows a mean difference below 0.5 mm (worst case for vastus lateralis: 0.45 mm). Since the average muscle thickness for the entire dataset is around 23 mm, the mean measurement error of the MUSA algorithm is around 2.5% of the nominal thickness value. A mean error equal to 2.5% can be considered negligible from a clinical point of view as a percentage reduction higher than 5-10% is needed to diagnose muscle pathological conditions such as the atrophy and hypotrophy [37].

The second method presented, named TRAMA, is an algorithm for the measurement of the cross-sectional area in musculoskeletal ultrasound images acquired in the transverse plane. The average absolute error between the manual and automatic measures is always below 45 mm² (around 4% of the nominal value), also confirmed by the excellent results of the Dice coefficient (DSC). Also in this case, the AE can be considered irrelevant as a CSA variation of more than 10% allows the correct identification of clinically relevant changes [37], [46].

The excellent performance of the MUSA and TRAMA is mainly due to the implementation of multiscale and scale filters like Frangi and FODG operators. These filters are able to selectively enhance all the regions/structures of interest with low computational cost. In fact, the parameters of FODG filter are optimized to enhance all the aponeurosis-like regions while the Frangi filter is capable of amplifying the anatomic size of the connective tissue. Both of the proposed techniques do not require any user interaction and they are able to correctly detect both the superficial and deep aponeuroses (100% segmentation success rate). Four skeletal muscles of the lower limb (tibialis anterior, vastus lateralis, medial gastrocnemius and rectus femoris) are investigated, due to the fact that they are the most informative muscles for detecting sarcopenia and neuromuscular disorders [15], [37], [30].

Additional studies are required to test the accuracy of MUSA and TRAMA in the extraction of muscle architectural parameters for other superficial muscles of the upper (e.g. biceps brachii) or lower limb (e.g. vastus medialis). From a technical point of view, MUSA and TRAMA can be applied to detect the aponeurosis of any muscle as long as it is represented in a transverse or longitudinal projection. Being totally automated, these methods could also be used

in future works to initialize muscle architectural parameters measurements in dynamic conditions.

5.5 Conclusion

Ultrasound imaging allows a non-invasive and reliable measurement of the skeletal muscle architecture. Nowadays, computer-aided approaches in musculoskeletal ultrasound imaging are semi-automatic or specifically optimized for a single muscle. In this Chapter, two methods (MUSA and TRAMA) are presented as first fully-automated systems capable for processing US images coming from different skeletal muscles of the lower limb. Statistical analysis shows that the manual and automatic measurements can be used interchangeably.

In the future, the MUSA and TRAMA algorithms could be integrated into ultrasound devices as fully automated tools for real-time acquisition session. In addition, we plan to use these algorithms as starting points to detect the muscle ROIs and automatize a multi-texture analysis to study the alteration of muscle architecture in physiological and pathological conditions [5], [47], [48] and its changes related to physiological and pathological condition. The proposed methods provide an excellent tool for the quantitative extraction of morphological parameters from muscle ultrasound images and, in the next future, they can be used as the first step in more complex texture-analysis studies.

References

- [1] P. Fish, *Physics and instrumentation of diagnostic medical ultrasound*. John Wiley & Sons Incorporated, 1990.
- [2] N. D. Reeves, C. N. Maganaris, and M. V. Narici, “Ultrasonographic assessment of human skeletal muscle size,” *Eur. J. Appl. Physiol.*, vol. 91, no. 1, pp. 116–118, Jan. 2004.
- [3] C. Caresio, F. Molinari, G. Emanuel, and M. A. Minetto, “Muscle echo intensity: reliability and conditioning factors,” *Clin. Physiol. Funct. Imaging*, vol. 35, no. 5, pp. 393–403, Sep. 2015.
- [4] A. Pretorius and J. L. Keating, “Validity of real time ultrasound for measuring skeletal muscle size,” *Phys. Ther. Rev.*, vol. 13, no. 6, pp. 415–426, Dec. 2008.
- [5] K. M. Meiburger, U. R. Acharya, and F. Molinari, “Automated localization and segmentation techniques for B-mode ultrasound images: A review,” *Comput. Biol. Med.*, vol. 92, pp. 210–235, Jan. 2018.
- [6] S. Pillen, “Skeletal muscle ultrasound,” *Eur. J. Transl. Myol.*, vol. 20, no. 4, pp. 145–156, 2010.
- [7] M. D. de Boer *et al.*, “Effect of 5 weeks horizontal bed rest on human muscle thickness and architecture of weight bearing and non-weight bearing muscles,” in *European Journal of Applied Physiology*, 2008.
- [8] S. Agyapong-Badu, M. Warner, D. Samuel, M. Narici, C. Cooper, and M. Stokes, “Anterior thigh composition measured using ultrasound imaging to quantify relative thickness of muscle and non-contractile tissue: A potential biomarker for musculoskeletal health,” *Physiol. Meas.*, 2014.
- [9] R. A. Atkinson *et al.*, “Effects of testosterone on skeletal muscle architecture in intermediate-frail and frail elderly men,” *Journals Gerontol. Ser. A Biomed. Sci. Med. Sci.*, vol. 65, no. 11, pp. 1215–1219, 2010.
- [10] M. V. Narici, C. N. Maganaris, N. D. Reeves, and P. Capodaglio, “Effect of aging on human muscle architecture,” *J. Appl. Physiol.*, vol. 95, no. 6, pp. 2229–2234, 2003.
- [11] M. V. Narici and N. Maffulli, “Sarcopenia: characteristics, mechanisms and functional significance,” *Br. Med. Bull.*, vol. 95, no. 1, pp. 139–159, Sep. 2010.
- [12] O. R. Seynnes *et al.*, “Effect of androgenic-anabolic steroids and heavy strength training on patellar tendon morphological and mechanical properties,” *J. Appl. Physiol.*, 2013.
- [13] Y. Takai *et al.*, “Applicability of ultrasound muscle thickness measurements for predicting fat-free mass in elderly population,” *J. Nutr. Health Aging*, vol. 18, no. 6, pp. 579–585, 2014.
- [14] Y. Takai *et al.*, “Validity of ultrasound muscle thickness measurements for predicting leg skeletal muscle mass in healthy Japanese middle-aged and older individuals,” *J. Physiol. Anthropol.*, 2013.
- [15] T. Abe, R. S. Thiebaud, J. P. Loenneke, M. Loftin, and T. Fukunaga, “Prevalence of site-specific thigh sarcopenia in Japanese men and women,” *Age (Omaha)*, 2014.
- [16] T. Abe, J. P. Loenneke, and R. S. Thiebaud, “Ultrasound assessment of hamstring muscle size using posterior thigh muscle thickness,” *Clin. Physiol. Funct. Imaging*, 2016.
- [17] Y. Takai, Y. Katsumata, Y. Kawakami, H. Kanehisa, and T. Fukunaga, “Ultrasound method for estimating the cross-sectional area of the psoas

- major muscle.," *Med. Sci. Sports Exerc.*, vol. 43, no. 10, pp. 2000–2004, 2011.
- [18] R. Akagi *et al.*, "Development of an equation to predict muscle volume of elbow flexors for men and women with a wide range of age," *Eur. J. Appl. Physiol.*, 2010.
- [19] J. I. Esformes, M. V. Narici, and C. N. Maganaris, "Measurement of human muscle volume using ultrasonography," *Eur. J. Appl. Physiol.*, 2002.
- [20] M. Ogawa, T. Yasuda, and T. Abe, "Component characteristics of thigh muscle volume in young and older healthy men," *Clin. Physiol. Funct. Imaging*, vol. 32, no. 2, pp. 89–93, 2012.
- [21] K. Hammond *et al.*, "Validity and reliability of rectus femoris ultrasound measurements: Comparison of curved-array and linear-array transducers," *JRRD*, vol. 51, no. 7, 2014.
- [22] J. M. Seymour *et al.*, "Ultrasound measurement of rectus femoris cross-sectional area and the relationship with quadriceps strength in COPD.," *Thorax*, vol. 64, no. 5, pp. 418–23, May 2009.
- [23] R. R. Estes *et al.*, "The Effect of High Intensity Interval Run Training on Cross-sectional Area of the Vastus Lateralis in Untrained College Students.," *Int. J. Exerc. Sci.*, vol. 10, no. 1, pp. 137–145, 2017.
- [24] M. N. M. Blue, A. E. Smith-Ryan, E. T. Trexler, and K. R. Hirsch, "The effects of high intensity interval training on muscle size and quality in overweight and obese adults.," *J. Sci. Med. Sport*, vol. 21, no. 2, pp. 207–212, Feb. 2018.
- [25] M. Noorkoiv, K. Nosaka, and A. J. Blazeovich, "Assessment of quadriceps muscle cross-sectional area by ultrasound extended-field-of-view imaging," *Eur. J. Appl. Physiol.*, vol. 109, no. 4, pp. 631–639, Jul. 2010.
- [26] S. Sipila and H. Suominen, "Ultrasound imaging of the quadriceps muscle in elderly athletes and untrained men," *Muscle Nerve*, vol. 14, no. 6, pp. 527–533, Jun. 1991.
- [27] L. Reider *et al.*, "Evaluating the relationship between muscle and bone modeling response in older adults," *Bone*, vol. 90, pp. 152–158, 2016.
- [28] F. Damas *et al.*, "Early resistance training-induced increases in muscle cross-sectional area are concomitant with edema-induced muscle swelling," *Eur. J. Appl. Physiol.*, vol. 116, no. 1, pp. 49–56, Jan. 2016.
- [29] K. D. Seymore, Z. J. Domire, P. DeVita, P. M. Rider, and A. S. Kulas, "The effect of Nordic hamstring strength training on muscle architecture, stiffness, and strength," *Eur. J. Appl. Physiol.*, vol. 117, no. 5, pp. 943–953, May 2017.
- [30] S. Pillen, I. M. P. Arts, and M. J. Zwarts, "Muscle ultrasound in neuromuscular disorders," *Muscle Nerve*, vol. 37, no. 6, pp. 679–693, Jun. 2008.
- [31] T. K. K. Koo, C. Wong, and Y. Zheng, "Reliability of Sonomyography for Pectoralis Major Thickness Measurement," *J. Manipulative Physiol. Ther.*, vol. 33, no. 5, pp. 386–394, Jun. 2010.
- [32] A. Wong, K. M. Gallagher, and J. P. Callaghan, "Computerised system for measurement of muscle thickness based on ultrasonography," *Comput. Methods Biomech. Biomed. Engin.*, vol. 16, no. 3, pp. 249–255, Mar. 2013.
- [33] Shan Ling, Yongjin Zhou, Ye Chen, Yu-Qian Zhao, Lei Wang, and Yong-Ping Zheng, "Automatic Tracking of Aponeuroses and Estimation of Muscle Thickness in Ultrasonography: A Feasibility Study," *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 6, pp. 1031–1038, Nov. 2013.

- [34] L. M. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, "Scale and the differential structure of images," *Image Vis. Comput.*, vol. 10, no. 6, pp. 376–388, Jul. 1992.
- [35] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man. Cybern.*, 1979.
- [36] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, 1972.
- [37] M. A. Minetto *et al.*, "Ultrasound-Based Detection of Low Muscle Mass for Diagnosis of Sarcopenia in Older Adults," *PM R*, 2016.
- [38] C. N. Maganaris and V. Baltzopoulos, "Predictability of in vivo changes in pennation angle of human tibialis anterior muscle from rest to maximum isometric dorsiflexion," *Eur. J. Appl. Physiol. Occup. Physiol.*, 1999.
- [39] L. Saba *et al.*, "What is the correct distance measurement metric when measuring carotid ultrasound intima-media thickness automatically?," *Int. Angiol. a J. Int. Union Angiol.*, vol. 31, no. 5, pp. 483–489, 2012.
- [40] F. H. Sheehan, E. L. Bolson, H. T. Dodge, D. G. Mathey, J. Schofer, and H. W. Woo, "Advantages and applications of the centerline method for characterizing regional ventricular function," *Circulation*, 1986.
- [41] J. M. Bland and D. G. Altman, "A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement," *Comput. Biol. Med.*, vol. 20, no. 5, pp. 337–340, 1990.
- [42] C. Caresio, M. Salvi, F. Molinari, K. M. Meiburger, and M. A. Minetto, "Fully Automated Muscle Ultrasound Analysis (MUSA): Robust and Accurate Muscle Thickness Measurement," *Ultrasound Med. Biol.*, vol. 43, no. 1, 2017.
- [43] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1998, pp. 130–137.
- [44] K. H. Zou *et al.*, "Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports," *Acad. Radiol.*, vol. 11, no. 2, pp. 178–189, 2004.
- [45] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, 1993.
- [46] M. V Franchi *et al.*, "Muscle thickness correlates to muscle cross-sectional area in the assessment of strength training-induced hypertrophy," *Scand. J. Med. Sci. Sports*, vol. 28, no. 3, pp. 846–853, 2018.
- [47] F. Molinari, C. Caresio, U. R. Acharya, M. R. K. Mookiah, and M. A. Minetto, "Advances in Quantitative Muscle Ultrasonography Using Texture Analysis of Ultrasound Images," *Ultrasound Med. Biol.*, vol. 41, no. 9, pp. 2520–2532, Sep. 2015.
- [48] G. J. R. Dubois, D. Bachasson, L. Lacourpaille, O. Benveniste, and J.-Y. Hogrel, "Local Texture Anisotropy as an Estimate of Muscle Quality in Ultrasound Imaging," *Ultrasound Med. Biol.*, vol. 44, no. 5, pp. 1133–1140, May 2018.

Conclusions and Final Remarks

As stated in the Introduction, medical research and healthcare are continuously seeking for computer-aided schemes that can offer non-invasive, reliable and fast image-based diagnosis. Moreover, looking at technological trends, future diagnostic procedures will be adopted to provide personalized medicine based on quantitative data extracted from patient-specific medical images. The work presented aims to develop quantitative and multimodal imaging techniques for the modeling and interpretation of physiopathological processes.

A new completely automated algorithm for the segmentation of relevant objects within medical images is developed and validated, and it is shown how the proposed technique is able to correctly detect biological structures in different imaging modalities and magnifications. Starting from this method, several fully automated algorithms are proposed for the reliable and quantitative analysis of medical images. The strategies proposed in this thesis are scalable and relevant in a wide range of applications (modeling of heart failure, cancer detection and classification, muscle architectural analysis) and modalities (fluorescent and optical microscopy, ultrasound imaging). The proposed algorithms are validated using expert manual segmentations and benchmarked with semi-automatic techniques, providing very promising results.

In the near future, the developed techniques will be used to process wide datasets, extract additional information from the images and help remove the subjectivity from the diagnosis process. These automated strategies combine the human cognitive procedures of structure recognition with image information revealed through the extraction of quantitative morphological and functional features, with the final aim to significantly improve the patient's diagnosis, prognosis and treatment.

Appendix A

Tuning of the ARCO input parameters

In all the proposed studies, the object-detection provided by ARCO is automatically optimized using the procedure sketched in Figure A.1.

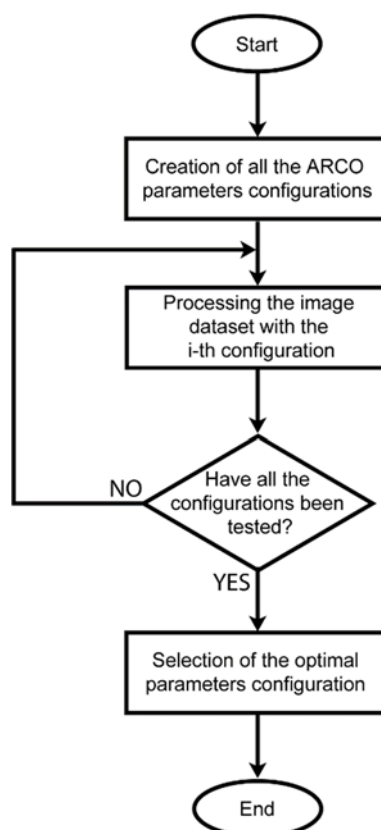


Fig. A.1 Workflow adopted for tuning of ARCO input parameters.

First of all, different combination of the ARCO input parameters (*imageLAYER*, *objectTYPE*, *polynomialORDER*, *objectPROPERTY*) are generated through the *creation of all the ARCO parameters configurations*. For each parameters configuration generated in the previous step, the ARCO algorithm is applied to the entire image dataset (*processing the image dataset with the i-th configuration*). This step is repeated until all the configurations are tested. The optimal parameters configuration of the ARCO algorithm is selected as the one with the highest

object-level $F1_{SCORE}$ (Eq. 2.2). If one or more configurations obtain the same performances, the one with the lower average computational time is chosen.

Figure A.2 shows an example of ultrasound image with different level of noise. For each of the three images, the ARCO algorithm is applied to identify muscle aponeuroses. Three different $polynomial_{ORDER}$ are used (3, 5, and 7) to evaluate the object-detection performed by the proposed method.

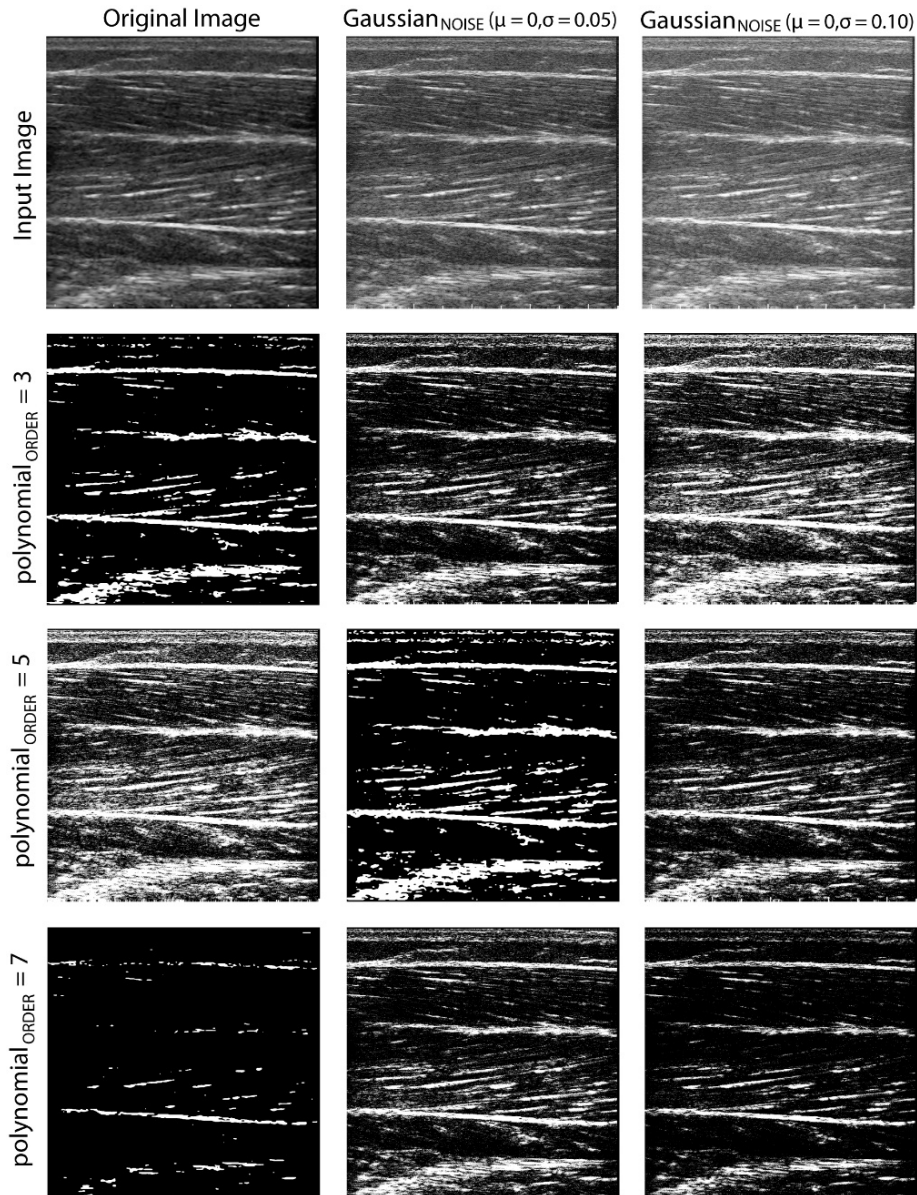


Fig. A.2 ARCO performance with different order of the polynomial function that fit the PWM_{CURVE} ($polynomial_{ORDER}$). Three ultrasound images with different noise-level are analyzed (columns). In the second row, the ARCO segmentation with a $polynomial_{ORDER}$ of 3 is illustrated. The third and last rows show the ARCO object-detection with a $polynomial_{ORDER}$ equal to 5 and 7 respectively. If the noise level increases (last column), a higher $polynomial_{ORDER}$ is required to achieve good segmentation.

The optimal $polynomial_{ORDER}$ for images with low-level noise (i.e. 1st row) is generally low (between 3rd and 5th). The increase in the noise level causes a higher "complexity" of the grayscale histogram and, therefore, a higher order of the polynomial function is required for a good fit of the PWM_{CURVE} . As can be seen in Figure A.2, a $polynomial_{ORDER}$ equal to 5 guarantees the best result for a mid-level noise image (2nd row) while a $polynomial_{ORDER} = 7$ is the best choice for a high-level noise image (3rd row).

The $object_{PROPERTY}$ generally depends on the number of expected objects within the image. The following figures show an example of images containing few (Figure A.3) and many objects (Figure A.4).

Generally, if the image contains few structures (≤ 10), the best performances are obtained by imposing a condition on the standard deviation of the detected objects intensity. In particular, the optimal segmentation is achieved by choosing the candidate threshold that has the objects' intensity with the lowest standard deviation (as the algorithm should identify homogeneous structures). Figure A.3 illustrates the results obtained with three candidate thresholds and the corresponding standard deviation of segmented objects. The same figure also shows how the best object-detection is achieved with the threshold that has the objects' intensity with the lowest standard deviation.

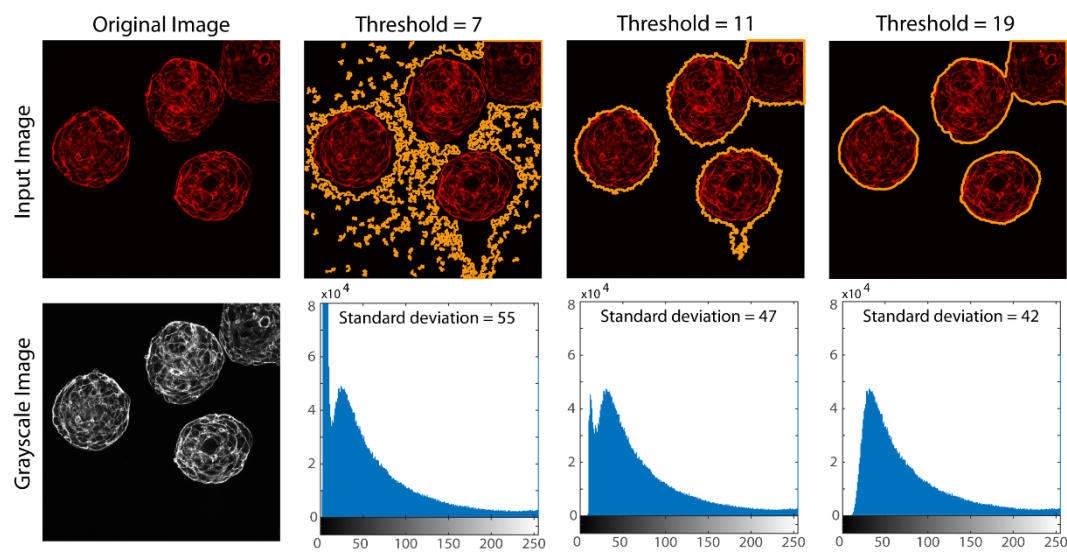


Fig. A.3 ARCO segmentations for a sample image with few expected objects. First column shows the original image and the corresponding grayscale image. Second, third and fourth columns illustrate the object-detection provided by ARCO using three different candidate thresholds. Last row shows the intensity distribution of segmented objects. The optimal segmentation is obtained with the candidate threshold that has the objects' intensity with the lowest standard deviation.

On the other hand, if the expected objects are higher than 10 (e.g. cell nuclei), the best performances are achieved by imposing a condition on the median area of the detected structures. In particular, the optimal segmentation is obtained by choosing the candidate threshold that has the objects with the highest median area. Figure A.4 illustrates the results obtained with three candidate thresholds and the

corresponding area distribution of segmented objects. The same figure also shows how the best object-detection is achieved with the threshold that has the objects with the highest median area.

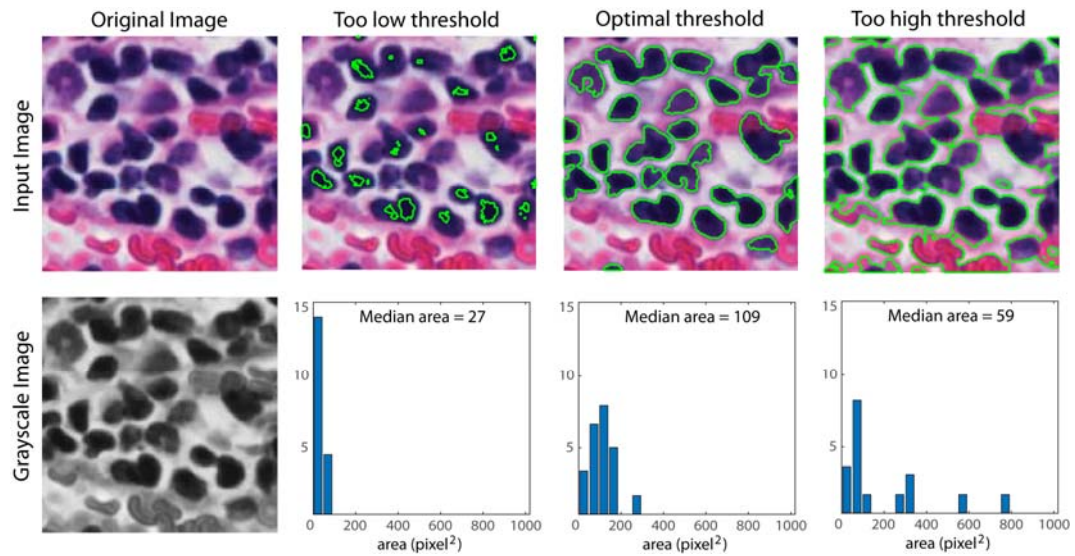


Fig. A.4 ARCO segmentations for a sample image with several expected objects. First column shows the original image and the corresponding grayscale image. Second, third and fourth column illustrate the object-detection provided by ARCO using three different candidate thresholds. Last row shows the area distribution of segmented objects. The optimal segmentation is obtained with the candidate threshold that has the objects with the highest median area.

Appendix B

Automated parameter tuning and algorithm configuration

In all the proposed studies, the algorithms are automatically optimized using the procedure sketched in Figure B.1.

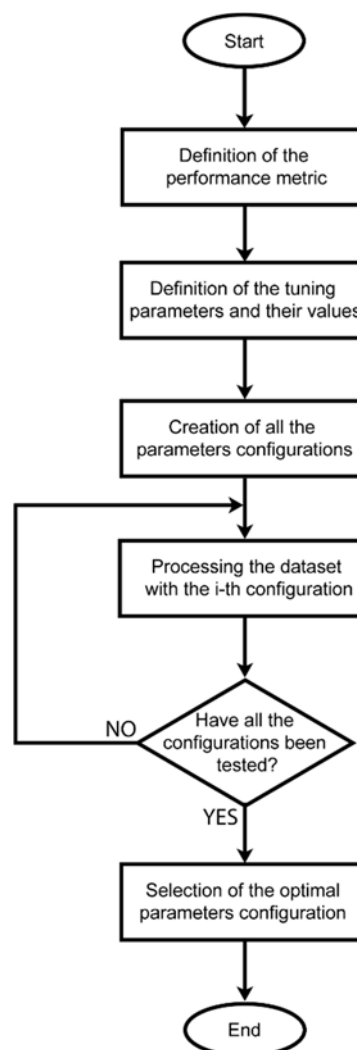


Fig. B.1 Workflow adopted for the algorithm tuning and optimization.

First of all, the performance metric for the algorithm optimization is chosen (*definition of the performance metric*). This metric can be an overlap index (e.g. Jaccard coefficient, $F1_{SCORE}$), a distance metric (e.g. Hausdorff distance) or a similarity index (e.g. Dice coefficient). For each of the algorithm parameter (e.g.

percentages, thresholds, etc.), we define all the reasonable values that the parameter can assume by creating a ‘parameter vector’ (e.g. $\text{parameter}_1 = [1,3,5,7]$). This step is called *definition of the tuning parameters and their values*. Then, all the possible combinations between the elements of the parameters vectors are generated through the *creation of all the configurations*. For each parameters configuration generated in the previous step, the algorithm is applied to the entire image dataset (*processing the dataset with the i-th configuration*). This step is repeated until all the configurations are tested. The optimal parameters configuration is selected as the one that maximizes the performance metric chosen (*selection of the optimal configuration*). If one or more configurations obtain the same performances, the one with the lower average computational time is chosen.

The optimal parameters and the performance metric used for each algorithm proposed in this thesis are listed below.

CARE algorithm (CARDiosphere Evaluation)

Performance metric: jaccard_{INDEX}

Parameter	Description	Possible values	Optimal value
<i>polynomial_{ORDER}</i>	Order of the polynomial function that fit the PWM _{CURVE} (ARCO algorithm)	[3;5;7;9;11;13;15]	7
<i>small_{COMPONENT}</i>	Minimum area of detected objects to be considered as cardiospheres (μm^2)	[600;1200;1800;2400]	1200
<i>overlap_{AREA}</i>	Percentage of overlapping between objects from realign frame and reference frame (%)	[60;65;70;75;80;85]	75
<i>shape_{SOLIDITY}</i>	Minimum solidity of a shape in the identification of the cut-off frame	[0.50;0.60;0.70;0.80]	0.60

MANA algorithm (Multiscale Adaptive Nuclei Analysis)

Performance metric: F1_{SCORE}

Parameter	Description	Possible values	Optimal value
<i>polynomial_{ORDER}</i>	Order of the polynomial function that fit the PWM _{CURVE} (ARCO algorithm)	[3;5;7;9;11;13;15]	5
<i>small_{OBJECTS}</i>	Maximum percentage with respect to the mean objects area for labelling a structure as 'small' (%)	[15;20;25;30;35;40]	25
<i>big_{OBJECTS}</i>	Multiplicative factor with respect to the mean objects area for labelling a structure as 'big'	[3;5;7;9;11;13;15]	5

Algorithm for Gleason Scoring

Performance metric: accuracy

Parameter	Description	Possible values	Optimal value
γ_{GABOR}	Spatial aspect ratio of the Gabor function support	[1;1.2;1.4;1.6;1.8;2]	1.2
θ_{GABOR}	Number of directions of the Gabor filter	[2;4;8;12;16;20]	8
$white_{THRESHOLD}$	Threshold to apply after the Gabor filter for lumen detection	[0.7;0.75;0.8;0.85;0.9;0.95]	0.90

Algorithm for Neoplastic Epithelium Detection in Breast Carcinoma

Performance metric: F1SCORE

Parameter	Description	Possible values	Optimal value
$percentage_{START}$	Initial percentage for the conversion of tumor and no-tumor cells (%)	[70;75;80;85;90;95]	95
$percentage_{STOP}$	End percentage for the conversion of tumor and no-tumor cells (%)	[50;55;60;65;70;75;80]	75
min_{POINTS}	Minimum number of neighborhood points used for the cluster expansion	[5;10;15;20]	5
ϵ_{DBSCAN}	Initial neighborhood size for the iterative DBSCAN (μm)	[50;100;150;200;250;300]	200

MUSA algorithm (Muscle UltraSound Analysis)

Performance metric: Absolute error

Parameter	Description	Possible values	Optimal value
$FODG_{SIZE}$	Kernel size of the FODG filter to enhance aponeurosis-like structures (mm)	[0.5;1;1.5;2;2.5;3]	1
$min_{ECCENTRICITY}$	Minimum eccentricity for considering a structure a candidate aponeurosis	[0.98;0.985;0.99;0.995]	0.995
$scan_{LINES}$	Number of scan lines for the detection of the actual aponeuroses in the heuristic search	[3;5;7;9;11;13]	7

TRAMA algorithm (TRAnverse Muscle ultrasound Analysis)

Performance metric: Absolute error

Parameter	Description	Possible values	Optimal value
$FODG_{SIZE}$	Kernel size of the FODG filter to enhance aponeurosis-like structures (mm)	[0.5;1;1.5;2;2.5;3]	1
β_{FRANGI}	First sensitivity parameter of the Frangi filter during the search of muscle connective tissue	[0.25;0.5;0.75;1;1.25;1.5]	0.5
C_{FRANGI}	Second sensitivity parameter of the Frangi filter during the search of muscle connective tissue	[5;10;15;20;25;30]	15
$scan_{LINES}$	Number of scan lines for the detection of the actual aponeuroses in the heuristic search	[3;5;7;9;11;13]	7

Appendix C

Ultrasound device settings

In the proposed studies, all ultrasound images are acquired using a MyLab™ Twice device (Esaote, Italy) with a 3-13 MHz linear array transducer. For all depths, the gain is set to 50%, the gain compensation is kept neutral and the dynamic image compression is switched off. All the ultrasound device settings are kept constant except for the image depth. For each subject, the image depth (initial value: 44 mm) is varied so as to be able to visualize the entire muscle (range:44 – 59 mm). The images are then exported to DICOM and transferred to a workstation for offline processing.

Appendix D

Acquisition protocol in skeletal muscle ultrasound images

During a single experimental session, the following muscles are investigated using ultrasounds: medial gastrocnemius, tibialis anterior, vastus lateralis and rectus femoris. These superficial skeletal muscles have been chosen since they are the most informative in the assessment of neuromuscular disorders and sarcopenia.

The following acquisition protocol is followed to ensure the best representation of the four muscles: i) medial gastrocnemius and vastus lateralis: the depiction of the two aponeuroses and of the muscle fascicles is optimized; ii) tibialis anterior: the representation of the muscle fascicles and of the bone boundary is maximized; iii) rectus femoris: the representation of the deep and superficial aponeuroses is optimized. Ultrasound images are acquired with the subjects in the supine position, except for the medial gastrocnemius where the subjects are positioned prone.

A suitable amount of ultrasound gel is used to minimize the transducer pressure on the skin and to ensure optimal image quality. During each acquisition, the lower limb joints are extended, and the subjects are asked to completely relax their muscles. All scans are performed by placing the transducer in correspondence of the largest muscle diameter at the following anatomic sites: the medial gastrocnemius from the mid-sagittal line of the muscle, midway between the proximal and distal tendon insertions; the tibialis anterior at one-quarter of the distance from the inferior border of the patella to the lateral malleolus; the vastus lateralis half-way along the line from the anterior-superior iliac spine to the superolateral border of the patella; and rectus femoris is measured half-way along the line from the anterior-superior iliac spine to the superior border of the patella.

Lists of Contributions

Journal Articles

1. C. Caresio, M. Salvi, F. Molinari, K.M. Meiburger and M.A. Minetto, **Fully Automated Muscle Ultrasound Analysis (MUSA): Robust and Accurate Muscle Thickness Measurement**, *Ultrasound in Medicine & Biology*, 2017, 43: 195–205.
2. M. Salvi, D. Rimini, F. Molinari, G. Bestente and A. Bruno, **Effect of Low-level Light Therapy on Diabetic Foot Ulcers: A Near-infrared Spectroscopy Study**, *Journal of Biomedical Optics*, 2017, 22 (3).
3. M. Salvi and F. Molinari, **Multi-Tissue and Multi-Scale Approach for Nuclei Segmentation in H&E Stained Images**, *Biomedical Engineering Online*, 2018, 17: 89-101.
4. K.M. Meiburger, M. Salvi, M. Giacchino, U.R. Acharya, M.A. Minetto, C. Caresio, F. Molinari, **Quantitative Analysis of Patellar Tendon Abnormality in Asymptomatic Professional “Pallapugno” Players: A Texture-Based Ultrasound Approach**, *Applied Sciences*, 2018, 8:660-673.
5. M. Albinì, M. Salvi, E. Altamura, S. Dinarelli, L. Di Donato, A. Lucibello, F. Mavelli, F. Molinari, U. Morbiducci and A. Ramundo-Orlando, **Movement of Giant Lipid Vesicles Induced by Millimeter Wave Radiation Change When They Contain Magnetic Nanoparticles**, *Drug Delivery and Translational Research*, 2018, 9:131-143.
6. M. Minetto, C. Caresio, M. Salvi, V. D’Angelo, N. Gorji, F. Molinari, G. Arnaldi and E. Arvat, **Ultrasound-based Detection of Glucocorticoid-induced Impairments of Muscle Mass and Structure in Cushing’s Disease**, *Journal of Endocrinological Investigation*, 2018, 1-12.
7. M. Salvi, C. Caresio, K.M. Meiburger, B. De Santi, F. Molinari and M.A. Minetto, **Transverse Muscle Ultrasound Analysis (TRAMA): Robust and Accurate Segmentation of Muscle Cross-Sectional Area**, *Ultrasound in Medicine & Biology*, 2018.

8. M. Salvi, U. Morbiducci, F. Amedeo, R. Santoro, F. Angelini, I. Chimenti, D. Massai, E. Messina, A. Giacomello, M. Pesce and F. Molinari, **Automated Segmentation of Fluorescence Microscopy Images for 3D Cell Detection in human-derived Cardiospheres**, under review on *Scientific Reports*, 2019.
9. M. Salvi, F. Molinari, N. Dogliani and M. Bosco, **Automatic discrimination of neoplastic epithelium and stromal response in breast carcinoma**, under review on *Computers in Biology and Medicine*, 2019.
10. M. Salvi, V. Cerrato, A. Buffo and F. Molinari, **Automated Segmentation of Brain Cells for Clonal Analyses in Fluorescence Microscopy Images**, paper in preparation.
11. M. Salvi, N. Michielli, L. Molinaro and F. Molinari, **A Novel Multi-Scale Method for Normalizing Histology Digital Slides**, paper in preparation.
12. J. Metovic, M. Salvi, F. Molinari and L. Molinaro, **Automated Quantitative Evaluation of Steatosis Before Liver Transplantation**, paper in preparation.
13. B. De Santi, M. Salvi, V. Giannini, D. Regge and F. Molinari, **Multimodal DWI and T2w Prostate Gland Segmentation**, paper in preparation.

Conference Papers

14. K.M. Meiburger, Z. Pan, M. Salvi, L. Molinaro, A. Sapino and F. Molinari, **Multi-Scale Automatic Roadmap for Studying Cancer Histopathological Images (MARS-CHI)**, *Enabling Technologies in 3-D Cancer Organoids*, March 8th-9th, 2016, Molecular Biology Center, Turin.
15. K.M. Meiburger, M. Salvi, C. Caresio, R. Santoro, F. Angelini, F. Amedeo, I. Chimenti, M. Pesce, U. Morbiducci and F. Molinari, **Automation in Quantitative Medical Imaging and its role as an Enabling Technology**, *29th Conference of the Society for Medical Innovation and Technology*, November 9th-10th, 2017, Lingotto Congress Center, Turin.
16. M. Salvi, K.M. Meiburger, B. De Santi, N. Michielli, L. Molinaro and F. Molinari, **Automated Gleason Grading in Prostate Cancer Histopathology**, *Imaging of Cancer Dynamics: CancerTO*, March 7th-9th, 2018, La Cavallerizza Hall University of Turin, Turin.
17. B. De Santi, M. Salvi, N. Michielli, K.M. Meiburger, N. Bonelli, R. Rossetto, R. Garberoglio and F. Molinari, **Classification of Ultrasound Thyroid Nodules using Artificial Neural Networks and Texture Analysis**, *Imaging of Cancer Dynamics: CancerTO*, 2018, March 7th-9th, 2018, La Cavallerizza Hall University of Turin, Turin.

18. K.M. Meiburger, M. Salvi, N. Michielli, B. De Santi, M.A. Minetto and F. Molinari, **The Role of Beamforming in Quantitative Muscle Ultrasonography: A Texture Analysis Approach**, *6th National Congress of Bioengineering*, June 25th-27th, 2018, Politecnico di Milano, Milan.
19. M. Salvi, K.M. Meiburger, B. De Santi, N. Michielli, L. Molinaro and F. Molinari, **Automated Gleason Grading in Prostate Cancer Histopathology Images**, *6th National Congress of Bioengineering*, June 25th-27th, 2018, Politecnico di Milano, Milan.
20. B. De Santi, M. Salvi, N. Michielli, K.M. Meiburger, N. Bonelli, R. Rossetto, R. Garberoglio and F. Molinari, **Ultrasound Thyroid Nodule Texture-based Classification with Artificial Neural Networks**, *6th National Congress of Bioengineering*, June 25th-27th, 2018, Politecnico di Milano, Milan.
21. N. Michielli, S. Seoni, B. De Santi, M. Salvi, K.M. Meiburger, A. Iadarola, A. Cicolin and F. Molinari, **Automatic Sleep Stage Classification: A Step Forward to Automated Assessment of Neurocognitive Performance**, *6th National Congress of Bioengineering*, June 25th-27th, 2018, Politecnico di Milano, Milan.
22. M. Pesce, M. Salvi, F. Amedeo, F. Angelini, I. Chimenti, M. Agrifoglio, F. Molinari, R. Santoro, E. Messina and U. Morbiducci, **YAP-based position sensing in myofibroblast evolution of cardiac stromal cells in human cardiospheres**, *European Society of Cardiology Congress*, August 25th-29th, 2018, Munich, Germany.
23. K. M. Meiburger, F. Veronese, M. Salvi, M. Fadda, E. Zavattaro, B. De Santi, N. Michielli, P. Savoia and F. Molinari, **Automatic Extraction of Dermatological Parameters from Nevi using an Inexpensive Smartphone Microscope: A Proof of Concept**, submitted for *Engineering in Medicine and Biology Congress*, 2019.
24. B. De Santi, M. Salvi, V. Giannini, K. M. Meiburger, N. Michielli, D. Regge and F. Molinari, **Multimodal T2w and DWI Prostate Gland Automated Registration**, submitted for *Engineering in Medicine and Biology Congress*, 2019.