Doctoral Dissertation
Doctoral Program in Bioengineering and Medical-Surgical Sciences (33rd cycle)

Innovative Strategies for Tissue Characterization in Multimodal Images

Bruno De Santi

* * * * * *

Supervisor
Prof. Filippo Molinari

Doctoral Examination Committee:
Prof. V. Bevilacqua, Referee, Polytechnic University of Bari
Prof. M. Casadio, Referee, University of Genoa
Prof. V. Corino, Referee, Polytechnic University of Milan
Prof. A. Cereatti, Polytechnic University of Turin

Politecnico di Torino
April 8th, 2021

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.

. . . .
Bruno De Santi
Turin, April 8th, 2021

Summary

In the last decade, multimodal medical imaging has been increasingly employed in the management of many pathologies thanks to its intrinsic ability to integrate complementary information from different imaging modalities. Although *radiomics* has been extensively studied for tissue characterization, relatively few strategies have been developed and optimized for tissue characterization in multimodal images. The aim of this thesis is to present a set of innovative multimodal-based strategies for accurate and robust tissue characterization in three different clinical contexts: thyroid cancer, neuroendocrine tumor and prostate cancer.

First, a novel strategy was developed to quantitatively describe the vascularization of thyroid nodules in terms of architecture, tortuosity and flow in three-dimensional (3D) power-doppler ultrasound (PDUS) images. The main novelty of this approach is the distinction between peripheral and central vascularization of the nodule in a 3D framework. A classifier was trained and cross-validated on a dataset of 30 thyroid nodules in order to differentiate benign from malignant thyroid nodules, achieving a sensitivity and specificity equal to 93.33 % and 73.33 %, respectively. Without considering distinction between peripheral and central vascularization, sensitivity of the predictive model drops to 66.67 %. This result confirms the importance of vascularity distribution (i.e., peripheral and central) in the differentiation between benign and malignant nodules. Integration with contrast-enhanced ultrasound is complex due to pitfalls associated with the image acquisition protocol but a strategy is under working for a more complete characterization of the nodule microvascularization.

Second, a semi-automatic algorithm for the segmentation of neuroendocrine tumors (NETs) in ^{68}Ga-DOTA-TOC PET/CT images is proposed. Although several radiomics approaches have been developed for tumor characterization in ^{18}F-FDG PET/CT images, investigation has to be made on the recent ^{68}Ga-DOTA-TOC PET imaging. The segmentation algorithm requires the user intervention only for the initial definition of the region-of-interest on PET and CT images, then, it exploits clustering techniques and active contours guided by image and edge information. Validation on 57 primary and metastatic neuroendocrine lesions with heterogeneous intensity and contrast demonstrate that the developed algorithm could be used to reduce analysis time and variability of manual segmentation. Furthermore, robustness analysis on 45 radiomics features extracted from the neuroendocrine tumors show that only 22 out of 45 are robust to both segmentation and discretisation. In particular, textural features based on low grey levels or large runs/zones emphasis are not robust, hence, their use is not recommended for multi-centric studies. Although CT images are fundamental for anatomic localization of these tumors, the current acquisition protocol makes difficult the extraction of quantitative data for segmentation and characterization in multimodal fashion.

Finally, a fully multimodal pipeline for the differentiation of healthy and cancerous prostatic tissue in multiparametric magnetic resonance images (mpMRI) is presented. The pipeline incorporates optimized algorithms for image registration of apparent diffusion coefficient/T2-weighted (ADC/T2w) sequences and whole-mount histology/T2w (WMH/T2w) images. The main advantage is that contouring of the tumor in the WHM directly defines the region-of-interest on the ADC and the T2w images, thus, limiting the subjectivity of manual tumor annotation on the mpMRI. Multivariate analysis is then performed using first-order textural features extracted from mpMRI (ADC and T2w) images. Results on 30 peripheral zone cancers demonstrate that the use of mpMRI improves diagnostic accuracy compared to single sequences. Moreover, ADC values are lower in cancer than in normal tissue ($P<0.0005$), whereas T2w image heterogeneity (standard deviation and entropy) is lower in cancer than in normal tissue ($P<0.005$). These findings confirm that multimodal integration is beneficial for prostate cancer characterization.

In conclusion, although multimodal integration has great potential in every clinical context analyzed, it is not always feasible due to pitfalls in image acquisition. The strategies proposed in this thesis obtained promising results in segmentation, registration and characterization of tissue. These methodologies will be extended to larger datasets and additional imaging modalities to create robust and high-throughput multimodal radiomics systems for the diagnosis of diseases with the ultimate goal of improving the patient's health and lifestyle.

Contents

Chapter 1

Introduction

1.1 Tissue characterization

It is difficult to give a single and complete definition to the term "tissue characterization", but the definition used by R.C. Chivers, almost 40 years ago, can serve as an introduction to the work of this thesis as it highlights some fundamental concepts that will be constantly used during the manuscript. He wrote [1]:

> *"Tissue characterization may be defined in its purest form as the identification of one or more physical parameters of a small volume of tissue that are sufficiently well correlated with the type or condition of the tissue"*

The keywords to highlight in this definition are: "physical parameters", "small volume of tissue", "correlated with the type or condition". The first keyword, "physical parameters", represents what we measure and implies the kind of interaction with the tissue, a physical interaction. The second highlighted term, "small volume of tissue", is where we extract these physical parameters from and the use of the adjective "small" gives the idea of a local measurement. Finally, the third term, "correlated with the type or condition", indicates the reason why these physical parameters are identified, that is, the correlation with the type or condition of the tissue.

This definition may encompass a wide spectrum of methodologies but it has a tight connection with medical imaging (after all, R.C. Chivers is an ultrasound imaging researcher). Medical imaging refers to a set of techniques able to create images of internal parts of the body for the purpose of diagnosising disease, monitoring disease and aiding the treatment strategy decision [2]. There are different medical imaging techniques, such as ultrasound imaging, radiography, magnetic resonance imaging, nuclear medicine imaging, and the common feature among these techniques is a physical interaction with the tissue. What differentiates one technique from the other is the way it interacts with the tissue. For example, in the case of ultrasound imaging the tissue is excited using acoustic waves [1] whereas in radiography, electromagnetic waves with frequency within a specific range radiate the tissue [3]. By changing the method of energy transmission and other parameters, such as the frequency of excitation, it is possible to interact with different tissues and to identify, using Chivers' words, different "physical parameters" of a tissue. Further, the excitation wavelength, which in case of electromagnetic waves is related to the energy by the Planck equation, regulates the scale at which the physical interaction with the tissue occurs. The adjective "small" in the definition of tissue characterization refers to this, and in medical imging terms it is known as image resolution. Therefore, tissue imaging can be definitely seen as a pure form of tissue characterization.

Over time, medical imaging has become a fundamental and necessary tool for the diagnosis, active suirvellance, and the treatment response assessment of the majority of existing pathologies. Furthermore, in the new era of *Big Data* and *Artificial Intelligence*, medical images are increasingly seen as data arrays from which quantitative parameters describing the tissue physiology or pathology. This practice is called *Radiomics* and was described by Gillies at al. in 2016 in the work *Radiomics: Images Are More than Pictures, They Are Data* [4] as:

> *"the development of processes for high-throughput extraction of quantitative features that result in the conversion of images into mineable data and the subsequent analysis of these data for decision support"*

In this respect, the interpretation that has been given in this thesis to the term "tissue characterization" is the cascade of Chivers' vision with that of Gillies; in other words, a process that starting from the local physical interaction leads to the extraction of features that are able to quantitatively and globally characterize a tissue.

1.2 Multimodal medical imaging

In the last decade, multimodal medical imaging has been showing significant improvements of imaging-based computer-aided systems for detection, diagnosis, treatment response quantification, and surgical guidance. The basic concept is the integration of medical images from different imaging modalities into a single system. This concept may be incorporated in the more general term of *data fusion*, i.e. the process of integrating data from different sources in order to obtain inferences that could not be obtained using a single source [5].

Since different imaging modalities reflect different intrinsic properties of a tissue which may be complementary to each other, it is reasonable that their integration may yield to more reliable and deepen characterization of a tissue. A typical combination of imaging modalities includes a functional and an anatomical imaging modality. The functional modality depicts the tissue functionality, such as its metabolism or blood flow, whereas anatomical imaging (also known as morphological or structural imaging) excels in depiction of the anatomy and the morphology of tissues [6]. For example multimodal positron-emission tomography/computed tomography (PET/CT) imaging, which is extensively used in clinical practice, combines the sensitivity of functional PET images to detect tissue with abnormal functionality with the high-resolution anatomical description of CT images. With the growing development of new hybrid systems for multimodal image acquisition and novel functional images, there is a growing need for novel and reliable strategies for multimodal medical image processing.

Multimodality pitfalls When applied to real clinical data, multimodal strategies have to cope with some critical points:

- large amount of images could require high computational cost for processing;
- in some cases, different imaging modalities need to be registered, and this may be very complex depending on the modalities;
- different imaging modalities have different artifacts which must be dealt before analysis;
- data extracted from different imaging modalities may be redundant.

Therefore, for some clinical problems multimodality may present more pitfalls than pearls.

1.3 Thesis objectives

The aim of this thesis is to present a set of innovative strategies for tissue characterization in multimodal medical images. Potential and pitfalls of multimodal integration are discussed in three different clinical applications: i.e. thyroid cancer, neuroendocrine tumor and prostate cancer. The proposed approaches include advanced techniques for image segmentation, registration and feature extraction and are optimized to work on ultrasound, nuclear medicine and magnetic resonance images. The ultimate goal is to identify and overcome current limitations of the traditional medical image processing techniques and pave the way towards the clinical use of computer-aided systems based on multimodal medical images.

1.4 Thesis outline

The thesis is divided into 6 chapters. The second chapter provides a technical background about the main multimodal imaging techniques and a description of current strategies for multimodal tissue characterization. In the third chapter, a novel strategy is proposed for the three-dimensional quantification and characterization of thyroid nodule vascularization in power doppler ultrasound images. In the fourth chapter, the clinical problem of neuroendocrine tumor is discussed, a semi-automatic algorithm for neuroendocrine tumor segmentation in positron emission tomography images is proposed, in addition, robustness of radiomics features extracted from neuroendocrine tumors is evaluated. In the fifth chapter, an innovative strategy based on advanced image registration algorithms for the differentiation between healthy and cancerous tissue in multiparametric magnetic resonance images is presented. Finally, the conclusions of the thesis and future works are reported in the last chapter.

1.5 References

[1] R. C. Chivers. “Tissue characterization”. In: *Ultrasound in Medicine and Biology* 7.1 (1981), pp. 1–20. ISSN: 03015629. DOI: `10.1016/0301-5629(81)90018-1`.

[2] Andreas Maier et al. *Medical Imaging Systems - An Introductory Guide*. Vol. 11111. 2018, p. 259. ISBN: 978-3-319-96519-2.

[3] Vadim Kuperman. *Magnetic Resonance Imaging: Physical Principles and Applications*. 1st. San Diego: Academic Press, 1999. ISBN: 0124291503.

[4] Robert J. Gillies, Paul E. Kinahan, and Hedvig Hricak. “Radiomics: Images are more than pictures, they are data”. In: *Radiology* 278.2 (Feb. 2016), pp. 563–577. ISSN: 15271315. DOI: `10.1148/radiol.2015151169`.

[5] David L. Hall and James Llinas. “An introduction to multisensor data fusion”. In: *Proceedings of the IEEE* 85.1 (1997), pp. 6–23. ISSN: 00189219. DOI: `10.1109/5.554205`.

[6] Stephanie N. Histed et al. *Review of functional/anatomical imaging in oncology*. Apr. 2012. DOI: `10.1097/MNM.0b013e32834ec8a5`.

Chapter 2

Background

This chapter summarizes the most used medical imaging modalities. Imaging modalities are described in terms of physical principles, technology and main applications. Subsequently, the main multimodal combinations are outlined. Finally, current strategies for tissue characterization in multimodal images are discussed.

2.1 Imaging modalities

2.1.1 Ultrasound imaging

Physical principles Ultrasounds are acoustic waves with a frequency above 20 kHz, hence, above the frequency band audible from humans. Acoustic waves propagate through a medium with a velocity v which depends on the coefficient of stifness C (or bulk modulus) and the density ρ of the medium through the following expression:

$$v = \sqrt{\frac{C}{\rho}} \tag{2.1}$$

An important medium property is the acoustic impedance Z which represents the resistance that the medium presents to the sound wave and depends on the density and the coefficient of stifness:

$$Z = \sqrt{\rho C} = \rho v \tag{2.2}$$

When propagating through a homogeneous medium, the sound wave is continously attenatued due to the absorption effect:

$$A(z) = A_0 e^{-\mu_0 z} \tag{2.3}$$

where A_0 is the initial amplitude of the wave before absorption, μ_0 is the absorption coefficient which depends on the medium material and on the wave frequency; z is the travelled distance through the medium. This equation is called Lambert-Beer law and was originally used to model attenuation of electromagnetical waves [1].

Absorption is not the only kind of interaction with the medium. Indeed, when a sound wave with amplitude A_0 meets a planar interface between two media with different acoustic impedances Z_1 and Z_2, generates a reflected sound wave (reflection), called echo, of amplitude A_r given by:

$$A_r = A_0 \left(\frac{Z_1 - Z_2}{Z_1 + Z_2} \right) \tag{2.4}$$

The energy which is not reflected is propagated to the second medium (transmission).

Technology In its simplest form, the technology consists of a ultrasound transducer which produces ultrasounds in response to an electrical impulse thanks to piezoelectric effect. An ultrasound probe contains, therefore, a set of piezoelectric crystals on its surface that emit ultrasound waves. Since the piezoelectric effect is reversible (a mechanical stress of a piezoelectric material produces a electrical field), the same crystals are able to receive ultrasound waves. Therefore, the ultrasound probe is used to insonate the tissue with an ultrasound pulse, then, the pulse propagates through the tissue and produces an echo when encounters an interface between two tissues with different acoustic impedance, the echo produced is received by the same probe.

Modes

Brightness-mode Brightness-mode, abbreviated as B-mode, is the most common ultrasound imaging mode. The amplitude of the echo and the time interval between the emission of the ultrasound pulse and the echo detection are measured. Assuming an average propagation velocity of ultrasounds in biological tissues equal to $c = 1540\,\mathrm{m\,s^{-1}}$, it is possible to find the depth d at which the echo is produced:

$$d = \frac{c\Delta t}{2} \tag{2.5}$$

On one hand, the amplitude of the echo which is related to the strongness of the reflection is converted in to brightness of a pixel (also called echogenicity). On the other hand, the time interval is needed for the spatial localization of the echo source. To have a two-dimensional image, a multitude of ultrasound scan lines produced by a string of piezoelectric crystals are combined, this process is called beamforming. B-mode is widely used in many clinical fields, such as vascular medicine [2], cardiology [3], gynecology [4], ophtalmology [5], otolaryngology [6], musculoskeletal [6] and urology [7]. B-mode is mostly used for visualization and size measurement of anatomical structures [5, 2], but also for detection of abnormalities [6, 8]. Figure 2.1 shows example of B-mode images of thyroid gland and testicle.

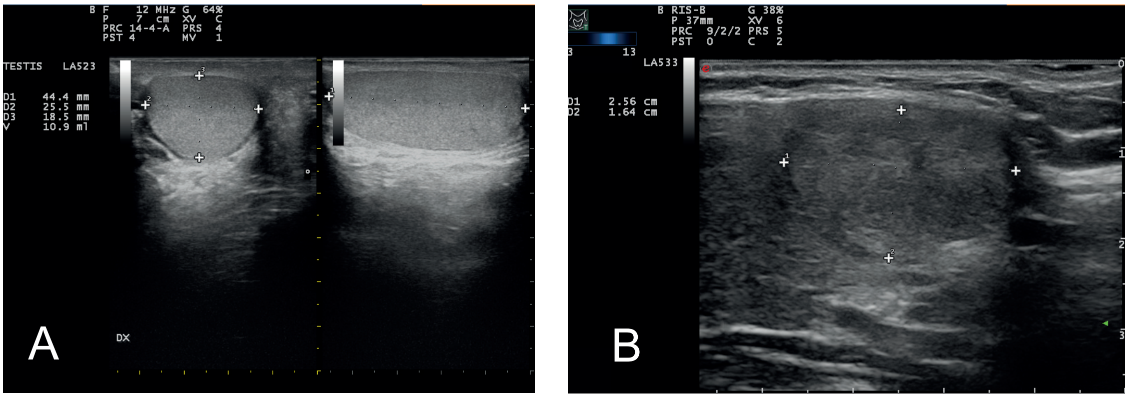


Figure 2.1: B-mode image of testicle (A) in transversal and longitudinal view. B-mode image of thyroid gland with a nodule (B). Testicular B-mode ultrasound can be used for testicular volume estimation and heterogenity evaluation for fertility evaluation [8]. On the other hand, ultrasonographic features extracted from B-mode images of thyroid nodules are used for risk stratification [9]

Color and power doppler ultrasound Color doppler ultrasound (CDUS) and power doppler ultrasound (PDUS) are two ultrasound-based functional modalities. They allow for visualization and measuring of blood flow thanks to the exploitation of the Doppler effect. Given a source and a observer, the doppler effect is the increase/decrease of the wave frequency detected by the observer when there is a non-null relative velocity of the source with respect to the observer. The frequency shift Δf is related to the wave frequency f to the relative velocity of the source v_r and the angle θ between the source velocity direction and the direction from the source to the observer through the following equation [10] :

$$\Delta f = \frac{2 f v_r \cos(\theta)}{c} \tag{2.6}$$

In Doppler ultrasound imaging, the source are the blood cells while the observer is the ultrasound probe. To maximize the Doppler signal, the ultrasound beam should have the same direction of the blood flow ($\cos(\theta) = 1$), but this is not feasible, hence, an angle between 30° to 60° is commonly used for a fair trade-off between sensitivity and ease of acquisition operation. Figure 2.2 is a graphical illustration of the Doppler image acquisition.

In CDUS, the Doppler shift frequency is used to compute the blood flow direction and velocity. The result is showed through a colormap, i.e. flow moving away from the probe the flow is represented in blue, otherwise in red. The color intensity represents the velocity of blood cells. In PDUS, the power of the Doppler signal, which is derived from the echoes amplitudes, is measured. This quantity is related to the number of blood cells moving in the sample volume. In this case, a monochromatic colormap is used. Figure 2.3 shows a CDUS and a PDUS image of a thyroid nodule.

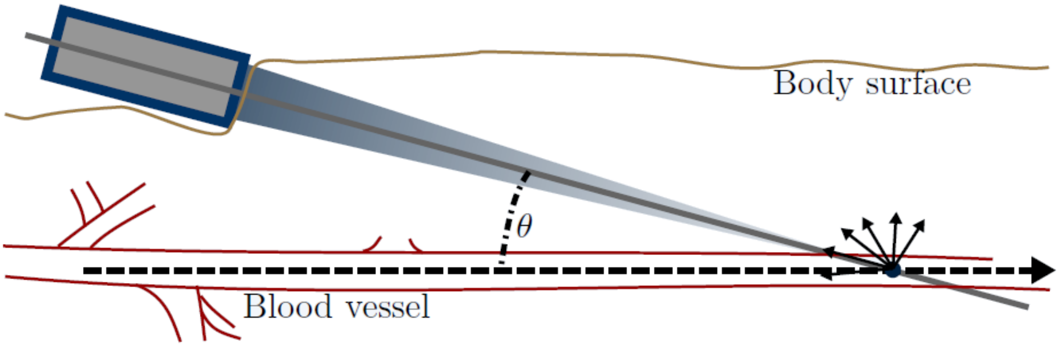


Figure 2.2: Graphical illustration of the Doppler ultrasound image acquisition [10].θ is the angle between the source velocity direction and the direction from the probe to the moving object.

Some of the clinical applications of doppler ultrasound imaging are transcranial doppler ultrasound [11], vascular doppler ultrasound [12], doppler echocardiography [13] and oncology [14, 15].

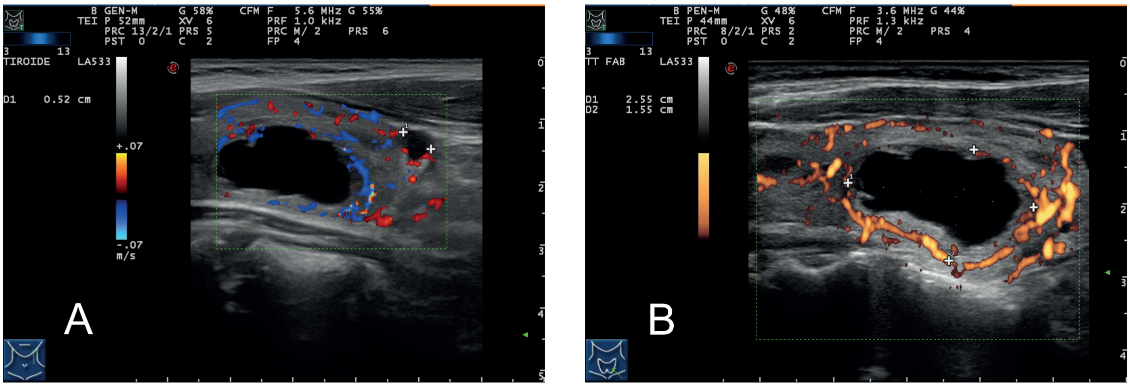


Figure 2.3: Color doppler ultrasound image (A) and power doppler ultrasound image (B) depicting the blood flow of a thyroid nodule. The arrangement of the vascular network is associated with the malignancy of the thyroid nodule [14]

Contrast-enhanced ultrasound In contrast-enhanced ultrasound (CEUS), a contrast agent which consists of microbubbles with a size from 1 µm to 8 µm filled with a low-solubility perfluorinated gas and stabilized in phospholipid membrane is used. The contrast agent is intravenously administrated to the patient when insonated the microbubbles which travel in the blood enhance the reflection of ultrasounds increasing the contrast of images and thus, creating images which depict the blood perfusion in tissues. Further, the small size allows the microbubbles to travel also through the microcirculation making this mode suitable for characterizing microvasculature [16]. In the clinic, CEUS imaging is used for liver and thyroid lesion vascularity characterization [17, 18], detection and characterization of cardiovascular diseases such as artherosclerosis [19].

Superb microvascular imaging Superb microvascular imaging (SMI) is a novel ultrasound mode which exploits the Doppler effect, as CDUS and PDUS, but uses advanced image filtering techniques for clutter (strong echoes generated by very slow moving tissue) suppression in order to detect with a better sensitivity blood flow of small vessels [20]. Due to its recent introduction, there is no evidence of its effectiveness in a clinical context. Anyway, preliminary research results on the microvascular characterization of thyroid nodules [20], breast lesions [21] and liver lesions [22] are promising. Figure 2.4 shows an example of superb microvascular image of a thyroid nodule.

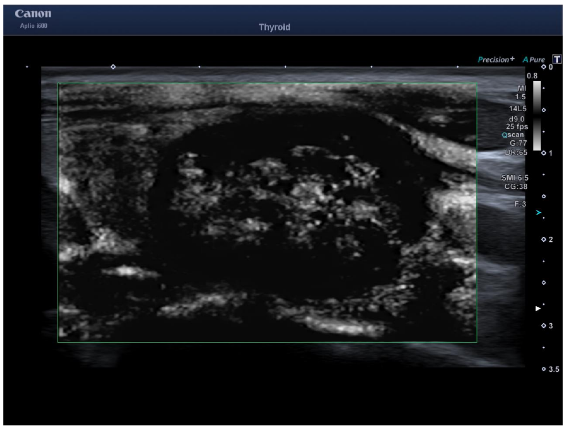


Figure 2.4: Superb microvascular image of a thyroid nodule.

Elastosonography Elastosonography exploits the physical concept of elastic strain, i.e. an elastic material under a mechanical stress produces a strain which is proportional to the mechanical stress applied. In this mode, ultrasound waves are used to stress mechanically a tissue, then, the strain field (strain elastosonography) or the shear wave (shear wave elastosonography) are using different techniques and the Young modulus which measures the stiffness of the tissue can be estimated [23]. Elastonography has been used and is most effective in the characterization of tissues which change their elastic properties when in pathological condition, e.g. musculoskeletal diseases [24], breast lesion differentiation [25] and thyroid nodule risk stratification [26] where generally malignant thyroid nodules have a higher stiffness than benign ones. Figure 2.5 shows the elastonographical examination for a benign and a malignant thyroid nodule.

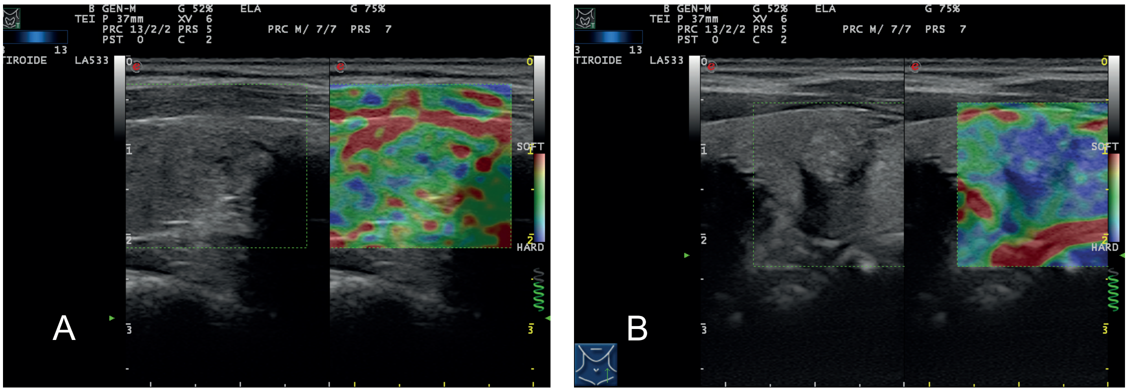


Figure 2.5: Strain elastosonographical image of a benign thyroid nodule (A) and a papillary caricinoma (B). The malingnant nodule had higher stiffness compared to the benign one, this feature may be associated with an increased risk of malignancy in thyroid nodules [26].

2.1.2 Magnetic resonance imaging

Nuclear magnetic resonance Magnetic resonance imaging (MRI) exploits the phaenomenon of nuclear magnetic resonance (NMR) which was discovered by Bloch and Purcell in 1946 [27, 28]. NMR involves three actors, nuclei with non-null magnetic moment, a static magnetic field and an oscillant magnetic field. When a nucleus with a gyromagnetic ratio, γ, under a static magnetic field, B_0, is excited with an oscillating magnetic field, B_1, with a frequency equal to $\omega = \gamma B_0$ (in usual applications, this frequency belongs to the radiofrequency spectrum) it experiences a transition between two energy states. In the lower energy state, i.e. B_1 is off, most part of the magnetic moments of the nuclei are parallel to the B_0 field; in the higher energy state, i.e. B_1 is on, most part of the magnetic moments of the nuclei are anti-parallel to the B_0. Transition from lower energy to higher energy state is called excitation, the inverse transition is called relaxation. Excitation gives rise to a rotation of the magnetization vector, M, which is the vectorial sum of the nuclei magnetic moments, producing a transversal component ($M_{tr} \neq 0$) of the magnetization vector[1]. In the relaxation phase, the transversal component of M decays and the longitudinal component M_z returns to its original value [29]. Bloch equations describe the dynamics of relaxation after a B_1 radiofrequency pulse through two characteristic times, i.e. T_1 and T_2. The first one regulates how fast the longitudinal component of M reaches its original value:

$$\frac{\mathrm{d}M_z}{\mathrm{d}t} = \frac{M_0 - M_z}{T_1} \tag{2.7}$$

---

[1] Using a laboratory coordinate system, where z is the longituginal direction parallel to B_0, x and y are the directions of the plane transveral to B_0, M_{tr} is used to indicate the transversal component of M

where M_0 is the initial value of M_z. T_1 is called spin-lattice relaxation time.

On the other hand, T_2 regulates the dynamics of the decay of the transversal components of M:

$$\frac{\mathrm{d}M_x}{\mathrm{d}t} = -\frac{M_x}{T_2} + \gamma M_y B_0 \tag{2.8}$$

$$\frac{\mathrm{d}M_y}{\mathrm{d}t} = -\frac{M_y}{T_2} + \gamma M_x B_0 \tag{2.9}$$

T_2 is called spin-spin relaxation time. Figure 2.6 shows the solutions of the two Bloch equations.

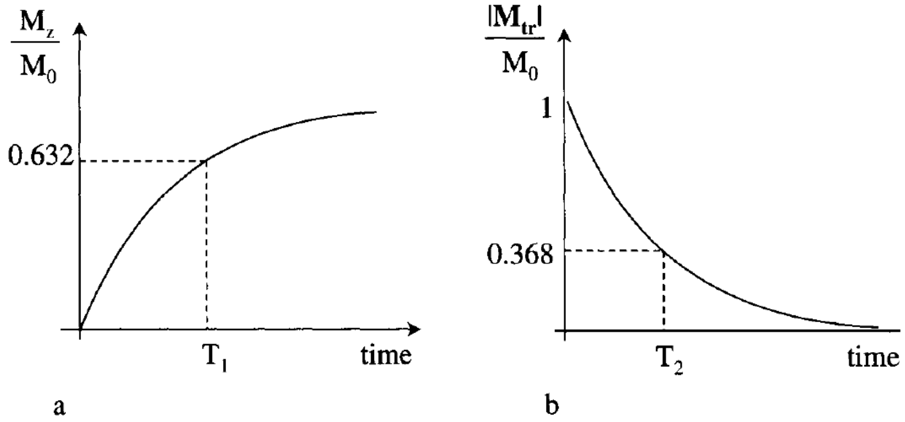


Figure 2.6: Solution of Bloch equations describing relaxation dynamics [29]

Technology Figure 2.7 graphically describes a magnetic resonance imaging system.

Briefly, a magnetic resonance system consists of:

- Superconductive magnet which produces the static magnetic field B_0;
- Radiofrequency coils which emit the radiofrequency magnetic field B_1 and receives the FID signal;
- Gradient coils which are responsible of the spatial localization.

Image contrast and sequences In MRI, hydrogen nuclei are the target nuclei thanks to their favorable gyromagnetic ratio and their abundant presence in tissues. Nuclei are excited by a series of radiofrequency pulses. Each pulse is characterized by an angle which represents the angle formed by the magnetization vector and the longitudinal direction, e.g. right after a 90° pulse $M_z = 0$ and $|M_{tr}| = M_0$. A receiver coil is used to measure the free induction decay (FID) signal which

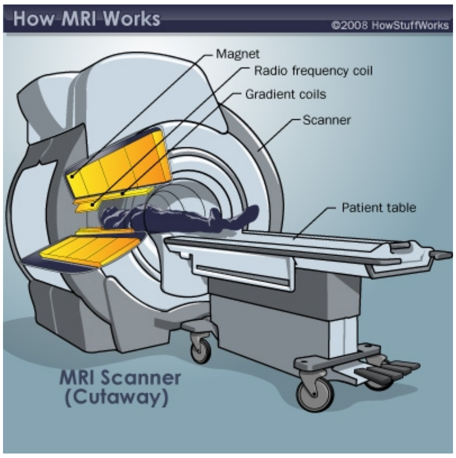
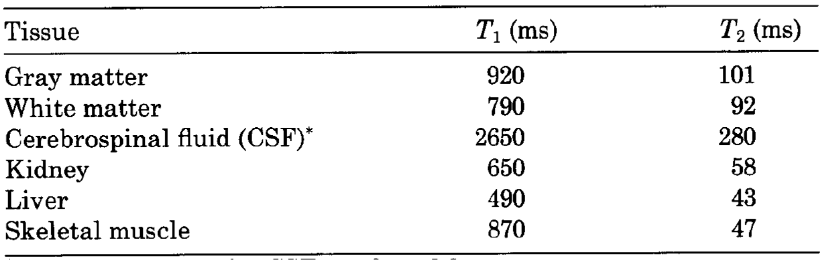


Figure 2.7: Graphical illustration of the components of a magnetic resonance imaging system, from https://science.howstuffworks.com/mri.htm

is proportional to the transversal component of the magnetization. By analyzing the Fourier spectrum of the FID signal, it is possible to estimate the dynamics of relaxation. Each tissue present different dynamics of relaxation (T1 and T2 values for different human tissues are reported in figure 2.8).



Tissue	T_1 (ms)	T_2 (ms)
Gray matter	920	101
White matter	790	92
Cerebrospinal fluid (CSF)*	2650	280
Kidney	650	58
Liver	490	43
Skeletal muscle	870	47

Figure 2.8: T1 and T2 values at 1.5 T of different tissues [29]

Main sources of image contrast in MRI are given by variations of local relaxation times, e.g. T1 and T2, and varations of hydrogen nuclei density (also called proton density, PD). Indeed image intensity I can be expressed as:

$$I \propto PD \cdot e^{-TE/T2} \cdot (1 - e^{-TR/T1}) \tag{2.10}$$

where TE is called echo time and TR is called repetition time and are characteristic timing parameters of the sequence used. A sequence consists of a series of pulses of different angles temporally spaced from each other. Depending on the tissue to be imaged a different sequence of pulses is chosen. Further, depending on the problem, more than one sequence can be selected. For example, a morphological sequence to study the anatomy of an organ, e.g. T2-weighted sequence, and a functional one to study the tissue physiology, e.g. diffusion-weighted.

Proton density Proton density (PD) images are obtained by using a $TE \ll T2$ and a $TR \gg T1$. The image obtained reflects the proton density in tissue. The use of this sequence is diffused in the anatomical imaging of articular cartilage [30] and brain tumor detection and characterization [31].

T1-weighted T1-weighted (T1w) images are acquired with a $TE \ll T2$ and $TR \sim T1$. These images are used for anatomical imaging of bones and brain, but also for detection of pathological conditions of myocardial tissue [32], brain tissue [33] and musculoskeletal tissue [34].

T2-weighted A T2-weighted (T2w) sequence is characterized by $TE \sim T2$ and $TR \gg T1$. The T2w sequence has been extensively used for anatomical description of pelvic area soft tissues [35], brain tumor imaging and characterization [34, 36].

Diffusion-weighted The image contrast in diffusion-weighted (DW) imaging is the diffusion of water molecules. Cell membranes act as microbarriers to the diffusion of water molecules in tissue. Different tissue can exhibit different diffusion coefficients D of water molecules, hence, image contrast can be generated. By applying a magnetic gradient in a specific direction, it is possible to estimate diffusion of water molecules along that direction using the Stejskal-Tanner equation which relates the diffusion coefficient D with the NMR signal attenuation with the following exponential e^{-bD}, where b depends on the gradient strength and timing [29]. The coefficient b is called b-value and typical values are in the range from $400\,\mathrm{s\,mm^{-2}}$ to $1400\,\mathrm{s\,mm^{-2}}$. Tissues with higher diffusion have lower diffusion-wieghted image intensity, and viceversa. In practice, there are many sources of signal attenuation and it is difficult to measure real diffusion. To reduce other signal attenuation sources, DW images acquired with different b are combined to build maps of the apparent diffusion, for example:

$$ADC_{1000} = \frac{\log DW_{b_{1000}}}{\log DW_{b_0}} \tag{2.11}$$

These maps are called apparent diffusion coefficient (ADC) maps and they showed excellent results for detection and characterization of pathologies in brain [37], prostate [38] and muscles [39].

Dynamic contrast-enhanced Dynamic contrast-enhanced imaging requires the injection of a contrast medium with particular paramagnetic properties, such as gadolinium, which interacts with hydrogen nuclei and enhances the NMR signal [29]. After the contrast agent administration, the MR image intensities dynamically change due to the relaxation times shortening. Hence, multiple acquisitions at different time points are performed and semi-quantitative parameters about the dynamics of perfusion of the contrast agent are extracted. This modality showed successful results in the angiogenesis characterization of brain tumor [40], breast cancer [41] and prostate cancer [42].

Quantitative magnetic resonance imaging So far, we discussed about what is called "conventional" MRI. Quantitative MRI (which is not to be counfounded with the quantitative analysis of MRI) consists of a set of imaging techniques which are able to acquire maps of physical or chemical variables that can be measured. Compared with "conventional" MRI, these maps are expressed in physical units and absolute values can be compared among different sequences and patients [43]. However, in most of usual applications, conventional MRI is preferred to quantitative MRI due to the reduction of acquision time.

2.1.3 Nuclear medicine imaging

Nuclear medicine imaging belongs to the family of functional imaging. Countrary to the imaging techniques described so far where the energy is transmitted by a device external to the patient, in nuclear medicine (NM) imaging the source of radiation is located within the patient.

Physical principles A radiotracer, which consists of a radioisotope tagged to a tracer which biochemically interacts with substances or cells inside the body, is administered to the patient [10]. Over time, the radioisotope begins ejecting particles from its nucleus to reach energy stability. This process is called radioactive decay and its dynamics depend on the isotope. The radioactivity (R) of a isotope is defined as the amount of atoms which decay per second and is measured in Becquerel (Bq). The process of radioactive decay can be modeled by the exponential law:

$$S(t) = S_0 e^{-\ln 2t/t_{1/2}} \tag{2.12}$$

where S_0 is the initial radioactivity and $t_{1/2}$ is the isotope's half-life which is the time that the isotope takes to lose half of its initial radioactivity [10]. There are multiple modes of decay, in each one different particles are emitted. For NM imaging, two particels are of interest: $\beta+$ and γ. The $\beta+$ particles are positrons which get annihilated when they encounter an electron; after annihilation two γ photons with an energy of 511 keV which travel in opposite directions are produced (figure 2.9). In γ decay, a γ particle is directly ejected from the nucleus (figure 2.9).

The two main NM imaging modalities are positron-emission tomography (PET) and single photon-emission computed tomography (SPECT). The former one exploits the $\beta+$ decay mode, while the latter one exploits the γ decay mode.

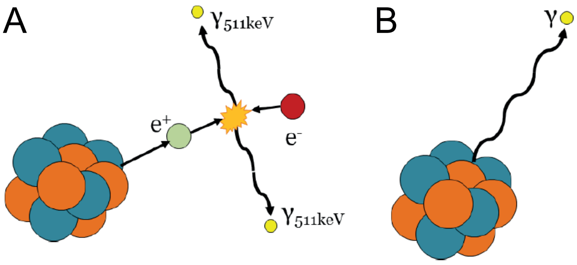


Figure 2.9: Illustration of the $\beta+$ (A) and γ (B) decay [10].

Modalities Both modalities allow the visualization of the distribution of the radiotrace throughout the body. Further, a huge advantage of these modalities is the possibility to extract quantitative parameters of the concentration of radiotracer which allow the comparison across different patients or time points, such as the standardized uptake value (SUV) parameter which is the ratio between radioactivity concentration and the whole-body concentration of the injected radioactivity [10]. Anyway, the technology behind PET and SPECT modalities is slightly different.

PET A PET imaging system consists of a ring with several detector blocks which are composed of a scintillator crystal, which transforms γ photons in visible photons, connected to a photomultiplier tube, which converts the photon energy to an electrical signal thanks to the photoelectric effect. Since the nature of the $\beta+$ decay and the opposite directions of the two γ photons, the ring geometry is required in order to count the number and the position of annihilations (figure 2.10). Indeed, through electronics, only γ photons that are detected simultaneously along the same scanning line (in practice, within a specified delay) are counted. Then, combining informations between different scanning lines, time-of-flight of photons and image reconstruction algorithms, an estimate of the annihilation sources location is obtained [10].

Depending on the target physiological or pathological process, a radiotracer is chosen. The most used radiotracer in PET imaging is the ^{18}F-fludeoxyglucose (^{18}F-FDG) where the isotope is ^{18}F and the tracer is a molecule (fludeoxyglucose) of the glucose family. ^{18}F-FDG-PET imaging are used to detect areas characterized by high glucose metabolism, hence, its use is diffused in oncological applications, since most part of tumoral cells have a glucose metabolism higher than normal cells [44]. ^{18}F-FDG-PET images can be used for detection, stratification, radiotherapy and treatment response evaluation of lung tumors, brain lymphomas and melanomas, breast tumors [45] but also for non-oncological applications [46], such as the evaluation of the progression rate of Parkinson's disease [47]. Although ^{18}FDG-PET is the most widely used radiotracer in NM imaging, other radiotracers were proposed. For example, new radiotracers based on ^{68}Ga-DOTA-peptides have been introduced in order to study the neuroendocrine tumors (NET). These tracers are able to bind

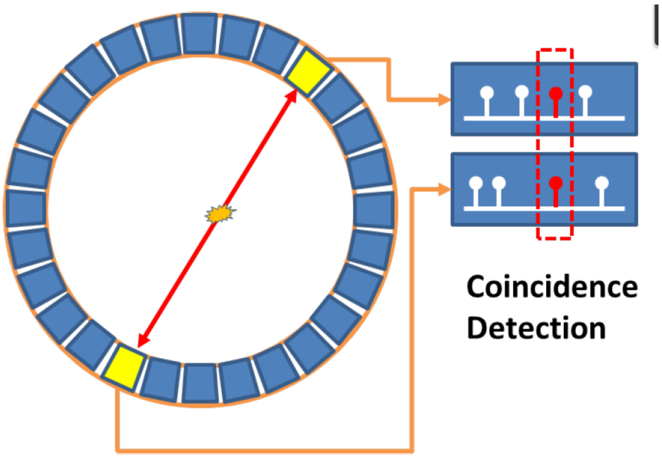


Figure 2.10: PET ring detector and the coincidence detection principle. Only γ photons that are detected simultaneously along the same scanning line (in practice, within a specified delay) are considered for image formation [10].

to somatostatin receptors which are expressed by NETs [48].

SPECT In SPECT imaging, photons are detected by a device called Gamma camera. This device is composed of three main parts: i) a collimator which is an array of holes separated by high attenuating material (such as lead) and used to limit the angle of acceptance of photons; ii) a scintillator and iii) an array of photomultiplier tubes. Advanced hardware and software systems are used to estimate position and the energy of the emitted photon.

Clinical applications of SPECT imaging are dopamine receptor imaging for the study of Parkinson's disease or multiple system artophies using the 123I-IBZM radiotracer [49], early detection of cardiac amyloids [50] or to detect bone metastasis using ^{99m}Tc bonded to phosphorous compounds [51].

2.1.4 Multimodal imaging modalities

It is important to note that the image modalities listed above are only a part of the existing ones. As discussed in the introduction, when it comes to multimodality, multiple modalities are integrated into one system. In reality, one can talk of multimodality even when integrating information from two types of images coming from the same imaging modality. To give an example, there are multimodal systems that provide for the simultaneous extraction of information from the elastonographic image and the B-mode image, in that case we talk of multimodal ultrasound imaging. Another classic example is multiparametric-MRI (mp-MRI) where anatomical MRI sequences are combined with functional MRI sequences (e.g. T2w + DW or T1w + DCE).

First multimodal modalities were *software fusion*, i.e. modalities were acquired using different acquisition system and fused using registration algorithms. In the late 1990s, the first hybrid acquisition system was proposed, it was a PET/CT scanner. Nowadays, hybrid acquisition systems which allow for the "simultaneous" acquisition of images of different modalities using the same acquisition system (*hardware fusion*) are increasingly used. The advantage of using hybrid acquisition technology is the intrinsic co-registration of modalities due to significantly lower changes in patient position. The most diffuse hybrid scanners are PET/CT and SPECT/CT scanners. Recently a new hybrid acquisition technology has been proposed for multimodal PET/MRI images. Further, efforts for combining MRI with ultrasound are currently under development [52, 53].

2.2 Multimodal image analysis strategies

Figure 2.11 shows a general pipeline of a system for the quantitative tissue characterization in unimodal images. The first step is the loading of the image, subsequently, there is the segmentation process, in which the regions-of-interest of the image are identified. At that point, feature extraction is performed to compute descriptive features of the tissue in the regions-of-interest, for example texture descriptors. The last step could be whether classification (in the case of differentiation systems, for example healthy/pathological differentiation) or statistical analysis (in the case of stratification systems, for example tumor aggressiveness evaluation).

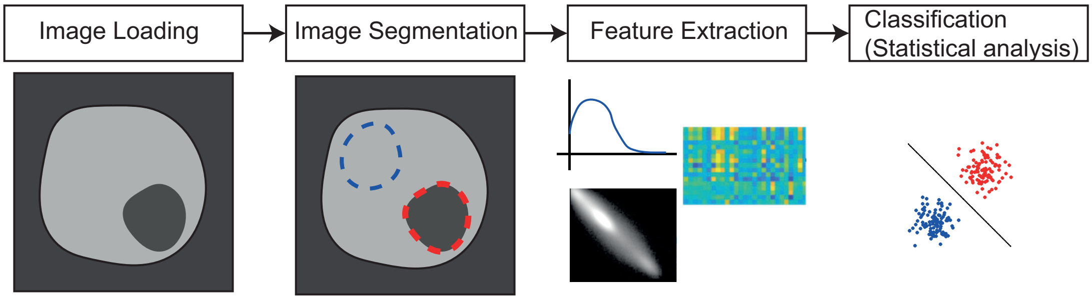


Figure 2.11: General pipeline of a unimodal system for tissue characterization

In a multimodal system, for simplicity let us consider only two images of different modality, multiple strategies are feasible based on which level data fusion occurs [54] (figure 2.12). Three levels of fusion can be identified: i) classifier level fusion; ii) feature level fusion; iii) image level fusion. As the level of data fusion affects the final outcome, it is important to choose the right strategy. This choice has to be made considering several factors, whether the two images can be registered or not, how much information do the images share, a-priori knowledge about significant patterns in images.

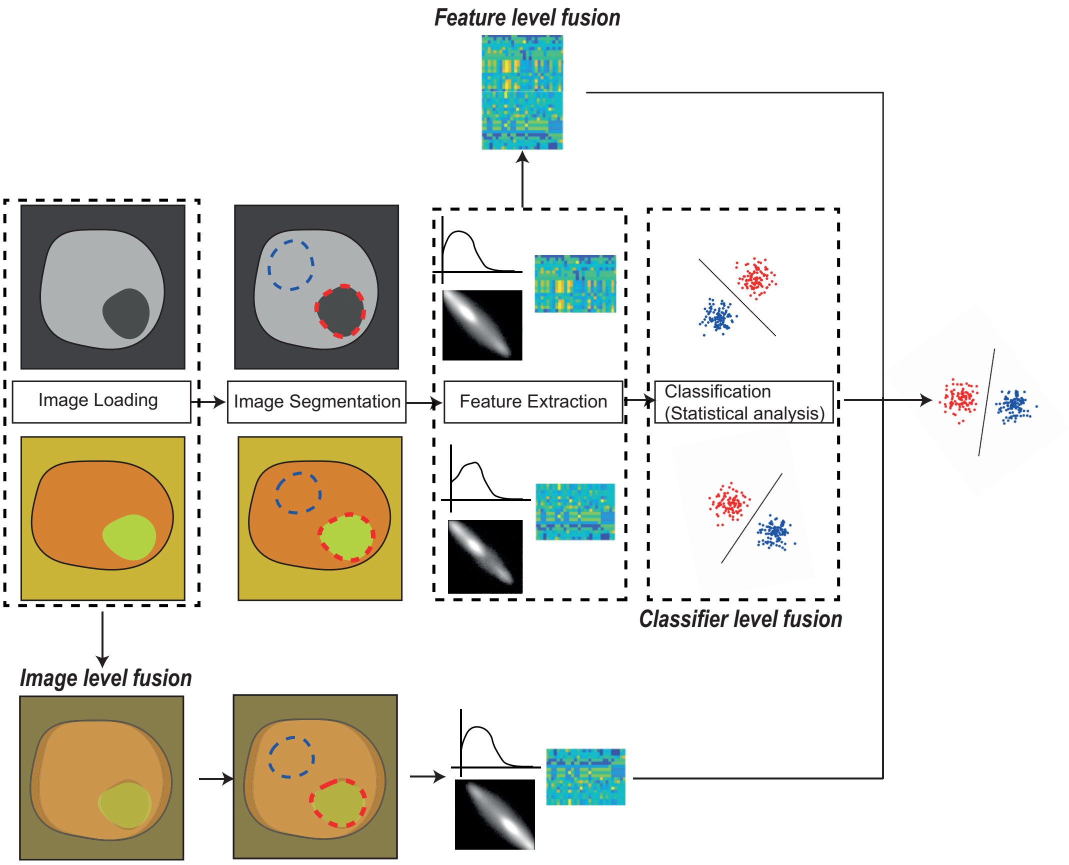


Figure 2.12: General pipeline of a multimodal system for tissue characterization

2.2.1 Classifier level fusion

In this strategy, both images are processed separately, two classifiers are trained separately using the unimodal feature sets, then, final output is obtained by fusing the output of both classifiers (for example majority voting or weighted sum of single outputs). In rare cases, this strategy may be preferred to others, e.g. when image registration between the two imaging modalities is unfeasible (different acquisition views, lack of anatomical reference) and when the nature of the extracted features makes it impossible or incorrect to merge the two modalities or more commonly, when the designed system must be able to produce a final outcome also when an imaging modality is missing for any reason. This strategy was applied to predict distant failure in lung cancer after stereotactic body radiation in PET/CT images [55]. Other examples of this strategy were only found in multimodal image segmentation studies [56, 57].

2.2.2 Feature level fusion

Feature level fusion means that features previously extracted from single imaging modalities are concatenated or combinated in some way. This feature set is then used to train a unique classifier model. Features may be combined using either principal component analysis (PCA) to reduce data redundancy or multivariate analysis or simple average. This strategy is more common than the classifier level fusion and its advantage is the ability to merge information from different modalities into a single set of features. Applications include prognostic prediction for nasopharyngeal carcinoma in ^{18}F-FDG-PET/CT images [58], radiotherapy treatment response quantification for lung cancer in ^{18}F-FDG-PET/CT images [59], diagnosis of non-palpable breast lesions in digital mammography/DCE-MRI images [60], progression evaluation of mild cognitive impairment and early Alzheimer's disease in ^{18}F-FDG-PET/MRI images [61].

2.2.3 Image level fusion

By adopting this strategy, the two images need to be fused in a single image which combine the information of both modalities. At that point, all the processing is performed on the fused image. Several studies showed that image fusion achieves excellent results thanks to the effects of image sharpening and significant feature enhancement [62]. Most common image fusion methods are based on PCA and Wavelet transform [62]. A limitation of the image level fusion strategy is that it requires accurate image registration, indeed, a misalignment between the images can produce blurring artifacts and image quality deterioration, and this can significantly affect the predictive performance. Anyway, with the development of novel fast and accurate registration algorithms, this strategy is increasingly used. Examples are the fusion of PET/CT images for the prediction of immunotherapy response of non-small cell lung cancer [63], quantification of local metabolic tumor volume in PET/CT fused images [64], the fusion of ^{18}F-FDG-PET/MRI images (wavelet-based) for prediction of lung metastases in soft-tissue sarcomas [65] and for the quantification of mild cognitive impairment progression [66].

2.2.4 Other strategies

Recently, novel strategies which are difficult to assign to the aforementioned categories have been proposed. In [67], authors developed a multiparametric imaging radiomic (mpRAD) framework. This framework was based on the concept of tissue signature which was the array of intensity values corresponding to the same voxel across all image modalities. From the tissue signature map, different matrices, similar to the classic texture matrices, were computed to finally obtain a set of multiparametric textural features. The strategy was tested in mpMRI images of breast cancer, the result was the improvement of performance for diagnosis of breast cancer compared to radiomics features extracted from single modalities.

Furthermore, with the fast growing development of deep learning techniques, efforts have been made on adapting convolutional neural networks (CNN) to multimodal images for tissue segmentation and classification. In [68], thee CNN-based strategies, by fusing layer outputs at different levels, were implemented for soft tissue sarcomas segmentation in PET, CT and MRI images. Regarding tissue classification, different CNN-based strategies especially in the neuroimaging field for Alzheimer's disease staging were proposed which differ in CNN architecture and level of fusion [69].

2.3 References

[1] Thomas G. Mayerhöfer and Jürgen Popp. "Beer's law derived from electromagnetic theory". In: *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* 215 (May 2019), pp. 345–347. ISSN: 13861425. DOI: `10.1016/j.saa.2019.02.103`.

[2] Eric de Groot et al. *Measurement of carotid intima-media thickness to assess progression and regression of atherosclerosis.* May 2008. DOI: `10.1038/ncpcardio1163`.

[3] Jayne Cleve and Marti L. McCulloch. "Conducting a Cardiac Ultrasound Examination". In: *Echocardiography.* Springer International Publishing, 2018, pp. 33–42. DOI: `10.1007/978-3-319-71617-6_2`.

[4] B. Caspi et al. "The contribution of transvaginal ultrasound in the diagnosis of acute appendicitis: An observational study". In: *Ultrasound in Obstetrics and Gynecology* 21.3 (Mar. 2003), pp. 273–276. ISSN: 09607692. DOI: `10.1002/uog.72`.

[5] Fiona Costello and James N. Scott. *Imaging in Neuro-ophthalmology.* Oct. 2019. DOI: `10.1212/CON.0000000000000783`.

[6] P. Vitti and T. Rago. *Thyroid ultrasound as a predicator of thyroid disease.* 2003. DOI: `10.1007/BF03347031`.

[7] Giorgia Spaggiari, Antonio R.M. Granata, and Daniele Santi. "Testicular ultrasound inhomogeneity is an informative parameter for fertility evaluation". In: *Asian Journal of Andrology* 22.3 (May 2020), pp. 302–308. ISSN: 17457262. DOI: `10.4103/aja.aja_67_19`.

[8] S. Lenz et al. "Ultrasonic testicular texture and size in 444 men from the general population: Correlation to semen quality". In: *European Urology* 24.2 (1993), pp. 231–238. ISSN: 03022838. DOI: `10.1159/000474300`.

[9] Franklin N. Tessler et al. “ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee”. In: *Journal of the American College of Radiology* 14.5 (May 2017), pp. 587–595. ISSN: 1558349X. DOI: 10.1016/j.jacr.2017.01.046.

[10] Andreas Maier et al. *Medical Imaging Systems - An Introductory Guide*. Vol. 11111. 2018, p. 259. ISBN: 978-3-319-96519-2.

[11] A. Alexandrov and M. Joseph. “Transcranial Doppler: An Overview of its Clinical Applications”. In: *The Internet Journal of Emergency and Intensive Care Medicine* 4.1 (June 2012). ISSN: 1092-4051. DOI: 10.5580/909.

[12] S. Battaglia et al. “Color Doppler ultrasonography of the abdominal aorta”. In: *Journal of Ultrasound* 13.3 (Sept. 2010), pp. 107–117. ISSN: 19713495. DOI: 10.1016/j.jus.2010.10.001.

[13] Abdul Latif Mohamed et al. “The prevalence of diastolic dysfunction in patients with hypertension referred for echocardiographic assessment of left ventricular function”. In: *Malaysian Journal of Medical Sciences* 11.1 (2004), pp. 66–74. ISSN: 1394195X.

[14] Jin Chung et al. “Clinical applications of doppler ultrasonography for thyroid disease: Consensus statement by the korean society of thyroid radiology”. In: *Ultrasonography* 39.4 (Oct. 2020), pp. 315–330. ISSN: 22885943. DOI: 10.14366/usg.20072.

[15] Jiacheng Niu et al. “Correlation between doppler ultrasound blood flow parameters and angiogenesis and proliferation activity in breast cancer”. In: *Medical Science Monitor* 25 (Sept. 2019), pp. 7035–7041. ISSN: 16433750. DOI: 10.12659/MSM.914395.

[16] Xi Wei et al. “The role of contrast-enhanced ultrasound (CEUS) in the early assessment of microvascularization in patients with aggressive B-cell lymphoma treated by rituximab-CHOP: A preliminary study (13-234R)”. In: *Clinical Hemorheology and Microcirculation* 58.2 (2014), pp. 363–376. ISSN: 18758622. DOI: 10.3233/CH-131773.

[17] Xin Sui et al. “Contrast-enhanced ultrasound and real-time elastography in the differential diagnosis of malignant and benign thyroid nodules”. In: *Experimental and Therapeutic Medicine* 12.2 (Aug. 2016), pp. 783–791. ISSN: 17921015. DOI: 10.3892/etm.2016.3344.

[18] Lukas Philipp Beyer et al. “Characterization of Focal Liver Lesions using CEUS and MRI with Liver-Specific Contrast Media: Experience of a Single Radiologic Center”. In: *Ultraschall in der Medizin* 38.6 (Dec. 2017), pp. 619–625. ISSN: 14388782. DOI: 10.1055/s-0043-105264.

[19] Arend F.L. Schinkel, Mathias Kaspar, and Daniel Staub. “Contrast-enhanced ultrasound: clinical applications in patients with atherosclerosis”. In: *International Journal of Cardiovascular Imaging* 32.1 (Jan. 2016), pp. 35–48. ISSN: 15730743. DOI: `10.1007/s10554-015-0713-z`.

[20] Priscilla Machado et al. “A novel microvascular flow technique: Initial results in thyroids”. In: *Ultrasound Quarterly* 32.1 (Mar. 2016), pp. 67–74. ISSN: 15360253. DOI: `10.1097/RUQ.0000000000000156`.

[21] Xiao Yun Xiao et al. “Superb microvascular imaging in diagnosis of breast lesions: A comparative study with contrast-enhanced ultrasonographic microvascular imaging”. In: *British Journal of Radiology* 89.1066 (2016). ISSN: 00071285. DOI: `10.1259/bjr.20160546`.

[22] Dong Ho Lee, Jae Young Lee, and Joon Koo Han. “Superb microvascular imaging technology for ultrasound examinations: Initial experiences for hepatic tumors”. In: *European Journal of Radiology* 85.11 (Nov. 2016), pp. 2090–2095. ISSN: 18727727. DOI: `10.1016/j.ejrad.2016.09.026`.

[23] J. L. Gennisson et al. *Ultrasound elastography: Principles and techniques.* May 2013. DOI: `10.1016/j.diii.2013.01.022`.

[24] Naomi Winn, Radhesh Lalam, and Victor Cassar-Pullicino. “Sonoelastography in the musculoskeletal system: Current role and future directions”. In: *World Journal of Radiology* 8.11 (2016), p. 868. ISSN: 1949-8470. DOI: `10.4329/wjr.v8.i11.868`.

[25] Ako Itoh et al. *Breast disease: Clinical application of US elastography for diagnosis.* May 2006. DOI: `10.1148/radiol.2391041676`.

[26] Chong-Ke Zhao and Hui-Xiong Xu. “Ultrasound elastography of the thyroid: principles and current status”. In: *Ultrasonography* 38.2 (Apr. 2019), pp. 106–124. ISSN: 2288-5919. DOI: `10.14366/usg.18037`.

[27] F. Bloch, W. W. Hansen, and Martin Packard. *Nuclear induction.* Feb. 1946. DOI: `10.1103/PhysRev.69.127`.

[28] E. M. Purcell, H. C. Torrey, and R. V. Pound. *Resonance absorption by nuclear magnetic moments in a solid.* Jan. 1946. DOI: `10.1103/PhysRev.69.37`.

[29] Vadim Kuperman. *Magnetic Resonance Imaging: Physical Principles and Applications.* 1st. San Diego: Academic Press, 1999. ISBN: 0124291503.

[30] Bhawan K. Paunipagar and D. D. Rasalkar. “Imaging of articular cartilage”. In: *Indian Journal of Radiology and Imaging* 24.3 (2014), pp. 237–248. ISSN: 19983808. DOI: `10.4103/0971-3026.137028`.

[31] M. Just and M. Thelen. "Tissue characterization with T1, T2, and proton density values: Results in 160 patients with brain tumors". In: *Radiology* 169.3 (1988), pp. 779–785. ISSN: 00338419. DOI: `10.1148/radiology.169.3.3187000`.

[32] Andrew J. Taylor et al. *T1 Mapping Basic Techniques and Clinical Applications*. Jan. 2016. DOI: `10.1016/j.jcmg.2015.11.005`.

[33] Manus J. Donahue et al. "Evaluation of human brain tumor heterogeneity using multiple T 1-based MRI signal weighting approaches". In: *Magnetic Resonance in Medicine* 59.2 (2008), pp. 336–344. ISSN: 15222594. DOI: `10.1002/mrm.21467`.

[34] E. Le Rumeur et al. "Multiparametric classification of muscle T1 and T2 relaxation times determined by magnetic resonance imaging. The effects of dynamic exercise in trained and untrained subjects". In: *British Journal of Radiology* 67.794 (1994), pp. 150–156. ISSN: 00071285. DOI: `10.1259/0007-1285-67-794-150`.

[35] Kelvin K. Lim et al. *Clinical applications of 3D T2-weighted MRI in pelvic imaging*. Apr. 2014. DOI: `10.1007/s00261-014-0124-y`.

[36] Kanji Nakai et al. "An analysis of T2 mapping on brain tumors." In: *Acta neurochirurgica. Supplement* 118 (2013), pp. 195–199. ISSN: 00651419. DOI: `10.1007/978-3-7091-1434-6_36`.

[37] R. N. Sener. "Diffusion MRI: Apparent diffusion coefficient (ADC) values in the normal brain and a classification of brain disorders based on ADC values". In: *Computerized Medical Imaging and Graphics* 25.4 (July 2001), pp. 299–326. ISSN: 08956111. DOI: `10.1016/S0895-6111(00)00083-5`.

[38] Alexey Surov, Hans Jonas Meyer, and Andreas Wienke. "Correlations between Apparent Diffusion Coefficient and Gleason Score in Prostate Cancer: A Systematic Review". In: *European urology oncology* 3.4 (Aug. 2020), pp. 489–497. ISSN: 25889311. DOI: `10.1016/j.euo.2018.12.006`.

[39] Ahmed A.K. Abdel Razek et al. "Performance of apparent diffusion coefficient of medial and lateral rectus muscles in Graves' orbitopathy". In: *Neuroradiology Journal* 30.3 (June 2017), pp. 230–234. ISSN: 19714009. DOI: `10.1177/1971400917691993`.

[40] Takashi Abe et al. "Diagnosis of brain tumors using dynamic contrast-enhanced perfusion imaging with a short acquisition time". In: *SpringerPlus* 4.1 (2015). ISSN: 21931801. DOI: `10.1186/s40064-015-0861-6`.

[41] Ingrid S. Gribbestad et al. "Comparative signal intensity measurements in dynamic gadolinium-enhanced MR mammography". In: *Journal of Magnetic Resonance Imaging* 4.3 (1994), pp. 477–480. ISSN: 15222586. DOI: `10.1002/jmri.1880040339`.

[42] Rose M. Berman et al. “DCE MRI of prostate cancer”. In: *Abdominal Radiology* 41.5 (May 2016), pp. 844–853. ISSN: 23660058. DOI: `10.1007/s00261-015-0589-3`.

[43] Carlo Pierpaoli. “Quantitative brain MRI”. In: *Topics in Magnetic Resonance Imaging* 21.2 (Apr. 2010), p. 63. ISSN: 08993459. DOI: `10.1097/RMR.0b013e31821e56f8`.

[44] Ronald Boellaard et al. “FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 42.2 (2015), pp. 328–354. ISSN: 16197089. DOI: `10.1007/s00259-014-2961-x`.

[45] James W. Fletcher et al. “Recommendations on the use of 18F-FDG PET in oncology”. In: *Journal of Nuclear Medicine* 49.3 (2008), pp. 480–508. ISSN: 01615505. DOI: `10.2967/jnumed.107.047787`.

[46] Hongming Zhuang and Ion Codreanu. *Growing applications of FDG PET-CT imaging in non-oncologic conditions.* 2015. DOI: `10.7555/JBR.29.20140081`.

[47] T. Brucke et al. *SPECT and PET imaging of the dopaminergic system in Parkinson’s disease.* 2000. DOI: `10.1007/pl00007769`.

[48] Michael M. Graham et al. “68Ga-DOTATOC imaging of neuroendocrine tumors: A systematic review and metaanalysis”. In: *Journal of Nuclear Medicine* 58.9 (Sept. 2017), pp. 1452–1458. ISSN: 2159662X. DOI: `10.2967/jnumed.117.191197`.

[49] T. Brucke et al. “Dopamine D2 receptor imaging with SPECT: Studies in different neuropsychiatric disorders”. In: *Journal of Cerebral Blood Flow and Metabolism* 11.2 (1991), pp. 220–228. ISSN: 0271678X. DOI: `10.1038/jcbfm.1991.53`.

[50] Sabahat Bokhari et al. *Nuclear imaging modalities for cardiac amyloidosis.* 2014. DOI: `10.1007/s12350-013-9803-2`.

[51] Takayoshi Uematsu et al. “Comparison of FDG PET and SPECT for Detection of Bone Metastases in Breast Cancer”. In: *American Journal of Roentgenology* 184.4 (Apr. 2005), pp. 1266–1273. ISSN: 0361-803X. DOI: `10.2214/ajr.184.4.01841266`.

[52] Haim Azhari, Robert R. Edelman, and David Townsend. *Multimodal imaging and hybrid scanners.* 2007. DOI: `10.1155/2007/45353`.

[53] Laura Curiel, Rajiv Chopra, and Kullervo Hynynen. “Progress in multimodality imaging: truly simultaneous ultrasound and magnetic resonance imaging.” In: *IEEE transactions on medical imaging* 26.12 (Dec. 2007), pp. 1740–6. ISSN: 0278-0062. DOI: `10.1109/tmi.2007.903572`.

[54] Wenbing Lv et al. “Multi-level multi-modality fusion radiomics: Application to PET and CT imaging for prognostication of head and neck cancer”. In: *IEEE Journal of Biomedical and Health Informatics* 24.8 (2020), pp. 2268–2277. ISSN: 21682208. DOI: 10.1109/JBHI.2019.2956354.

[55] Zhiguo Zhou et al. “Constructing multi-modality and multi-classifier radiomics predictive models through reliable classifier fusion”. In: *arXiv* (2017), pp. 1–13.

[56] Hongmin Cai et al. “Probabilistic Segmentation Of Brain Tumors Based on Multi-modality magnetic resonance images”. In: *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 2007, pp. 600–603. ISBN: 1-4244-0671-4. DOI: 10.1109/ISBI.2007.356923.

[57] Bjoern H. Menze et al. “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”. In: *IEEE Transactions on Medical Imaging* 34.10 (Oct. 2015), pp. 1993–2024. ISSN: 1558254X. DOI: 10.1109/TMI.2014.2377694.

[58] Wenbing Lv et al. “Radiomics Analysis of PET and CT Components of PET/CT Imaging Integrated with Clinical Parameters: Application to Prognosis for Nasopharyngeal Carcinoma”. In: *Molecular Imaging and Biology* 21.5 (Oct. 2019), pp. 954–964. ISSN: 18602002. DOI: 10.1007/s11307-018-01304-3.

[59] Manushka Vaidya et al. “Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer”. In: *Radiotherapy and Oncology* 102.2 (Feb. 2012), pp. 239–245. ISSN: 01678140. DOI: 10.1016/j.radonc.2011.10.014.

[60] Shujun Chen et al. “A new application of multimodality radiomics improves diagnostic accuracy of nonpalpable breast lesions in patients with microcalcifications-only in mammography”. In: *Medical Science Monitor* 25 (2019), pp. 9786–9793. ISSN: 16433750. DOI: 10.12659/MSM.918721.

[61] Marco Lorenzi et al. “Multimodal Image Analysis in Alzheimer’s Disease via Statistical Modelling of Non-local Intensity Correlations”. In: *Scientific Reports* 6.1 (Apr. 2016), p. 22161. ISSN: 2045-2322. DOI: 10.1038/srep22161.

[62] Samaneh Mazaheri et al. “Hybrid pixel-based method for cardiac ultrasound fusion based on integration of PCA and DWT”. In: *Computational and Mathematical Methods in Medicine* 2015 (2015). ISSN: 17486718. DOI: 10.1155/2015/486532.

[63] Wei Mu et al. "Radiomic biomarkers from PET/CT multi-modality fusion images for the prediction of immunotherapy response in advanced non-small cell lung cancer patients". In: *Medical Imaging 2018: Computer-Aided Diagnosis*. Ed. by Kensaku Mori and Nicholas Petrick. Vol. 10575. SPIE, Feb. 2018, p. 136. ISBN: 9781510616394. DOI: 10.1117/12.2293376.

[64] Sadegh Riyahi et al. "Quantification of Local Metabolic Tumor Volume Changes by Registering Blended PET-CT Images for Prediction of Pathologic Tumor Response". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11076 LNCS (Aug. 2018), pp. 31–41. arXiv: 1808.08312.

[65] M. Vallières et al. "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities". In: *Physics in Medicine and Biology* 60.14 (July 2015), pp. 5471–5496. ISSN: 13616560. DOI: 10.1088/0031-9155/60/14/5471.

[66] Hucheng Zhou et al. "Dual-model radiomic biomarkers predict development of mild cognitive impairment progression to Alzheimer's disease". In: *Frontiers in Neuroscience* 13.JAN (2019). ISSN: 1662453X. DOI: 10.3389/fnins.2018.01045.

[67] Vishwa S. Parekh and Michael A. Jacobs. "Multiparametric radiomics methods for breast cancer tissue characterization using radiological imaging". In: *Breast Cancer Research and Treatment* 180.2 (Apr. 2020), pp. 407–421. ISSN: 0167-6806. DOI: 10.1007/s10549-020-05533-5.

[68] Zhe Guo et al. "Deep Learning-Based Image Segmentation on Multimodal Medical Imaging". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.2 (2019), pp. 162–169. ISSN: 2469-7311. DOI: 10.1109/trpms.2018.2890359.

[69] Yan Xu. "Deep Learning in Multimodal Medical Image Analysis". In: vol. 11837. April. 2019, pp. 193–200. ISBN: 978-3-030-32961-7. DOI: 10.1007/978-3-030-32962-4_18.

Chapter 3

Thyroid nodule characterization

3.1 Introduction

According to American Cancer Society, the incidence of thyroid cancer is rapidly increasing and the chance of being diagnosed with thyroid nodule has tripled in the past three decades [1]. Thyroid nodules are pathologies characterized by the formation of solid or liquid lumps within the thyroid gland caused by anomalous growth of thyroid cells [2]. Most thyroid nodules are benign and do not show any sign of cancerous cell growth. Previous studies showed a malignancy rate of thyroid nodules from 5 % to 10 % [3, 4].

With the high incidence of thyroid nodules and the low malignancy rate, the risk of overdiagnosis is important [5].

The standard method for diagnosing a thyroid nodule is by ultrasound-guided fine-needle aspiration (US-FNAB), where a fine-needle is guided by ultrasound imaging to collect samples of tissue from the thyroid nodule. The result of FNAB is the cyological classification. Different cytological classification systems exist (e.g. Thy classification proposed by the British Thyroid Association, Bethesda classification [6]). In Italy, the standard classification is guided by the SIAPEC-IAP Italian Consensus Working Group whose most recent guideline was proposed in 2014 [7]. According to the SIAPEC-IAP, the tissue sample is classified in seven cytological subtypes:

- **TIR 1** malignancy risk can not be quantified due to low quality sample, a repeated FNAB is recommended.

- **TIR 1C** the sample contains fluid from a cyst but low presenza of epithelial cells or colloid to confirm cyst type. Low malignancy risk but a repeated FNAB could be recommended depending on the clinical context.

- **TIR 2** goiter with malignancy risk lower than 3%. Follow up is recommended.

- **TIR 3A** high cellular density, not enough microfollicular structures to assess follicular neoplasm. Malignancy risk is lower than 10%. Recommendations are follow up or repeated FNAB.
- **TIR 3B** high cellular density and presence of microfollicular trabecular structures which suggest follicular neoplasm or presence of cellular nuclei alterations suggesting papillary carcinoma. Malignancy risk higher than 15-30%. Recommendation is surgical treatment.
- **TIR 4** high suspect of papillary carcinoma. Malignancy risk higher than 60-80%.
- **TIR 5** samples with malignancy risk higher than 95% including papillary carcinoma, medullary carcinoma, anaplastic carcinoma, lymphoma and metastatic neoplasia.

Although cytological examination does not provide a definitive diagnostic assessment, each cytological class leads to different treatment strategies. The classes with most diagnostic uncertainty are TIR3A and TIR3B. In the previous Italian Society of Pathological Anatomy and Diagnostic Citology - International Academy of Pathology (SIAPEC-IAP) classification (2007), these two classes were merged into a unique undetermined cytology TIR3 which in most cases was treated with surgery, leading over-treatment since only half of TIR3 nodules were actually malignant [8]. Although the new cytological classification provides better detection accuracy of thyroid nodule malignancy, diagnostic efficiency of TIR3B class is still under debate due to its variability in malignancy rate. Improving the characterization of these nodules with alternative techniques may lead to a reduction of clinical cost thanks to avoidance of unnecessary surgeries while improving patients' lifestyle.

It is well known that tumor vascularization plays a fundamental role in cancer growth, invasion, and metastasis [9]. In fact, angiogenesis (i.e., the formation of new vessels) is a key process during the early or middle stages of cancer, since they are highly oxygen demanding and therefore produce a large quantity of angiogenetic growth factors [10]. Indeed, papillary carcinomas have a vessel density that is three-fold higher than the adjacent normal thyroid tissue [11]. Further, also morphology of intratumoral vessels in papillary carcinoma is relevant: these vessels typically present a more tortuous vascularization with respect to those in benign nodules [11]. Regarding follicular carcinomas, it was previously shown that whereas degree of vascularity between the follicular adenomas and the follicular carcinomas is relatively similar, vascular distribution, meaning the spatial distribution of vessels throughout the nodule, is a relevant factor for their differentiation [11]. In overall, more than 40% of nodules that present a central hypervascularity are found to be malignant, whereas a benign marker is a predominant peripheral flow [12].

Numerous imaging techniques are clinically used to assess tumor vascularization, such as magnetic resonance imaging (MRI) [13], contrast-enhanced computed tomography (CT) [14], and ultrasonography (US). Compared to MRI and CT, ultrasonography has emerged as an optimal imaging modality, despite its lower image resolution, thanks to being inexpensive and not using ionizing radiation. Furthermore, the use of the Power-Doppler ultrasound (PDUS) and Contrast-Enhanced ultrasound (CEUS) gives forth important information of vasculature. The additional use of a mechanical 3D ultrasound probe can further improve sensitivity in vascularization quantification, as the probe does not require movement and the acquisition of the entire nodule volume allows a more complete vascularization analysis and allows the differentiation between central and peripheral blood flow in a 3D context.

Recent research has focused greatly on the quantification of the vascular pattern in numerous imaging modalities, using quantitative vascular parameters that can be employed to differentiate between benign and malignant or healthy and diseased tissue [15, 16]. In this respect, in the specific field of thyroid nodule differentiation using ultrasonography, many studies have focused on the use of only 2D images. Specifically, Lyshchik et al. [17] analyzed 86 thyroid nodules on 2D PDUS images and extracted parameters that quantify the vascular density and the strength of flow. Sultan et al [18] quantified the vascular network and pattern in three different regions of the nodule (central, rim and surrounding) in 2D Color Doppler US images, demonstrating that the vascular density and the flow velocity in the central region were significant ($P < 0.05$) for thyroid nodule differentiation while not in the whole nodule. Finally, Baig et al. [19] showed that vascular density computed in the central and peripheral regions combined with other sonographic features can increase the diagnostic accuracy of thyroid nodule. Some authors proposed the use of an additional architectural feature, number of penetrating vessels, which quantifies the number of vessels that cross the border between the central and peripheral region and resulted to be associated with malignancy [20, 21]. However, these methods were limited to 2D images of the nodule. Previous studies using 3D CEUS and or PDUS images can be found in the works by Molinari et al. [22], where the entire 3D vascular network of the thyroid nodule was reconstructed using 3D CEUS images and by Caresio et al. [23] ,who extended the work and analyzed both 3D CEUS and PDUS images of the nodule. Both of these studies that worked on 3D images computed the vasculature skeleton from the acquired images and further calculated various quantitative features on the skeleton that take into consideration the morphology and the tortuosity of the nodule vasculature, which were proven able to discriminate between malignant and benign nodules. The features in these studies, however, were calculated in volumes-of-interest throughout the entire nodule, without distinguishing between the central and the peripheral vasculature.

Objectives The aim of this chapter is to differentiate between TIR3A with low suspect of malignancy and TIR3B with high suspect of papillary carcinoma by quantitative characterization of the nodule vascularization in terms of flow, vascular architecture, and vessel tortuosity using 3D PDUS images. The novelty in this study is the use of a fully 3D approach which takes into consideration the entire volume of the nodule, and the subsequent distinction between central and peripheral vasculature which is determinant in diagnostical differentiation between benign thyroid nodules and papillary carcinomas. An innovative method to quantify the number of penetrating vessels in 3D images was proposed. Although in this research phase the use of CEUS images has not yet been introduced and malignancy of 10 out 15 TIR3B nodules has not yet been histologically confirmed, this is the first methodology proposed for the 3D quantitative characterization of the vascularization of the nodules and could have a strong impact in the improvement of diagnostic differentiation between benign and malignant nodules. This study is a part of a project which was born through the collaboration with the endocrinology division of the Molinette and Mauriziano hospitals (Torino, Italy) with the final aim of improving the diagnosis of thyroid nodules with undetermined TIR3 cytology through the quantitative description of the vascular network from functional ultrasound images.

3.2 Materials and methods

The proposed method were developed and validated on a 2.21 GHz quad-core and 16GB RAM.

Patients A total of 62 patients with a thyroid nodule underwent PDUS and CEUS examination. Only patients whose nodules were cytologically examined and with TIR3A and TIR3B outcomes were included in this study. The final database was made up of 30 thyroid nodule PDUS volumes acquired on 30 patients (15 TIR3A and 15 TIR3B). High suspiciousness of papillary carcinoma thyroid nodules for TIR3B nodules were evaluated looking to the presence of nuclear changes suggestive of papillary carcinoma by an expert pathologist. For 5 out of the 15 TIR3B nodules, papillary carcinoma was histologically confirmed. The mean age of the patients was 55.5 ± 14.7 years, and there was a prevalence of females compared to males (females = 18, 60%). The average dimensions of the nodules in terms of longitudinal diameter (LD), anteroposterior diameter (AD) and transversal diameter (TD) were as follows: TIR3A, LD = 21.77 ± 11.25 mm, AD = 14.55 ± 7.5 mm, TD = 17.91 ± 9.47 mm; TIR3B: LD = 25.35 ± 14 mm, AD = 18.25 ± 9.89 mm, TD = 21.58 ± 13.65 mm. The ultrasound volumes were acquired at two hospitals in Torino, Italy, specifically the Molinette A.O.U. Città della Salute e della Scienza and the A.O.U. Ordine Mauriziano di Torino. The study complies with the Declaration of Helsinki and all patients were informed of the study and signed an informed consent before being included.

Image acquisition Three-dimensional PDUS scans were acquired on all patients using the same experimental protocol. Two expert operators (R.G. with more than 30 years of experience, M.D. with more than 20 years of experience) acquired all volumes. The same model ultrasound device was used at both centers, specifically a MyLab 70 (Esaote, Genova, Italy) equipped with a 4-13 MHz 3D linear-volumetric array transducer (BL433). This transducer allows a 3D scan to be done without physically moving the probe. The scanning angle was set to 50°, using a scanning step within the range from 0.24° to 0.30°. The B-mode gain setting was fixed at 50% and the time gain compensation was maintained neutral. The Power Doppler gain was adjusted and standardized by initial increasing the gain until noise appeared and then gradually decreasing the gain until the noise disappeared. Finally, the central frequency was kept at 5 MHz, the wall filter was set to 4 Hz, and the pulse repetition frequency was fixed at 1 kHz.

Thyroid nodule segmentation Segmentation of the thyroid nodule was essential for two reasons: i) to extract features only within the vascular network of the nodule; ii) to allow the separation of the peripheral and the central part of the nodule. Each thyroid nodule was segmented as following:

1. The initial PDUS volume was sub-sampled reducing the number of frames to 30% of the total number of frames (on average about 40 frames per volume), so that the manual segmentation could be performed on a reasonable number of images;
2. the sub-sampled volume was manually segmented frame-by-frame by an expert operator (R.G.) creating a 3D binary mask;
3. the mask was up-sampled to the original number of frames with a cubic interpolation.

3.2.1 Vascular network extraction

The method adopted to extract the vascular network of the nodule consisted of the following processing steps:

1. The Doppler signal, which is represented in light blue scale, was extracted from the PDUS volume (figures.3.1A and 3.1E) by computing the standard deviation of the volume across the color channels (figures 3.1B and 3.1F).

2. Vessel-like structures were enhanced using the filter proposed by Caresio et al. [23]. Similarly to the Frangi filter [24], this filter computes the Hessian of the image by convolving the image with a Gaussian kernel at different σ. For each σ, the vesselness response is computed as:

   $$VR_\sigma(i,j,k) = \begin{cases} 0, & S(i,j,k) = 0 \\ \frac{\lambda_1(i,j,k)}{S(i,j,k)}, & S(i,j,k) \neq 0 \end{cases} \tag{3.1}$$

   where

   $$S(i,j,k) = \sqrt{\lambda_1^2(i,j,k) + \lambda_2^2(i,j,k) + \lambda_3^2(i,j,k)} \tag{3.2}$$

   The final filtered image is obtained as follows:

   $$VR(i,j,k) = \max_{\sigma_{min}<\sigma<\sigma_{max}} VR_\sigma(i,j,k)) \tag{3.3}$$

   In other words, the output image represents the vesselness response at different scales σ (i.e. size of vessels). As suggested by Caresio et al., a range from 1 to 5 was used as σ. The final result is an image where the blood information is enhanced and noise is reduced (Fig. 3.1C and 3.1G).

3. The filtered image was thresholded using Otsu's method to obtain a binary volumetric image where true voxels contain a vessel (Fig. 3.1D and 3.1H).

4. Skeletonization was performed to the binarized volume to extract a topological description of the vessel network. As the nodule vascularization might be very complex, a ridge tracking approach such as the height ridge transversal algorithm [25] would be computationally expensive and fail in complicated branching points. The skeletonization algorithm used in this work is based on the medial axis thinning which is one of the most used in this kind of application [26, 27]. This algorithm iteratively removes border voxels from the object while preserving the topology and the Euler number, until the centerline of the object is obtained. In order to correct some imperfections of the skeleton, isolated voxels were removed. Figure 3.2 shows the result of the skeletonization.

3.2.2 Central and peripheral segmentation

In order to obtain the masks of the peripheral (M_p) and central (M_c) region of the nodule, 3D morphological operators were used. In particular, M_c was obtained by eroding the mask of the nodule (M_n) with a spherical structuring element while M_p was computed as the set subtraction of M_c from M_n. The radius of the spherical structuring element was set to obtain a thickness of the peripheral crown equal to 10% of the radius of the equivalent sphere of the thyroid nodule. This specific

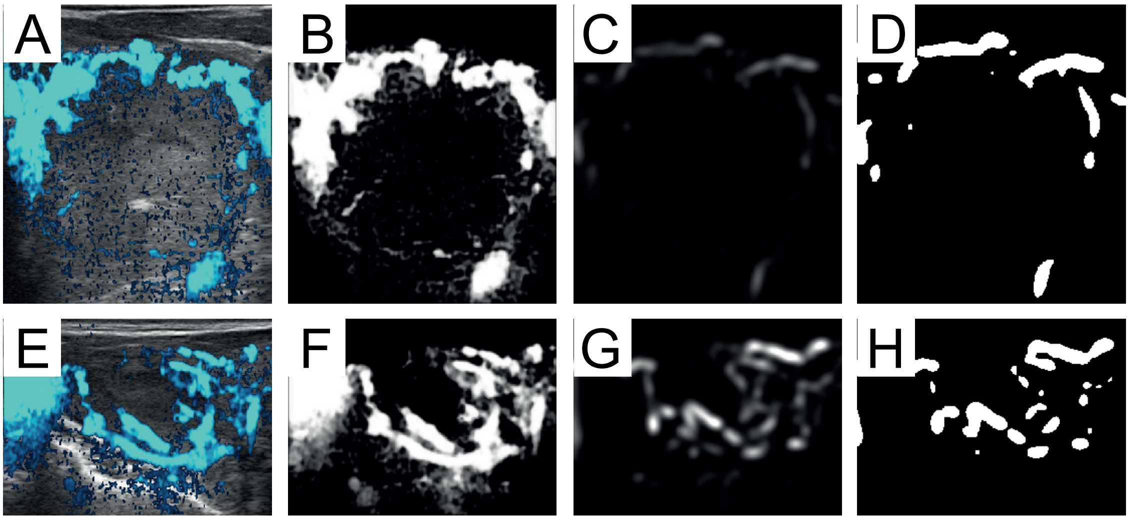


Figure 3.1: Vascular extraction algorithm steps. (A,E) Original PDUS volume frames of a TIR3A and a TIR3B nodule, respectively. (B,F) Extraction of Doppler signal. (C,G) Result of the vesselness filter. (D,H) Image binarization with Otsu thresholding.

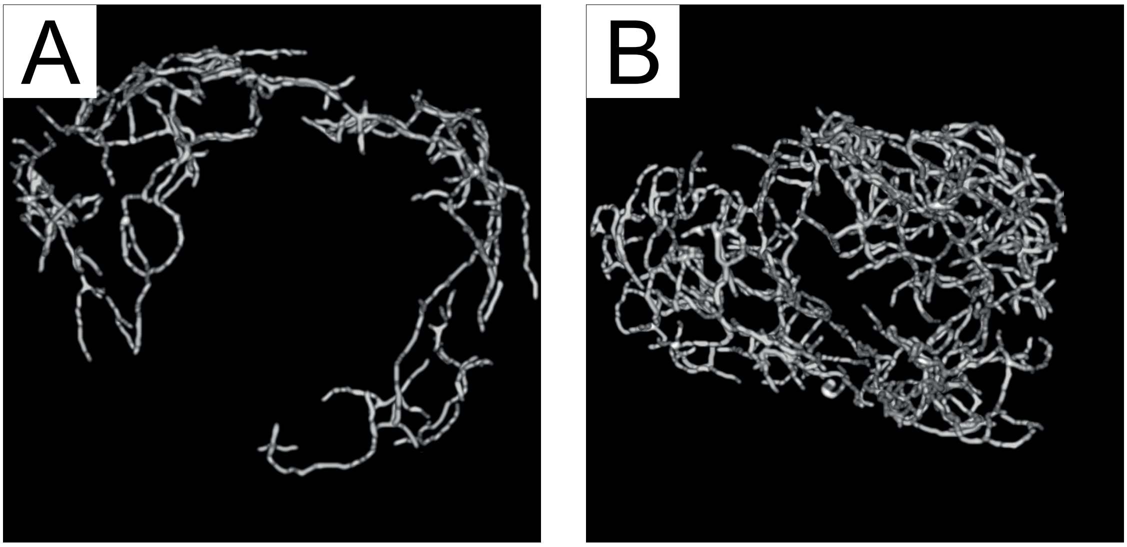


Figure 3.2: Three-dimensional rendering of the final skeleton of the vascular network for a TIR3A nodule (A) and a TIR3B nodule (B).

percentage (Percentage Ratio of the Equivalent Radius, PRER) was determined using an optimization procedure (described in 3.2.4). Figure 3.3 shows an example of the vascularization segmentation in central (in red) and peripheral (in green) regions of a TIR3A (panel A) and a TIR3B nodule (panel B).

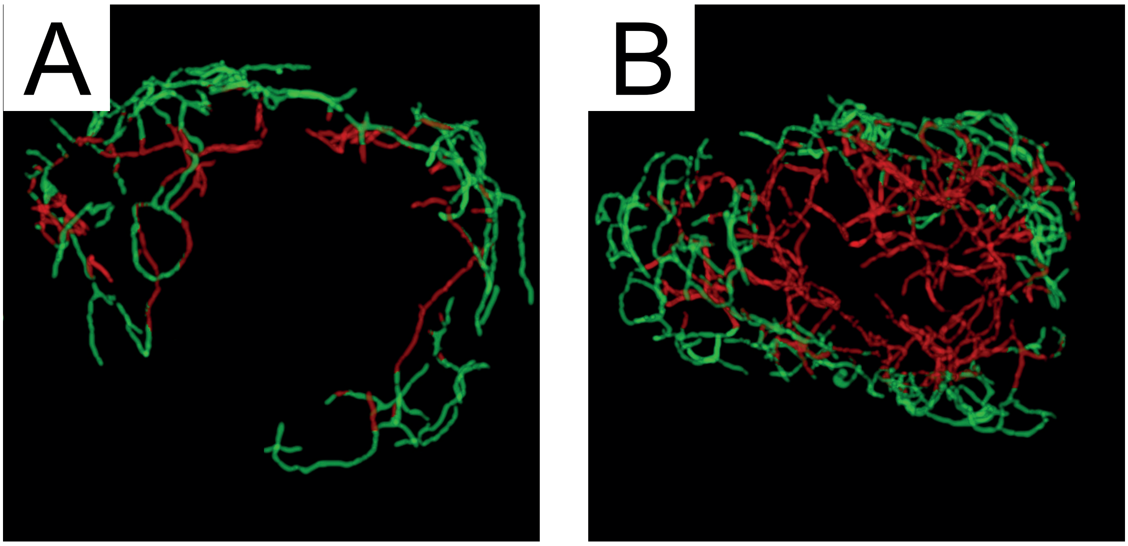


Figure 3.3: Central and peripheral segmentation of the vascular network of a TIR3A (panel A) and TIR3B (panel B) nodule. Green and red lines represent the centerlines of peripheral and central vessels, respectively.

3.2.3 Quantitative features extraction

The extracted quantitative features can be classified in three groups:

Architectural features Cancer angiogenesis process determines changes of the vascular architecture such as increase of vessels and intersections. Furthermore, papillary carcinomas can present vascular invasion from the peripheral to the central area [11]. In order to measure architectural changes the following three features were computed from the vascular network model:

- **Number of vascular Trees (NT)**, which is the number of vessel trees of the skeleton. This feature was computed as the number of 26-connected objects of the skeleton.
- **Number of vascular Branches (NB)**, which is the number of vessels branching from trees of the skeleton; This feature was computed as the number of intersections between the main path (longest path between every pair of endpoints) of each tree and branches.
- **Number of Penetrating Vessels (NPV)**, number of vessels crossing the interface between the peripheral and central regions of the nodule. To compute this feature, the central mask M_c was dilated using a spherical structuring element with radius equal to 5 voxels. The interface mask was computed as the difference between the eroded central mask and the original central mask. Subsequently, a binary volume was obtained by logically intersecting

the interface mask with the skeleton volume. The number of penetrating vessels was defined as the number of 26-connected objects present in this new binary volume.

Tortuosity features Since vascularity of malignant thyroid nodules present higher tortuosity than that of benign ones, tortuosity was measured using the following metrics:

- **Distance Metric (DM)**, which is defined as the ratio between the arc length of the vessel and the Euclidean distance between its endpoints:

$$DM = \frac{\sum_{k=1}^{N-1} \|P_{k+1} - P_k\|}{\|P_N - P_1\|} \tag{3.4}$$

- **Inflection Count Metric (ICM)**, which is the product between the number of inflections, N_{inf}, and the distance metric plus 1:

$$ICM = N_{inf} \cdot \frac{\sum_{k=1}^{N-1} \|P_{k+1} - P_k\|}{\|P_N - P_1\|} + 1 \tag{3.5}$$

- **Sum of Angles Metric (SOAM)**, which is computed as:

$$SOAM = \frac{\sum_{k=1}^{N-3} CP_k}{\sum_{k=1}^{N-2} \|P_{k+1} - P_k\|} \tag{3.6}$$

  where CP_k represents the total angle (torsional and in-plane) at point k; hence, this metric is the total curvature along the curve normalized by the arc length and is measured as radians/voxel.

Tortuosity metrics were extracted from the principal paths of the trees, i.e. excluding the branches. Further, the vascular network skeleton always contains more than one vessel, whereas the previously defined metrics are intended for one single vessel. In order to effectively assign each nodule a vascularization tortuosity descriptor, it is therefore necessary to combine the metrics computed on various vessels into one final value that correctly describes the vessel cluster tortuosity. A simple average of the metric values would erroneously give the same weight to long and short vessels. In order to assign a tortuosity value to a cluster of vessels, the three metrics were computed for each vessel. Then, the final DM and ICM were obtained by summing numerators of DM and ICM, respectively, and dividing by the sum of denominators. For the SOAM, the angles of each vessel were summed and divided by the sum of the arc length of each vessel [15].

Flow features Peripheral and central regions of thyroid nodules present different patterns in terms of blood flow and their quantification is useful for the differentiation between benign and malignant thyroid nodules. Then, blood flow was quantitatively characterized using the following conventional indices:

- **Vascularization Index (VI)**, which describes the vascularization of tissue and is defined as the ratio between the number of voxels with non-zero doppler signal (N_c) and the total number of voxels of the region under analysis (N_t):

$$VI = \frac{N_c}{N_t} \tag{3.7}$$

- **Flow Index (FI)**, which is proportional to the average number of blood cells and is computed as:

$$FI = \frac{\sum_i^{N_c} gl_i}{N_t} \tag{3.8}$$

  where gl_i is the intensity of the i-th voxel with non-zero doppler signal.

Every feature, except for NT, NB and NPV, was calculated for each of the three regions segmented (peripheral, central and total nodule). NT and NB were calculated only for the total nodule since in most cases the vascular trees are extended throughout the nodule and the segmentation in peripheral and central region could create spurious trees and branches. In overall, 18 features were extracted for each nodule. Features extracted from the peripheral and central region will be referred using the subscript “p” and “c”, while for features extracted from the entire nodule, the subscript “t” will be used. Figure 3.4 graphically describes the extraction of architectural and tortuosity features.

3.2.4 Statistical analysis and classification

In this preliminary study, we considered TIR3B nodules which showed high suspectness of papillary carcinoma at the cytological examination as malignant nodules. TIR3A nodules were considered benign nodules. Mean and standard deviation values of features were computed for both groups. Mann-Whitney U-test was performed to test the null hypothesis that data extracted from the two groups were from continuous distributions with equal medians. The level of significance was set to 0.05.

One-way multivariate analysis of variance (MANOVA) was performed to compare the group means. MANOVA combines linearly the initial features to create new variables called canonical variables which maximize the variance among groups. Canonical variables were ordered in decreasing explained variance. Level of significance was set to 0.05.

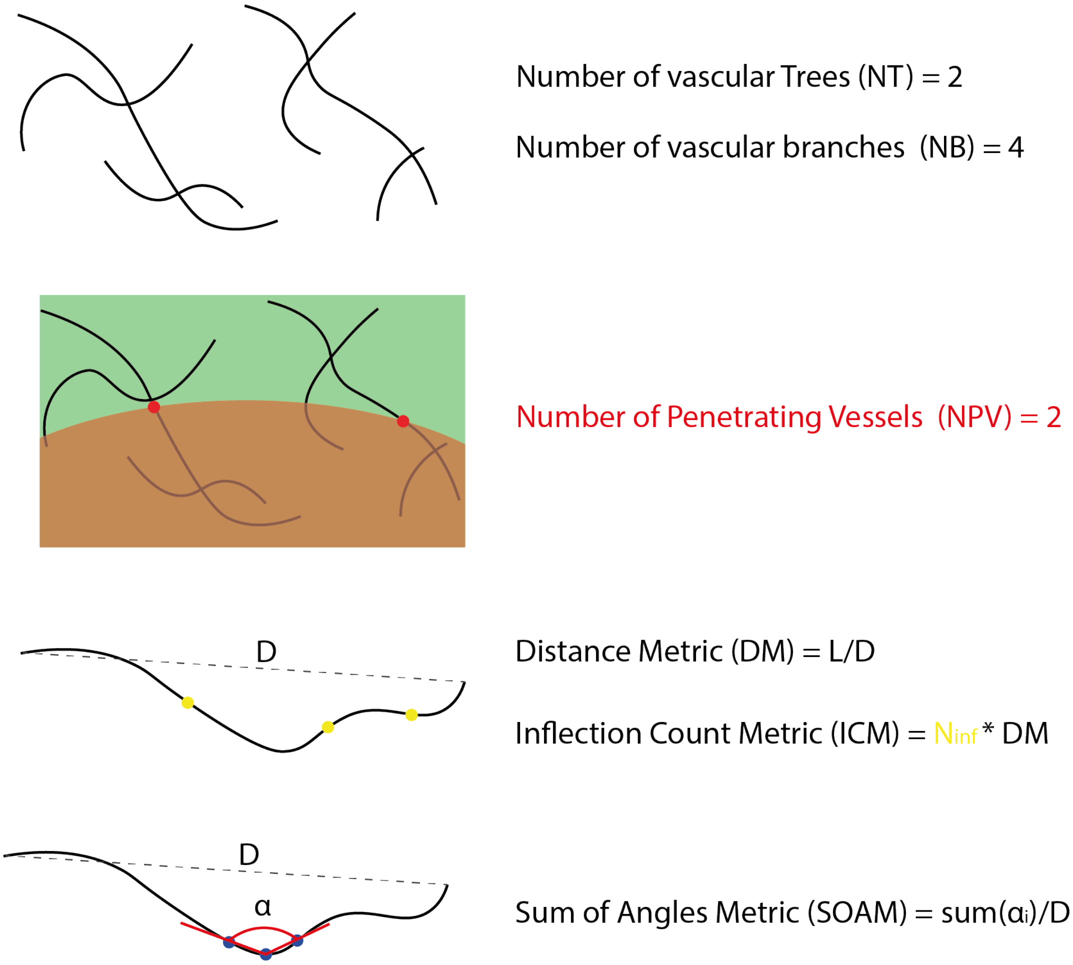


Figure 3.4: Graphical illustration of architectural and tortuosity features extraction.

Automatic classification was performed using fuor different models: i) Multivariate linear regression (MLR), ii) Linear support vector machine (lSVM), iii) Gaussian support vector machine (gSVM) and iv) artificial neural network (ANN). Models were trained and validated using the quantitative features extracted to evaluate their differentiation capability [28]. Five-fold cross-validation was performed to avoid over-fitting. Two feature selection methods were compared: i) Principal component analysis (PCA) which transforms the initial variables into a reduced number of principal components which explain a portion of the total variance; the first four principal components were retained as they always explained at least 90% of the total variance as empirically showed best performance ii) minimum redundance maximum relevance (mrMR) algorithm which calculates a subset of the initial features by maximizing the relevance with respect to the response variable and minimizing the redundancy among variables; the ten features with highest ranking were selected. The five-fold cross-validation scheme works as follows:

1. Randomly split the dataset in five stratified folds. Each fold will include 3 TIR3A nodules and 3 TIR3B nodules;
2. select four folds as training set. Perform feature selection;
3. Classify the test set with the learned model and save the confusion matrix;
4. go back to step 2 until each fold is used as test set;
5. sum up the confusion matrices to obtain a unique confusion matrix. The same classification experiment was performed using only features extracted from the total nodule to test whether if the peripheral/central division could increase the differentiation performance. Traditional accuracy (ACC), sensitivity (SENS) and specificity (SPEC) were computed from the confusion matrix.

Tuning the percentage ratio of the equivalent radius In order to optimize the percentage ratio of the equivalent radius (PRER) for dividing the nodule mask in peripheral and central mask, the experiment was repeated five times using five different values for the percentage parameter: 5%, 10%, 15%, 20% and 30%.

3.3 Results

The computational time required for vascular network extraction and the calculation of the features was about 2.5 ± 0.4 min for nodule. Tables 3.1, 3.2 and 3.3 report the mean, standard deviation and P-values of the features extracted from the vascular network of the total nodule, the peripheral part and the central part. The Mann-Whitney U-test showed that malignant nodules had higher values of NT_t ($P<0.005$), NPV ($P<0.005$), DM_t ($P<0.05$), ICM_t ($P<0.05$), ICM_p ($P<0.005$), $SOAM_p$ ($P<0.05$) and ICM_c ($P<0.05$). No flow parameters showed statistical differences between TIR3A and TIR3B nodules.

MANOVA assessed the existence of one canonical variable (c_1) which differs among the two groups, dimension of the space containing the group means was equal to 1 with a p-value lower than 0.005. Figure 3.5 shows the scatter plot of the two most discriminant canonical variables (c_1 and c_2). PCA showed best performance compared to mrMR algorithm (Table 3.4 and Table 3.5). The lSVM and the ANN classifiers yielded the best classification accuracy equal to 83.33%, a sensitivity of 93.33% and a specificity of 73.33% (Table 3.4). When the model was trained using only features from the whole nodule, the classification accuracy and sensitivity dropped to 66.67% and 60%, respectively, while specificity remained unchanged.

Feature	TIR3B nodules	TIR3A nodules	P-value
$\mathbf{NT_t}$	**10.31 ± 8.12**	**4.31 ± 4.41**	**<0.005**
NB_t	16.12 ± 14.22	11.33 ± 12.92	0.26
NPV	**5.44 ± 2.21**	**3.17 ± 1.97**	**<0.005**
$\mathbf{DM_t}$	**2.67 ± 0.19**	**2.48 ± 0.19**	**<0.05**
$\mathbf{ICM_t}$	**101.43 ± 27.62**	**78.87 ± 24.25**	**<0.05**
$SOAM_t$	1.49 ± 0.02	1.48 ± 0.03	0.20
VI_t	0.71 ± 0.24	0.61 ± 0.25	0.20
FI_t	0.38 ± 0.17	0.31 ± 0.11	0.25

Table 3.1: Mean and standard deviation values of the features extracted from the vascular network of the whole nodule. Every feature is dimensionless apart from SOAM which is measured in radians/voxel. NT = number of trees; NB = number of branches; NPV = number of penetrating vessels; DM = distance metric; ICM = inflection count metric; SOAM = sum of angles metric; VI = vascularization index; FI = flow index. Subscript "t" indicates that the feature is extracted from the total nodule.

Feature	TIR3B nodules	TIR3A nodules	P-value
DM_p	1.84 ± 0.17	1.76 ± 0.17	0.26
$\mathbf{ICM_p}$	**63.18 ± 18.52**	**36.79 ± 15.34**	**<0.005**
$\mathbf{SOAM_p}$	**1.43 ± 0.03**	**1.38 ± 0.06**	**<0.05**
VI_p	0.76 ± 0.23	0.66 ± 0.22	0.25
FI_p	0.42 ± 0.16	0.35 ± 0.12	0.28

Table 3.2: Mean and standard deviation values of the features extracted from the peripheral vascular network of the nodule. Subscript "p" indicates that the feature is extracted from the peripheral region of the nodule.

Best percentage ratio of the equivalent radius Regarding PRER tuning, five-fold cross validation with the lSVM classifier revealed that the percentage equal to 10% obtained the best accuracy, sensitivity and specificity. Table 3.6 shows the classification performance for each value of PRER tested.

3.4 Discussion

Vascularization is known to be an important biomarker for the characterization and differentiation of thyroid nodules. Several studies confirmed that malignant thyroid nodules present a denser and more complex vascular network than benign ones. Vascularity pattern, i.e. how the vessels are distributed spatially throughout the nodule, plays a relevant role in the differentiation of benign and malignant nodules [11]. More than 40% of nodules which show central hypervascularity are malignant, while a predominant peripheral flow is a benign marker [12]. Further, each histological subtype of thyroid cancer have different characteristics in terms of

Feature	**TIR3B nodules**	**TIR3A nodules**	**P-value**
DM_c	2.11 ± 0.17	2.09 ± 0.16	0.46
ICM$_c$	**88.23 ± 23.98**	**59.62 ± 23.40**	**<0.05**
$SOAM_c$	1.46 ± 0.02	1.45 ± 0.03	0.16
VI_c	0.68 ± 0.26	0.58 ± 0.27	0.25
FI_c	0.36 ± 0.18	0.28 ± 0.11	0.38

Table 3.3: Mean and standard deviation values of the features extracted from the central vascular network of the nodule. Subscript "c" indicates that the feature is extracted from the central region of the nodule.

Model	**ACC**	**SENS**	**SPEC**
MLR	23/30	11/15	12/15
lSVM	**25/30**	**14/15**	**11/15**
gSVM	20/30	10/15	10/15
ANN	**25/30**	**14/15**	**11/15**

Table 3.4: Classification performance with Multivariate Linear Regression (MLR), linear support-vector machine (lSVM), gaussian support-vector machine (gSVM) and artificial neural network (ANN) using principal component analysis as feature selection method. Performance was measured using accuracy (ACC), sensitivity (SENS) and specificity (SPEC).

vascular density, distribution and morphology [11]. For these reasons, quantitative description of the vascular network in terms of flow, architecture and tortuosity in functional modalities such as PDUS could be a powerful tool for aiding in the correct diagnosis of thyroid nodules, which is known to present pitfalls in both ultrasound and cytological evaluation especially in TIR3 nodules [29, 6].

In this study, 30 thyroid nodules with an undetetermined cytology were retrospectively analyzed, 15 TIR3A nodules with low presence of microfollicular/trabecular structures and 15 TIR3B nodules with high suspect of papillary carcinoma due to nuclei alterations.

Quantitative analysis considering the total nodule showed that in agreement with previous studies [22, 23] architectural feature NT_t resulted to be higher in TIR3B nodules ($p<0.005$) suggesting that papillary carcinoma had higher number of vessel trees due to tumor angiogenesis whereas NB_t did not reach level of significance in contrast to previous studies. This may be possible for two likely reasons: i) TIR3B nodules present lower number of branches with respect to TIR4 nodules which show more distinctive features of papillary carcinoma; ii) PDUS is not the best imaging modality for detection small branches formation, indeed, for microvascularization characterization CEUS imaging is the preferred imaging modality [22]. In this study, the number of penetrating vessels, which showed to be higher in malignant nodules in previous studies based on 2D CDUS images [20], was introduced in a 3D analysis of the nodule vascular network for the first time. As expected,

Model	ACC	SENS	SPEC
MLR	21/30	11/15	10/15
lSVM	21/30	10/15	11/15
gSVM	20/30	10/15	10/15
ANN	21/30	12/15	9/15

Table 3.5: Classification performance with Multivariate Linear Regression (MLR), linear support-vector machine (lSVM), gaussian support-vector machine (gSVM) and artificial neural network (ANN) using minimum redundancy maximum relevance as feature selection method. Performance was measured using accuracy (ACC), sensitivity (SENS) and specificity (SPEC).

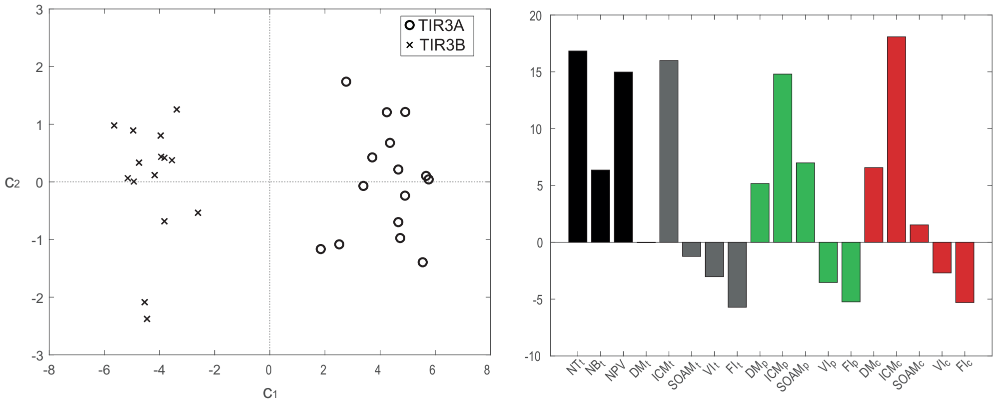


Figure 3.5: (Left) Scatter plot of the canonical variables computed by MANOVA grouped by TIR3A and TIR3B cytological classification; (Right) Weights of original features into the canonical variable c_1. The first canonical variable (c_1) was able to discriminate the two groups with a p-value lower than 0.005. Architectural features NT_t and NPV, and ICM computed from total, peripheral and central region were the features with highest weights in the first canonical variable.

NPV_t resulted to be higher in suspicious malignant nodules ($p<0.005$) suggesting that the two types of nodules have differences in the perfusion of the central region of the nodule confirming the importance of vascularity distribution in the characterization of TIR3 nodules. It is known that global morphology of intratumoral vessels in papillary carcinoma is relevant [11, 23], however morphology has never been described in function of the region of the nodule. Both global and regional (peripheral and central) tortuosity resulted to be higher in malignant nodules, indeed DM_t, ICM_t, ICM_p, $SOAM_p$ and ICM_c resulted to be higher in malignant nodules ($p<0.05$). Interestingly, SOAM, which is specialized in detecting vessels forming high frequency coils [15], resulted to be significantly higher only when computed in the peripheral region suggesting that peripheral network of TIR3B nodules contain vessels arranged in a more serpiginous way than benign ones. Surprisingly, no flow parameters differed significantly between the two groups, suggesting that vascular density and strength of flow properties are similar in these cytological borderline nodules. In previous studies [22, 23], every architectural and tortuosity feature was

PRER	**ACC**	**SENS**	**SPEC**
5%	66.67%	80%	53.33%
10%	**83.33%**	**93.33%**	**73.33%**
15%	70%	73.33%	66.67%
20%	56.67%	53.33%	60%
30%	53.33%	53.33%	53.33%

Table 3.6: Five-fold cross-validation PCA-SVM model accuracy, sensitivity and specificity varying the percentage ratio of the equivalent radius. Ten percent revealed to be the best choice as it yielded the best classification performance. PRER = percentage ratio of the equivalent radius; ACC = accuracy; SENS = sensitivity; SPEC = specificity.

resulting higher in malignant nodules, anyway, in these other studies the tortuosity was calculated in volumes-of-interest which were typically located in the central region of the nodule and not on the entire nodule, as done in this study. Moreover, the two other studies considered a dataset containing also TIR2 and TIR4 nodules with different histological cancer subtypes (cystic lesions, papillary carcinoma, follicular carcinoma and Hurtle cell carcinomas) whereas in the current study only TIR3B nodules which presented typical features of papillary carcinoma were included. This study is a preliminary quantitative analysis of vascular network of undetermined TIR3A and TIR3B nodules. Compared to other quantitative studies of thyroid nodule vascularization, the main novelty of this study is the quantitative analysis of the 3D model of the thyroid nodule vascular network in peripheral and central region (table 3.7).

Authors	**Image Modality**	**Dataset**	**3D**	**Per/Cen Division**	**Features**	**Classifier**	**Performance**
Lyshchik et al. 2007 [17]	PDUS	86 nodules (46 B, 40 M)	X	X	VI, FI	Cutoff selection;	VI: 72.4% SENS, 100% SPEC, 86.2% ACC; FI: 72.4% SENS, 100% SPEC, 86.2% ACC
Molinari et al. 2010 [22]	CEUS	20 nodules (10 B,10 M)	V	X	NT, vascular density, NB, average vessel radius, DM, ICM, SOAM	Mann-Whitney U-test	N.D.
Sultan et al. 2015 [18]	CDUS	100 nodules (58 B, 42 M)	X	V	Velocity index, FI, blood flow volume index	Cutoff selection;	VI central region: 90% SENS, 88% SPEC, 89% ACC
Baig et al. 2017 [19]	B-mode, CDUS	111 nodules (62 B, 49 M)	X	V	Velocity Index, sonographic features (microcalcifications, echogenicity, margin regularity, tall to width ratio)	Cutoff selection, at least one suspicious sonographic feature;	66.7% SENS, 83.3% SPEC, 79.3% ACC
Caresio et al. 2018 [23]	PDUS, CEUS	20 nodules (10 B, 10 M);	V	X	NT, NB, DM, ICM, SOAM, spatial vascularity pattern	MANOVA-Linear Discriminant Analysis;	100% SENS, 100% SPEC, 100% ACC
Cansu et al. 2019 [30]	PDUS	81 nodules (56 B, 25 M)	V	X	VI, FI, FVI	Cutoff selection;	VI: 72% SENS, 55.4% SPEC, 68% AUC; FI: 68% SENS, 57.1% SPEC, 61% AUC; FVI: 68% SENS, 67.9% SPEC, 67% AUC
Proposed	PDUS	30 nodules (15 B, 15 M),	V	V	NT, NB, DM, ICM, SOAM, VI, FI, NPV	PCA-lSVM;	93.3% SENS, 73.3% SPEC, 83.3% ACC

Table 3.7: Quantitative studies of thyroid nodule vascularization for differentiation between benign (B) and malignant (M) nodules. In every study reported, every nodule underwent FNA-biopsy and malignancy was surgically confirmed. PDUS = Power Doppler Ultrasound; CEUS = Contrast Enhanced Ultrasound; CDUS = Color Doppler Ultrasound; VI = Vascularization Index; FI = Flow Index; NT = number of trees; NB = number of branches; DM = distance metric; ICM = inflection count metric; SOAM = sum of angles metric; FVI = Vascularization Flow Index; NPV = number of penetrating vessels; SENS = sensitivity; SPEC = specificity; ACC = accuracy.

A classification experiment was performed using different classifier models and different feature selection methods. PCA was used in order to reduce the data dimensionality, as some features presented high correlation (ICM_t-ICM_c and NT_t-NPV, r equal to 0.86 and 0.81). mrMR feature selection method performed worse than PCA, likely to the difficuly of choosing the number of features to be kept in the final subset and the exclusion of significant variables which did not show high relevance. In a 5-fold cross-validation scheme, lSVM and ANN yielded best classification performances in terms of accuracy, sensitivity and specificity equal to 83.33%, 93.33% and 73.33% , respectively. Anyway, it is not appropriate to claim the superiority of the two classifiers over the others, given the small performance differences and the dataset numerosity. Training and validating the same classifier model using only features extracted from the whole nodule produced a decrease in the performance (ACC = 66.67%, SENS = 60%, SPEC = 73.33%). This confirms that the division in peripheral and central region increases the diagnostic accuracy of the system.

Some limitations have to be acknowledged. First of all, papillary carcinoma was histologically confirmed for only 5 out of 15 TIR3B thyroid nodules. For the rest of TIR3B nodules histological confirmation will be available in short time. Second, for TIR3A nodules surgical removal is not recommended, follow-up repeated biopsies will be considered.

Pitfalls in multimodal ultrasound evaluation of thyroid nodules Multimodal ultrasound has great potential in thyroid nodule differentiation: B-mode images have the ability of depicting tissue microstructure hetereogeneity and several studies confirm that echogenicity and echotexture are significantly different between benign and malignant nodules [31, 32, 33]. In this study, we demonstrated the potentiality of PDUS in the characterization of the vascularity pattern of thyroid nodules. Other studies claimed the usefulness of CEUS and SMI modes for a more deepen characterization of vascularity [22, 23, 34] (figure 3.6). Ideally, integration of these modalities in a unique characterization system may have a great potential in the diagnosis of thyroid nodules. However, co-registration between these modalities is very challenging. The main reason is the 3D ultrasound image acquisition protocol: using the 3D probes, orientation of slices and field of view of these volumes may change significantly between different acquisitions due to slight differences in probe positioning. Due to these limitations, so far, the only feasible multimodal strategy is the feature level or classifier level fusion where features are extracted separately from each modality. These strategies would have limited ability to fuse complementary multimodal features. We are currently working to a method to co-register B-mode, PDUS and CEUS. The basic idea is: i) excluding the Doppler signal from the PDUS images to align B-mode and PDUS images, ii) co-registering the PDUS and CEUS images through a structural-based registration between the segmented vascular networks.

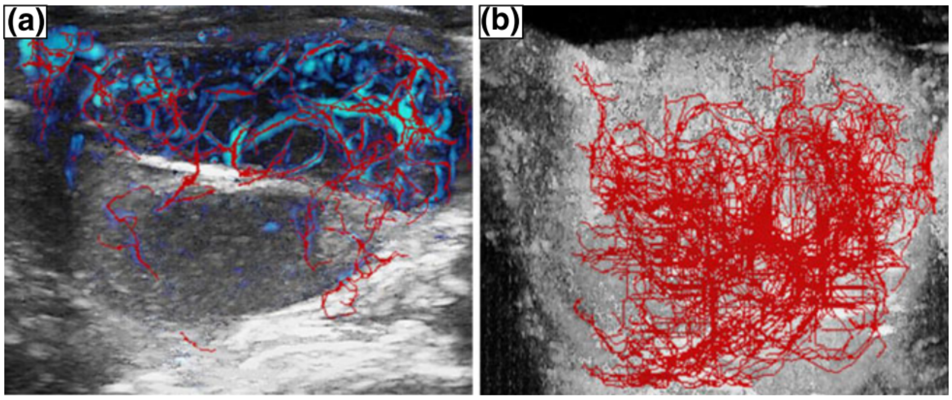


Figure 3.6: Vascular network extraction of a thyroid nodule in a PDUS image (a) and a CEUS image (b). Due to the contrast-enhancement effect and the reduced size, microbubbles allow for depiction of microcirculation of the central part of the nodule [23]

3.5 Conclusions

A novel method for quantitative analysis of vascularization in 3D Power Doppler images of thyroid nodules was proposed. Quantitative features describing the architecture, tortuosity and flow were extracted from the total nodule and from the peripheral and central region of the nodule. A dataset including 30 thyroid nodules, 15 TIR3A benign and 15 TIR3B with high suspect of papillary carcinoma, was used. A classification experiment with support vector machine confirms that the division in peripheral and central region increased the diagnostic accuracy (from 66.67% to 83.33%) of the differentiation between the two groups. Histological examination will be considered for increase results reliability. Furthermore, the methodology will be extended to contrast-enhanced ultrasound imaging and superb micro-enhanced imaging for a deepen characterization of thyroid nodule vascularization.

3.6 References

[1] American Cancer Society. *Key Statistics for Thyroid Cancer*. 2017.

[2] P. Vitti and T. Rago. *Thyroid ultrasound as a predicator of thyroid disease*. 2003. DOI: 10.1007/BF03347031.

[3] Hamdan Ahmed Pasha et al. "Malignancy rate in thyroid nodules with atypia or follicular lesion of undetermined significance". In: *International Archives of Otorhinolaryngology* 24.2 (July 2020), E221–E226. ISSN: 18094864. DOI: 10.1055/S-0039-1698784.

[4] Carlo La Vecchia et al. "Thyroid cancer mortality and incidence: a global overview." In: *International journal of cancer* 136.9 (May 2015), pp. 2187–95. ISSN: 1097-0215. DOI: 10.1002/ijc.29251.

[5] D. N. Poller and S. J. Johnson. “Recent Developments in the Pathology of Thyroid Cancer”. In: *Clinical Oncology* 29.5 (May 2017), pp. 278–282. ISSN: 14332981. DOI: 10.1016/j.clon.2017.01.006.

[6] Silvia Trombetta et al. “THY3 cytology: What surgical treatment? Retrospective study and literature review”. In: *International Journal of Surgery* 28 (2016), S59–S64. ISSN: 17439159. DOI: 10.1016/j.ijsu.2015.05.060.

[7] Francesco Nardi et al. “Italian consensus for the classification and reporting of thyroid cytology”. In: *Journal of Endocrinological Investigation* 37.6 (2014), pp. 593–599. ISSN: 17208386. DOI: 10.1007/s40618-014-0062-0.

[8] F. Tartaglia et al. “Fine needle aspiration cytology of 650 thyroid nodules operated for multinodular goiter: A cyto-histological correlation based on the new Italian cytological classification (Siapec 2014)”. In: *Journal of Biological Regulators and Homeostatic Agents* 30.4 (2016), pp. 1187–1193. ISSN: 0393974X.

[9] Jake Forster et al. “A review of the development of tumor vasculature and its effects on the tumor microenvironment”. In: *Hypoxia* Volume 5 (Apr. 2017), pp. 21–32. DOI: 10.2147/hp.s133231.

[10] Naoyo Nishida et al. “Angiogenesis in cancer”. In: *Vascular Health and Risk Management* 2.3 (Aug. 2006), pp. 213–219. ISSN: 1176-6344. DOI: 10.2147/vhrm.2006.2.3.213.

[11] Matvey Vladimi Sprindzuk. “Angiogenesis in Malignant Thyroid Tumors”. In: *World Journal of Oncology* 1.6 (2010), p. 221. ISSN: 19204531. DOI: 10.4021/wjon263e.

[12] Gopinathan Anil, Amogh Hegde, and F. H Vincent Chong. “Thyroid nodules: Risk stratification for malignancy with ultrasound and guided biopsy”. In: *Cancer Imaging* 11.1 (Dec. 2011), pp. 209–223. ISSN: 14707330. DOI: 10.1102/1470-7330.2011.0030.

[13] Xiangyu Yang and Michael V. Knopp. “Quantifying Tumor Vascular Heterogeneity with Dynamic Contrast-Enhanced Magnetic Resonance Imaging: A Review”. In: *Journal of Biomedicine and Biotechnology* 2011 (2011), pp. 1–12. ISSN: 1110-7243. DOI: 10.1155/2011/732848.

[14] K. A. Miles et al. “Current status and guidelines for the assessment of tumour vascular support with dynamic contrast-enhanced computed tomography”. In: *European Radiology* 22.7 (July 2012), pp. 1430–1441. ISSN: 09387994. DOI: 10.1007/s00330-012-2379-4.

[15] Elizabeth Bullitt et al. “Measuring tortuosity of the intracerebral vasculature from MRA images”. In: *IEEE Transactions on Medical Imaging* 22.9 (Sept. 2003), pp. 1163–1171. ISSN: 0278-0062. DOI: 10.1109/TMI.2003.816964.

[16] Vyacheslav Kalchenko et al. "In vivo characterization of tumor and tumor vascular network using multi-modal imaging approach". In: *Journal of Biophotonics* 4.9 (June 2011), n/a–n/a. ISSN: 1864063X. DOI: 10.1002/jbio.201100033.

[17] Andrej Lyshchik et al. "Quantitative Analysis of Tumor Vascularity in Benign and Malignant Solid Thyroid Nodules". In: *Journal of Ultrasound in Medicine* 26.6 (June 2007), pp. 837–846. ISSN: 02784297. DOI: 10.7863/jum.2007.26.6.837.

[18] Laith R. Sultan et al. "Vascularity assessment of thyroid nodules by quantitative color doppler ultrasound". In: *Ultrasound in Medicine and Biology* 41.5 (2015), pp. 1287–1293. ISSN: 1879291X. DOI: 10.1016/j.ultrasmedbio.2015.01.001.

[19] Faisal N. Baig et al. "Computer-aided assessment of regional vascularity of thyroid nodules for prediction of malignancy". In: *Scientific Reports* 7.1 (2017), pp. 1–9. ISSN: 20452322. DOI: 10.1038/s41598-017-14432-7.

[20] Jin Chung et al. "Clinical applications of doppler ultrasonography for thyroid disease: Consensus statement by the korean society of thyroid radiology". In: *Ultrasonography* 39.4 (Oct. 2020), pp. 315–330. ISSN: 22885943. DOI: 10.14366/usg.20072.

[21] Yi Cheng Zhu et al. "A prospective study to compare superb microvascular imaging with grayscale ultrasound and color doppler flow imaging of vascular distribution and morphology in thyroid nodules". In: *Medical Science Monitor* 24 (2018), pp. 9223–9231. ISSN: 16433750. DOI: 10.12659/MSM.911695.

[22] Filippo Molinari et al. "Characterization of Single Thyroid Nodules by Contrast-Enhanced 3-D Ultrasound". In: *Ultrasound in Medicine and Biology* 36.10 (Oct. 2010), pp. 1616–1625. ISSN: 03015629. DOI: 10.1016/j.ultrasmedbio.2010.07.011.

[23] Cristina Caresio et al. "Quantitative analysis of thyroid tumors vascularity: A comparison between 3-D contrast-enhanced ultrasound and 3-D Power Doppler on benign and malignant thyroid nodules". In: *Medical Physics* 45.7 (2018), pp. 3173–3184. ISSN: 00942405. DOI: 10.1002/mp.12971.

[24] Alejandro F. Frangi et al. "Multiscale vessel enhancement filtering". In: *Medical Image Computing and Computer-Assisted Interventation — MICCAI'98*. Ed. by William M. Wells, Alan Colchester, and Scott Delp. Vol. 1496. Springer Berlin Heidelberg, 1998, pp. 130–137. ISBN: 978-3-540-65136-9. DOI: 10.1007/BFb0056195.

[25] S.R. Aylward and Elizabeth Bullitt. “Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction”. In: *IEEE Transactions on Medical Imaging* 21.2 (Feb. 2002), pp. 61–75. ISSN: 02780062. DOI: `10.1109/42.993126`.

[26] K. M. Meiburger et al. “Skeletonization algorithm-based blood vessel quantification using in vivo 3D photoacoustic imaging”. In: *Physics in Medicine and Biology* 61.22 (Nov. 2016), pp. 7994–8009. ISSN: 0031-9155. DOI: `10.1088/0031-9155/61/22/7994`.

[27] T.C. Lee, R.L. Kashyap, and C.N. Chu. “Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms”. In: *CVGIP: Graphical Models and Image Processing* 56.6 (Nov. 1994), pp. 462–478. ISSN: 10499652. DOI: `10.1006/cgip.1994.1042`.

[28] Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press, 2000. ISBN: 9780511801389. DOI: `10.1017/CBO9780511801389`.

[29] Daniel T. Ginat et al. *Pearls and pitfalls of thyroid nodule sonography and fine-needle aspiration*. Sept. 2010. DOI: `10.1097/RUQ.0b013e3181efa710`.

[30] Ayşegül Cansu et al. “Diagnostic value of 3D power Doppler ultrasound in the characterization of thyroid nodules”. In: *Turkish Journal of Medical Sciences* (Jan. 2019), pp. 723–729. ISSN: 1303-6165. DOI: `10.3906/sag-1803-92`.

[31] Franklin N. Tessler et al. “ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee”. In: *Journal of the American College of Radiology* 14.5 (May 2017), pp. 587–595. ISSN: 1558349X. DOI: `10.1016/j.jacr.2017.01.046`.

[32] Hanung Adi Nugroho et al. “Texture analysis for classification of thyroid ultrasound images”. In: *2016 International Electronics Symposium (IES)*. IEEE, Sept. 2016, pp. 476–480. ISBN: 978-1-5090-1640-2. DOI: `10.1109/ELECSYM.2016.7861053`.

[33] Ali Abbasian Ardakani, Akbar Gharbali, and Afshin Mohammadi. “Classification of Benign and Malignant Thyroid Nodules Using Wavelet Texture Analysis of Sonograms”. In: *Journal of Ultrasound in Medicine* 34.11 (Nov. 2015), pp. 1983–1989. ISSN: 02784297. DOI: `10.7863/ultra.14.09057`.

[34] Priscilla Machado et al. “A novel microvascular flow technique: Initial results in thyroids”. In: *Ultrasound Quarterly* 32.1 (Mar. 2016), pp. 67–74. ISSN: 15360253. DOI: `10.1097/RUQ.0000000000000156`.

Chapter 4

Neuroendocrine tumor characterization

Part of this chapter is currently under review for journal publication:

- Brief Research Report: V. Liberini, O. Rampado, E. Gallio, B. De Santi, F. Ceci, B. Dionisi, P. Thuillier, L. Ciuffreda, A. Piovesan, F. Fioroni, A. Versari, F. Molinari, D. Deandreis, **^{68}Ga-DOTATOC PET/CT based radiomic analysis and PRRT outcome: a preliminary evaluation based on an exploratory radiomic analysis.**, *Frontiers in Medicine*, 2020
- Journal Article: V. Liberini, B. De Santi, O. Rampado, E. Gallio, B. Dionisi, F. Ceci, G. Polverari, P. Thuillier, F. Molinari, D. Deandreis, **Impact of Segmentation and Discretization on Radiomic Features in ^{68}Ga-DOTA-TOC PET/CT images of Neuroendocrine Tumor**, *European Journal of Nuclear Medicine and Molecular Imaging Physics*, 2021
- Journal Article: P. Thuillier, V. Liberini, O. Rampado, E. Gallio, , B. De Santi, F. Ceci, J. Metovic, M. Papotti, M. Volante, F. Molinari, D. Deandreis, **Diagnostic Value of Conventional PET Parameters and Radiomic Features Extracted from 18F-FDG-PET/CT for Histologic Subtype Classification and Characterization of Lung Neuroendocrine Neoplasms**, *Biomedicines*, 2021

Part of this work was funded by Fondazione Cassa di Risparmio di Torino (CRT).

4.1 Introduction

According to the National Cancer Institute, the incidence rate of neuroendocrine neoplasms increased by 6.4-fold over the last two decades [1]. NENs are a heterogeneous group of malignancies represented by different histological subtypes and different primary locations, frequently localized in the lung and gastroenteropancreatic (GEP) site. These neoplasms arise from neuroendocrine cells which are cells that can produce neuroamines and peptide hormones [2]. NENs are divided in well-differentiated neuroendocrine tumors (NETs), which are mainly indolent neoplasms, and poorly differentiated neuroendocrine carcinomas (NECs), which are highly aggressive cancer [3]. A remarkable sign of NENs is the expression of somatostatin receptors (SSTRs), with SSTRs 1 and 2 present in the vast majority of GEP-NENs, SSTRs 3 and 5 expressed by approximately 60% of cases and SSTR 4 rarely represented; moreover, the poorly differentiated NECs appeared to express lower percentage of SSTR [4].

An accurate characterization of these tumors is fundamental, as their biological heterogeneity is strongly connected with their prognosis and treatment response [5]. In this scenario, PET/CT imaging is a powerful tool for in-vivo characterization, monitoring and treatment response quantification thanks to the possibility of extracting semi-quantitative parameters from images [6, 5]. Moreover, integration of PET imaging with CT imaging gives a complete picture of the functionality and the anatomical location of the NENs. Although ^{18}F-FDG-PET is the most widely used radiotracer in nuclear medicine, in the specific case of NETs, new radiotracers based on ^{68}Ga-DOTA-peptides have been introduced. These are able to bind to STTRs which are expressed by NETs. While ^{18}F-FDG-PET is the preferred choice for NECs, ^{68}Ga-DOTA-peptides PET is useful for characterization of NETs. Multiple studies confirmed the diagnostic and prognostic role of ^{68}Ga-DOTA-peptides PET, as SUV values of the lesion are correlated with survival rates [7] and treatment response [8, 9, 6, 5].

In this scenario, radiomics analysis could be a powerful tool as more advanced features than conventional SUV-based parameters, e.g. maximum and minimum SUV value of the lesion, could turn out to be imaging biomarkers related to the aggressiveness grade of the tumor. This would allow to better characterize the tumor phenotype in-vivo and to optimally choose the treatment strategy or the need for other diagnostic tests, such as ^{18}F-FDG-PET exam. Recent studies showed that advanced textural features extracted from ^{68}Ga-DOTA-peptides PET images are able to characterize NETs in terms of lymph nodal involvement [10] and to predict response to peptide receptor radionuclide therapy (PRRT) in patients affected by NET [11].

As the field of radiomics grows, there is an increasing need for a standardized protocol for the extraction of radiomics features (RFs). Furthermore, researchers are increasingly investigating about automatic tumor segmentation methods, PET image reconstruction algorithms and RFs robustness. Regarding tumor segmentation, automatic and semi-automatic segmentation methods are becoming more used as manual segmentation is time-consuming and highly subjective [12]. Automatic segmentation of NETs is a complex task because of partial volume effects, high variability in tumor size and in some cases low contrast of the tumor with neighbouring tissue [12, 13]. Thresholding is very used for PET tumor segmentation where an operator define an initial region-of-interest including the tumor and then voxels with a value above a percentage of the maximum SUV, usually 20% or 40%, are considered for the tumor mask [12]. Other thresholding methods, such as OTSU thresholding, are able to automatically determine a threshold from the image properties [14]. However, so far, thresholding methods were validated on ^{18}F-FDG-PET where high SUVs directly implies a pathology or an inflammatory condition [12], whereas SUVs of NETs in ^{68}Ga-DOTA-peptides PET images show high variability due to the inter-tumoral heterogeneity in the expression of SRRTs [15]. Further, thresholding techniques have the following drawbacks: i) segmentation result is strongly dependent on the choice of the region of interest; ii) possible exclusion of voxels belonging to the NET for their low SUVs (e.g. necrotic tissue) iii) possible inclusion of voxels belonging to healthy tissue when the site of NET have high physiological SUVs (e.g. liver NETs) [15, 16]. For these reasons, more advanced PET tumor segmentation methods have been proposed such as clustering techniques such as the fuzzy C-means (FCM) clustering algorithm, watersheds, deep learning and active contours [17]. Active contours are deformable models able to evolve in the image by optimizing an energy functional. Their use is widespread thanks to their flexibility, i.e. energy functional can be modified depending on the segmentation task, and excellent performance for many medical segmentation tasks. Most part of used active contours are either region-based, e.g. the Chan-Vese active contour (CVAC) [18] or edge-based, e.g. geodesic active contours [19]. However, the aforementioned active contours are not directly employable for NET segmentation as NETs show high variability in terms of edges and foreground/background contrast ratios. Hybrid active contours (HACs) are active contours which combine a region-based and an edge-based term in a unique energy functional [20]. This feature confers reproducibility of segmentation results in cases where the object-of-interest present diverse intensities and gradient values image-by-image, which is the case of NETs in ^{68}Ga-DOTA-peptides PET images. Nevertheless, so far, no attempts have been made on automatic segmentation of NETs.

Given the diagnostic and prognostic potential role of a radiomic model based on ^{68}Ga-DOTA-peptides PET imaging, model robustness is fundamental. In other words, it is necessary that the extracted imaging biomarkers are robust under the variation of some fundamental factors of the radiomic model such as segmentation, discretization settings and PET image reconstruction algorithms. Bailly et al. [21] studied the robustness of 2 conventional PET parameters, 6 GLCM-based features, 3 GLRLM-based features and 6 GLSZM-based features in ^{68}Ga-DOTA-NOC PET images as a function of image reconstruction settings on a cohort of 29 patients diagnosed with NET. The main limitations of their study are the limited number of patients and texture features analysed. In another study, robustness of intensity histogram and shape features among two 68Ga-DOTA-TATE PET/CT scans acquired two days apart was measured [22]. To the best of our knowledge, these are the only examples of published robustness studies in ^{68}Ga-DOTA-peptides PET images, while in literature there are several examples of RFs robustness analysis in ^{18}F-FDG-PET images [23, 24, 25].

Objectives Given the relatively recent introduction of ^{68}Ga-DOTA-peptides radiotracers in the clinical imaging and the lack of comparative studies on (semi-) automatic neuroendocrine tumor segmentation methodologies, in this chapter, a novel semi-automatic segmentation algorithm based on fuzzy C-means clustering and hybrid active contours (FCM-HAC) is proposed for the NET segmentation in ^{68}Ga-DOTA-peptides PET images. The algorithm is validated on a dataset of 60 NET lesions and compared with other methods. Subsequently, robustness analysis of RFs extracted from NETs is performed in function of segmentation and discretisation settings to identify a set of robust radiomics features for the characterization of NETs.

4.2 Materials and methods

Every algorithm proposed in this section was developed and tested in MATLAB (MathWorks, Natick, MA, USA, version r2019a) on a 2.21 GHz quad-core and 16 GB RAM.

Patients A total of 270 consecutive patients were enrolled between February 2017 and July 2019. The inclusion criteria were: i) presence of NET at histological examination; ii) treatment-naïve patients who underwent ^{68}Ga-DOTA-TOC PET/CT for staging or restaging after surgery; iii) patients who signed an informed consent form. Finally, 49 patients with a total of 60 lesions were found to be compatible with the inclusion criteria and were included in the study. Fourty-two lesions were primary, 18 were metastastis. Primary tumors were pancreatic NET, gastro-enteric NET, lung NET and others NET in 57.9%, 42.1%, 18.4% and 4.1% of cases, respectively. The NET histological sub-types were G1 (67.4%), G2 (12.2%), G3 in (2.0%), atypical carcinoid (8.2%) and typical carcinoid (10.2%).

Image acquisition and reconstruction The PET/CT scanner was an analog 3D PET scanner (Philips Gemini Dual-slice EXP scanner – PET AllegroTM system with Brilliance CT scanner – Philips Medical Systems, Cleveland, OH). As suggested by the procedure guidelines for PET imaging [26], the injected tracer activity was 145.1 ± 25.3 MBq of 68Ga-DOTA-TOC. After 60 minutes of uptake, PET data were acquired in 3D mode, covering the same anatomical region of the CT. Ordered subset expectation maximization (OSEM) algorithm (3D-RAMLA) was used for

PET image reconstruction using the following settings: 4 iterations, 8 subsets and field of view (FOV) of 576 mm. PET image resolution was 4 mm × 4 mm × 4 mm, CT image resolution was 1.1719 mm × 1.1719 mm × 6.5 mm.

SUV conversion The standardized uptake value is a semi-quantitative measure of the radiotracer activity concentration which can be extracted from PET images and is defined as:

$$\mathrm{SUV(kg/cc)} = \frac{\text{Activity Concentration}_{\mathrm{(Bq/cc)}}}{\frac{\text{Injected dose(Bq)}}{\text{Body weight (kg)}}} \tag{4.1}$$

In this study, standardized uptake value (SUV) was obtained by multiplying the image intensity values by the SUV scale factor which is included in the DICOM header of the PET image in Allegro and Gemini PET image acquisition systems (DICOM tag: 7053,1000) [27].

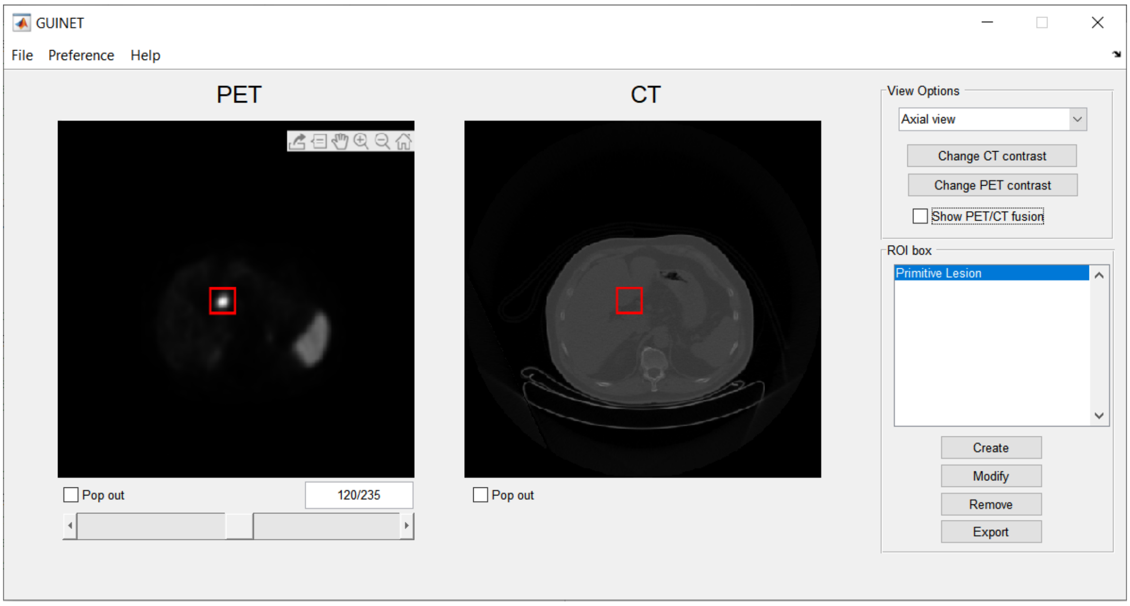
Manual annotation Each lesion was manually delineated on the PET image by two independent observers, both nuclear medicine physicians (FC, VL with 10 and 7 years of expertise respectively), by using the software LIFEx v 4.81 [28].

Image registration In this study, the CT volume will be used in conjunction with the PET volume, as structural reference for identifying the anatomical location of the NET and define an initial region-of-interest. For this reason, there was no need to correct image artifacts such as physiological motion (peristaltic and breathing) which would require non-rigid transformations. In order to register PET and CT images, transformation from Image-Based Coordinate System (ICS) to Patient-Based Coordinate System (PCS) was performed in both modalities, as described in A.5. Since in this study all image processing, from segmentation to radiomics analysis, will take place on the PET image, the CT image was interpolated to match the PET image size and resolution.

4.2.1 Tumor segmentation using fuzzy C-means and hybrid active contours

The segmentation algorithm contours the NET in 3D PET/CT images and requires the user intervention to define an initial region-of-interest (ROI) including the tumor. The general pipeline consists of three steps: i) ROI definition; ii) Tumor mask initialization by fuzzy C-means clustering; iii) Tumor segmentation by hybrid active contour.

ROI definition The user chooses the patient to be analyzed, a graphical-user-interface (GUI) shows the patient's PET and CT images on the screen and gives the possibility to define a parallelepiped ROI centered on the tumor. Starting from the central slice, the user can easily shrink/enlarge the ROI on the xy-plane by using either the mouse wheel or by dragging the sides of the ROI using the left mouse button. Then, the user can extend the ROI in the z-direction by pressing CTRL + left mouse click on the new slice. Figure 4.1 shows a screenshot of the developed GUI.


View Options: Axial view

ROI box: Primitive Lesion

120/235

Figure 4.1: Screenshot of the graphical user-interface (GUI) developed for the tumor ROI definition.

Tumor mask initialization by fuzzy C-means clustering As tumoral tissue shows higher SUV compared to normal tissue, the idea is to identify an initial mask positioned within the lesion and then refine the contours using an active contour sensitive to the image edges. The clustering algorithm used is the fuzzy C-means (FCM) [29] which allows the partitioning of voxels into classes by iteratively minimizing the objective function:

$$O = \sum_{i=1}^{C} \sum_{j=1}^{N} w_{ij}^{m} \left\| x_j - c_i \right\|^2 \tag{4.2}$$

where C is the number of clusters, N is the number of voxels belonging to the ROI, x_j is the intensity of jth voxel, m is a fixed parameter which controls the fuzzyness of clusters, c_i is the centroid of the ith cluster computed as the mean of the elements of the cluster weighted by their degree of membership to the cluster w_{ij} which is defined as:

$$w_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_j - c_i\|}{\|x_j - c_k\|} \right)^{\frac{2}{m-1}}} \tag{4.3}$$

The algorithm iteratively updates c_i and w_{ij} in order to minimize the objective function O until the maximum absolute difference of centroids between two consecutive iterations is below a threshold (th_{dc}). In this application the number of clusters (C) was set to 3 in order to partition the ROI in background (low SUV), tumor edges and tissue with a physiological uptake (medium SUV) and the core of the tumor (high SUV). Figure 4.2 shows the FCM membership maps in a volume slice. Other parameters set were $q = 2$ and $th_{dc} = 1\text{e}{-3}$ whose values were chosen after experimental tuning. Then, the membership map of the cluster with the highest centroid value corresponding to the tumor core is isolated. Based on the assumption that the user positioned the ROI so that the tumor is approximately in its center and since the ROI could include tissue with uptake values similar to the tumor, in order to exclude these regions from the initial mask an anisotropic gaussian kernel is used. This gaussian kernel is centered at the central voxel of the ROI and has the same size of the ROI and standard deviations along x, y and z equal to a quarter of the columns, rows and slices of the ROI, respectively. Subsequently, the filtered membership map is thresholded using a global threshold equal to 0.5. Final result of the tumor mask initialization is shown in figure 4.2.

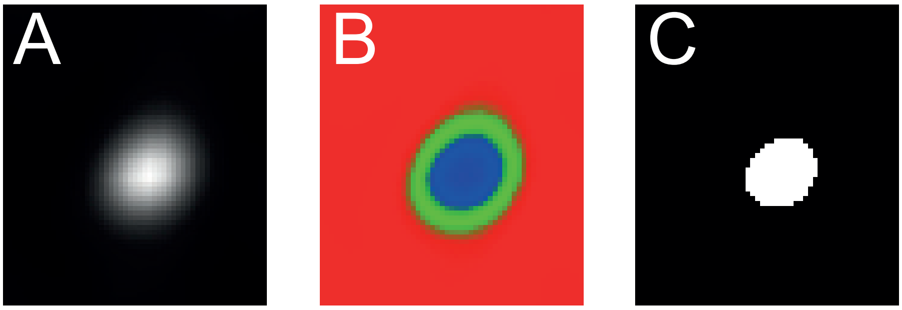


Figure 4.2: Tumor mask initialization by fuzzy C-means (FCM) clustering. Panel (A) shows the central slice of the volume after ROI cropping. Panel (B) shows the FCM membership maps codified in RGB (Each color channel value is the membership value to a cluster, R is the first cluster, G is the second cluster and B is the third cluster). The third cluster (highest average intensity value) is then isolated, multiplied by the gaussian kernel and thresholded. Final result is shown in panel (C).

Tumor segmentation by hybrid active contour Active contours are one of the most used segmentation methods. They are deformable models able to evolve on an image by minimizing an energy functional. Two categories of active contours can be distinguished: parametric and geometrical active contours. The former use an explicit formulation of the curve (2D) or of the surface (3D) and are often recognized as "snakes" [30]. Geometrical active contours instead use an implicit formulation: the curve or surface are represented by the zero level-set of a higher-order a function [18]. This property makes the geometrical active contour more effective in the case of irregular shapes and topological changes [18]. The Chan-Vese active contour (CVAC) is an example of geometrical active contour which uses a level-set formulation and a region-based energy functional [18]. Let $\Omega\epsilon\mathbb{R}^n$ be the spatial domain of the image I, where n is the dimension ($n = 3$ for volumetric images), the function $\phi : (\Omega, t) \rightarrow \mathbb{R}$ is the level-set function. The curve C can be defined implicitly as the zero level-set of ϕ:

$$C = \{(x, y, z, t)|\phi(x, y, z, t) = 0\} \tag{4.4}$$

We also define C_{in} and C_{out} are the regions inside and outside the curve C, respectively:

$$C_{in} = \{(x, y, z, t)|\phi(x, y, z, t) > 0\} \tag{4.5}$$

$$C_{out} = \{(x, y, z, t)|\phi(x, y, z, t) < 0\} \tag{4.6}$$

The energy functional proposed by Chan and Vese is:

$$\begin{aligned} F_{\text{CVAC}} = {} & \mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(C_{in}) \\ & + \lambda_1 \int_{C_{in}} |I(x,y,z) - c_1|^2 \, \mathrm{d}x\mathrm{d}y\mathrm{d}z + \lambda_2 \int_{C_{out}} |I(x,y,z) - c_2|^2 \, \mathrm{d}x\mathrm{d}y\mathrm{d}z \end{aligned} \tag{4.7}$$

where $\mu \geq 0$, $\nu \geq 0$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ are fixed constants, c_1 and c_2 are the average grey value of the image inside and outside the curve, respectively.

The same energy functional (4.7) can be written in a level-set formulation using the Heaviside H function[1]:

$$\begin{aligned} F_{CVAC}(c_1, c_2, \phi) = {} & \mu \int_\Omega |\nabla H(\phi)| \, \mathrm{d}\Omega + \nu \int_\Omega H(\phi) \mathrm{d}\Omega \\ & + \lambda_1 \int_\Omega (I - c_1)^2 H(\phi) \mathrm{d}\Omega + \lambda_2 \int_\Omega (I - c_2)^2 (1 - H(\phi)) \mathrm{d}\Omega \end{aligned} \tag{4.9}$$

In a level-set framework, evolving the curve means minimizing the functional energy with respect to ϕ. Mathematically, the curve evolution ($\frac{\partial \phi}{\partial t}$) is derived by solving the Euler-Lagrange equation associated to ϕ. For the numerical implementation,

---

[1]
$$H(\phi) = \begin{cases} 0, & \text{if } \phi < 0 \\ 1, & \text{if } \phi \geq 0 \end{cases} \tag{4.8}$$

the solution is discretized defining a time step Δt and using a finite difference scheme:

$$\frac{\phi^{n+1}_{i,j,k} - \phi^{n}_{i,j,k}}{\Delta t} = \delta_\epsilon(\phi^n)\left[\mu \cdot \mathrm{div}\left(\frac{\nabla\phi^n}{|\nabla\phi^n|}\right) - \nu - \lambda_1(I_{i,j,k} - c_1(\phi^n))^2 + \lambda_2(I_{i,j,k} - c_2(\phi^n))^2\right] \tag{4.10}$$

where δ_ϵ is the Dirac function using a spatial step ϵ for numerical implementation, $\mathrm{div}(\cdot)$ is the divergence operator, n is the iteration number. As already mentioned, the CVAC exhibits good performances when applied to images where the object-of-interest and the background are two homogenous and distinct regions. This is not the case of NETs mainly for two reasons: (i) some NETs appear heterogeneous on the PET image due to heterogenous tracer distribution of the radiotracer throughout the lesion, for example necrotic areas of the tumoral mass do not uptake radiotracer [13]; (ii) depending on the tumor location, the neighbouring tissue, physiologically, can exhibit uptake values similar to the NET [13]. To solve these problems, an hybrid active contour (HAC) model which involves the combination of a regional-based and a edge-based term in the energy functional [20], was used. The functional energy was defined as:

$$F_{\mathrm{HAC}} = -\alpha \int_\Omega (I(x,y,z) - \mu_m)\,\mathrm{d}\Omega + \beta \int_\Omega g\,|\nabla H(\phi)|\,\mathrm{d}\Omega \tag{4.11}$$

where $\alpha \geq 0$, $\beta \geq 0$ are fixed parameters which regulate the importance of the regional and edge terms, respectively, μ_m is an estimate of the lower gray-level of the object to be segmented and g is a feature map which is a function of the image gradient:

$$g(\nabla I) = \frac{1}{1 - |\nabla I|^2} \tag{4.12}$$

When ϕ is a signed distance function, the HAC discretized curve evolution is the following:

$$\frac{\phi^{n+1}_{i,j,k} - \phi^{n}_{i,j,k}}{\Delta t} = \alpha(I - \mu_m) + \beta\mathrm{div}(g\nabla\phi_n) \tag{4.13}$$

Compared to the CVAC curve evolution 4.10, the most important difference is the presence of the image gradient ∇I which forces the active contour to stop propagating in correspondence of edges. In this application, μ_m was automatically determined for each image as the average intensity of the second FCM cluster which corresponded approximately to the tumor edges, $\alpha = 0.05$, $\beta = 0.005$, $\Delta t = 0.05$ and number of iterations was set to 10. Morphological operation (erosion and dilatation with a spherical structural element of radius of 4 mm) was performed in order to remove mask imperfections.

Comparison with other methods In order to benchmark the proposed approach, four segmentation methods were implemented. The first method was Otsu global thresholding which works by minimizing intra-class intensity variance [14]. The second one was Fuzzy C-means (FCM) clustering, which is described in 4.2.1. Then, two active contour models were implemented: the Chan-Vese active contour (CVAC) model and the Caselles geodesic active contour (GAC) which is an edge-based active contour[2] [19]. Experimental tuning was performed for optimal configuration of parameters of each method [3]. For each method, segmentation result was compared to a common ground truth calculated from Op1 and Op2 segmentations using the STAPLE algorithm [31].

Performance metrics Dice-similarity coefficient (DSC), relative volume difference (RVD) and average boundary distance (ABD) were adopted as performance metrics for segmentation. Let V_1 and V_2 be the two volumetric logical masks to be compared, DSC is defined as:

$$\mathrm{DSC} = \frac{2\,|V_1 \cap V_2|}{|V_1| + |V_2|} \tag{4.14}$$

where $|V|$ is the number of true voxels of V. The DSC measures the overlap between two masks and ranges from 0 (no overlap) to 1 (total overlap). The RVD measures the difference between the number of true voxels of the masks:

$$\mathrm{RVD} = 100 \cdot \left(\frac{|V_1|}{|V_2|} - 1 \right) \tag{4.15}$$

A positive RVD value means that V_1 has a greater number of true voxels than V_2 and vice versa. In our scenario, if V_1 is the algorithm segmentation and V_2 is the ground-truth (manual segmentation), then, this metric is an index of under- (negative RVD values) or over- (positive RVD values) segmentation. If RVD is equal to zero, then, the two segmentations have the same number of true voxels. When evaluating comparison between operator and FCM-HAC algorithm, V_1 is the algorithm mask and V_2 is the operator mask. When evaluating the inter-operator variability (Op1 vs Op2), V_1 is the Op2 mask and V_2 is the Op1 mask.

The ABD measures the average distance in millimeters between the segmentation surfaces and is defined as:

$$\mathrm{ABD} = \frac{\sum_{x \in \delta V_1} d(x, V_2) + \sum_{x \in \delta V_2} d(x, V_1)}{|\delta V_1| + |\delta V_2|} \tag{4.16}$$

where d is the Euclidean distance operator, δV_1 and δV_2 are the set of surface points of V_1 and V_2, respectively.

---

[2] Curve evolution equation: $\frac{\partial \phi}{\partial t} = \alpha \mathrm{div}(g \nabla \phi / |\nabla u|)|\nabla \phi| + \beta g |\nabla \phi|$

[3] For CVAC model: $\lambda_1 = 0.95 \cdot 10^{-4}$, $\lambda_2 = 1.05 \cdot 10^{-4}$, $\nu = 0$, $\mu = 1 \cdot 10^{-7}$, $\Delta t = 0.05$
For GAC model: $\alpha = 0.01$, $\beta = 0.01$, $\Delta t = 0.05$

4.2.2 Tumor characterization using radiomics features

A total of fourty-five radiomics features (RFs) were extracted from PET images considering only voxels belonging to the volume-of-interest (VOI) of the NET. In particular, for each NET, four VOIs were considered, which corresponded to the two experts segmentation (Op1 and Op2), the semi-automatic algorithm segmentation proposed (FCM-HAC) and the STAPLE mask obtained by collecting Op1, Op2 and FCM-HAC segmentations. Additionally, to analyze the impact of SUVmax thresholding on the robustness of RFs, a 40% SUVmax thresholding was applied to VOI obtained after each segmentation. The choice of the intensity discretisation settings is fundamental and strongly affects the final value of the calculated feature. With the aim of evaluating which RFs are robust to the discretization settings variation, the images were discretized using different number of levels (32, 64 and 128). A maximum SUV of 80 was chosen for the discretisation, as the nearest integer to the 99th percentile of the maximum SUVs among lesions. Every RF was implemented in MATLAB using the formulation described in the Image Biomarker Standardisation Initiative document which aims to propose standardized nomenclatures and definitions of radiomics features for the whole radiomics-field research community [32]. The computed RFs are listed in table 4.1.

Family	**Description**	**Nomenclature**
Conventional PET parameters	maximum SUV	SUVmax
	mean SUV	SUVmean
	minimum SUV	SUVmin
	mean SUV value of the 26-connected neighbourhood of maximum SUV voxel	SUVpeak
	standard deviation of SUVs	SUVstd
	total lesion somatostatin receptor expression	TLSRE
Intensity histogram features	variance of the grey level histogram	H_Variance
	skewness of the grey level histogram	H_Skewness
	kurtosis of the grey level histogram	H_Kurtosis
	energy of the grey level histogram, squared sum of histogram counts	H_Energy
	entropy of the grey level histogram, using the base 2 logarithm	H_Entropy
Morphological features	volume of the VOI in milliliter	M_Volume_mL
	solidity of the VOI	M_Solidity
	sphericity of the VOI	M_Sphericity
Grey level co-occurence based features	inverse difference	GLCM_ID
	angular second moment	GLCM_ASM
	contrast	GLCM_Contrast
	correlation	GLCM_Correlation
	joint entropy	GLCM_Entropy
	dissimilarity	GLCM_Dissimilarity
Grey level run length based features	short runs emphasis	GLRLM_SRE
	long runs emphasis	GLRLM_LRE
	low grey level run emphasis	GLRLM_LGRE
	high grey level run emphasis	GLRLM_HGRE
	short run low grey level emphasis	GLRLM_SRLGE
	short run high grey level emphasis	GLRLM_SRHGE
	long run low grey level emphasis	GLRLM_LRLGE
	long run high grey level emphasis	GLRLM_LRHGE
	grey level non-uniformity	GLRLM_GLNU
	run length non uniformity	GLRLM_RLNU
	run percentage	GLRLM_RP
Grey level size zone based features	short zone emphasis	GLSZM_SZE
	long zone emphasis	GLSZM_LZE
	low grey level zone emphasis	GLSZM_LGZE
	high grey level zone emphasis	GLSZM_HGZE
	small zone low grey level emphasis	GLSZM_SZLGE
	small zone high grey level emphasis	GLSZM_SZHGE
	large zone low grey level emphasis	GLSZM_LZLGE
	large zone high grey level emphasis	GLSZM_LZHGE
	grey level non-uniformity	GLSZM_GLNU
	zone size non uniformity	GLSZM_ZSNU
	zone percentage	GLSZM_ZP
Neigbourhood grey tone difference based features	coarseness	NGTDM_Coarseness
	contrast	NGTDM_Contrast
	busyness	NGTDM_Busyness

Table 4.1: List of radiomics features extracted

Each family of features quantitatively describes different textural properties of the image. For example, conventional PET parameters are the most used quantitative parameters used in PET imaging for tracer uptake quantification thanks to their simple description and high concreteness. Intensity histogram features describe the SUV values distribution of the NET in terms of asymmetry, peakdness or degree of randomness. Morphological features describe the shape geometry and complexity of the VOI. Features extracted from the GLCM, GLRLM, GLSZM and NGTDM have a less direct meaning but have the ability to investigate spatial relationships between voxels with similar SUV values, and therefore are those able to provide additional and advanced information about tumor heterogeneity compared to the eye of a expert radiologist.

4.2.3 Radiomics features robustness to segmentation and discretisation

As already discussed, several factors such as segmentation method/operator, discretisation setting and image reconstruction method have a strong impact on feature values. Robustness analysis has the purpose of measuring consistency of features when one or more factors are perturbed. The syntax: "robustness of feature i to factor k" will be used for indicating the robustness of feature i when perturbating factor k. For example, one can measure the robustness of GLCM_Correlation to segmentation by repeating measurements of GLCM_Correlation changing the segmentation mask for each repetition. In this study, robustness was measured by varying segmentation (Op1, Op2 and FCM-HAC) and number of discretisation levels. RFs were calculated for a total of 12 combinations by changing segmentation VOI, number of levels and SUVmax threshold as described in table 4.2.

Robustness to	**Number of levels**	**VOI**	**SUVmax threshold**
Segmentation	64	Op1, Op2, FCM-HAC	\
	64	Op1, Op2, FCM-HAC	40%
Number of levels	32, 64, 128	STAPLE	\
	32, 64, 128	STAPLE	40%

Table 4.2: Combinations of VOI and discretisation settings for robustness analysis.

Robustness can be evaluated under two different aspects: consistency and agreement. In this study, only consistency was evaluated, as agreement is not the critical point for modern classification techniques such as machine learning which extrapolates higher-level relation between features rather than using cut-off on the feature absolute values. Consistency concerns if different measurements of the same RF are correlated in an additive manner [33]. In order to measure consistency, the intraclass correlation coefficient (ICC) was used. This metric measures the proportion of a variance that is attributable to objects of measurements [34]. There are 10 forms of ICC which involve different assumptions on the model (one-way or two-way, random or mixed effects), the type (single measurement or mean of multiple measurements) and the relationship to measure (agreement or consistency) [33]. Based on the practical guideline proposed by Koo et al. [33], the most appropriate form of ICC for measuring consistency of RFs to segmentation and discretisation was the two-way random-effects, consistency, single measurement ICC which is defined as

$$ICC = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E} \tag{4.17}$$

where k is the number of measurements, MS_R is the mean square for rows[4] and MS_E is the mean square error. ICC ranges from 0 to 1, where 0 indicates no consistency and 1 indicates total consistency.

4.3 Results

4.3.1 Segmentation results

Three tumors were excluded from the segmentation and robustness analysis due to a number of voxels lower than 16 voxels. Hence, a total of 57 tumors were studied.

FCM-HAC algorithm vs manual segmentation Figure 4.3 shows three different examples of NET segmentation using the FCM-HAC algorithm proposed. The algorithm was able to succesfully segment every primary and metastatic lesions of the dataset. Quantitative comparison between FCM-HAC algorithm and manual segmentations is shown in table 4.3. DSC analysis showed an agreement between FCM-HAC and manual segmentation comparable to inter-operator agreement. RVD values show a tendency of the FCM-HAC to slightly over-segment the NET.

	DSC	RVD	ABD (mm)
FCM-HAC vs Op1	0.68 ± 0.13	22.54 ± 31.05	2.91 ± 1.47
FCM-HAC vs Op2	0.71 ± 0.14	13.23 ± 40.67	2.68 ± 1.21
Op1 vs Op2	0.70 ± 0.09	-6.97 ± 48.69	2.06 ± 1.07

Table 4.3: FCM-HAC performance evaluation: mean and standard deviation values of Dice-Similarity Coefficient (DSC), Relative Volume Difference (RVD) and Average Boundary Distance (ABD).

---

[4] rows correspond to observations and columns to measurements

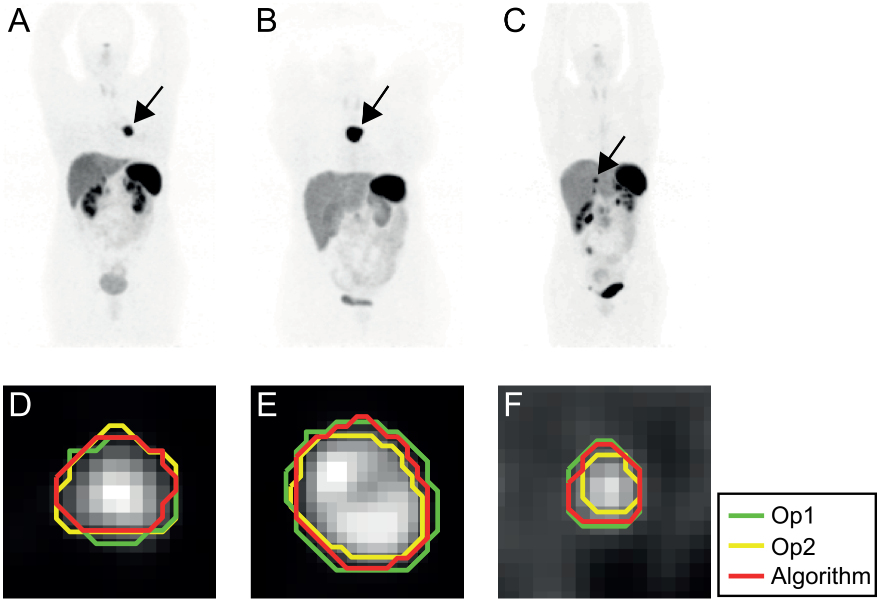


Figure 4.3: Example of segmentation of NETs extracted from three different patients. (A,D) lung primary NET; (B,E) metastatic lesion in a mediastinal lymph node; (C,F) metastatic liver lesion. The first row shows the maximum intensity projection (MIP) of the PET volumes, black arrows point at the segmented lesions. The second row shows the contours of manual segmentation (Op1 in green, Op2 in yellow) and the FCM-HAC segmentation contour in red

Comparison with other methods Table 4.4 shows segmentation performance of FCM-HAC and the other implemented segmentation methods. FCM-HAC achieved the highest DSC and lowest RVD compared to other standard segmentation methods.

Method	DSC	RVD	ABD (mm)
OTSU	0.53 ± 0.25	17.68 ± 73.27	5.98 ± 3.43
FCM	0.54 ± 0.14	-72.03 ± 60.29	4.11 ± 3.09
CVAC	0.54 ± 0.17	29.84 ± 124.36	4.62 ± 3.44
GAC	0.61 ± 0.11	-47.45 ± 86.34	3.35 ± 1.59
HAC	0.66 ± 0.14	19.52 ± 35.48	3.01 ± 1.22

Table 4.4: Quantitative comparison between methods using the Dice similarity coefficient (DSC), relative volume difference (RVD) and average boundary distance (ABD). STAPLE mask of manual segmentations was used as reference ground-truth. Methods were: OTSU thresholding, fuzzy C-means clustering (FCM), Chan-Vese active contour (CVAC), geodesic active contour (GAC), hybrid active contour (HAC). Initialization for CVAC, GAC and HAC was the same. Parameters of each method were tuned to reach best mean DSC.

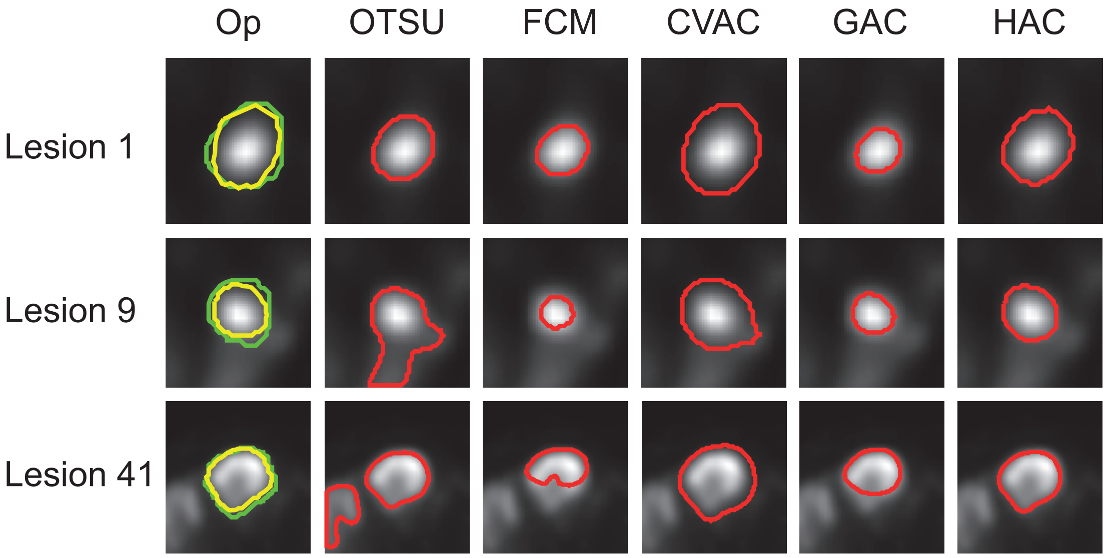


Figure 4.4: Qualitative comparison between FCM-HAC and other methods for three different lesions. The methods were OTSU thresholding, Fuzzy C-Means (FCM), Chan-Vese Active Contour (CVAC), Geodesic Active Contour (GAC). Green line is the Op1 contour, yellow line is the Op2 contour, red line is the FCM-HAC algorithm contour. Lesion 1 and lesion 9 are duodenum primary NETs, lesion 41 is a ileum primary NET. In lesion 1 image, lesion and background are two well-distinct regions. Lesion 9 image has a heterogeneous background due to adjacent uptaking tissue. Lesion 41 has a necrotic core and the background is heterogeneous.

4.3.2 Robustness analysis results

Figures 4.5 and 4.6 show ICC values in the form of diagram bars for consistency of radiomics features to segmentation and number of levels respectively. In overall, 22 out of 45 RFs resulted highly robust to segmentation (ICC > 0.9) using no threshold and 39 out of 45 when using 40% SUVmax threshold. Regarding discretisation, 16 out of 45 RFs resulted highly robust using no threshold and 13 out of 45 when using 40% SUVmax threshold. In order to have a graphical representation of which RFs were consistent to both segmentation and discretisation, figure 4.7 shows a scatter plot where the x-axis is the segmentation ICC and the y-axis is the discretisation ICC. Eight RFs were highly robust to both segmentation and discretisation using no threshold, whereas using 40% SUVmax threshold led to thirteen highly robust RFs.

4.4 Discussion

Neuroendocrine tumor segmentation Since radiomics has a strong potential in the diagnosis and prognosis of neuroendocrine tumor in ^{68}Ga-DOTA-TOC PET/CT images, accurate semi-automatic segmentation of these tumors is useful

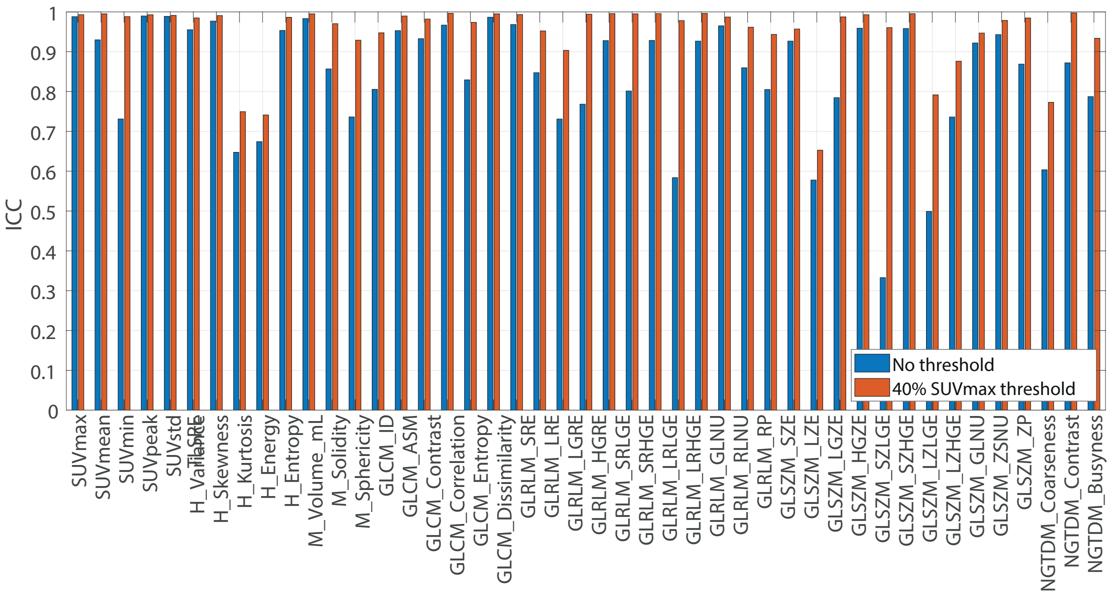


Figure 4.5: Diagram bars of intra-class correlation coefficient (ICC) for consistency to segmentation (Op1, Op2, FCM-HAC). In this experiment, number of levels was fixed at 64. Blue bars are ICC values without thresholding while red ones are ICC values obtained using a 40% SUVmax threshold.

for saving time and increase reproducibility of results with respect to manual segmentation. However, neuroendocrine tumor segmentation is not a trivial task due to high variability in size, low contrast with neigbouring tissue, and partial volume effects of PET images. A novel semi-automatic segmentation algorithm based on fuzzy c-means clustering and hybrid (image and edge-based) level set (FCM-HAC) was proposed. Quantitative comparison between FCM-HAC and manual segmentations by two expert operators was performed using Dice similarity coefficient (DSC) and relative volume difference (RVD). The FCM-HAC was able to correctly segment lesions from different locations and with different background (figure 4.3). Agreement between FCM-HAC and manual segmentations was comparable to the inter-operator agreement (FCM-HAC vs Op1 DSC $= 0.68 \pm 0.13$; FCM-HAC vs Op2 DSC $= 0.71 \pm 0.14$; Op1 vs Op2 DSC $= 0.70 \pm 0.09$) suggesting that the algorithm might substitute/aid an expert operator for the tumor segmentation task. RVD analysis showed that FCM-HAC tended to over-segment the tumor (FCM-HAC vs Op1 RVD $= 22.54 \pm 31.05$; FCM-HAC vs Op2 RVD $= 13.23 \pm 40.67$; Op1 vs Op2 RVD $= -6.97 \pm 48.69$) with respect to manual segmentation. This occured mainly in the initial and final slices of the tumor region of interest where the contrast between tumor and adjacent tissue was even less evident and therefore the algorithm has difficulty in detecting the tumor boundaries. However, the comparison

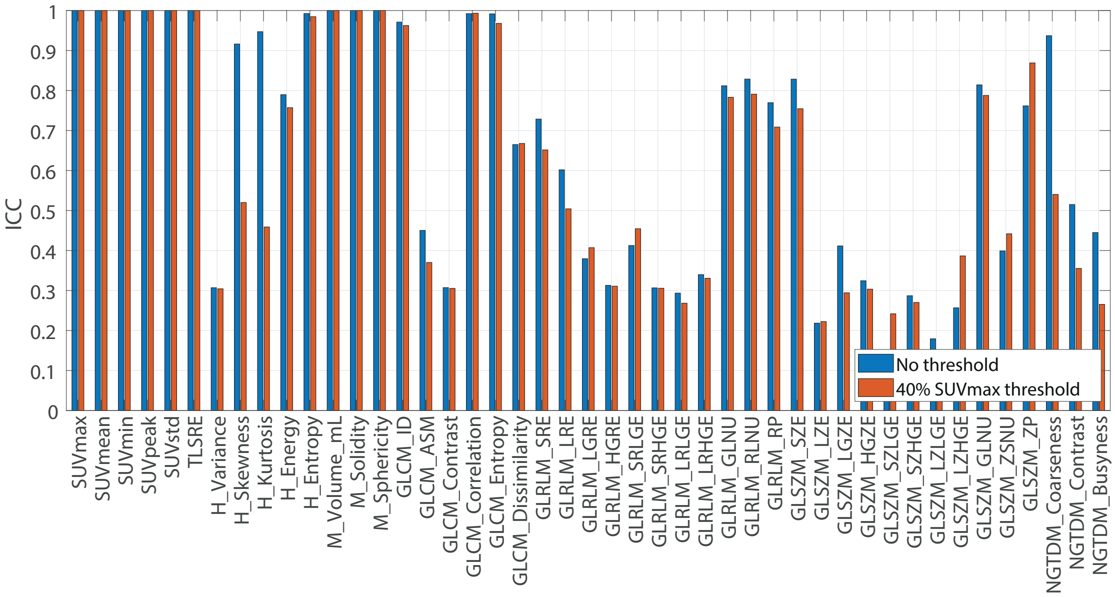


Figure 4.6: Diagram bars of intra-class correlation coefficient (ICC) for consistency to number of levels (32, 64, 128). In this experiment, the VOI was obtained by using STAPLE. Blue bars are ICC values without thresholding while red ones are ICC values obtained using a 40% SUVmax threshold.

of the FCM-HAC algorithm with other methods highlighted the superior performance of FCM-HAC for NET tumors in terms of accuracy (highest mean DSC and lowest mean RVD) and reproducibility (second-lowest standard deviation of DSC and lowest standard deviation of RVD) (table 4.4). This may be motivated by the hybrid nature of the active contour used. Indeed, the curve evolution equation of the HAC (equation 4.13) provided for the contribution of a region-based term and a edge-based term, giving stability to the method both in images where the lesion and the background were two distinct homogeneous regions (region-based term predominance) and in images in which the lesion and the background had similar SUVs and therefore the image gradient was fundamental for the detection of tumor contours (edge-based term predominance). As shown by figure 4.4, the OTSU thresholding method was able to accurately segment the tumor when the image histogram was bimodal (lesion and background were two distinct peaks in the image histogram) but failed when the ROI included uptaking tissue other than the lesion. This led to high variability of performances and a tendency to over-segment majority of tumors (OTSU vs STAPLE RVD = 17.68±73.27), hence, good sensitivity but low specificity. The FCM, using 3 clusters, produced slightly more accurate and reproducible segmentations than OTSU, but the method under-segmented the NET as most part of the lesions present lower SUV in the periphery. Merging the two highest intesity clusters or using 2 FCM clusters did not improve the results. Among active contour methods, the Chan-Vese active contour (CVAC) had the

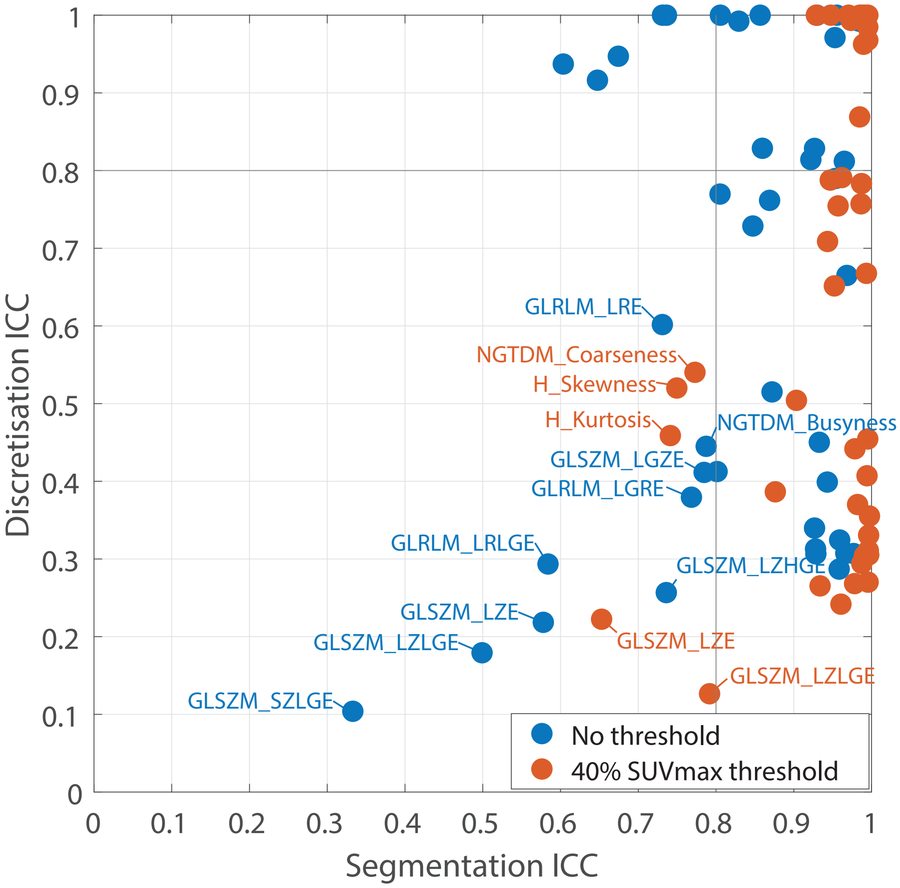


Figure 4.7: Scatter plot of segmentation ICC vs discretisation ICC. Blue points represent ICC values obtained without thresholding while red ones are ICC values obtained using a 40% SUVmax threshold.

worst performance in images with a medium SUV background where the average intensity outside the curve c_2 did not provide a good estimate of the background (equation 4.7), hence, the algorithm failed to stop evolution in correspondence of tumor boundaries. On the other hand geodesic active contours (GAC) which are edge-based active contours, showed high specificity but low sensitivity due to early stopping in correspondence of high image gradient. Average boundary distance (ABD) analysis confirmed the superiority of HAC among other methods showing the best ABD both in terms of mean value and standard deviation.

Multimodal PET/CT NET segmentation An advantage of active contours is their versatility. In fact, it is possible to add other terms in the evolution equation, for example a term related to the texture of the image that is being segmented or, even more interesting for the purpose of the thesis, some terms calculated on a second image such as the CT image in our case. In this regard, as additional experiments, other two active contours were tested: i) FCM-HAC$_{\mathrm{mm1}}$ where the region-based term of the HAC evolution was computed on the CT image and the edge-based on the PET image ii) FCM-HAC$_{\mathrm{mm2}}$ where the region-based term of the HAC evolution was computed on the PET image and the edge-based on the CT image. FCM-HAC$_{\mathrm{mm2}}$ resulted to perform better with respect to FCM-HAC$_{\mathrm{mm1}}$, however, both multimodal active contours performed worse than unimodal segmentation methods for two main reasons: firstly, CT images are acquired at low resolution to minimize patient exposure to radiations; secondly, most part of NETs do not exhibit differences in CT image intensity, especially GEP metastasis. Figure 4.8 show some representative examples of NET in both PET and CT images.

Radiomics features robustness Robustness of radiomics features (RFs) is fundamental for the creation of convincing and reliable radiomics model which in the future may be used in clinical practice. Results showed that 31 RFs were robust (ICC > 0.8) to segmentation, 22 of them were highly robust (ICC > 0.9). Every conventional RF was highly robust except from SUVmin which is the minimum SUV inside the VOI. This is motivated by the high variability between segmentations in the tumor edges where generally uptake is lower. Morphological features describe volume and shape of the tumor, hence, among them M_Volume_mL showed highest robustness (ICC $= 0.86$) followed by M_Sphericity (ICC $= 0.81$) and M_Solidity (ICC $= 0.74$). Indeed, no morphological features yielded high robustness, reasoned by the high influence of the segmentation on these features. However, results show that M_Sphericity has to be preferred compared to M_Solidity for its higher consistency. Among advanced textural features, GLCM showed better consistency. In line with previous studies [24, 25], textural features measuring runs or zones with low grey level, such as GLRLM_SRLGE, GLSZM_SZLGE, GLSZM_LZLGE were not robust to segmentation. This is due to the lower uptake values in lesions edges, where operator and FCM-HAC segmentation showed higher variability. For the same reason, NGTDM_Coarseness and NGTDM_Busyness resulted to be not robust to segmentation. Therefore, the main effect of changing VOI is the inclusion/exclusion of low SUV voxels belonging to the lesion edges leading to high variability of some RFs describing morphology and tumor heterogeneity in terms of runs and zones with low grey level. In overall, 5/6 conventional features, 3/5 intensity histogram features, 0/3 morphological features, 5/6 GLCM based features, 4/11 GLRLM based features, 5/11 GLSZM based features and 0/3 NGTDM based features resulted highly robust to segmentation. As expected, using a 40% SUVmax threshold had a positive impact on the robustness of RFs to segmentation: from 22 to 39 highly robust features. Indeed, the following RF became highly robust: SUVmin, M_Volume_mL, M_Sphericity, M_Solidity, GLCM_Correlation, GLRLM_LGRE, GLRLM_LRE, GLRLM_LRLGE, GLRLM_RLNU, GLRLM_RP, GLRLM_SRE, GLRLM_SRLGE, GLSZM_LGZE, GLSZM_SZLGE, GLSZM_ZP.

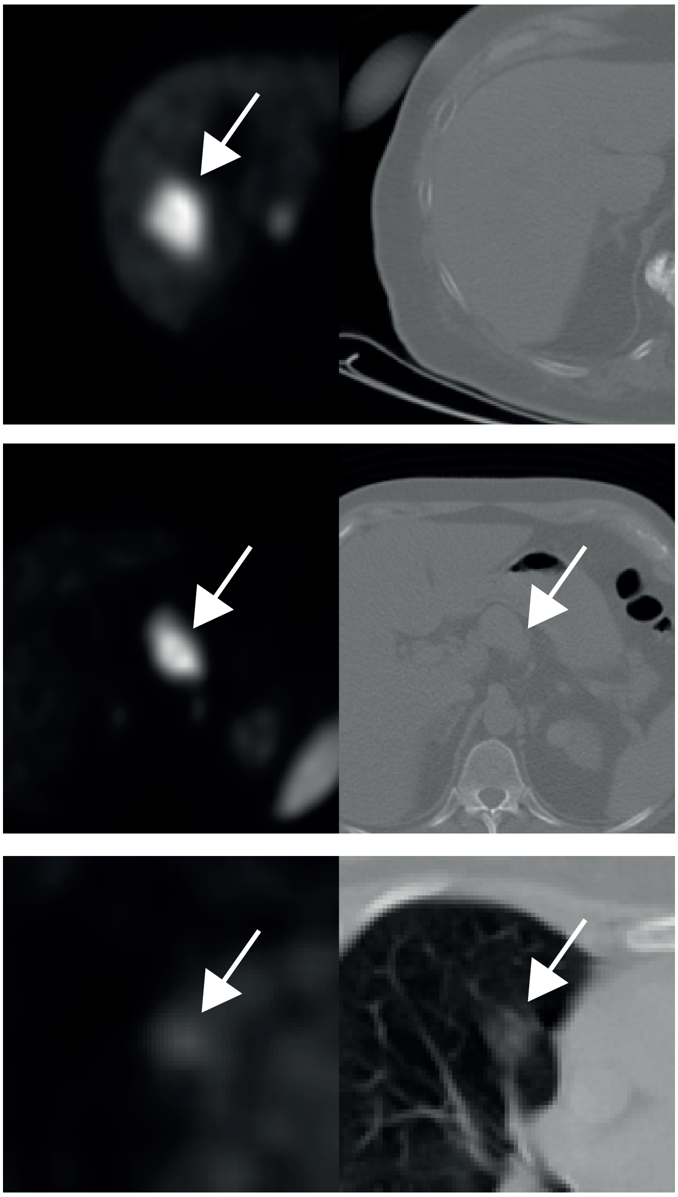


Figure 4.8: Examples of primary and metastatic neuroendocrine lesions in three patients in PET (left) and CT (right) images. First row represents a metastatic liver lesion (first row): the lesion is visible on the PET image (white arrow) but not on the CT image. Second row represents a primary pancreatic NET visible on the CT image but with blurry edges due to low image resolution. Finally, a lung primary NET is shown in the last row. In this case lesion is visible on both PET and CT images and multimodal segmentation could be useful.

Since some of these features may provide useful information for diagnosis and prognosis, the suggestion is to compute them after applying a SUVmax threshold for better reliability. Anyway, applying SUVmax threshold could lead to esclusion of voxels which contribute to the heterogeneity.

Previous studies showed that the choice of discretisation has a huge impact on RFs values [24, 25]. One of the most important discretisation parameter is the number of levels which determines the intensity discretisation resolution and, hence, affecting values of histogram-based features and advanced textural features which require discrete intensity levels. Only 20 out of 45 resulted robust to number of levels, among them 16 RFs were highly robust. Obviously, discretisation does not affect conventional and morphological features (ICC = 1). Regarding histogram-based features, only H_Skewness, H_Kurtosis and H_Entropy resulted to be highly robust histogram with ICC equal to 0.92, 0.95 and 0.98. These finding are in accordance with other studies performed on ^{18}F-FDG-PET images [35]. It is known that discretisation has great impact on advanced textural features computation. Indeed, only 8 of them resulted robust: GLCM_ID, GLCM_Correlation, GLCM_Entropy, GLRLM_GLNU, GLRLM_RLNU, GLSZM_GLNU and NGTDM_Coarseness. The low number of robust features is due to the construction of the GLCM, GLRLM, GLSZM and NGTDM matrices which is strongly dependent on the number of levels. Interestingly, among robust advanced textural features, no features measuring emphasis of high or low grey level as the definition of low and gray is strictly connected to the number of levels used. In this case, applying a 40% SUVmax threshold had a negative impact on the robustness of H_Skewness (ICC from 0.92 to 0.52), H_Kurtosis (ICC from 0.95 to 0.46) and NGTDM_Coarseness (ICC from 0.93 to 0.54). For the rest of RFs the SUVmax threshold had not a significant impact on robustness to number of levels. It has to be noted that whereas segmentation is a predominantly stochastic factor (especially if manual), number of levels is a factor which can be controllated by the user who performs the radiomics analysis, hence is a less critical point.

Overall, when considering robustness to both segmentation and discretisation (Figure 4.7), only 15 out of 45 RFs resulted robust, 8 of them were highly robust using no threshold: SUVmax, SUVmean, SUVpeak, SUVstd, TLSRE, H_Entropy, GLCM_ID, GLCM_Entropy. Four additional RFs were highly robust after the 40% SUVmax threshold: SUVmin, M_Volume_mL, M_Solidity, M_Sphericity and GLCM_Correlation mainly due to increased similarity between segmentations and, therefore, higher consistency to segmentation variation. On the other hand, the following RFs showed an ICC < 0.8 for both segmentation and discretisation: GLSZM_LZE, GLSZM_LZLGE whatever the threshold. Hence, their use is not recommended due to unconsistency.

The results reported in this chapter, from the semi-automatic segmentation to RFs robustness, are a fundamental preliminary step towards the clinical translation of radiomis analysis of NET in ^{68}Ga-DOTA-peptides PET/CT images. However, some limitations have to be acknowledged. Despite the good performance in the NET segmentation by HAC, as every active contour model, the choice of parameters was fundamental. In this study HAC parameters were tuned by testing different combination of parameters and choosing the one with highest DSC, anyway, there is the need of a larger dataset for validating the FCM-HAC, as NETs present a high inter-tumoral heterogeneity. In the dataset analysed, only 2% of cases were G3 as they are rare and generally evaluated with ^{18}F-FDG-PET imaging. Moreover, most part of NETs were G1 (67.4% of cases). Therefore, it was not possible to evaluate the correlation of RFs with tumor NET grade. No robustness analysis in function of the image reconstruction was performed due to technical difficulties with the PET/CT scanner, even if it is well known that PET image reconstruction algorithm has an impact on the RFs values [36].

In the next future, the methodology will be tested with a larger and more balanced dataset to evaluate RFs correlation with tumor histopathological subtypes. Finally, work is already underway on the design of antropomorphic phantoms containing NET prototypes modeled using ^{68}Ga-DOTA-peptides PET/CT images and created by a 3D printer. This gives the possibility of testing each piece of the radiomics model on lesions with different customized features.

4.5 Conclusions

Although the topic of radiomics for tumor characterization was extensively discussed for ^{18}F-FDG-PET images, its application in ^{68}Ga-DOTA-peptides PET/CT images was barely examined due to relatively recent introduction of ^{68}Ga-DOTA-peptides-based radiotracers in clinics. A novel algorithm for neuroendocrine tumor segmentation in ^{68}Ga-DOTA-peptides PET/CT images was proposed. Agreement between algorithm and operators was comparable to inter-operator variability suggesting that the algorithm could substitute/aid an expert operator in the segmentation task saving time and reducing variability. Robustness analysis of radiomics features revealed a higher impact of discretisation settings on radiomics features than segmentation. Using a SUVmax threshold improved robustness to segmentation significantly. Overall, only 15 out of 45 radiomics features resulted robust (ICC $>$ 0.8) to both segmentation and discretisation. Advanced textural features based on low grey level values or large runs/zones emphasis resulted to be not robust to both discretisation and segmentation, hence, their use is not recommended in ^{68}Ga-DOTA-peptides-based radiomics model construction. Current technology of PET/CT scanners limits the use of CT images for multimodal segmentation and characterization of these tumors due to low image resolution, especially along the longitudinal direction. These results are a step towards the clinical translation of radiomics models for the characterization of neuroendocrine tumors in ^{68}Ga-DOTA-peptides PET/CT images.

4.6 References

[1] Arvind Dasari et al. "Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States". In: *JAMA Oncology* 3.10 (Oct. 2017), pp. 1335–1342. ISSN: 23742445. DOI: 10.1001/jamaoncol.2017.0589.

[2] Irvin M. Modlin et al. *Gastroenteropancreatic neuroendocrine tumours*. Jan. 2008. DOI: 10.1016/S1470-2045(07)70410-2.

[3] Silvia Uccella et al. *Immunohistochemical Biomarkers of Gastrointestinal, Pancreatic, Pulmonary, and Thymic Neuroendocrine Neoplasms*. June 2018. DOI: 10.1007/s12022-018-9522-y.

[4] M. Papotti et al. "Expression of somatostatin receptor types 1-5 in 81 cases of gastrointestinal and pancreatic endocrine tumors: A correlative immunohistochemical and reverse-transcriptase polymerase chain reaction analysis". In: *Virchows Archiv* 440.5 (2002), pp. 461–475. ISSN: 09456317. DOI: 10.1007/s00428-002-0609-x.

[5] C. Kratochwil et al. "SUV of [68Ga]DOTATOC-PET/CT Predicts Response Probability of PRRT in Neuroendocrine Tumors". In: *Molecular Imaging and Biology* 17.3 (June 2015), pp. 313–318. ISSN: 18602002. DOI: 10.1007/s11307-014-0795-3.

[6] Alexander R. Haug et al. "68Ga-DOTATATE PET/CT for the early prediction of response to somatostatin receptor-mediated radionuclide therapy in patients with well-differentiated neuroendocrine tumors". In: *Journal of Nuclear Medicine* 51.9 (Sept. 2010), pp. 1349–1356. ISSN: 01615505. DOI: 10.2967/jnumed.110.075002.

[7] Akira Toriihara et al. "Prognostic value of somatostatin receptor expressing tumor volume calculated from 68Ga-DOTATATE PET/CT in patients with well-differentiated neuroendocrine tumors". In: *European Journal of Nuclear Medicine and Molecular Imaging* 46.11 (2019), pp. 2244–2251. ISSN: 16197089. DOI: 10.1007/s00259-019-04455-9.

[8] Luca Filippi et al. "Molecular response assessed by 68Ga-DOTANOC and survival after 90Y microsphere therapy in patients with liver metastases from neuroendocrine tumours". In: *European Journal of Nuclear Medicine and Molecular Imaging* 43.3 (Mar. 2016), pp. 432–440. ISSN: 16197089. DOI: 10.1007/s00259-015-3178-3.

[9] Michael Gabriel et al. "68Ga-DOTA-Tyr3-octreotide PET for assessing response to somatostatin-receptor-mediated radionuclide therapy". In: *Journal of Nuclear Medicine* 50.9 (Sept. 2009), pp. 1427–1434. ISSN: 01615505. DOI: 10.2967/jnumed.108.053421.

[10] Paola Mapelli et al. “Dual tracer 68Ga-DOTATOC and 18F-FDG PET/computed tomography radiomics in pancreatic neuroendocrine neoplasms: An endearing tool for preoperative risk assessment”. In: *Nuclear Medicine Communications* (2020), pp. 896–905. ISSN: 14735628. DOI: 10.1097/MNM.0000000000001236.

[11] Christoph Wetz et al. “Predictive Value of Asphericity in Pretherapeutic [111In]DTPA-Octreotide SPECT/CT for Response to Peptide Receptor Radionuclide Therapy with [177Lu]DOTATATE”. In: *Molecular Imaging and Biology* 19.3 (2017), pp. 437–445. ISSN: 18602002. DOI: 10.1007/s11307-016-1018-x.

[12] Brent Foster et al. *A review on segmentation of positron emission tomography images*. July 2014. DOI: 10.1016/j.compbiomed.2014.04.014.

[13] Michael M. Graham et al. “68Ga-DOTATOC imaging of neuroendocrine tumors: A systematic review and metaanalysis”. In: *Journal of Nuclear Medicine* 58.9 (Sept. 2017), pp. 1452–1458. ISSN: 2159662X. DOI: 10.2967/jnumed.117.191197.

[14] Nobuyuki Otsu. “A Threshold Selection Method from Gray-Level Histograms”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (Jan. 1979), pp. 62–66. ISSN: 0018-9472. DOI: 10.1109/TSMC.1979.4310076.

[15] A. Kroiss et al. “68Ga-DOTA-TOC uptake in neuroendocrine tumour and healthy tissue: differentiation of physiological uptake and pathological processes in PET/CT”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 40.4 (Apr. 2013), pp. 514–523. ISSN: 1619-7070. DOI: 10.1007/s00259-012-2309-3.

[16] Amit Tirosh and Electron Kebebew. “The utility of 68Ga-DOTATATE positron-emission tomography/computed tomography in the diagnosis, management, follow-up and prognosis of neuroendocrine tumors”. In: *Future Oncology* 14.2 (Jan. 2018), pp. 111–122. ISSN: 17448301. DOI: 10.2217/fon-2017-0393.

[17] Mathieu Hatt et al. “The first MICCAI challenge on PET tumor segmentation”. In: *Medical Image Analysis* 44 (Feb. 2018), pp. 177–195. ISSN: 13618423. DOI: 10.1016/j.media.2017.12.007.

[18] Tony F Chan and Luminita A Vese. “Active contours without edges”. In: *IEEE Transactions on Image Processing* 10.2 (2001), pp. 266–277. ISSN: 10577149. DOI: 10.1109/83.902291.

[19] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. “Geodesic active contours”. In: *IEEE International Conference on Computer Vision* February 1997 (1995), pp. 694–699. DOI: 10.1109/iccv.1995.466871.

[20] Zhang Yan et al. "Medical image segmentation using new hybrid level-set method". In: *Proceedings - 5th International Conference BioMedical Visualization, Information Visualization in Medical and Biomedical Informatics, MediVis 2008* 1 (2008), pp. 71–76. DOI: 10.1109/MediVis.2008.12.

[21] Clément Bailly et al. "Revisiting the Robustness of PET-Based Textural Features in the Context of Multi-Centric Trials". In: *PLOS ONE* 11.7 (July 2016). Ed. by Konradin Metze, e0159984. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0159984.

[22] D. Huizing et al. "First step towards radiomics in 68Ga-DOTATATE PET/CT: feature reproducibility". In: *Journal of Nuclear Medicine* 59.Supplement 1 (2018).

[23] Alberto Traverso et al. "Repeatability and Reproducibility of Radiomic Features: A Systematic Review". In: *International Journal of Radiation Oncology Biology Physics* 102.4 (Nov. 2018), pp. 1143–1158. ISSN: 1879355X. DOI: 10.1016/j.ijrobp.2018.05.053.

[24] Ralph T.H. Leijenaar et al. "The effect of SUV discretization in quantitative FDG-PET Radiomics: The need for standardized methodology in tumor texture analysis". In: *Scientific Reports* 5.1 (Aug. 2015), pp. 1–10. ISSN: 20452322. DOI: 10.1038/srep11075.

[25] Baderaldeen A. Altazi et al. "Reproducibility of F18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms". In: *Journal of Applied Clinical Medical Physics* 18.6 (Nov. 2017), pp. 32–48. ISSN: 15269914. DOI: 10.1002/acm2.12170.

[26] Ronald Boellaard et al. "FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0". In: *European Journal of Nuclear Medicine and Molecular Imaging* 42.2 (2015), pp. 328–354. ISSN: 16197089. DOI: 10.1007/s00259-014-2961-x.

[27] Han Back Shin et al. "Digital Imaging and Communications in Medicine (DICOM) information conversion procedure for SUV calculation of PET scanners with different DICOM header information". In: *Physica Medica* 44 (2017), pp. 243–248. ISSN: 1724191X. DOI: 10.1016/j.ejmp.2017.05.063.

[28] Christophe Nioche et al. "Lifex: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity". In: *Cancer Research* 78.16 (Aug. 2018), pp. 4786–4789. ISSN: 15387445. DOI: 10.1158/0008-5472.CAN-18-0125.

[29] James C. Bezdek, Robert Ehrlich, and William Full. "FCM: The fuzzy c-means clustering algorithm". In: *Computers and Geosciences* 10.2-3 (Jan. 1984), pp. 191–203. ISSN: 00983004. DOI: 10.1016/0098-3004(84)90020-7.

[30] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. “Snakes: Active contour models”. In: *International Journal of Computer Vision* 1.4 (Jan. 1988), pp. 321–331. ISSN: 09205691. DOI: `10.1007/BF00133570`.

[31] Simon K. Warfield, Kelly H. Zou, and William M. Wells. “Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation”. In: *IEEE Transactions on Medical Imaging* 23.7 (July 2004), pp. 903–921. ISSN: 02780062. DOI: `10.1109/TMI.2004.828354`.

[32] Alex Zwanenburg et al. “Image biomarker standardisation initiative”. In: *arXiv* 41.2 (Dec. 2016), pp. 366–373. ISSN: 09544879. DOI: `10.1148/radiol.2020191145`. arXiv: `1612.07003`.

[33] Terry K. Koo and Mae Y. Li. “A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research”. In: *Journal of Chiropractic Medicine* 15.2 (2016), pp. 155–163. ISSN: 15563707. DOI: `10.1016/j.jcm.2016.02.012`.

[34] Kenneth O. McGraw and S. P. Wong. “Forming inferences about some intraclass correlations coefficients”. In: *Psychological Methods* 1.4 (1996), pp. 390–390. ISSN: 1082-989X. DOI: `10.1037//1082-989x.1.4.390`.

[35] Alex Zwanenburg. “Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 46.13 (2019), pp. 2638–2655. ISSN: 16197089. DOI: `10.1007/s00259-019-04391-8`.

[36] Jianhua Yan et al. “Impact of image reconstruction settings on texture features in 18F-FDG PET”. In: *Journal of Nuclear Medicine* 56.11 (Nov. 2015), pp. 1667–1673. ISSN: 2159662X. DOI: `10.2967/jnumed.115.156927`.

Chapter 5

Prostate cancer characterization

Part of this chapter has been published as:

- Conference proceeding: B. De Santi, M. Salvi, V. Giannini, K.M. Meiburger, N. Michielli, S. Seoni, D. Regge, F. Molinari **Multimodal T2w and DWI Prostate Gland Automated Registration**, *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019
- Conference proceeding: B. De Santi, M. Salvi, V. Giannini, K.M. Meiburger, F. Marzola, F. Russo, M. Bosco, F. Molinari **Comparison of Histogram-based Textural Features between Cancerous and Normal Prostatic Tissue in Multiparametric Magnetic Resonance Images**, *2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2020

Part of this work was funded by Fondazione Cassa di Risparmio di Cuneo (CRC).

5.1 Introduction

According to the American Cancer Society, prostate cancer (PCa) is the most common cancer in American men, other than skin cancer, with about 190000 new cases of PCa estimated for 2020. For men over 65 years old, there is an incidence of 6 cases in 10 men. In 2018, PCa was also the most diagnosed cancer among european men with about 450000 new cases [1]. The estimated number of deaths in 2020 for PCa in Europe is 78800 making PCa the third cancer with highest number of deaths, after lung and colorectum cancer [2].

Prostate gland anatomy The prostate is a gland of the male reproductive system. Its main function is to produce and release semen fluid during the ejaculation. The gland is located adjacent to the rectum and below the bladder and protects

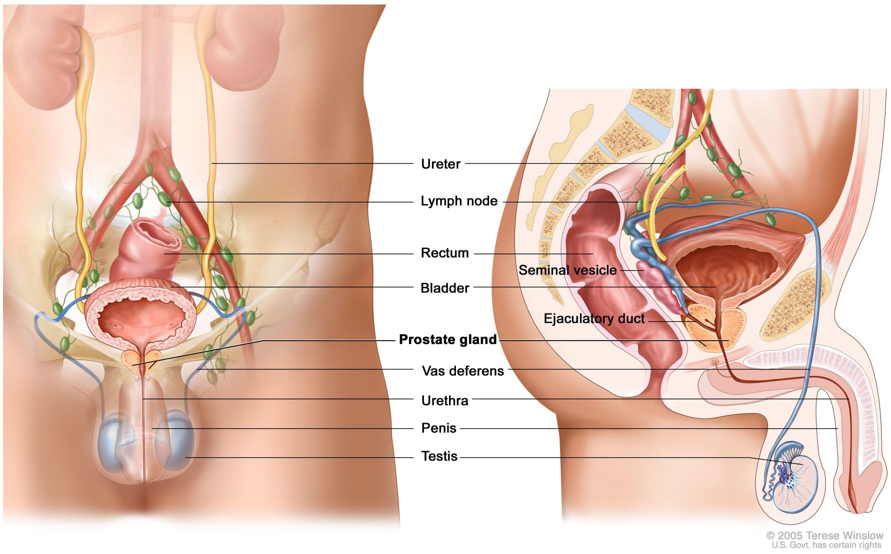


Figure 5.1: Graphical illustration of the prostate gland morphology and location in coronal (left) and sagittal (right) view. Figure available at `https://www.cancer.gov/types/prostate/`

the urethra which crosses the prostate gland (PG) from the bladder to the penis. The morphology and size is often compared to that of a nut (figure 5.1).

Conventionally, the PG can be divided in three parts along the longitudinal direction: base, midgland and apex. Further, from a histological point of view four different zones are identified: transition zone (TZ), central zone (CZ), peripheral zone (PZ) and anterior fibromuscolar stroma (AFS). As shown in figure 5.2, the base of the gland includes the AFS, CZ, the TZ and a small part of the PZ, the midgland includes the AFS, the PZ and the TZ and the apex includes the AFS, the PZ and the TZ.

Prostate cancer As most part of glandular tissue is confined in the central and the peripheral zone, PCa mostly originate in these zones. Indeed, the most diffused type of PCa is adenocarcinoma which originates from gland cells. Although PCa generally grows slowly, cancer cells may multiply and form tumoral masses which can invade neighbouring tissues such as bones and lymph nodes [4]. Figure 5.3 shows a typical example of tumoral growth in the PG. Other PCa types less common than adenocarcinoma are:

- Small cell carcinoma
- Transitional cell carcinoma
- Sarcoma
- Neuroendocrine tumor

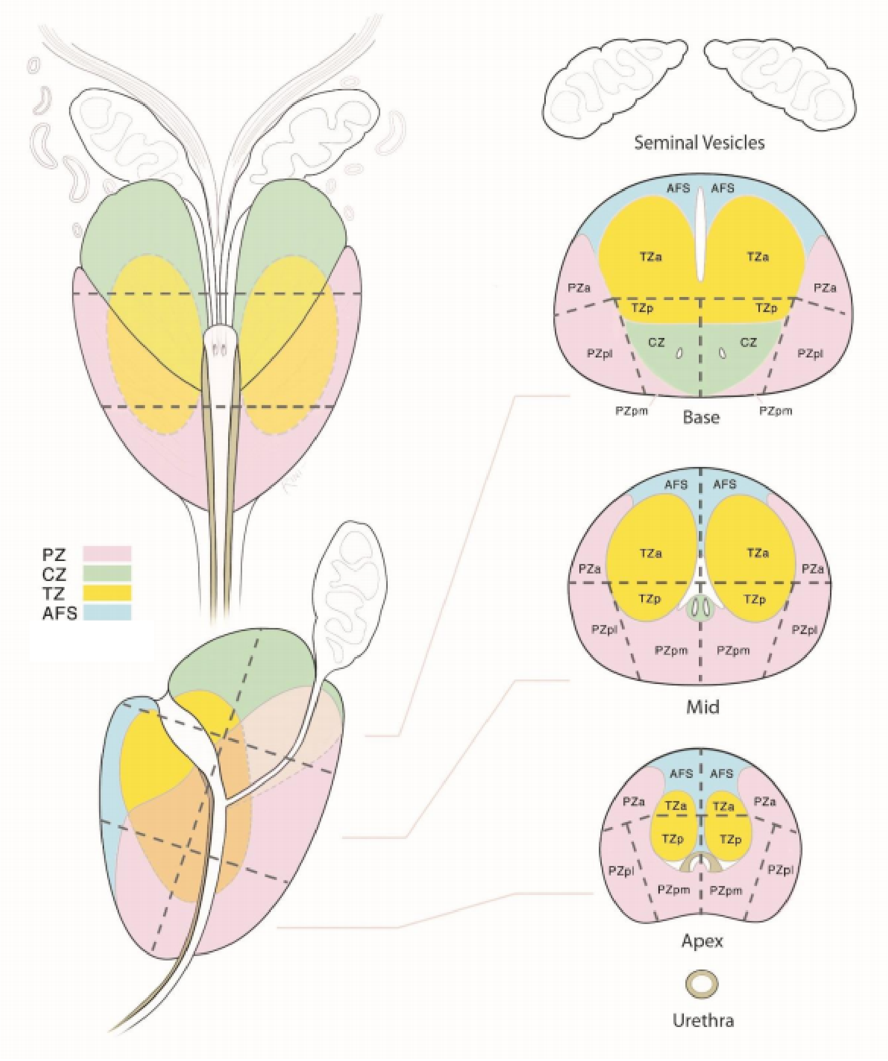


Figure 5.2: Graphical illustration of the prostate gland zones in coronal view (upper-left), sagittal view (bottom-left) and in base, mid and apex slices (right side, from top to bottom) [3].

Early detection, diagnosis and staging PCa begins to show its symptoms only after a time when the tumor invades other parts of the genitourinary system causing for example disturbances during urination, pain during ejaculation, or even more serious, the presence of blood in the urine or semen. The functionality of other systems can also be altered such as the skeletal one, when the tumor invades the bones, causing pain in the hips, thigs and pelvis [5]. Currently, digital rectal examination and prostate-specific antigen are the most used tests for early diagnosis of PCa. In the digital rectal examination (DRE) test, the urologist inserts a finger into the patient's rectum to identify irregularities of the rectal wall by palpation.

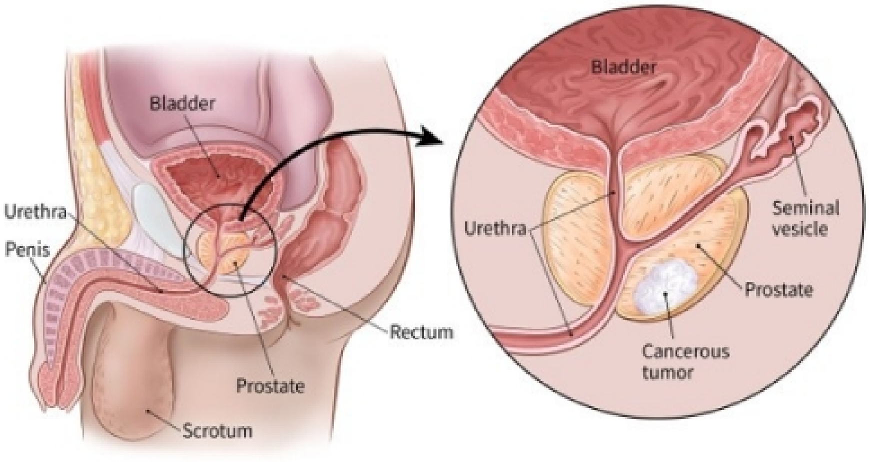


Figure 5.3: Graphical example of a cancerous tumor growth. Figure available at `cancer.org`

Texture abnormalities can be associated to the presence of adenocarcinoma. Although this procedure is non-expensive and low invasive, its diagnostic accuracy is strongly dependent on the examiner's experience. Further, i) only peripheral zone cancer with a significant size can be palpated through the rectal wall, ii) a part from size and solidity evaluation, this examination does not allow to assess cancer aggresiveness [5]. Unlike DRE, the prostate-specific antigen test is a quantitative test. In fact, the amount of the prostate-specific antigen protein, which is responsible for improving sperm flowability after ejaculation, in the blood is measured. A high concentration of PSA in the blood could be associated with a pathology of the PG. PSA concentration is measured in $\mathrm{ng\,mL^{-1}}$ and a concentration greater than $4\,\mathrm{ng\,mL^{-1}}$ is considered abnormal [6]. However, this test has an important drawback: PSA is not specific for PCa, in fact, even benign prostatic hypertrophy or prostatitis can raise the PSA concentration in the blood. According to the American Cancer Society, only 25% of patients with a PSA level between 4 and $10\,\mathrm{ng\,mL^{-1}}$ are diagnosed with PCa, and PSA values above $10\,\mathrm{ng\,mL^{-1}}$ increase the percentage to 50%. There is also a sensitivity problem, indeed, 15% of patients with a PSA level below $4\,\mathrm{ng\,mL^{-1}}$ have PCa. These statistics have been at the center of the discussion of the scientific community that questions the actual advantage of screening and early diagnosis of PCa through PSA, which for some may only lead to over-diagnosis and therefore to excessive clinical cost [7].

Gold-standard technique for PCa diagnosis is fine-needle aspiration biopsy (FNA biopsy). The procedure consists in the insertion of a needle into the prostatic tissue in order to extract small portions of the cancerous tissue to be observed under the lens of a microscope. Obviously, FNA biopsy is performed once the presence of PCa had been confirmed either by DRE or by PSA test [5]. Usually, the entire procedure is guided by trans-rectal ultrasound (TRUS) imaging, hence, an ultrasound probe is coupled with the fine-needle and inserted through the rectum, as shown in figure 5.4. In order to increase sensitivity, the test is repeated from 6 to 16 times to be sure to collect enough cancerous tissue [5].

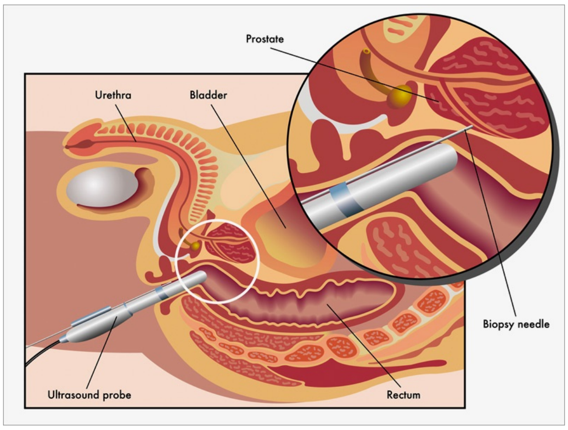


Figure 5.4: Graphical illustration of the transrectal ultrasound-guided fine-needle aspiration biopsy. Available at `https://www.cancerresearchuk.org/`

The analysis of the tissue under the microscope is entrusted to a pathologist who evaluates the size and regularity of the glands, the arrangement of the nuclei, the presence of fused glands, and many other image patterns. The gold-standard system for grading PCa in terms of aggresiveness is the Gleason system: a Gleason pattern/score which ranges from 1 to 5 is assigned to the tumoral tissue, where 1 represesents well-differentiated and not aggressive tumoral cells and 5 represents poorly differentiated or anaplastic tumoral cells which are more aggressive and can rapidly invade nearby tissues [8]. Figure 5.5 shows examples of tissue appearance in histological images of PCas with different Gleason patterns. Usually, the sample includes multiple tumoral areas with different features. For this reason, a Gleason group is used which sums up the Gleason patterns of the two most extended tumoral areas [8].

However, FNA biopsy has some important drawbacks [9]:

- The procedure is invasive and may be not comfortable for the patient and could lead to infections and damages to the rectal wall.
- The aspirated tissue is only a small part of the tumor, therefore, it is difficult to have a complete characterization of the tumor, therefore, potentially leading to wrong diagnosis.

Tumor characterization and detection using multiparametric MRI In the last decade, multiparameter magnetic resonance imaging (mp-MRI) is emerging as a fundamental tool for the characterization of PCa [10]. Figure 5.6 shows the significant increase in PCa studies using mp-MRI.

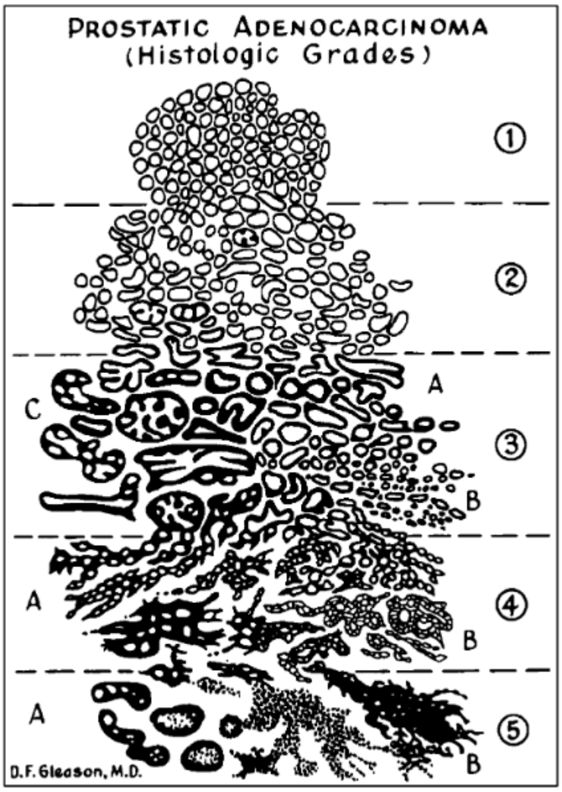


Figure 5.5: Drawing by Dr. Gleason of the five Gleason patterns. [8]

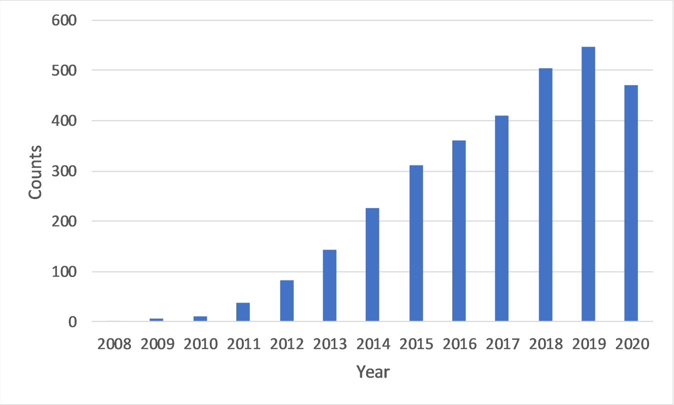


Figure 5.6: PubMed results by year for "Multiparametric MRI Prostate Cancer"

The term "multiparametric" indicates the combination of one or more morphological images, for example T1w and T2w, with or one or more functional images, for example ADC and DCE. In the particular case of PCa, several studies have shown that combining multiple images increases the performance in the detection and characterization of PCa tissue, both in diagnostic and prognostic terms [10].

Particular attention is given to the use of the T2w image to study the morphology of the gland, while as regards the tissue functionality the most used modality is the ADC. Figure 5.7

The ADC image adds information regarding the diffusion of water molecules within the tissue. Figure 5.7 shows a case of prostate with a PZ adenocarcinoma in a T2w, DW_0 (b-value=$0\,\mathrm{s\,mm^{-2}}$), DW_{1000}(b-value=$1000\,\mathrm{s\,mm^{-2}}$) and ADC image. In this case, even if the tumor is visible on the T2w image as a partial darkening of the pixels, on the ADC image the distinction between tumor tissue and healthy tissue is more definite. This is due to the fact that the cancerous tissue is a tissue with high density and complexity which significantly hinders the diffusion of water molecules, and for this reason, the tumor areas on the ADC image have a lower intensity value. On the other hand, on the T2w image it is easy to observe different tissues such as bladder, seminal vesicles, pelvic muscles, rectum, in addition, the clear distinction of the PZ with respect to the central gland (CZ + TZ).

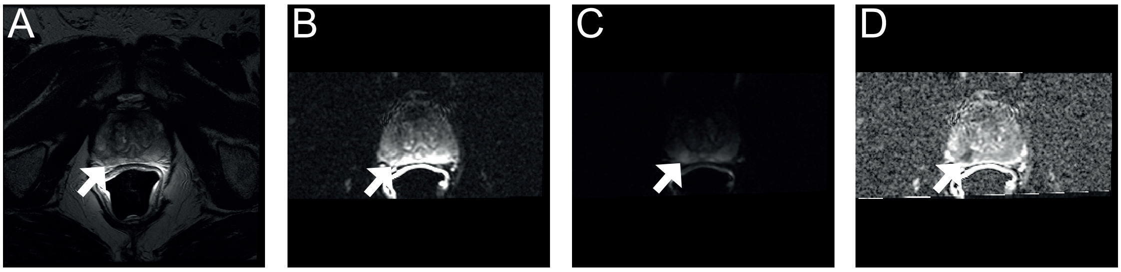


Figure 5.7: Example of peripheral zone (PZ) adenocarcinoma in a T2w image (A), DW image with b-value=$0\,\mathrm{s\,mm^{-2}}$ (B), DW image with b-value = $1000\,\mathrm{s\,mm^{-2}}$ (C), and finally, ADC image (D) which is computed from the two previous DW images. Although the PZ adenocarcinoma (indicated by the white arrows) is visibile on the T2w image as a zone with lower image intensities, this difference is most evident on the ADC image which is extracted by a log-difference operation from the DW images. Moreover, one can notice the different background features between the T2w image, where one can see for example the obturator internus muscle, the psoas muscle or the ischial tuberosities, and the DW images where background is predominantly noisy.

Over the past decade, interest in the development of automatic or semi-automatic systems for the quantitative analysis of PCa in mpMRI has grown in the research community. These studies follow a process consisting essentially of: i) the *ex-vivo* histological image analysis to determine the histological sub-type of the cancer or evaluate its aggressiveness; ii) definition of a region-of-interest (ROI) on the mpMRI image, *in-vivo* or in some cases *ex-vivo*, by visual comparison with the annotations of the histological image; iii) texture analysis on the defined ROI; iv) evaluation of the correlation of textural features with the presence or with the phenotype of the cancer. For example, an automatic computer-aided detection sysytem for the detection of PCa was proposed using first-order and GLCM-based textural features extracted from *ex-vivo* T2w images [11]. Quantitative differentation between normal and cancerous tissue was also evaluated using first-order and textural features extracted from *in-vivo* T2w, ADC together with pharamacokinetic parameters from DCE images [12, 13]. Tumor aggresiveness was also predicted using GLCM-based textural features [14]. Prognosis, in terms of Gleason score upgrading or biochemical recurrence following radiotherapy, resulted to be correlated with textural features extracted from the histogram of PZ cancers in ADC images [15] and in T2w images [16]. However, there is a critical point that is worth raising: in all these studies the ROI on the cancerous tissue was traced on the mpMRI images by operators through visual comparison with histological images. This introduces a dependence of the ROI tracing on the experience of the operator, and therefore also a high inter-operator variability especially in the case of tumors that appear heterogeneous on the mpMRI image. The (semi-) automatic registration of the histological and mpMRI image could partially solve the problem because the definition of the tumor area on the histological image would uniquely determine the ROI on the MR image. The only studies that used a registration-based approach are studies that work with *ex-vivo* MRI images, however, to pave the way for the future clinical use of quantitative mpMRI for PCa it is necessary to develop quantitative systems using *in-vivo* images.

Image registration To conduct a mpMRI-based PCa study, the first problem to be addressed is the registration between the mpMRI images, the morphological and the functional one. This is not a trivial task for several reasons

- patient motion between acquisitions
- physiological motion of organs
- presence of geometric distortion in the phase encode direction caused by the air-filled balloon surrounding the endorectal coil and magnetic field inhomogeneities.

So far, the most used methods to co-register the T2 image and the ADC are semi-automatic, in which an experienced operator defines a series of corresponding points, which represent anatomical landmarks such as the urethra, between the two images. However, these methods are time-consuming and suffer from high inter-operator variability. In order to cope with these limitations, automatic registration methods are beginning to spread in PCa research studies. The common pipeline of these methods consists of a first affine registration phase where motion misalignments are corrected followed by a non-rigid registration phase to correct non-linear geometric distortions. Buerger et al. [17] compared five standard non-rigid registration algorithms in a dataset of T2w and ADC images of 20 patients diagnosed with PCa. In the era of multimodal integration, several medical image registration toolboxes have been developed which allow easy use of a wide range of different registration algorithms. One such toolbox is Elastix which is an ITK-based open-source software package [18]. However, while these toolboxes offer a wide variety of transformations, metrics and registration parameters, their effectiveness depends on the images to be registered. In the particular case of mpMRI prostate registration, a simple and direct application of these toolboxes may lead to registration inaccuracy due to the following main reasons:

- the background characteristics of the T2w image are completely different from those of the DW image. In the T2w image, anatomical structures are visible that are not present in the functional image. This can adversely affect the quality of the registration when the registration metric used is based on intensities (e.g. mutual information) or correlation (e.g. cross-correlation) between images. For this reason, extracting a ROI within which the registration metric is calculated can significantly improve the performance of the registration algorithm.

- initial deformation between the ADC and T2w images is highly variable due to patient or organs motion which has a stochastic nature and the effectiveness of the metrics implemented in these toolboxes (e.g. mutual information, mean square error, cross-correlation) is highly dependent on the initial deformation In this regard, again ROI extraction and an initialization of the non-rigid transformation based on other metrics adapted to the mpMRI prostate images could improve the registration results.

If registration between T2w and ADC is a difficult task, registration between whole-mount histology (WMH) and MRI is even more challenging. Indeed, in addition to the problems introduced by the artifacts present on the MR images, registration between *in-vivo* MR and *ex-vivo* WMH images is an extremely difficult task for several reasons [19]: i) the histological image suffers from artifacts caused for example by dehydration and sectioning operations; ii) while the MR image is acquired *in-vivo*, the histological image is acquired *ex-vivo*, therefore, the gland is deformed because it is no longer supported by the adjacent anatomical structures; iii) the WMH is a color image while the mpMRI are grey scale images iv) different resolutions (WMH: units of µm at 20 x scale, MR: hundreds of µm; v) depending on the image acquisition protocol, the orientation of the histological slices may not coincide with that of the MR slices. These reasons make the automated registration of these images difficult, therefore, often the image acquisition protocols must be chosen in order to minimize orientation errors of the slices, in addition, the algorithm is supported by expert annotations, such as defining anatomical reference points in both images and correspondences between WMH and MR slices.

Objectives In this chapter, a methodology is proposed to differentiate healthy prostate tissue from cancerous tissue in mpMRI images (ADC and T2w). The main novelty of this methodology is the use of advanced and optimized multimodal registration algorithms. Particular attention will, therefore, be given to the automatic algorithm developed for the registration of DW and T2w images and, subsequently, to the registration between T2w and WMH images. Finally, the texture analysis performed on mpMRI images will reveal the presence of imaging biomarkers capable of differentiating, in a multivariate fashion, cancerous tissue from healthy tissue.

5.2 Materials and methods

Each method was developed and validated on a 2.21 GHz quad-core and 16GB RAM.

Patients A total of 25 patients who performed mpMRI examination at the Candiolo Cancer Institute - IRCCS with a biopsy-proven PZ prostate adenocarcinoma greater than 0.5 mL were included in this study. Twenty-one patients underwent a radical prostatectomy within two months of the mpMRI examination. Eight patients showed multiple foci of PCa (3 patients with 3 foci and 5 patients with 2 foci). In overall, 37 tumors were diagnosed: Gleason group 2 (13 cases), Gleason group G3 (13 cases), Gleason group 4 (11 cases). All the participants signed an informed consent form and the study was approved by the local Ethics Committee.

Only for the validation of the ADC/T2w registration algorithm, another 26 patients were included in the study whose images were acquired at the National Institutes of Health (NIH), Bethesda, Maryland between 2008 and 2010. These patients were also diagnosed with PCa. The dataset is public and accessible from the website `cancerimagingarchive.net` under the name of "PROSTATE-MRI".

Image acquisition

Multiparametric magnetic resonance In the IRCCS dataset, mpMRI images were acquired using a 1.5 T scanner (Signa Excite HD, General Electrics Helathcare, USA) with a phased-array coil and an endorectal coil (Medrad, Indianola, USA). T2w axial volumes were acquired using a field of view of 160 mm × 160 mm, TR/TE of 3020/85 ms, acquisition and reconstruction matrix of 384 x 288 and 512 x 512 pixels, slice thickness of 3 mm. DW images were acquired at two different b-values: 0 and 1000 $\mathrm{s\,mm^{-2}}$ using the following acquisition parameters: field of view of 160 mm × 160 mm, TR/TE of 7000/101 ms, acquisition and reconstruction matrix of 128 x 128 and 256 x 256 pixels, slice thickness of 3 mm. Voxel resolution was 0.3125 mm × 0.3125 mm × 3 mm and 0.625 mm × 0.625 mm × 3 mm for the T2w and the DW volumes respectively.

Regarding the NIH dataset, a 3 T scanner was used (Philips Achieva, Best, The Netherlands) with the following acquisition parameters: field of view of 140 mm × 140 mm, TR/TE of 8870/120 ms, slice thickness 3 mm, voxel resolution of

0.27 mm × 0.27 mm × 3 mm for the T2w and field of view of 160 mm × 180 mm, TR/TE of 4584/59 ms, slice thickness of 3 mm, voxel resolution of 1.03 mm × 1.03 mm × 2.73 mm for the DW volumes. The DW volumes were acquired at different b-values: 0, 188, 375, 563 and 750 $\mathrm{s\,mm^{-2}}$.

Whole-mount histology Whole-mount histological (WMH) images were acquired after radical prostatectomy at the San Lazzaro Hospital, Alba, Cuneo by an expert pathologist (M.B. wtih 15 years of experience in uropathology). The prostate specimen was sectioned along the longitudinal direction (apex-base axis) using a slice spacing of 3 mm, thus, reproducing the MR image acquisition. Tissue slices were thick 5 µm and were mounted onto adhesive slides and stained with haematoxylin and eosin.

Apparent diffusion coefficient calculation Apparent diffusion coefficient maps were computed from DW images acquired at b-value = 0 $\mathrm{s\,mm^{-2}}$, DW_0 and b-value = 1000 $\mathrm{s\,mm^{-2}}$, DW_{1000}, using equation 2.11.

Manual annotations The PG was segmented by two expert radiologist (D.R. and F.R. with 9 and 11 years of experience in mpMRI prostate examination) in T2w and DW_0 volumes using the 3D Slicer software [20]. The radiologist D.R. also annotated a set of corresponding anatomical landmarks between T2w and DW_0 images. For each volume, from 30 to 60 landmarks points were defined in correspondence of relevant anatomical references such as urethra, ejaculatory ducts, cysts, peripheral zone boundaries and eventually adenomas by the radiologist (D.R.).

T2w image pre-processing Image pre-processing of the T2w image is fundamental for both the registration step and the texture analysis. In particular, two corrections are important: i) bias field correction: bias field is a low-frequency illumination field produced by the endorectal coil which significantly alters the signal intensity especially near the coil. This artefact was corrected using the N4ITK algorithm[1] which adopts a strategy based on smoothing Bsplines [21]. ii) intensity standardization: in this study, the T2w acquired image was not quantitative, hence, the signal intensity values have no a direct physical meaning and are not consistent among patients. Image standardization means reducing the inter-volume variability of signal intensities of same body regions. Intensity standardization was performed using an algorithm which estimates from a training set the parameters of a standardized histogram [22]. Subsequently, the test set images are transformed

---

[1] The source code available to the public and is implemented in the 3DSlicer software. Bsline grid resolution was set to 3 mm × 3 mm × 3 mm and order was set to 3

in order to match the intensity standardized histogram. As the T2w images presented bimodal histograms, the histogram landmark chosen was the local minimum between the two peaks by detecting zero-crossing points on the derivative of the histogram[2]. In our implementation, since histogram landmarks had similar values between volumes, the standardized histogram parameters were estimated using the whole dataset as training test. The transformation was then applied to the whole dataset of volumes.

5.2.1 Automatic ADC/T2w image registration

In order to align the ADC image and the T2w image, a fully automatic algorithm was developed. The algorithm was able to: i) transform both T2w and ADC images from Image Coordinate System (ICS) to Patient Coordinate System (PCS); ii) define a volume-of-interest (VOI) on both T2w and the DW (b-value = $0\,\mathrm{s\,mm^{-2}}$), DW_0, volumes that includes the PG; iii) detect relevant keypoints in both images; iv) identify correspondences between keypoints between the T2w and DW_0 images; v) estimate the deformation field as a cascade of an affine and a Bspline non-rigid registration using a multimetric registration cost function (euclidean distance between corresponding keypoints and advanced Mattes mutual information).

Coordinates transformation Slice origins and orientations of T2w and DW images were adjusted by transforming image coordinates in the patient coordinate system, as described in the Appendix section A.5. DW and T2w images were interpolated in order to have an isotropic voxel resolution equal to $0.5\,\mathrm{mm} \times 0.5$ mm $\times$ $0.5\,\mathrm{mm}$.

Automatic VOI definition Five central slices 3 mm a part of the T2w and DW_0 volumes were processed. In the DW image, a $3\,\mathrm{mm} \times 3\,\mathrm{mm}$ is applied 2-D median filtering was applied. Subsequently, an object-based detection algorithm, which was originally delevoped for histological image analysis [23], was adapted and applied in order to identify the PG and the endorectal coil. PG medio-lateral width and antero-posterior height were estimated by using two scanning lines passing by the mask centroid. At this point, it is possible to define a rectangle with height equal to the PG antero-posterior height and width equal to the PG medio-lateral width. This rectangle is enlarged by a factor equal to 1.2 to ensure the total inclusion of the PG in the ROI. The T2w image is median filtered ($3\,\mathrm{mm} \times 3\,\mathrm{mm}$) and OTSU thresholded. The rectum is segmented using the region growing algorithm in order to identify the lower bound of the PG. The rectangle obtained in the DW_0 image

---

[2] Other parameters chosen were $pc_1 = 2$, $pc_2 = 99.5$, $s1 = 0$, $s2 = 255$. For a detailed description of these parameters refer to the original paper [22]

is positioned in the T2w image to match the lower bound of the PG. The rectangle is enlarged by a factor equal to 1.2. Figure 5.8 graphically describes the image processing steps for the automatic ROI definition. Once the procedure was repeated for each central slice, a volume-of-interest (VOI) with height and weight equal to the height and weight of the largest ROI and depth equal to the number of slices of the volumes was defined.

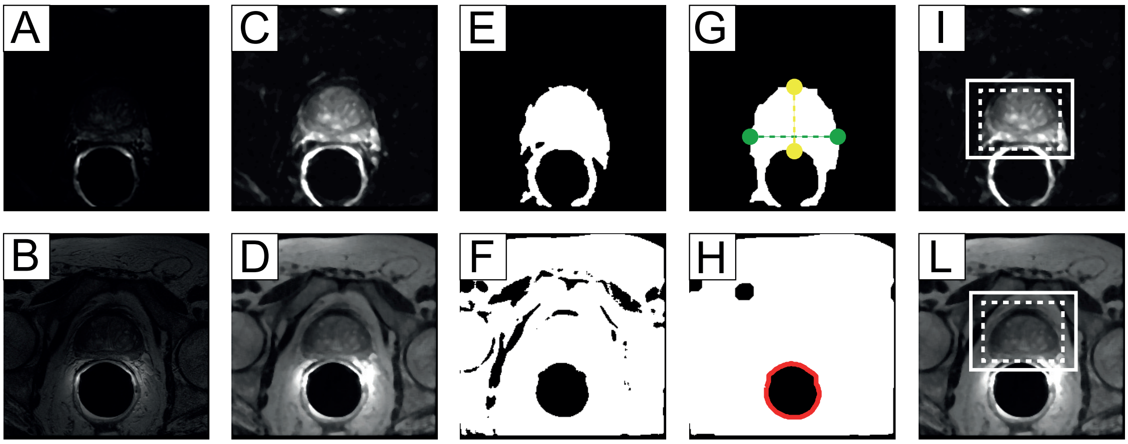


Figure 5.8: Automatic ROI definition in a DW_0 central slice (first row) and in a T2w central slice (second row). Panels A and B are the original images; C and D are the filtered images; E and F are the thresholded images using the object-based detection algorithm in DW_0 and OTSU in T2w; G represents the antero-posterior (yellow points) and medio-lateral (green points) bounds estimation; H is the rectum segmentation (red line); I and L represent the initial rectangle (dashed line) and the final ROI (solid line).

SIFT extraction and matching The T2w and the DW_0 volumes were cropped in correspondence of the VOI automatically obtained. The cropped volumes were filtered using Laplacian of Gaussian (LoG) kernels at different scales. Being the Gaussian kernel with standard deviation:

$$G(x, y, z) = \frac{1}{\sigma^3 (2\pi)^{3/2}} \exp\left[-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right] \tag{5.1}$$

the LoG kernel is defined as the laplacian of $G(x, y, z)$. The LoG kernel is known as a blob detector because it enhances blob-like structures with size proportional to the standard deviation (σ) of the Gaussian kernel. Thanks to the application of the gaussian kernel which is a smoothing kernel, the filtering is also robust to noise. In this study, LoG filtering was applied using different scales (σ from 1 to 4 mm using a step equal of 0.5 mm). Then, the filtered images were summed to obtain a unique image containing image features at different scales. For this reason, this filtering approach was called multiscale LoG filtering. Figure 5.9 (C,D) show examples of the multiscale LoG filtering result for a T2w and aDW_0 image. The most notable effect of the filter is to highlight common structures in both images and in the same time reducing low-scale textural features which are peculiar to the image modality.

Subsequently, the scale-invariant feature transform (SIFT) algorithm was applied to detect keypoints in both filtered images. This algorithm has been extensively used in image registration, also in the medical field [24]. Firstly, candidate key points at different scales using a difference-of-Gaussian function are detected, then low contrast points with respect to the neighborhood or points localized along an edge are discarded. Each key point is descripted by a 128-element vector which is computed from the local image gradient [25]. Figure 5.9 (E,F) show the SIFT key points detected in a DW_0 and in a T2w image.

Let A be the set of SIFT descriptors in the DW image and B be the set of SIFT descriptors in the T2w image, the key point matching task can be addressed as a linear assignment problem which means finding a bijection $f : A \to B$ that minimizes the cost function:

$$\sum_{a \in A} C(a, f(a)) \tag{5.2}$$

where $C(a, f(a))$ is the cost of assigning the descriptor $a \in A$ to a descriptor $b = f(a)$ which belongs to B. The Kuhn-Munkres algorithm, also known as Hungarian algorithm, provides an optimal solution to the linear assignment problem in polynomial runtime complexity by operating on the cost matrix [26]. For our application, we used the Kuhn-Munkres algorithm by defining the cost of assigning a SIFT descriptor a to a SIFT descriptor b as the Euclidean distance between the two descriptor vectors. Matches where partner key points were distant more than 3 mm along the phase-encode direction were discarded considering that geometric distortion mainly occur along this direction [27]. Figure 5.9 (G,H) show an example of matching between key points.

Outlier matches detection In order to discard further false matches, the random sample consensus (RANSAC) algorithm was applied [28]. Given a model and a dataset, the RANSAC algorithm randomly selects k sample data points, estimates the model parameters by fitting the model, then apply the model to the rest of data points and measures the fitting error. If the fitting error is below a threshold t the data point is considered an inlier. This procedure is iterated n times, for each iteration, the number of inliers is saved. The parameter models that yield the highest number of inliers is the optimal model. In this study, we adopted an affine model and the RANSAC pseudocode is reported in Algorithm 1.

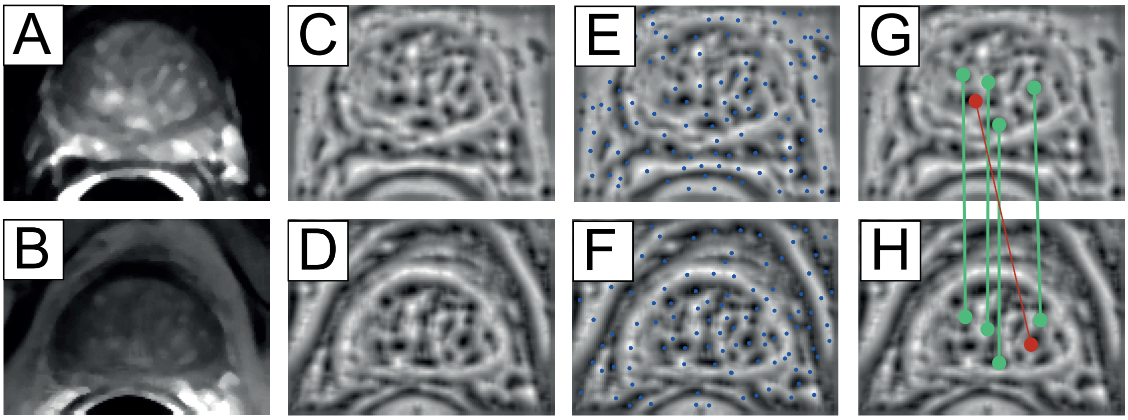


Figure 5.9: SIFT extraction and matching. A and B are the DW_0 and T2w slices cropped; C and D are the results of the multiscale Log filtering; E and F represent the keypoints (blue points) detected by the SIFT algorithm; in G and H the keypoints are matched using the Kuhn-Munkres algorithm, green points are some of the approved matches while red points are matches discarded by the algorithm because they did not meet the heuristic criteria because the partner key points were distant more than 3 mm along the phase-encode direction

Algorithm 1 Random sample consensus for affine transformation estimation

```
for $iteration = 1; iteration < n, iteration + + \ldots$ do
    Randomly select $k$ pairs of matching points
    Estimate affine transformation parameters $(A)$ by minimizing least squares error
    for each remaining pair of matching points do
        Compute affine transformation of the $DW_0$ key point using $A$ parameters
        Measure error $(\epsilon)$ as the distance between the $DW_0$ transformed key point and the matching T2w key point
        if $\epsilon < t$ then
            $nInliers = nInliers + 1$
        end if
    end for
    if $nInliers > Best\_nInliers$ then
        $BestAffine \leftarrow A$
        $Best\_nInliers \leftarrow nInliers$
    end if
end for
```

In this application, $k = 15$, $t = 3\,\text{mm}$, $n = 30000$.

At the end of this stage, a set of corresponding keypoints were defined in the T2w and DW_0 volumes.

Non-rigid registration The final deformation field was estimated using the open-source Elastix software package (version 4.9) which is based on ITK. A toolbox capable of interfacing Matlab and Elastix, ElastixFromMatlab[3], was used to be able to manage both image processing and non-rigid registration on the MATLAB environment. The T2w cropped volume (fixed image) and the DW_0 cropped (moving image) were used as input images for the registration. Both volumes were first filtered with the multiscale log kernel described in 5.2.1. A two-step registration strategy was adopted: affine, as initialization for motion correction, followed by a Bspline (order 3) for non-linear geometric distortions correction. A multi-metric registration cost function was used:

$$C(T_\mu, T2w, DW_0) = w_1 TRE(T_\mu, T2w, DW_0) + w_2 amMI(T_\mu, T2w, DW_0) \quad (5.3)$$

where T_μ is the deformation fields, TRE is the target registration error, computed as mean euclidean distance, between the corresponding SIFT keypoints automatically identified as previously described and amMI is the advanced Mattes mutual information computed using a number of histogram bins equal to 64 [29]. The latter metric was used given the excellent effectiveness in multimodal image registration and the reduced computational cost compared to the standard mutual information [29]. The weights w_1 and w_2 were set to 0.25 and 0.75, respectively. Multi-resolution was adopted using 2 pyramid levels: control point grid spacing equal to 32 and 16 mm. Random sampler was adopted using a number of samples equal to 4096. Adaptive stochastic gradient descent optimizer was used. For final interpolation, a Bspline interpolator with order equal to 3 was chosen. Each of these choices were made after an experimental tuning. This approach will be called **mm_Affine_Bspline** where mm stays for "multimetric".

Other experiments Further experiments were carried out to verify that the registration pipeline chosen was the optimal one. In particular, three other approaches were tested:

- only a non-rigid Bspline transformation was calculated, using the $amMI$ as the only metric without any initialization to correct motion artifacts. This approach will be called **amMI_Bspline**
- an affine transformation estimated from the corresponding SIFT points was calculated using least square error minimization, and subsequently, a Bspline transformation using the amMI as the only metric. This approach will be called **SIFT_Affine + amMI_Bspline**

---

[3] Copyright (C) CNRS and Riverside Research Contributors: Alain Coron, Jonathan Mamou (2010). Website: `https://sourcesup.renater.fr/www/elxfrommatlab/`

- a two-step transformation was calculated: Affine + Bspline, using the amMI as the only metric. This approach will be called **amMI_Affine_Bspline**.

An experimental tuning was carried out to choose the best parameters for the amMI_Affine and the amMI_Bspline. In both cases, best parameters were: two resolutions, grid spacing equal to 32 and 16 mm, number of random samples equal to 4096, maximum number of iterations to 35 and number of histogram bins equal to 64.

Performance metrics and statistical analysis Alignment between volumes was evaluated pre- and post-registration considering using three different metrics. The first metric was the target registration error which is the mean euclidean distance between the corresponding anatomical landmarks annotated by the expert radiologist. The measurement unit is mm. A decrease in TRE after registration indicates that the same corresponding anatomical structures are closer to each other and therefore the images are more aligned. In addition, the overlap between the manual segmentation mask of the PG contoured on the T2w image and the one drawn on the DW_0 image was evaluated pre- and post-registration. Overlap was measured using the Dice-Similarity Coefficient (DSC), defined in equation 4.14 and ranges from 0 (no overlap) to 1 (total overlap). The better the alignment, the higher the DSC value. Finally, mutual information (MI) which measures the amount of information in common between the images [30], was measured within a VOI manually defined by the expert radiologist in order to avoid bias due to the different background properties of the $T2w$ and DW_0 images. For a detailed description of TRE and MI, it is possible to refer to Appendix A.

A statistical test, two-tailed paired t-test, was performed to compare values of metrics pre- and post-registration using a level of significance of 0.05.

5.2.2 Comparison of first-order textural features between cancerous and normal tissue

In this study, first-order textural features were extracted from cancerous and normal tissue in mpMRI and compared in order to identify whether these features or the combination of these features were associated to the presence of PCa. The pipeline consisted of six steps: i) T2w and DWI ($b = 0$ and $b = 1000\,\mathrm{s\,mm^{-2}}$ volumes acquisition and ADC map calculation; ii) Automatic ADC/T2w image registration; iii) Semi-automatic whole-mount histological slice/T2w slice image registration; iv) Tumoral and normal tissue region-of-interest extraction; v) Tissue characterization using first-order textural features; vi) Statistical analysis.

Step i) was already described in this chapter under the paragraphs "Image acquisition" and "Apparent diffusion coefficient calculation". Step ii) was described in 5.2.1. The following paragraphs will describe in detail steps from iii) to vi).

Semi-automatic whole-mount histological slice/T2w slice image registration For each whole-mount histological (WMH) slice, the pathologist (M.B.) and the radiologist (F.R.) manually selected the corresponding T2w slice by visually identifying anatomical landmarks. Further, a set of approximately 20 corresponding points were annotated in correspondence of identifiable anatomical landmarks in both T2w and WMH slices (e.g. urethra, ejaculatory ducts, adenomas). The selected T2w image was cropped as described in 5.2.1. The T2w cropped image and the red channel of the WMH slice were used as fixed and moving image, respectively. A multi-metric and multi-resolution (two pyramid levels) registration was adapted: advanced Mattes mutual information and Euclidean distance between corresponding points with equal weights. An affine transformation followed by a B-spline with grid spacing equal to 32 pixels were computed. Number of iterations was equal to 10. The final registration parameters were tuned by visual assessment and minimization of TRE.

Tumoral and normal tissue region-of-interest extraction Once the deformation field between WMH and T2w was estimated, the tumor mask annotated on the WMH slice was registered to the T2w. Subsequently the same registered mask is positioned on the ADC registered image. This mask will be used as ROI for the cancerous tissue.

In order to reduce bias in the comparison of textural between normal and tumoral tissue due to ROI with either different sizes or slices or zones, the ROI on the normal tissue was obtained automatically by reflecting the ROI of the tumor across the minor axis of an ellipse approximating the PG mask (figure 5.10).

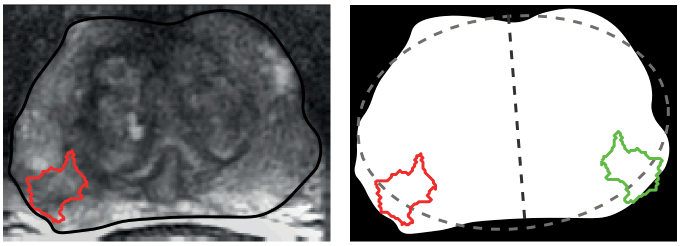


Figure 5.10: Normal tissue region-of-interest extraction. The tumoral mask (contour in red line) is reflected across the minor axis (black dotted line) of the ellipse (grey dotted line) with the same second moment of the prostate gland mask (white mask). The green line are the final normal tissue region-of-interest boundaries.

Tissue characterization using first-order textural features The texture of tumoral and normal tissue was analyzed in the T2w and ADC images using six first-order features: the mean image intensity in the ROI (I_Mean), the standard deviation of image intensities in the ROI (I_SD), the kurtosis of image intensities in the ROI ($I\_Kurtosis$), the skewness of image intensities in the ROI

(*I_Skewness*), the entropy of the grey level histogram (*H_Entropy*) and energy of the grey level histogram (*H_Energy*). While the former four listed features were computed direcyly from the images, the latter two features were computed from the grey level histograms. The histograms were computed using 256 number of levels and the 1^{st} percentile and 99^{th} percentile of grey values in the image dataset as minimum and maximum values, separately for the T2w and the ADC images. It is worthy note that the T2w images were not quantitative, hence, the absolute values of the features have not a direct physical meaning, but they were normalized as described in 5.2 to ensure consistency of image intensities between different volumes. On the other hand, ADC image intensities have a direct physical meaning which is expressed in $10^{-6}\,\mathrm{mm}^2\,\mathrm{s}^{-1}$.

Statistical analysis The texture features extracted from healthy and tumor tissue were compared in terms of mean and standard deviation. Two-tailed paired t-test was used to test the null hypothesis that the mean of the distribution of the difference between the two observations, healthy and tumoral tissue, was zero. The level of significance was set to 0.05. Variables collinearities were measured in a multiple linear regression using the Belsley collinearity. After removing collinear variables, multivariate analysis of variance (MANOVA) was used to compare the means of the two data groups in three conditions: i) using only features extracted from the T2w image; ii) using only features extracted from the ADC image; iii) combining features extract from both modalities (T2w + ADC). A standard cut-off analysis was performed for each single feature and for the MANOVA canonical variables in order to compute the receiver operating characteristics (ROC) curve and the area under it (AUC).

5.3 Results

5.3.1 ADC/T2w registration results

The proposed algorithm took 108.81 ± 29.42 seconds to register each pair of volumes (ADC and T2w). Six volumes from the NIH dataset were discarded as they presented strong image corruption, as was done in a previous study [17].

Mean and standard deviation of metrics pre- and post-registration using the proposed approach (mm_Affine_Bspline) are reported in table 5.1 for the two datasets analysed (IRCCS and NIH). For each metric, paired t-test analysis showed statistical significant improvement of each metric between pre- and post-registration for both datasets. For all patients, registration has always improved image alignment.

Figure 5.11 shows two qualitative examples of registration results.

Quantitative comparison between the proposed algorithm and the other tested approaches is shown in table 5.2. The proposed algorithm achieved the best TRE

Dataset	**Metric**	**Pre-registration**	**Post-registration**	**Δ(%)**
IRCCS, 1.5 T	TRE (mm)	2.34 ± 0.80	1.33 ± 0.31	-40.12 ± 14.42
	DSC	0.87 ± 0.04	0.90 ± 0.04	3.60 ± 2.28
	MI	0.38 ± 0.51	0.49 ± 0.09	30.75 ± 13.71
NIH, 3 T	TRE (mm)	2.04 ± 0.65	1.39 ± 0.17	-25.14 ± 13.97
	DSC	0.91 ± 0.04	0.93 ± 0.02	1.89 ± 2.22
	MI	0.45 ± 0.15	0.59 ± 0.13	40.87 ± 30.13
Total	TRE (mm)	1.93 ± 0.51	1.35 ± 0.26	-33.46 ± 15.95
	DSC	0.89 ± 0.04	0.91 ± 0.03	2.84 ± 2.39
	MI	0.44 ± 0.14	0.53 ± 0.12	35.24 ± 22.81

Table 5.1: Mean and standard deviation values of target registration error (TRE), dice-similarity coefficient (DSC) and mutual Information (MI) pre- and post-registration using the proposed approach (mm_Affine_Bspline) for the IRCCS and the NIH dataset. Relative difference (Δ) is the mean increase (if positive) or decrease (if negative) of the metric normalized to the metric value pre-registration and multiplied by 100.

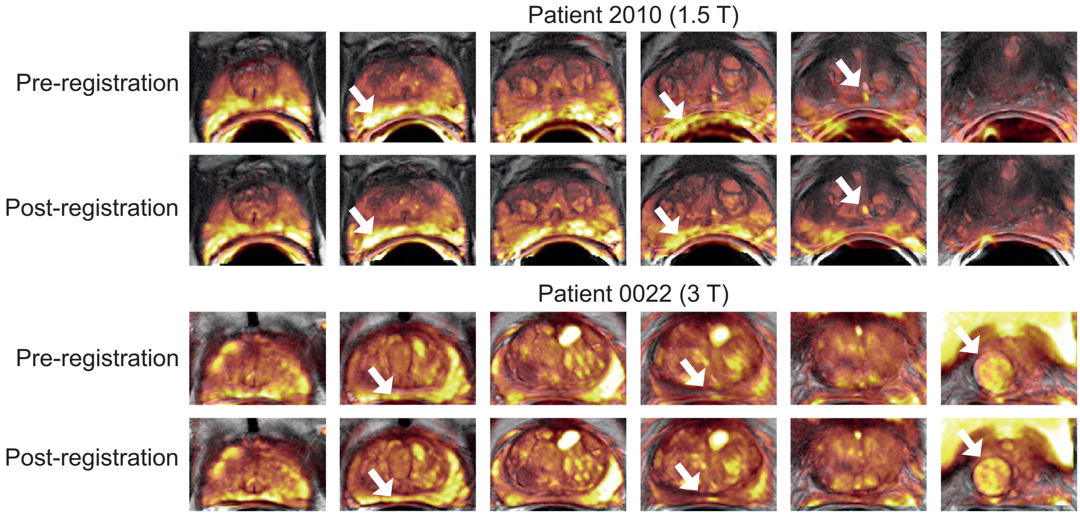


Figure 5.11: Registration results for Patient 2010 and 0022, acquired with the 1.5 T and the 3 T scanner, respectively. For each patient, the first row are the fused images (T2w in grey colormap, DW_0 in red colormap) pre-registration while the second row are the fused images post-registration. Only six equidistant slices from the volumes, from apex to base, are shown here. White arrows indicate misalignents which were corrected after the proposed registration algorithm. The algorithms was able to correctly align the prostate gland in every zone, from apex to base.

among other methods. However, the approach amMI_Affine_Bspline yielded best results in terms of DSC, MI and computational time.

Method	Metric	Post-registration	$\Delta(\%)$	Δt
amMI_Bspline	TRE (mm)	1.52 ± 0.42	-25.61 ± 21.98	$\mathbf{4.72 \pm 0.25}$
	DSC	0.91 ± 0.03	2.67 ± 2.35	
	MI	0.57 ± 0.13	43.25 ± 27.21	
SIFT_Affine + amMI_Bspline	TRE (mm)	1.43 ± 0.43	-28.61 ± 25.81	72.66 ± 29.71
	DSC	0.91 ± 0.03	2.25 ± 3.05	
	MI	0.58 ± 0.13	46.32 ± 28.81	
amMI_Affine_Bspline	TRE (mm)	1.47 ± 0.42	-26.95 ± 25.27	7.40 ± 0.37
	DSC	$\mathbf{0.92 \pm 0.03}$	$\mathbf{3.58 \pm 3.08}$	
	MI	$\mathbf{0.59 \pm 0.13}$	$\mathbf{50.28 \pm 29.03}$	
mm_Affine_Bspline (proposed)	TRE (mm)	$\mathbf{1.35 \pm 0.26}$	$\mathbf{-33.46 \pm 15.95}$	108.81 ± 29.42
	DSC	0.91 ± 0.02	2.84 ± 2.39	
	MI	0.53 ± 0.12	35.24 ± 22.81	

Table 5.2: Mean and standard deviation values of target registration error (TRE), dice-similarity coefficient (DSC) and mutual Information (MI) post-registration for the other registration approaches: Bspline using advanced Mattes mutual information as metric (amMI_Bspline); Affine between SIFT matching keypoints + Bspline using advanced Mattes mutual information as metric (SIFT_Affine + amMI_Bspline); Affine + Bspline using advanced Mattes mutual information as metric (amMI_Affine_Bspline) on the total dataset. Finally, the proposed algorithm, mm_Affine_Bspline, is the multimetric registration using target registration error between SIFT keypoints and advanced Mattes mutual information. Relative difference (Δ) is the mean increase (if positive) or decrease (if negative) of the metric normalized to the metric value pre-registration and multiplied by 100. Δt is the computational time on a 2.21 GHz quad-core and 16GB RAM.

5.3.2 Comparison of first-order features between cancerous and normal tissue

Seven tumors were discarded from the study, since no slice correspondence (WMH slice - T2w slice) could be established by the experts with certainty. In overall, a total of 19 patients with 30 lesions were texture analysed (3 patients with 3 foci, 5 patients with 2 foci, 11 patients with single foci). Regarding WMH-T2w registration, post-registration TRE was $2\,\mathrm{mm} \pm 1.2\,\mathrm{mm}$ compared to an average equivalent lesion diameter of $9.3\,\mathrm{mm} \pm 6.1\,\mathrm{mm}$.

Figure 5.12 shows the comparison between grey level histograms computed in a normal and a cancerous region for a representative case of the study.

Table 5.3 shows the quantitative comparison between first-order features extracted from normal and cancerous tissue in T2w and ADC in terms of mean and standard deviation values, AUC, and paired t-test.

Pearson correlation analysis highlighted high correlation ($r > 0.9$) for the following pairs of features: $H\_Entropy_{T2w}$ - I_SD_{T2w}, $H\_Entropy_{T2w}$ - H_Energy_{T2w}, $H\_Entropy_{ADC}$ - I_SD_{ADC}, $H\_Entropy_{T2w}$ - H_Energy_{ADC}. Indeed, Belsley collinearity assessed multicollinearity introduced by $H\_Entropy_{T2w}$ and $H\_Entropy_{ADC}$ with a proportion of variance-decomposition equal to 0.94 for both modalities. These two features were removed before performing MANOVA analysis.

The MANOVA null hypothesis was rejected with the dimension of the space containing the group means equal to 1 for every subset of features (T2w with

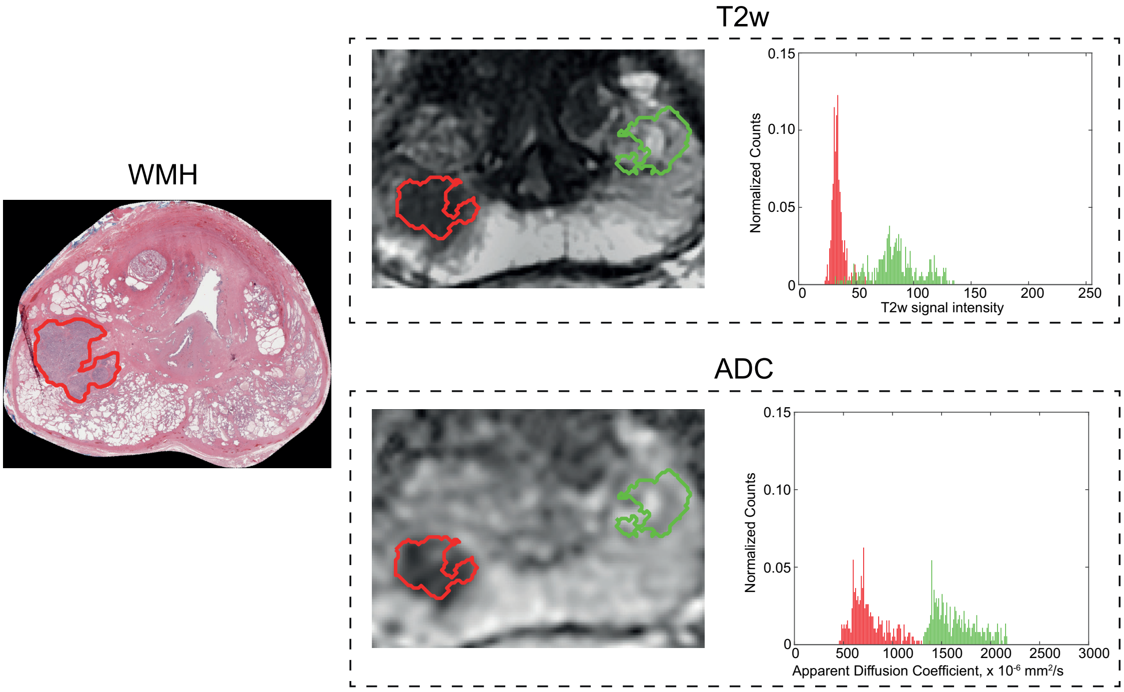


Figure 5.12: Comparison of grey level histograms between normal (green line) and cancerous (red line) tissue in T2w and ADC images of a peripheral zone cancer. Although mean signal intensity appears to be significantly different from the two regions in both images, mean is not the only histogram descriptor which seems to be significant. Indeed, for this case, the T2w histogram of the cancerous region had lower standard deviation than the normal one. The same pattern cannot be observed on the ADC image in which, however, the two distributions are clearly separated from each other.

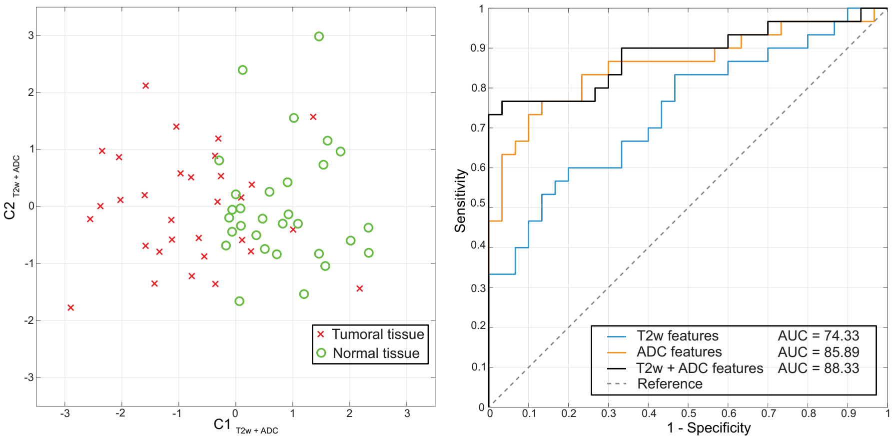
$P < 0.05$, ADC with $P < 0.005$, and T2w + ADC features with $P < 0.005$). The scatter plot of T2w + ADC features for normal and canceours tissue in the MANOVA canonical variables projection is shown in figure 5.13 panel (a). Figure 5.13 panel (b) shows the three ROC curves obtained for the first canonical variables obtained from the T2w, ADC and T2w + ADC features. AUC values were 74.33, 85.89 and 88.33, respectively for T2w, ADC and T2w + ADC.

5.4 Discussion

The use of multiparametric magnetic resonance imaging (mpMRI) for the characterization of prostate cancer is increasing over the last decade. The proposed methodology was able to register T2w, ADC and histological images in order to compare texture of cancerous and normal tissue in *in-vivo* mpMRI images of patients diagnosed with peripheral zone prostate cancer.

Modality	Feature	Normal tissue	Cancerous tissue	AUC
T2w	I_Mean $**$	145.98 ± 67.42	106.61 ± 52.75	0.68
	I_SD $**$	43.38 ± 23.28	27.71 ± 14.94	0.71
	I_Kurtosis	0.44 ± 0.67	0.82 ± 0.66	0.66
	I_Skewness	3.25 ± 1.16	4.20 ± 1.47	0.68
	H_Entropy $**$	5.26 ± 0.77	4.67 ± 0.81	0.71
	H_Energy $*$	0.04 ± 0.03	0.06 ± 0.03	0.71
ADC	I_Mean $***$	1278.68 ± 297.17	994.25 ± 291.26	0.85
	I_SD	248.41 ± 90.23	207.29 ± 64.45	0.65
	I_Kurtosis $*$	-0.07 ± 0.84	0.29 ± 0.59	0.64
	I_Skewness	3.14 ± 1.43	3.33 ± 0.85	0.65
	H_Entropy	5.92 ± 0.56	5.88 ± 0.44	0.58
	H_Energy	0.02 ± 0.01	0.02 ± 0.01	0.56

Table 5.3: Mean and standard deviation values of first-order features for normal and cancerous tissue in T2w and ADC images. AUC is the area under the curve. Number of asterisks indicates the P-value (e.g. ** means P<0.005). I_Mean and I_SD in ADC are measured in $10^{-6}\,\mathrm{mm}^2\,\mathrm{s}^{-1}$.



(a) (b)

Figure 5.13: Projection of T2w + ADC features in the MANOVA space (panel a) using the first two canonical variables ($C1_{T2w+ADC}$, $C2_{T2w+ADC}$ and receiver-operating characteriscs curve (panel b) for the first canonical variable for T2w, ADC and T2w + ADC features).

Image Registration Image registration is an essential step and a low accuracy in the registration between structures in images may affect final diagnostic accuracy. The proposed automatic algorithm for ADC/T2w registration provided excellent results for the private (IRCCS) and public (NIH) datasets acquired using different scanners at 1.5 T and 3 T respectively. The proposed algorithm exploits image processing techniques for automatic region-of-interest extraction, computer vision algorithms (scale-invariant feature transform) to detect corresponding points between images, a multiscale laplacian-of-gaussian (LoG) image processing filter and a two-step registration approach (Affine + Bspline) using a registration cost function formed by an intensity-based term as the advanced mattes mutual information and a geometric term as the Euclidean distance between the corresponding SIFT points. Performance was measured by three metrics: target registration error (TRE), dice-similarity coefficient (DSC) between manual prostate gland segmentation masks and mutual information (MI). TRE is decreased from pre- to post- registration for each case analyzed. Although TRE is the most reliable metric because it measures the true misalignment between reference anatomical structures, DSC and MI were also used. The DSC can be useful in a context where registration can serve as an initialization for segmentation of the prostate in the functional image, whereas the MI is one of the most used metric for multimodal registration. However, DSC is affected by error and variability of the manual segmentations and values of MI strongly depends on the intensity distributions of the involved images [30]. Post-registration TRE in the total dataset resulted 1.35 mm ± 0.25 mm compared to a pre-registration TRE of 1.93 mm ± 0.51 mm with an average decrease of −33.46 % ± 15.95 %. The other metrics also showed an increase in alignment between images ($\Delta DSC = 2.78\,\% \pm 2.53\,\%$, $\Delta MI = 36.53\,\% \pm 21.63\,\%$).

The algorithm has been shown to achieve excellent results when evaluated in datasets acquired with different scanners (IRCCS: TRE = 1.33 ± 0.31, NIH: TRE = 1.39 ± 0.17, Wilcoxon test: P = 0.10). The substantial difference between the two scanners (the one used in IRCCS and the one in NIH) is the intensity of the static magnetic field, in the case of IRCCS equal to 1.5 T and in the case of NIH equal to 3 T It is well known that this parameter has a heavy influence on the image quality and the presence of artifacts in both sequences [31]. Also, the two datasets had significant differences in inital deformation (IRCCS: initial TRE of 2.34 ± 0.80, NIH: inital TRE of 2.04 ± 0.65, Wilcoxon test: P = 0.048) and this motivates the inferior value ΔTRE obtained in the NIH compared to IRCCS. It has to be highlighted that tuning of parameters led to the same configuration of parameters for both datasets. These results suggest that this registration algorithm may be used also in multi-centric datasets.

In comparison with the other registration approaches, the proposed method reported the best value of TRE. Looking at the other metrics, the amMI_Affine_Bspline approach showed the best results in terms of DSC and MI (table 5.2). Nevertheless, priority was given to the TRE metric because aligment of internal structures is the most important factor for a tissue characterization system. Regarding computational time, amMI_Bspline and amMI_Affine_Bspline had the lowest computational time (table 5.2). Introducing image processing for the extraction of SIFT points and above all the RANSAC method to detect outliers between the corresponding points significantly increases the computational time (amMI_Affine_Bspline: $\Delta t = 7.40\,\mathrm{s} \pm 0.37\,\mathrm{s}$, (mm_Affine_Bspline: $\Delta t = 108.81\,\mathrm{s} \pm 29.42\,\mathrm{s}$). However, by reducing the number of iterations of RANSAC it is possible to reduce computational times at the expense of robustness registration. The choice of the parameter must be faced according to the application scenario; in this study, computational time was not a critical point, hence, we preferred high accuracy and robustness with respect to lower computational time.

The issue of multimodal registration between MR images of the prostate has already been addressed in the literature [17]. In their experiment, the Bspline registration method based on the normalized correlation coefficient (NCC) metric showed the largest post-registration TRE equal to 2.05 mm ± 1.05 mm concluding that the Bsplines appear not be suitable for DW/T2w image registration. Although this conclusion seems to be in contrast with the results of this study, it should be noted that in this study important preprocessing steps (multiscale LoG filtering), region-of-interest extraction, detection of corresponding points through the SIFT algorithm, correction of the movement through affine and a multimetric registration cost function were introduced. This motivates the significant improvement in the performance of the final registration compared to the use of the non-rigid Bspline transformation based onthe single NCC metric. This is also confirmed by the decrease in performance obtained using the amMI_Bspline approach, which showed the worst post-registration TRE equal to 1.59 mm ± 0.55 mm.

Compared to ADC/T2w image registration, registration between histological image and MR is even more challenging. It is difficult to find corresponding points automatically since the two images present different image patterns and the resolutions are orders of magnitude apart. Further, in the *in-vivo* MR images the prostate gland is subject to deformations due to the presence of surrounding organs and the endorectal coil. To cope with these critical points, the user intervention was fundamental to reach a satisfactory registration accuracy (post-registration TRE was 2 mm ± 1.2 mm compared to an average equivalent lesion diameter of 9.3 mm ± 6.1 mm).

Comparison of first-order features between cancerous and normal tissue

The ultimate goal of this work was to evaluate the ability of mpMRI, through first-order characteristics and multivariate analysis, to differentiate cancerous tissue from healthy tissue. The methodology was developed with the aim of minimizing user intervention that could introduce errors and variability due to the interpretation and experience of the radiologist. The proposed methodology is innovative since, unlike other studies in the literature in which either the *ex-vivo* mpMRI images are used or the tumor is segmented on the mpMRI images by visual comparison, here the tumor mask annotated in the histological images was semi-automatically registered on the mpMRI image, which was previously obtained by automatic registration of T2w and ADC images. It is important to note that the only user interventions in this methodology are to establish the correspondence between T2w slice and WHM slice and to insert corresponding points on the pair of T2w/WMH images in order to make the registration, hence the mpMRI tumor segmentation, as accurate as possible. Furthermore, image registration is fundamental for the development of systems for the automated detection of prostate cancer in *in-vivo* mpMRI images.

The choice to extract only first-order features from images has two main reasons:

- more advanced features such as GLCM-based or GLRLM-based features require a larger number of voxels to estimate the textural characteristics of the image. The analyzed dataset included small tumors of which the ROI had a few tens of pixels.
- first-order features have a greater robustness to the variation of the shape of the ROI than the advanced features.

Results obtained showed the presence of 4 features extracted from the T2w image able to distinguish the two classes: I_Mean, I_SD, H_Entropy and H_Energy. As expected and as confirmed by other studies [13, 32, 11], cancerous tissue had lower T2w signal intensity values (table 5.3).

The other two features describe the heterogeneity of the image and reported the same AUC value equal to 0.71. Symmetry and tailness of the T2w image did not appear to be associated with the differentiation between healthy and cancerous tissue. In ADC image, I_Mean had $AUC = 0.85$ and $P < 0.0005$. This is in total agreement with the literature, in fact, the high cell density in the tumor areas is an obstacle to the diffusion of water molecules, causing a clear darkening of the image. I_Kurtosis extracted from ADC images was lower in normal tissue (-0.07 ± 0.84) than in cancerous tissue (0.29 ± 0.59), meaning that the cancerous tissue ADC distribution had heavier tails with respect to normal tissue, however, significance and AUC value were low and further investigation is needed.

Analyzing in general the results obtained considering the single features, it can be concluded that if on the one hand, the average value of the ADC was a very important parameter, on the other hand, the T2w image was able to describe in greater detail the heterogeneity of the tissue, due to its better resolution, which was correlated with the presence of cancerous tissue.

As shown by figure 5.12, multi-variate analysis of variance (MANOVA) highlighted the superior performance of the functional ADC image in the tissue differentation with respect to the morphological T2w image (ADC: $P = 0.0003$ and $AUC = 0.86$; T2w: $P = 0.03$ and $AUC = 0.74$). Other studies claimed that the combination of ADC and T2w modalities increases the diagnostic accuracy of PCa [33]. This is confirmed by the results, indeed, combinining ADC and T2w features increased the differentiation accuracy ($AUC = 0.88$). An interesting result, especially in a PCa screening scenario, is the increase in sensitivity introduced by the combination of the two modalities which can be observed on the ROC curve (figure 5.13) with specificity equal to 1 (T2w sensitivity = 0.33, ADC sensitivity =0.47, T2w + ADC sensitivity = 0.73). Interestingly, the highest MANOVA weights of the first canonical variables were I_Mean_{ADC} with $w = 1.3$, $I\_Kurtosis_{T2w}$ with $w = -1$ and $I\_Skewness_{T2w}$ with $w = 0.7$ suggesting that: i) the main contribution of the T2w modality is the heterogeneity characterization and ii) features which did not show high statistical significance in a univariate analysis can be useful in a multivariate framework.

Limitations and future works Some limitations have to be acknowledged. First of all, the advantage of introducing SIFT points matching could be also tested using other nonrigid registration algorithms such as Demons [34], fast elastic image registration [35] and hierarchical local affine registration [36] which showed registration performance comparable to the mm_Affine_Bspline proposed here [17]. Qualitative assessment of corresponding WHM/MR slices is prone to errors due to slight slice orientation difference between WMH and MR images. In future, quantitative approaches will be implemented to find slice correspondences. Finally, the methodology will be extended to other functional imaging modality such as dynamic-contrast enhanced and to central zone cancers which exhibit different image patterns [10].

5.5 Conclusions

In this chapter, a complete methodology for prostate cancer characterization was proposed. The methodology was based on automatic and semi-automatic innovative registration algorithms in order to reduce user intervention which could be prone to errors and high variability. First-order features extracted from mpMRI images resulted to be correlated with presence of prostate cancer. While apparent diffusion coefficient values are lower in cancerous tissue than in normal tissue, T2-weighted images add discriminating features regarding tissue heterogeneity. To conclude, multimodal analysis is a powerful tool for *in-vivo* prostate cancer characterization.

5.6 References

[1] Prashanth Rawla. "Epidemiology of Prostate Cancer". In: *World Journal of Oncology* 10.2 (2019), pp. 63–89. ISSN: 1920-4531. DOI: 10.14740/wjon1191.

[2] G. Carioli et al. "European cancer mortality predictions for the year 2020 with a focus on prostate cancer". In: *Annals of Oncology* 31.5 (May 2020), pp. 650–658. ISSN: 09237534. DOI: 10.1016/j.annonc.2020.02.009.

[3] A.G Fallis. "PI-RADS v2 Prostate Imaging and Report and Data System: Version 2". In: *Journal of Chemical Information and Modeling* 53.9 (2013), pp. 1689–1699. ISSN: 1098-6596. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[4] Louise Clare Brown et al. "Multiparametric MRI to improve detection of prostate cancer compared with transrectal ultrasound-guided prostate biopsy alone: the PROMIS study". In: *Health Technology Assessment* 22.39 (2018), pp. 1–176. ISSN: 1366-5278. DOI: `10.3310/hta22390`.

[5] Yan Kit Fong and Bob Djavan. "Early detection of prostate cancer." In: *Reviews in urology* 7.1 (2005), pp. 63–4. ISSN: 1523-6161.

[6] J. E. Oesterling. *Prostate specific antigen: A critical assessment of the most useful tumor marker for adenocarcinoma of the prostate.* 1991. DOI: `10.1016/S0022-5347(17)38491-4`.

[7] Stefano Ciatto. *Reliability of PSA testing remains unclear.* 2003. DOI: `10.1136/bmj.327.7417.750`.

[8] Peter A Humphrey. "Gleason grading and prognostic factors in carcinoma of the prostate". In: *Modern Pathology* 17.3 (2004), pp. 292–306. ISSN: 0893-3952. DOI: `10.1038/modpathol.3800054`.

[9] Shahrokh F Shariat and Claus G Roehrborn. "Using biopsy to detect prostate cancer." In: *Reviews in urology* 10.4 (2008), pp. 262–80. ISSN: 1523-6161.

[10] John V. Hegde et al. "Multiparametric MRI of prostate cancer: An update on state-of-the-art techniques and their performance in detecting and localizing prostate cancer". In: *Journal of Magnetic Resonance Imaging* 37.5 (May 2013), pp. 1035–1054. ISSN: 10531807. DOI: `10.1002/jmri.23860`. arXiv: `NIHMS150003`.

[11] Anant Madabhushi et al. "Automated detection of prostatic adenocarcinoma from high-resolution Ex vivo MRI". In: *IEEE Transactions on Medical Imaging* 24.12 (Dec. 2005), pp. 1611–1625. ISSN: 02780062. DOI: `10.1109/TMI.2005.859208`.

[12] Farzad Khalvati, Alexander Wong, and Masoom A. Haider. "Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models". In: *BMC Medical Imaging* 15.1 (2015), pp. 1–14. ISSN: 14712342. DOI: `10.1186/s12880-015-0069-9`.

[13] Emilie Niaf et al. "Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric MRI". In: *Physics in Medicine and Biology* 57.12 (June 2012), pp. 3833–3851. ISSN: 0031-9155. DOI: `10.1088/0031-9155/57/12/3833`.

[14] Andreas Wibmer et al. "Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores". In: *European Radiology* 25.10 (2015), pp. 2840–2850. ISSN: 14321084. DOI: `10.1007/s00330-015-3701-8`.

[15] Radu Rozenberg et al. "Whole-Tumor Quantitative Apparent Diffusion Coefficient Histogram and Texture Analysis to Predict Gleason Score Upgrading in Intermediate-Risk 3 + 4 = 7 Prostate Cancer". In: *American Journal of Roentgenology* 206.4 (Apr. 2016), pp. 775–782. ISSN: 0361-803X. DOI: `10.2214/AJR.15.15462`.

[16] Khémara Gnep et al. "Haralick textural features on T2-weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer". In: *Journal of Magnetic Resonance Imaging* 45.1 (2017), pp. 103–117. ISSN: 15222586. DOI: `10.1002/jmri.25335`.

[17] C. Buerger et al. "Comparing nonrigid registration techniques for motion corrected MR prostate diffusion imaging". In: *Medical Physics* 42.1 (Jan. 2015), pp. 69–80. ISSN: 00942405. DOI: `10.1118/1.4903262`.

[18] Stefan Klein et al. "Elastix: A toolbox for intensity-based medical image registration". In: *IEEE Transactions on Medical Imaging* 29.1 (Jan. 2010), pp. 196–205. ISSN: 02780062. DOI: `10.1109/TMI.2009.2035616`.

[19] Hyunjin Park et al. "Registration Methodology for Histological Sections and In Vivo Imaging of Human Prostate". In: *Academic Radiology* 15.8 (Aug. 2008), pp. 1027–1039. ISSN: 10766332. DOI: `10.1016/j.acra.2008.01.022`.

[20] Andriy Fedorov et al. "3D Slicer as an image computing platform for the Quantitative Imaging Network". In: *Magnetic Resonance Imaging* 30.9 (Nov. 2012), pp. 1323–1341. ISSN: 0730725X. DOI: `10.1016/j.mri.2012.05.001`. arXiv: `0402594v3 [arXiv:cond-mat]`.

[21] Nicholas J Tustison et al. "N4ITK: Improved N3 Bias Correction". In: *IEEE Transactions on Medical Imaging* 29.6 (June 2010), pp. 1310–1320. ISSN: 0278-0062. DOI: `10.1109/TMI.2010.2046908`.

[22] J. K. Udupa, P. K. Saha, and R. A. Lotufo. ""Relative Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation"". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.2 (2004), p. 287. ISSN: 01628828. DOI: `10.1109/TPAMI.2004.1265723`.

[23] Massimo Salvi and Filippo Molinari. "Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images". In: *BioMedical Engineering Online* 17.1 (June 2018), pp. 1–13. ISSN: 1475925X. DOI: `10.1186/s12938-018-0518-0`.

[24] Xuesong Lu et al. "SIFT and shape information incorporated into fluid model for non-rigid registration of ultrasound images". In: *Computer Methods and Programs in Biomedicine* 100.2 (Nov. 2010), pp. 123–131. ISSN: 01692607. DOI: `10.1016/j.cmpb.2010.03.005`.

[25] David G. Lowe. “Distinctive image features from scale-invariant keypoints”. In: *International Journal of Computer Vision* 60.2 (Nov. 2004), pp. 91–110. ISSN: 09205691. DOI: 10.1023/B:VISI.0000029664.99615.94.

[26] H. W. Kuhn. “The Hungarian method for the assignment problem”. In: *Naval Research Logistics Quarterly* 2.1-2 (Mar. 1955), pp. 83–97. ISSN: 00281441. DOI: 10.1002/nav.3800020109.

[27] Yousef Mazaheri et al. “Image artifacts on prostate diffusion-weighted magnetic resonance imaging: Trade-offs at 1.5 tesla and 3.0 tesla”. In: *Academic Radiology* 20.8 (2013), pp. 1041–1047. ISSN: 10766332. DOI: 10.1016/j.acra.2013.04.005. arXiv: 15334406.

[28] Shamsundar Kulkarni et al. “RANSAC Algorithm for Matching Inlier Correspondences in Video Stabilization”. In: *Advances in Image and Video Processing* 5.1 (Feb. 2017). DOI: 10.14738/aivp.51.2692.

[29] David Mattes et al. “Nonrigid multimodality image registration”. In: *Medical Imaging 2001: Image Processing*. Ed. by Milan Sonka and Kenneth M. Hanson. Vol. 4322. July 2001, pp. 1609–1620. DOI: 10.1117/12.431046.

[30] Frederik Maes, Dirk Vandermeulen, and Paul Suetens. “Medical image registration using mutual information”. In: *Proceedings of the IEEE* 91.10 (2003), pp. 1699–1721. ISSN: 00189219. DOI: 10.1109/JPROC.2003.817864.

[31] Jacob Sosna et al. “MR imaging of the prostate at 3 tesla: Comparison of an external phased-array coil to imaging with an endorectal coil at 1.5 tesla”. In: *Academic Radiology* 11.8 (Aug. 2004), pp. 857–862. ISSN: 10766332. DOI: 10.1016/j.acra.2004.04.013.

[32] Yu Sun et al. “Multiparametric MRI and radiomics in prostate cancer: a review”. In: *Australasian Physical & Engineering Sciences in Medicine* 42.1 (Mar. 2019), pp. 3–25. ISSN: 0158-9938. DOI: 10.1007/s13246-019-00730-z.

[33] Yu Sun et al. *Multiparametric MRI and radiomics in prostate cancer: a review.* Mar. 2019. DOI: 10.1007/s13246-019-00730-z.

[34] Tom Vercauteren et al. “Diffeomorphic demons: efficient non-parametric image registration.” In: *NeuroImage* 45.1 Suppl (Mar. 2009), S61–S72. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2008.10.040.

[35] Jan Kybic and Michael Unser. “Fast parametric elastic image registration”. In: *IEEE Transactions on Image Processing* 12.11 (2003), pp. 1427–1442. ISSN: 10577149. DOI: 10.1109/TIP.2003.813139.

[36] Christian Buerger, Tobias Schaeffter, and Andrew P. King. “Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation”. In: *Medical Image Analysis* 15.4 (Aug. 2011), pp. 551–564. ISSN: 13618415. DOI: 10.1016/j.media.2011.02.009.

Chapter 6

Conclusions

With the rapid development of novel imaging modalities and new hybrid scanners, multimodal medical imaging has become a fundamental tool for the diagnosis and prognosis of a large part of existing pathologies. However, due to the complexity of these images and pitfalls associated with the image acquisition, accurate and efficient strategies for quantitative characterization of tissue are needed. This work aims to present innovative approaches for tissue characterization in multimodal images with the purpose of improving diagnosis and prognosis in three important oncological applications: thyroid cancer, neuroendocrine tumors and prostate cancer. Since different imaging modalities explains different physical/biochemical information, strategies were developed and optimized for each single clinical context. The proposed algorithms work on functional ultrasound images, nuclear medicine images and multiparametric magnetic resonance images.

Regarding thyroid cancer, multimodal ultrasound imaging is a powerful tool for in-vivo and non-invasive characterization of thyroid nodules, thanks to the wide spectrum of available modes. For example, B-mode imaging can accurately describe morphology and tissue microstructure of the thyroid nodule in terms of echogenicity and heterogeneity. Power doppler ultrasound (PDUS), contrast-enhanced ultrasound (CEUS) and the novel superb microvascular imaging (SMI) allow the visualization and the quantitative characterization of the nodule vascularization at different scales. In this work, an innovative strategy for the three-dimensional and quantitative description of the vascular network of thyroid nodules in PDUS images was presented. The system showed promising results in the differentiation of cytological borderline nodules. Integration with other functional modalities such as CEUS and SMI, or with anatomical B-mode can be performed only in a feature or classifier level fusion strategy due to difficulties in co-registering ultrasound modes.

Multimodal imaging is already widely used in the field of nuclear medicine imaging. The integration of functional positron emission tomography (PET) and single photon emission computed tomography images which are able to target physiological or pathological activities, with anatomical images, such as computed tomography and more recently magnetic resonance images, offers a complete and quantitative in-vivo mapping of tissue functionality. Furthermore, with the recent developments of novel radiotracers it is possible to trace proteins which are tumor-specific. One example is the use of ^{68}Ga-DOTA-peptides radiotracers targetting somatostatin receptors which is a protein expressed by neuroendocrine tumors. In this study, with the final aim of establishing an optimized and reliable radiomics model for NETs characterization, a novel semi-automatic pipeline for the segmentation of neuroendocrine primary and metastatic lesions was developed and validated, showing that the methodology was able to correctly segment lesions from different sites with high heterogeneity variability. Furthermore, as first step towards translation of radiomics model for NET into clinical use, robustness analysis of radiomics features extracted from the lesions was performed in order to identify a subset of features which were reproducible under different segmentation and intensity discretisation settings. Future works will focus on correlation of radiomics features with clinical outcomes of patients. Although CT images are fundamental for anatomical localization of these tumors, their quantitative analysis is limited by the low image resolution. However, in the perspective of automatic detection of NETs, CT imaging will be used for increasing specificity by removing physiological uptaking tissues by means of anatomic atlas registration.

High incidence makes prostate cancer one of the most expensive clinical problems. Development of hardware and software technologies for improvement of diagnosis and prognosis of prostate cancer could have a huge clinical impact. Over the last decade, multiparametric magnetic resonance imaging is increasing its role in the clinical pathway of prostate cancer thanks to its broad spectrum of anatomical and functional sequences. In this work, a full pipeline for differentiation between healthy and cancerous tissue in multiparametric magnetic resonance images (T2w and ADC images) using advanced and optimized image registration methods and texture analysis. The results from the validation on thirty peripheral zone cancer lesions demonstrated the improvement of the diagnostic accuracy when integrating information from T2w and ADC sequences. In future, methodology will be applied to larger dataset and extended for automatic detection of prostate cancer in multiparametric magnetic resonance images.

In conclusion, multimodal imaging was evaluated and discussed in three different clinical contexts. Although multimodal integration in thyroid cancer and neuroendocrine tumors present important pitfalls associated with the image acquisition, quantitative analysis of multiparametric magnetic resonance images is yet feasible and useful for prostate cancer characterization. Nonetheless, the (semi-) automatic approaches proposed in this study resulted to be accurate and robust and can be used for reducing current clinical costs, and improving patient's health and lifestyle.

Appendix A

Image registration

This appendix introduces basic principles of image registration[1]. As discussed throughout this thesis, image registration is fundamental in many clinical situations. Image registration is the process of finding a spatial transformation mapping one image to another.

A.1 Image registration framework

Image registration requires at least two images to be registered. Let $I_M(x)$ be the moving image and $I_F(x)$ be the fixed image, defined on their spatial domains Ω_M and Ω_F of dimension d, image registration is finding the displacement $u(x)$ that makes $I_M(x + u(x))$ spatially aligned to $I_F(x)$. A tranformation T can be defined as the function $T : \Omega_F \subset \mathbb{R}^d \to \Omega_M \subset \mathbb{R}^d$. Finding T the spatially aligns the fixed and the moving image is an optimisation problem which can be written as:

$$T^* = \arg\min_T C(T; I_F, I_M) \tag{A.1}$$

where C is a cost function that has to be minimized. In most applications this cost function includes two terms: a similarity term S which which measures the similarity between the two images through a registration metric and a regularization term R which penalizes irregular transformations:

$$C(T; I_F, I_M) = -S(T; I_F, I_M) + \gamma R(T) \tag{A.2}$$

Two approaches are possible for solving the minimization problem of cost function C: parametric and non-parametric approach. In this appendix only the parametric approach is discussed as it was the only one used for the ADC/T2w prostate gland registration (section 5.2.1).

---

[1] For writing this appendix the main reference was the Elastix manual [1]

A.2 Parametric registration algorithms

In the parametric approach, a parametrisation of the transformation is introduced to rewrite the initial minimization problem (equation A.1) as:

$$\mu^{*} = \arg\min_{\mu} C(\mu; I_F, I_M) \tag{A.3}$$

where μ is the vector of the transformation parameters. In this new form, C is minimized over the elements of the vector μ and not over the entire space of transformations T, thus, reducing the search space. The next section will provide a summary of the most used transformation models and explain how to parametrise them.

A.3 Transformation models

- **Translation** The image is translated by a translation vector as follows:

  $$T(x) = x + t \tag{A.4}$$

  where t is the translation vector. The registration parametrization is $\mu = t$

- **Rigid** A rigid transformation implies rotation and translation of the image, and can be defined as:

  $$T(x) = Rx + t \tag{A.5}$$

  where R is the rotation matrix. R can be parametrised using Euler angles, hence, rigid transformation parameters are $u = (\theta, t)$

- **Similarity** A similarity transformation is a rigid transformation with isotropic scaling, and is defined as:

  $$T(x) = sRx + t \tag{A.6}$$

  where s is the scaling factor. A possible parametrization is obtained by using versors q for the rotation angles so that the transformation parameters are $u = (s, q, t)$.

- **Affine** An affine transformation includes translation, rotation, scaling and shearing of the image. This can be defined as:

  $$T(x) = Ax + t \tag{A.7}$$

  where A is the affine matrix and is composed by elements a_{ij} which do not present any restrictions. The transformation parameters are $u = (a, t)$.

- **B-spline** This transformation belongs to the non-rigid transformations. B-splines are piecewise polynomial functions defined on a grid of control points x_k. Transformation is computed as:

  $$T(x) = x + \sum_{x_k \in X_k} p_k \beta^n \left(\frac{x - x_k}{\sigma} \right) \tag{A.8}$$

  where X_k is the local support of the control point x_k, p_k are the coefficients of each B-spline, β is the B-spline polynomial, n is the order of the polynomial, and σ is control point spacing. In words, the deformation field is computed locally for each control point, then, the image deformation is obtained by interpolation using polynomial functions. Usually control points are distributed on a regular grid, spacing between points is regulated by σ: a large σ produces more regular deformation field but with lower registration accuracy in the control points, small σ produces irregular deformation fields but accurate on control points. The transformation can be parametrised using the order n of the polynomials and B-spline coefficients p_k, which represent the control points displacement.

Depending on the application, a transformation model may be preferred to others. For example, rigid and affine transformations are commonly used for intra-patient image registration (e.g. patient motion), whereas a non rigid transformation may be applied in case of non-linear image distortion, inter-patient alignment or atlas matching. Figure A.1 shows an example of registration between two brain images using different transformation models.

A.4 Registration metrics

The choice of registration metric is fundamental for a successful image registration. One can compute a custom registration metric depending on the images to be registered or use a combination of standard metrics, as described in 5.2.1. Most common image registration metrics are the following:

- **Mean square difference** Mean square difference (MSD) measures a difference between intensities of the two images and is defined as:

  $$MSD(u, I_F, I_M) = \frac{1}{|\Omega_F|} \sum_{x \in \Omega_F} \left[I_F(x) - I_M(x + u(x)) \right]^2 \tag{A.9}$$

  The minimum value of MSD is ideally 0 when the two images are perfectly aligned. This metric may be preferred when corresponding structures have similar intensity values in the fixed and the moving image, e.g. registration of unimodal images.

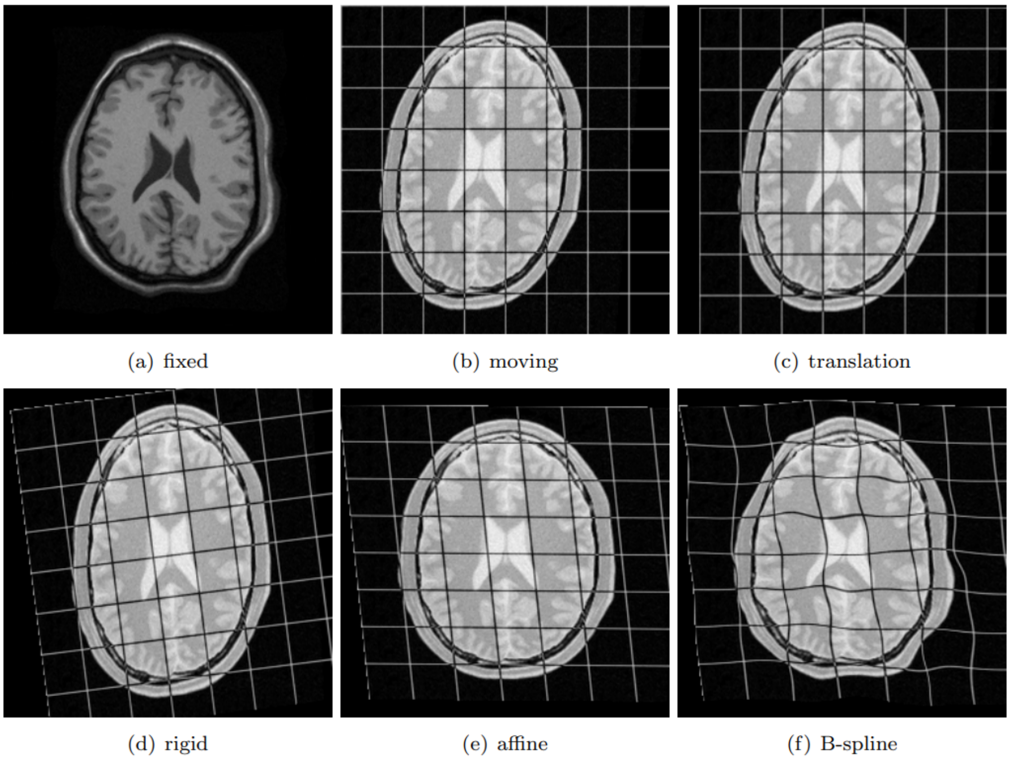


Figure A.1: Image registration between two brain magnetic resonance images using different transformation models. (a) is the fixed image, (b) is the moving image, (c) deformed moving image using translation model. (d) rigid model, (e) affine model and (f) B-spline. Available at https://elastix.lumc.nl/

- **Normalised correlation coefficient** Normalised correlation coefficient (NCC) is a measure of correlation between images and is computed as:

$$NCC(u, I_F, I_M) = \frac{\sum_{x \in \Omega_F} (I_F(x) - \mu_F)(I_M(x + u(x)) - \mu_M)}{\sqrt{\sum_{x \in \Omega_F} (I_F(x) - \mu_F)^2 \cdot \sum_{x \in \Omega_F} (I_M(x + u(x)) - \mu_M)^2}} \tag{A.10}$$

  where μ_F and μ_M are the average intensities of image I_F and I_M in the domain Ω_F. NCC ranges between -1 and 1. Higher the absolute value of NCC better is the alignment, the sign indicates if positive or negative correlation. This metric is used when exists a linear relation between the intensities of the images to be registered.

- **Mutual information** Mutual information measures the amount of information shared by two images and is defined as:

$$MI(u, I_F, I_M) = H(I_F(x)) + H(I_M(x + u(x))) - H(I_F(x), I_M(x + u(x))) \tag{A.11}$$

where $H(I_F(x))$ and $H(I_M(x+u(x)))$ are the marginal entropies of the fixed and the moving image, respectively, and $H(I_F(x), I_M(x+u(x)))$ is their joint entropy. MI ranges from 0 to infinite. This metric assumes that intensities of the two images are related. Its generality make this metric suitable for multimodal image registration where the relation between intensities is generally unknown.

Several variants of MI have been introduced. For example, the **normalized mutual information** (NMI) is a normalized version of the classic MI which is computed as:

$$NMI(u, I_F, I_M) = \frac{H(I_F(x)) + H(I_M(x+u(x)))}{H(I_F(x), I_M(x+u(x)))} \tag{A.12}$$

The NMI ranges from 0 to 2 (maximum alignment). Compared to MI, the NMI has the advantage of allowing direct comparison of registration accuracies.

Another variant is the **advanced Mattes mutual information** (amMI). This metric is substantially equal to MI but entropies are not computed considering each histogram bin, hence, reducing computational cost [2].

- **Target registration error** The target registration error measures the mean euclidean distance between corresponding points in the two images:

$$TRE(u, P_F, P_M) = \frac{1}{N_P} \sum_{k=1}^{N_P} \|x_{Fk} - T(x_{Mk})\| \tag{A.13}$$

where N_P is number of corresponding points, $x_{Fk} \in P_F$ and $x_{Mk} \in P_M$ are the point coordinates of the k-th correspondence in the fixed and the moving image, respectively. $T(x_{Mk})$ is the transformed moving point. Corresponding points can be defined manually or automatically by searching similar image features in both fixed and moving images [3].

A.5 DICOM coordinate systems

Medical images are generally stored in DICOM (Digital Imaging and COmmunications in Medicine) format. The DICOM file includes the image data and a header which reports several information about the image. The image data is a matrix containing intensity values of voxels[2]. The position of each voxel can be identified using the Image Coordinate System (ICS). In the ICS, the three axes i,

---

[2] Assuming volumetric images for consistency with images studied in this thesis

j and k have their origin at the first voxel transmitted and increase their value to the right for the i-axis, the bottom for the j-axis, and backwards for the k-axis. Another possible coordinate system is the patient coordinate system (PCS) which is based on the anatomical directions, hence, x-axis increases from the right to the left hand side of the patient, y-axis increases from the anterior to the posterior side of the patient and z-axis increases from the feet to the head of the patient. Moreover, in PCS the coordinates are expressed in mm. In case of multimodal images, each modality respects its ICS, whereas the PCS is unique for all. Therefore, it is common practice to transform image coordinates of every volume into patient coordinates to have a unique reference system and to be consistent with voxel resolutions.

The transformation from ICS to PCS follows an affine model whose parameters can be obtained by the DICOM header. The affine matrix is:

$$A_{\mathrm{ICS}\rightarrow\mathrm{PCS}} = \begin{bmatrix} F_{11}\Delta r & F_{12}\Delta c & \frac{T_1^1 - T_1^N}{1-N} & T_1^1 \\ F_{21}\Delta r & F_{22}\Delta c & \frac{T_2^1 - T_2^N}{1-N} & T_2^1 \\ F_{31}\Delta r & F_{32}\Delta c & \frac{T_3^1 - T_3^N}{1-N} & T_3^1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.14}$$

where F is a 3x2 matrix containing the direction cosines which define the orientation, using cosine directions, of the first row and the first column of the image in the PCS (first column of F is the orientation of the first row and the second column of F is the orientation of the first column), T^1 and T^N are 3 element vectors specifying the coordinates in the PCS of the upper left hand corner of the first slice and the last slice, respectively. F and T can be found in the DICOM header under the name of Image Orientation (tag: 0020,0032) and Image Position (tag: 0020,0037) in the Image Plane module. The symbols Δr and Δc indicate the row and column pixel resolutions, respectively, and can are contained in the Pixel Spacing (tag: 0028,0030). Assuming two volumes V_1 and V_2, to transform[3] image coordinates of V_1 and V_2 in the PCS:

$$\begin{bmatrix} P_x^{V_1} \\ P_y^{V_1} \\ P_z^{V_1} \\ 1 \end{bmatrix} = M_{\mathrm{ICS}\rightarrow\mathrm{PCS}}^{V_1} \begin{bmatrix} i^{V_1} \\ j^{V_1} \\ k^{V_1} \\ 1 \end{bmatrix} \tag{A.15}$$

and

---

[3] In equations A.15 and A.16 homogeneous coordinates were used to convert the affine transformation in a linear transformation

$$\begin{bmatrix} P_x^{V_2} \\ P_y^{V_2} \\ P_z^{V_2} \\ 1 \end{bmatrix} = M_{\mathrm{ICS}\to\mathrm{PCS}}^{V_2} \begin{bmatrix} i^{V_2} \\ j^{V_2} \\ k^{V_2} \\ 1 \end{bmatrix} \tag{A.16}$$

Once voxel coordinates of both volumes are in the same reference system (PCS), it is possible to decide whether to interpolate the first image to match the second image resolution or viceversa.

This Ph.D. thesis has been typeset by means of the TeX-system facilities. The typesetting engine was pdfLaTeX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TeX-system installation.