FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Image-based Statistical Shape Modelling for Prostate Shape Analysis Towards Radiotherapy Treatments of Prostate Cancer

Ana Isabel Azevedo Cunha

Mestrado em Engenharia Biomédica

Supervisor: João Manuel R.S. Tavares

October 18, 2022

© Ana Isabel Azevedo Cunha, 2022

Image-based Statistical Shape Modelling for Prostate Shape Analysis Towards Radiotherapy Treatments of Prostate Cancer

Ana Isabel Azevedo Cunha

Mestrado em Engenharia Biomédica

Faculdade de Engenharia da Universidade do Porto

October 18, 2022

Resumo

A próstata é uma glândula onde ocorre o segundo cancro mais frequente e a quinta principal causa de morte por cancro entre os homens. Assim, facilitar o diagnóstico médico, criando algoritmos capazes de construir automaticamente modelos tridimensionais da próstata, além de ajudar no diagnóstico médico, poderiam ajudar a reduzir estes números.

Em estudos anteriores, a criação de uma reconstrução tridimensional da próstata a partir de uma imagem tomografia computadorizada não foi alcançado. Os resultados obtidos não foram positivos, uma vez que a segmentação é uma tarefa difícil, devido ao facto de este órgão estar rodeado por estruturas com densidades semelhantes. Além disso, a próstata varia muito entre indivíduos e no mesmo indivíduo, devido ao facto de estar rodeada por estruturas que variam a sua morfologia.

A segmentação da próstata não é uma tarefa fácil e não pode ser feita utilizando técnicas simples. Para auxiliar este processo têm sido desenvolvidos cada vez mais algoritmos computacionais que auxiliem este processo de forma a não criar ambiguidades. Para averiguar se existe algum tipo de diferença na precisão de como as segmentações são realizadas na próstata, foram desenvolvidos dois modelos médios. O *Statistical Shape Model* (SSM) ajudou nesta tarefa, devido às particularidades deste algoritmo.

Os modelos médios foram criados usando dois *datasets*:

- DatasetP: *dataset* onde tem imagens segmentadas da próstata, e além desta, tem imagens onde também está segmentada o recto, o bolbo e os femores;
- DatasetP_GG: possuí imagens onde está segmentada a próstata e os gânglios linfáticos.

Foi utilizada manipulação matemática de forma a analisar as variações de forma de cada um dos conjuntos de imagens. Além disto, os volumes dos modelos médios produzidos foram calculados, e, para analisar visualmente a forma recorreu-se à sobreposição dos modelos.

Abstract

The prostate is a gland that is the site of the second most frequent cancer and the fifth leading cause of cancer death among men. Thus, facilitating medical diagnosis by creating algorithms capable of automatically building three-dimensional models of the prostate, in addition to aiding medical diagnosis, could help reduce these numbers.

In previous studies, creating a three-dimensional reconstruction of the prostate from a CT scan image was not achieved. The results obtained were not positive, since segmentation is a challenging task, due to the fact that this organ is surrounded by structures with similar densities. Furthermore, the prostate varies a lot between individuals and within the individual, due to the fact that it is surrounded by structures that vary its morphology.

The segmentation of the prostate is not an easy task and cannot be performed using simple techniques. To assist this process, more and more computer algorithms have been developed to help this process so as not to create ambiguities. In order to investigate whether there is any difference in the accuracy of how segmentations are performed on the prostate, two average models were developed. The Statistical Shape Model (SSM) helped in this task, due to the particularities of this algorithm.

The average models were created using two datasets:

- DatasetP: dataset where it has segmented images of the prostate, and in addition to this, it has images where it is also segmented the rectum, bulb and femora;
- DatasetP_GG: it has images where the prostate and the lymph nodes are segmented.

Mathematical manipulation was used in order to analyse the shape variations of each of the image sets. In addition, the volumes of the average models produced were calculated, and, to analyse the shape visually, the models were superimposed.

Acknowledgements

The realisation of this project was only possible thanks to the support and trust of several people.

First of all, I thank Prof. João Manuel R. S. Tavares. for his availability and guidance throughout the project.

My sincere thanks to the PhD student Gonçalo Almeida for his help, guidance and continuous support.

I also thank my family and friends who have accompanied and encouraged me during these years. Last but not least I thank Pedro. I would like to thank everyone for their trust and words of encouragement in the best and worst moments.

Thank you all.

Ana Cunha

"... Eu não meti o barco ao mar
Pra ficar pelo caminho..."

José Mário Branco

Contents

List of Figures

List of Tables

Abbreviations

ASM	Active Shape Model
BPH	Benign Prostatic Hyperplasia
CT	Computed Tomography
CZ	Central zone
MRI	Magnetic Ressonance Imaging
PC	Primary Component
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PZ	Peripheral zone
RF	Radiofrequency
SSM	Statistical Shape Models
TZ	Transition Zone

Chapter 1

Introduction

This first chapter aims to contextualise the reader to the issues to be addressed.

1.1 Background

The association established between Imaging and Engineering has allowed the opening of new horizons in medical diagnosis, influencing the clinical approach with increased precision and accuracy of radiological diagnosis. Thus, it is imperative to use Imaging techniques and image enhancement, in order to diagnose earlier the disease, to avoid more serious possible complications.

The evolution of digital imaging techniques has made possible the improvement of disease diagnosis through image processing techniques in the medical area. One of the greatest advantages of image analysis is the elimination of the intrinsic subjectivity present in human evaluation. The evolution of these techniques has allowed the extraction of information which would be difficult or even impossible to obtain otherwise, such as measurements of volume, area, diameter or analysis of the shape and texture of structures.

Today's medical imaging covers a number of areas, which are indispensable in assisting medical diagnosis, from image acquisition, storage and processing, contributing to the improvement of medical diagnoses.

The huge demand of technological advancement for medical diagnosis in certain areas is explained by the numbers indicating the presence of diseases, such as prostate cancer. The process of segmentation is important during the radiotherapy process, so the waves impinge as little as possible on healthy tissue. The goal of this dissertation is then to evaluate in various ways, both visual and statistical, this segmentation process.

Many algorithms have been developed for segmentation in a variety of ways. Statistical Shape Models aim to create the average of shapes from a given dataset, thus forming a model of the desired shape. In this way, through the analysis of statistical models developed from different patients, the way prostates are segmented will be analyzed.

1.2 Objectives

Segmentations vary considerably from observer to observer. Performing correct segmentations in the radiotherapy treatment process becomes even more important, because the more accurate the segmentations are, the lower the risk of exposing healthy tissue to radiation. Thus, we intend to investigate whether there is any difference in the accuracy of how the segmentations are performed when a patient receives treatment in the prostate, for those who, in addition to the prostate, also receive treatment in the lymphatic glanglia .

The main goal of this document is to outcome statistical shape models of the prostate from CT images using surface principal components, do a systematic review of the anatomical parts of the male pelvis, as well as the prostate disorders and compare the differences between the prostate segmentation in two different datasets.

1.3 Report Outline

Apart from this introduction this document has more 5 chapters:

- Chapter 2 is a review of the anatomy of the prostate, as well as the diseases associated with this gland. Understanding the principles behind imaging is important. In this way a review of the most commonly used imaging methods in the treatment of prostate cancer will be made.
- Chapter 3 reviews past work in order to show the reader how difficult prostate segmentation can be. This chapter comtemplates a systematic literature review. It is also mentioned in this chapter important information about the dataset to be used and some of the fundamentals of the applied techniques.
- Chapter 4 shows the methodology, results and discussion.
- Chapter 5, shows the conclusion of the developed work and future work.

Chapter 2

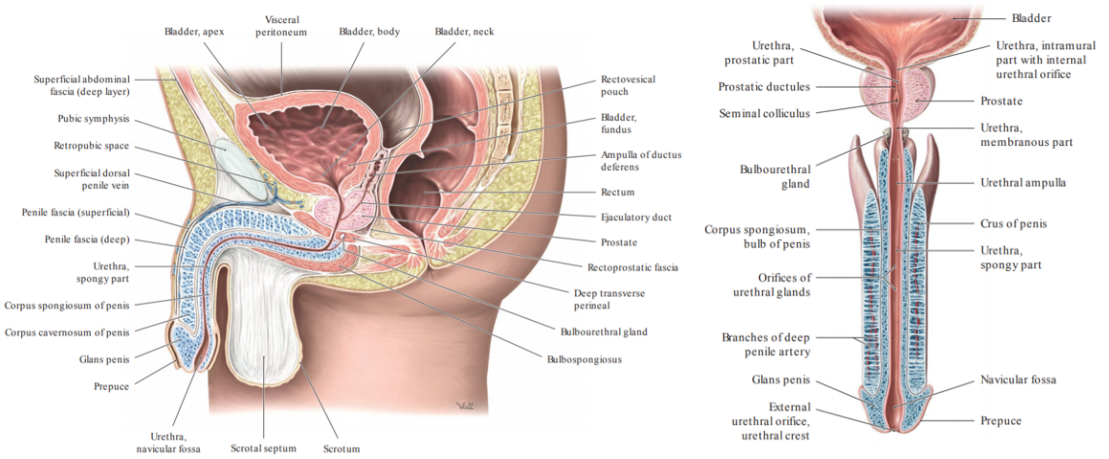
Prostate

The prostate gland is a component of the male reproductive system with approximately the size of a walnut. This gland is the site of several problems, including prostate cancer. Due to its high incidence this is a site of much investigation by researchers.

This chapter aims to investigate the anatomy, structure and function of the prostate, as well as some disorders of this gland. Due to the epidemiology of prostate cancer, its risks, factors and treatment will also be discussed. The imaging modalities most used on prostate cancer will be also reviewed.

2.1 Anatomy, Function and Structure

During the third month of gestation the prostate gland appears. The prostate is a fibromuscular gland with a dense structure. From a histological perspective, the prostatic gland it's a multilayered connective tissue, being composed mainly for smoth muscle [1].



(a) Sagittal section through a male pelvis

(b) Male urethra in longitudinal section

Figure 2.1: Anatomy of a male pelvis, adapted from [2].

In Figure 2.1, it's possible to see the male pelvis anatomy. The prostate is located inferiorly to the bladder, anterior to the rectum, and surrounds the proximal prostatic urethra in the lower pelvis. The normal prostate gland measures $3 \times 3 \times 5\ cm$ approximately, or a volume of 25 cm^3 [3]. Although this is the average size, the figure 2.2 shows the variation of prostate volume with age.

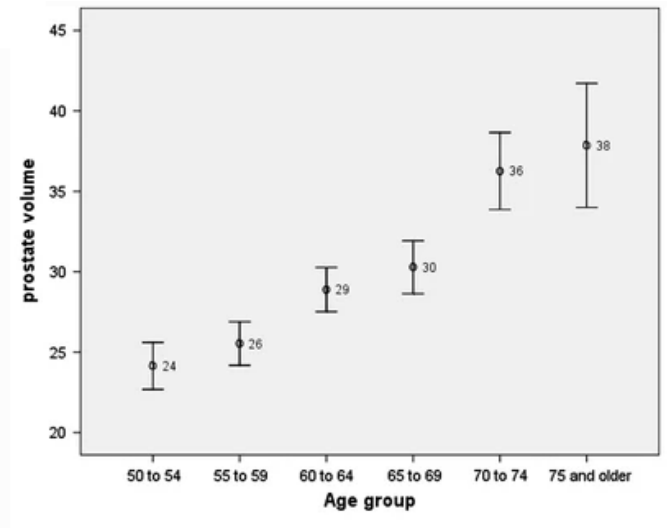


Figure 2.2: Age-stratified normal values for prostate volume. From [4].

Coming from the testicles, the vas deferens join the seminal vesicles and with the ejaculatory ducts cross the prostate. Figure 2.3 illustrates perfectly how the prostate is the place of junction of the seminal duct with the urinary duct. It's primary role is to secrete the components of sperm and drive them into the urethra during orgasm, resulting in ejaculation [5]. This gland secrete an alkaline solution that protects sperm from the acidic vaginal environment. The fluid helps to regulate the vaginal acidity, which extends the sperm's lifetime and allows them to fertilize an ovule for the longest possible period [6].

It is divided into four anatomic regions (Figure 2.3):

- Central zone (CZ): encompasses the ejaculatory duct [7]. Its location is between peripheral and transition zones, at the base of the prostate covering 25% of the gland [8].

- Transition zone (TZ): two small lobules of glandular tissue that surrounds the urethra [7].The TZ represents only 5% of the glandular tissue [9].

- Peripheral zone (PZ): it covers 70% of the prostate, involving the urethra [9].

- Anterior fibromuscular stroma: non-glandular part with fibrous and smooth muscular elements [8].

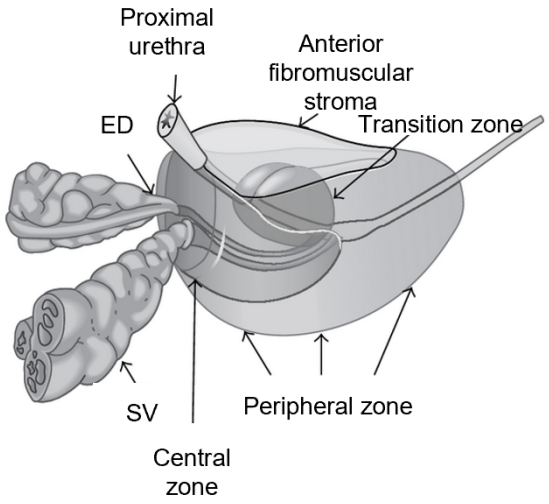


Figure 2.3: Prostate anatomic structures. (ED: ejaculatory ducts; SV: seminal vesicles). Adapted from [8].

2.2 Prostate Disorders

The three main forms of prostatic illness are benign prostatic hyperplasia (BPH), prostatitis and prostate cancer [10, 11]. In the following subsections these pathologies will be addressed.

2.2.1 Benign Prostatic Hyperplasia

The American Urological Association describes benign prostatic hyperplasia (BPH) as a histologic diagnostic that corresponds to the proliferation of smooth muscle and epithelial cells inside the prostatic transition zone [12].

Non-malignant prostate growth is frequent in ageing men [13]. In practically all males, hyperplasia of the prostate gland occurs as a completely age-related phenomena beginning around the age of 40. In practically all males, hyperplasia of the prostate gland occurs as a completely age-related phenomena beginning around the age of 40. In reality, according to various autopsy studies throughout the world, the histologic prevalence of BPH is around 10% for men in their 30s, 20% for men in their 40s, 50% to 60% for men in their 60s, and 80%-90% for men in their 70s and 80s. Most men will, without a doubt, develop certain histologic findings that are compatible with BPH if they live long enough [13].

BPH is caused by an overgrowth of smooth muscle, Figure 2.4, tissue and glandular epithelial tissue, which can be caused by a variety of circumstances including age, late activation of cell development, hereditary factors, and hormone changes [14]. As shown in the Figure 2.4, this disorder can lead to obstruction of the bladder outflow. Causing symptoms such as reduced urinary stream force, intermittency, occasional urinary retention and irritative symptoms like pain in urination. The symptoms of this disease are usually related to the obstruction [15].

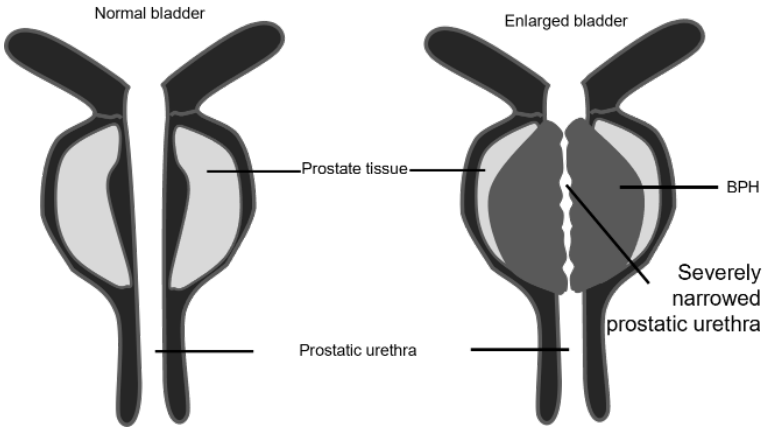


Figure 2.4: Normal prostate compared to a prostate with BPH.

Although BPH is diagnosed histologically, clinicians evaluate men using a multi-faceted approach. Physicians typically employ symptom rating, patient history, diagnostic imaging, such as ultrasound or prostate MRI, and laboratory investigations [12].

If the patient's symptom load necessitates it, drug therapy should be considered [16]. The type of medication used is determined on the symptoms. When deciding which pharmaceuticals to employ, the most significant clinical outcomes of each drug class, their respective adverse effect profiles, and suggested follow-up schedules should all be taken into account.

2.2.2 Prostatitis

Prostatitis is a term that refers to several types of prostate inflammation. This disorder is a common urinary tract infection that many clinicians find difficult to adequately treat. Nearly half of all men are believed to experience prostatitis symptoms at some point in their lives [17].

This disease can be classified into 4 categories [17, 18]:

1. Type I (Category I) or Acute bacterial Prostatitis: this type is characterized by severe prostate symptoms and systemic infection; contains bacteria, which causes an acute bacterial urinary tract infection including the prostate gland. The patient may have symptoms such as fever, local pelvic/perineal pain, irritative symptoms and other generalised symptoms of an acute infection.

2. Type II (Category II) or Chronic bacterial Prostatitis: patients usually have a history of recurrent urinary tract infections. Many people are asymptomatic between infections and react well to antibiotic therapy. Others are still experiencing pelvic and perineal discomfort. The diagnosis is suggested by a history of recurrent urinary tract infections and/or a past response to medication; Prostate inflammation is brought on by a bacterial infection that has been present for a long time.

3. Type III (Category III) or Chronic Prostatitis/Chronic Pelvic pain Syndrome: this prostatitis is distinguished by the presence of chronic pelvic pain sensations and the absence of urogenital infections, making its diagnosis unique.

4. Type IV (Category IV) or Asymptomatic Inflammatory prostatitis: there is inflammation of the prostate in this kind of prostatitis, but no genital urinary tract infection symptoms.

The diagnosis of prostatitis can be complex, from the use of medical imaging, urine analysis and culture, physical exam to semen tests and culture. Treatment is carried out using the correct medication depending on the type of prostatitis.

2.2.3 Prostate cancer

Prostate gland cells regularly give rise to tumors, most commonly in the mid-to-late stages of life. The peripheral zone contributes the most to normal prostate function and is the most prevalent location of neoplasms in the aged prostate, accounting for over 80% of all prostate tumors. [19].

Prostate cancer is the second most frequent cancer and the fifth leading cause of cancer death among men in 2020 [20], as can be seen in Figure 2.5. Every year, around 1-3 million new cases of prostate cancer are detected around the world, making it a serious public health concern [21].

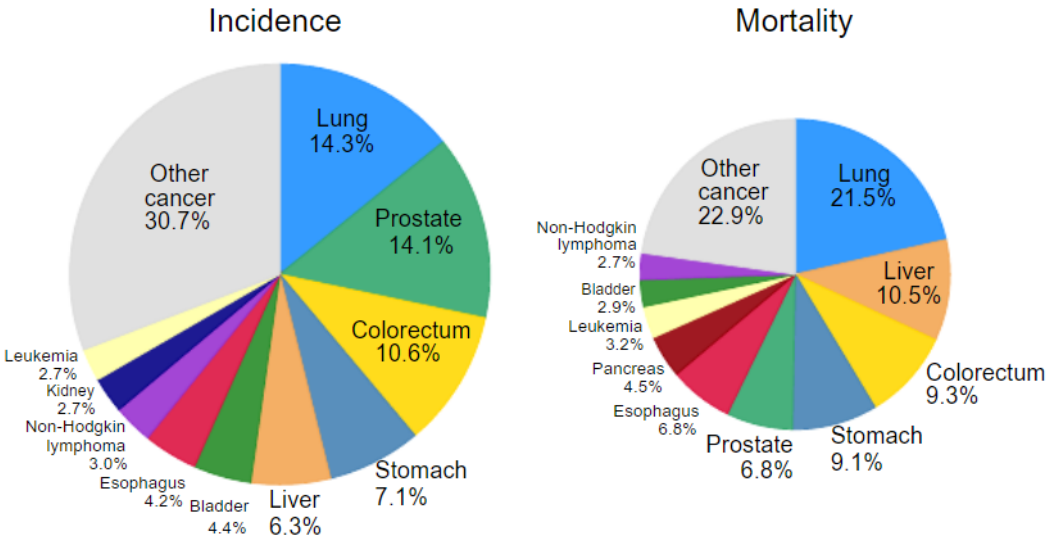


Figure 2.5: Distribution of Cases and Deaths for the Top 10 Most Common Cancers in 2020 for men, adapted from [20].

The most significant risk factor for the development of prostate cancer is the age. This type of cancer is uncommon in men under the age of 40, and its prevalence rises with age [22]. **The average age at diagnosis of prostate cancer is 66 years** [23]. Other factors, which are also important to mention, are a diet rich in polyunsaturated fatty acids, family predisposition and ethnicity, with the higher frequency of prostate cancer in African Americans likely due to environmental as well as genetic and socio-economic factors [24].

The prostate cancer can be discovered in a variety of methods in men [25]. The majority of men who have been diagnosed with prostate cancer early on do not have any symptoms. This tumor has lower genitourinary tract symptoms that are similar to those seen in other more prevalent diseases such benign prostatic hyperplasia, making a symptomatic diagnosis even more improbable. These symptoms may be: urinary urgency, pelvic discomfort or pain, erectile dysfunction, painful ejaculation and incontinence. Disseminated tumours may cause bone pain, spinal cord compression or paresthesia in the lower limbs [26].

To categorise prostate cancer is used the 2002 TNM Classification for Adenocarcinoma of the Prostate based on T (primary tumour), N (lymph nodes) and M (metastases) [27]:

- T categories: clinical examination, imaging, endoscopy, biopsy, and biochemical testing are used to determine;
- N categories: clinical examination or imaging;
- M categories: clinical examination, imaging, skeletal examinations, and biochemical tests.

In Table 2.1 it is possible to analyse the stages of prostate cancer in more detail.

Table 2.1: Staging of prostate cancer, adapted from [27].

Staging of prostate cancer	
Tx	Primary tumour cannot be assessed
T0	No evidence of primary tumour
T1	Clinically inapparent tumour, neither palpable nor visible by imaging
Ta1	Tumour (non-palpable) as incidental histological finding at transurethral resection of prostate in 5% tissue resected
T1b	Tumour (non-palpable) as incidental histological finding at transurethral resection of prostate in > 5% of tissue resected
T1c	Tumour (non-palpable) identified by needle biopsy (for elevated serum PSA): includes bilateral nonpalpable tumour on needle biopsy
T2	Tumour confined within prostate (including prostatic apex, prostate capsule) that is either palpable or visible on imaging or (with p-prefix) demonstrated in radical prostatectomy specimen
T2a	Tumour involving one-half of one lobe or less
T2b	Tumour involving more than one-half of one lobe but not both lobes
T2c	Tumour involving both lobes
T3	Tumour extends through prostatic capsule
T3a	Extra-capsular extension (ECE)
T3b	Invasion of seminal vesicle(s)
T4	Tumour fixed or invades adjacent structures: bladder neck, external sphincter, rectum, levator muscles and pelvic wall
Nx	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastases
N1	Regional lymph node metastases within true pelvis, below common iliac artery bifurcation, either unilateral or bilateral
Mx	Distant metastases cannot be assessed
M0	No distant metastases
M1a	Non-regional lymph node metastasis
M1b	Metastasis to bone(s)
M1c	Other site(s) of metastasis

A person's prognosis for prostate cancer is very unpredictable and dependent on the tumor grade and stage at the time of original diagnosis. If identified early enough, men with localized prostate cancer can live for up to ten years and have a 99% chance of surviving. This extended life can be due in great part to advancements in medical diagnosis. Men diagnosed with late-stage disease (distant metastases) have a poor overall survival of only 30% at 5 years.

Radiation therapy is essential in the treatment of cancer, like the prostate cancer. The calculation of the radiation dose depends on the electronic densities of the tissues, thus, the definition of the anatomical characteristics in medical exams is fundamental. For this, Computed Tomography (CT) is used, since it is the only imaging modality that provides these densities. However, due to the particular advantages of other modalities such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) scans can be, and are, also used in clinical practice to help

and assist with contouring. Nevertheless, CT is indispensable to radiation oncology [28]. The most often used imaging modality for determining the degree of local dissemination of prostate cancer is computed tomography of the pelvis [22].

2.3 Image modalities

Since the discovery of X-rays in 1895 by Wilhelm Conrad Röntgen, medical imaging has contributed significantly to the progress of medicine [29]. Imaging has become an important tool, in establishing diagnoses and monitoring pathologies.

Medical image acquisition consists in obtaining images of the human body through capture devices which vary according to the type of images of interest. Initially, after the discovery of the X-ray, the imaging representation of the human body limited the field of action to three areas:

- skeleton: by the contrast of the bones with the soft tissues;
- thorax: by the contrast provided by the air in the lungs;
- radiopaque foreign bodies [30].

Today the ”new” diagnostic modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) provide 3D images, revealing both anatomical and functional features of the human body [31].

2.3.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a diagnostic technique non-invasive for acquiring information about the inside of the body. The biggest advantage of MRI is the creation of 3D images of the object under study without the use of any ionising radiation, since it uses a magnetic field to produce images of the inside of the body. This exam differs from CT as it does not use X-rays.

The patient is put in a powerful uniform magnetic field to get a magnetic resonance image. Within the patient, the magnetic field aligns hydrogen nuclei in the direction of the field. An external radiofrequency (RF) pulse "disturbs" the nuclei from their current alignment. The hydrogen nuclei re-align inside the externally imposed magnetic field once the RF pulse is discontinued, emitting RF signals as they lose energy [32].

As it is possible to see in Figure 2.6 hydrogen atoms spin in their normal condition with their axes of rotation randomly aligned. They align in a consistent direction when put in a magnetic field. The H-atoms are knocked out of their magnetic field alignment by an RF pulse. The atoms revert to their previous orientation once the RF pulse is discontinued, emitting a signal that is then utilised to construct the image.

The intensity of the field determines the frequency of the RF signal generated by the hydrogen nucleus as it returns to its original orientation inside the field. As a result, the position of each hydrogen nucleus’s RF signal may be estimated. The computer examines each RF wave for strength

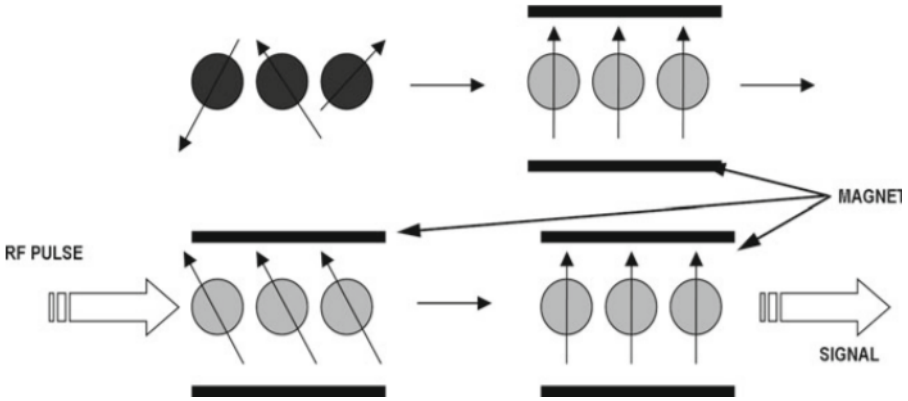


Figure 2.6: Magnetic Resonance Imaging-Hydrogen nuclei, from [32].

and other factors and then the computer assigns gray-scale values (white to black) to the signals on the detector. MRI images can display different sorts of diseases since the method of forming an image based on tissue features is fundamentally different from the absorption of X-rays by different tissues. Soft tissue abnormalities such as herniated discs, ligament tears, and soft tissue cancers in the spine are typically diagnosed with MRI, which can discern soft tissue changes better than CT images [32].

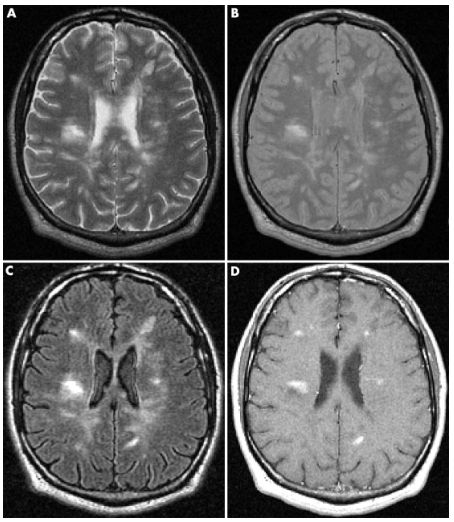


Figure 2.7: Axial magnetic resonance imaging of a 30 year old man with relapsing remitting multiple sclerosis (MS) showing multiple periventricular lesions: (A) T2 weighted image; (B) proton density weighted image; (C) fluid attenuated inversion recovery image; (D) T1 weighted image following administration of gadolinium demonstrating enhancing lesions, from [33]

The images may be more sensitive to particular tissue qualities depending on how and how long we excite the atoms. T2-weighted scans, for example, show fluid, demyelination, and regions of enlargement in the brain tissue with greater clarity - high signal. The white matter is lighter than the grey matter on T1-weighted imaging, and locations with high protein content and adipose tissue in general have stronger signal - clearer images [34]. Figure 2.6 shows axial MRI of the

brain from a 30 years old man.

2.3.2 Positron Emission Tomography

Positron Emission Tomography (PET) is a type of nuclear medicine imaging, which measures the emissions of radioactive substances, which are metabolically active, that have been previously injected into the bloodstream. Nuclear medicine imaging uses small amounts of radioactive material to diagnose, evaluate or treat a variety of diseases. PET has the ability to allow earlier diagnosis, also giving information related to prognosis and therapy. It is therefore acquiring a primary role in the diagnosis and evaluation of many disease states, with main reference to cancer. PET uses 'positron' emitting radiotracers to provide images of the human body [31].

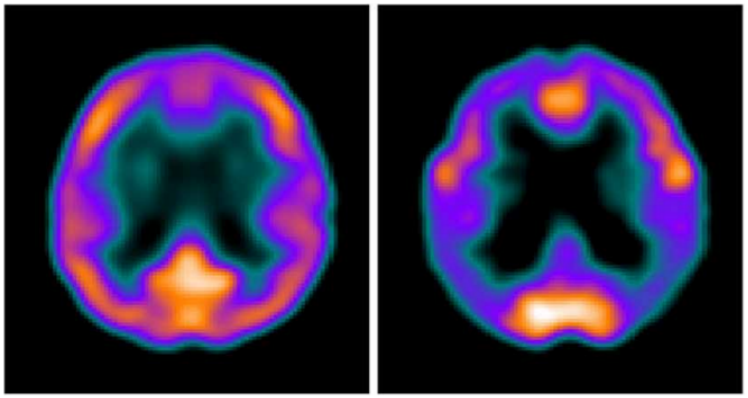


Figure 2.8: These PET images show normal brain activity (left) and reduced brain activity as a result of Alzheimer's disease (right). The decrease in the intense red and yellow areas on the scan on the right indicates the presence of slight Alzheimer's disease. The increase in the blue colour tone shows a decrease in brain activity., from [31]

This technique is used in a variety of diseases, including many types of cancers, heart disease, gastrointestinal, endocrine or neurological diseases and other abnormalities. As nuclear medicine scans can identify molecular activity, they have the potential to identify diseases in their earliest stages. They can also show whether a patient is responding to treatment. Nuclear medicine imaging procedures are non-invasive. With the exception of intravenous injections, they are usually painless.

Figure 2.8 illustrates an example of the results that can be obtained with this technique.

2.3.3 Computed Tomography

CT was the first method that made it possible to generate nonoverlapping axial slices of the internal structure of a human body without opening it [35]. Computed Tomography uses ionizing radiation to create a cross-sectional image [36], allowing a visualization of a greater variety of tissue structures apart from the four basic densities (air, bone, soft tissue, and fat). Contrary to conventional x-ray, which utilizes just one projection to make the image, this modality employs many projections that cross the body to generate the picture, combining them to create a picture

that enables for more detailed viewing of soft tissues [36]. Each individual 2D image from a CT scan is called a slice, since these images can be interpreted as if the patient has been completely sliced throughout a plane.

The main components of a conventional diagnostic CT scanner are an X-ray source and a detector, positioned on opposing ends of the patient and set on a rotatable gantry that can spin this imaging chain at very high speeds. The recorded X-ray intensities received by the detector at different places around the patient in a circular orbit are mathematically reconstructed to provide a cross-sectional picture. The X-ray attenuation coefficient (μ) of various tissues is varied [37]. Beer–Lambert law can be used to calculate the attenuation coefficients:

$$I = I_0 e^{-\mu x}, \tag{2.1}$$

where I is the intensity of the detected X-ray, I_0 is the intensity of the incident X-ray, and x is the thickness of the tissue that the X-ray beam traverses

The attenuation coefficients for the materials may be normalized with regard to the water coefficient (μ_w), and the result is represented in Hounsfield Units (HU) [37]:

$$\mathrm{HU} = 1000 \times \frac{\mu - \mu_\mathrm{w}}{\mu_\mathrm{w}} \tag{2.2}$$

The HU values provides information about the nature of the tissues. In Table 2.3 are schematized some of the HU values assigned for the different tissues.

Table 2.3: Typical HU values for different materials, adapted from [36].

Substance	Hounsfield units
Air	−1000
Fat	−100
Water	0
Muscle/Soft tissue	+40
Contrast	+130
Bone	+1000

2.4 Summary

In this chapter it was possible to cover the anatomy of the male pelvis and also the structure of the prostate, in order to locate the prostate in an image. It was also possible to survey the three diseases that most affect the prostate. One of these diseases was prostate cancer, which is the second most prevalent cancer in men.

The treatment of prostate cancer often involves radiotherapy. Medical imaging plays a fundamental role in the diagnosis and planning of radiotherapy. The most commonly used medical modality is CT, as it allows the planning of the radiation dosage for treatment. However, other modalities are used in order to complement the diagnosis and treatment.

Chapter 3

3D Reconstruction of the Prostate in CT Images - State of the Art

This chapter focuses on the current approaches for the 3D reconstruction of the prostate in CT images. Although the focus of the thesis is not on three-dimensional reconstruction of the prostate, it is important to understand how segmentation is performed at this point. This organ is difficult to segment, as will be possible to see in the work previously developed in this section. This occurs due to the intrinsic nature of this organ, and the organs that surround it.

This chapter will also contemplate some important fundamentals for the realization of the work.

3.1 Methodology

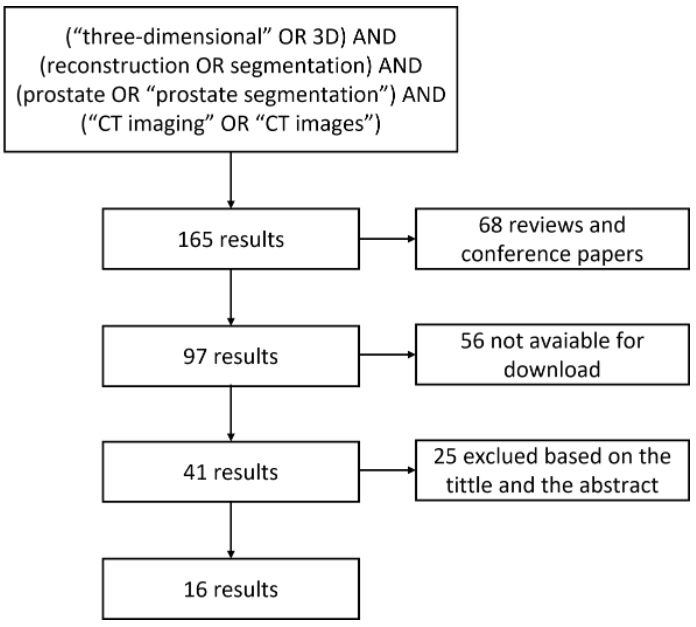


Figure 3.1: PRISMA diagram showing the results of the performed literature search.

A systematic literature search was performed, Figure 3.1, in Scopus using the following keywords combination: (“three-dimensional” OR 3D) AND (reconstruction OR segmentation) AND (prostate OR “prostate segmentation”) AND (“CT imaging” OR “CT images”). This research as allowed a total of 165 results. From those, 16 articles where useful for the topic. The remaining 148 results were eliminated for different reasons: either because it was not an article, or because the title or abstract did not fit with the search, or because they were not available for download. In spite of the keywords "CT image" or "CT imaging", some of the results are from other medical modalities since the articles would be useful for the literature review.

3.2 Prostate Segmentation - A Review

Table 3.1 is a resume of the works that where found addressing prostate segmentation.

Limited datasets can compromise the work results, a solution for this issue is Data Augmentation. Data Augmentation is a term that refers to a range of strategies for increasing the amount and quality of training datasets so that stronger Deep Learning models may be generated using them [38]. Data Augmentation was used in some of the articles reviewed due to the lack of data in some of the developed works [39, 40, 41, 42, 43].

In addition to performing augmentation, in [40], they proposed and developed a semi-supervised adversarial learning framework supported by data augmentation using GAN.

Many of the authors have used machine or deep learning techniques, to produce results. In [39], the authors proposed an extension of 3D U-Net on unsupervised domain adaptation, applying adversarial networks to the segmentation of the prostate, and other near organs. A 3D U-Net based segmentation was taken by [44]. Based on deep learning convolutional neural network [41] developed their work. The authors [42] establish an algorithm based on a customized U-Net architecture. Founded on 3D convolutional network the [43] segmented the prostate.

From a deep learning approach, in [45], was construct a voxel-wise segmentation from training data with 3D bounding box annotations.

The authors [46] frame the editing problem as a semisupervised learning problem that may take advantage of a priori training data knowledge as well as relevant information from user interactions. In particular, a region of interest around the specified user interactions.

The developers of [47] use an online-learning and location-adaptive classification method.

A learning-based segmentation scheme that combines the population learning and patient-specific learning together for segmenting the prostate on 3D CT images, was made in [48].

In [49] it was adopted an U-Net architecture. Their method can be divide in three parts:

1. Organ localization used to focus on the candidate segmentation region for each organ for better performance;

2. Boundary sensitive representation aimed at representing the semantic boundary information with robust and accurate soft labels;

3. Multi-label segmentation introducing a multi-label cross-entropy loss function to reinforce the network with semantic boundary information for segmentation [49].

In [50] the authors opted for an approach via a novel boundary detection method. In this approach they firstly, on the organ boundary, automatically detected several key landmarks to initialise the shape of the organ. Based on this initialization it was possible to know the organ boundaries. Then a local boundary regression by the near-boundary image patches is made. Thus,a boundary voting map can be obtained to enhance the entire organ boundary. In order to improve the method they combine them regression forest with the auto-context model to improve the quality of the boundary voting map. Finally, for the final organ segmentation, a deformable model learned during the training step is applied to the resulting border voting map.

The authors in [51] developed an approach where in a training dataset, using the prostate gland centroid point as the origin of a coordination system, they defined an inter-subject correspondence between the prostate surface points based on the spherical coordinates. Then they applied this correspondence to generate a point distribution model for prostate shape using principal component analysis and to study the local texture difference between prostate and non-prostate tissue close to the different prostate surface sub-regions. It was used the learned shape and texture characteristics of the prostate in CT images and then combined them with user inputs to segment a new image.

The authors of the article, [52], presented a novel framework, at the time, for building a linked statistical shape model (LSSM), a statistical shape model (SSM) that links the shape variation of a structure of interest (SOI) across multiple imaging modalities. In [53] they estimate a displacement field for guiding deformable segmentation.

The Bayesian method, [54], combines a deformable prostate model, learned from examples, and a geometrical likelihood strategy that maps actual observation into the space of model organs and selects the best shape to superimpose it in the original CT image

The overlap index, also known as the Dice Similarity Coefficient (DSC), is the most often used statistic for evaluating medical volume segmentations [44]. DSC can be defined as:

$$\mathrm{DSC} = \frac{2 \mid True \cap Pred \mid}{\mid True \mid + \mid Pred \mid} \tag{3.1}$$

where the number of foreground voxels in the ground truth is "True" and "Pred" is the predicted segmentations

All authors obtained good DSC values, however the best result was obtained by the authors [49] (DSC = 94%).

Table 3.1: Works found addressing prostate segmentation

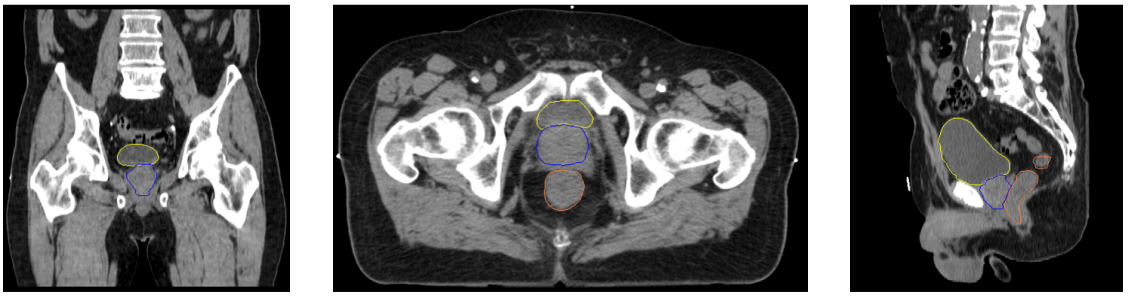
Authors	Year	Imaging Modality	Data Dimension	Data Source	Data Size	Developed Approach	DSC(%)
[43]	2020	MRI	3D	PROMISE12	50	SIP-Net: 3D CNN with densely connected residual blocks (DRBs) and attention-focused modules (AMs)	91.42
[40]	2021	CT	3D	Private	120	Semi-supervised adversarial learning framework supported by data augmentation using GAN	90.00
[39]	2021	CT	3D	Private	134	3D U-Net on unsupervised domain adaptation, applying adversarial networks to the segmentation	71.90
[41]	2020	MRI	3D	Private	326	Deep learning convolutional neural network	-
[45]	2020	CT	3D	Private	313	Voxel-wise segmentation from training data with 3D bounding box annotations	80.31
[49]	2019	CT	3D	Private	313	Convolutional networks with boundary sensitive representation	94.00
[44]	2018	CT	3D	Private	136	3D U-Net based segmentation	90.00
[52]	2012	MRI-CT	2D	Private	20	Builded a linked statistical shape model (LSSM), a statistical shape model (SSM) that links the shape variation of a structure of interest (SOI) across multiple imaging modalities	82.83
[47]	2012	CT	3D	Private	11	Online-learning and location-adaptive classification method	90.80

[54]	2014	CT	3D	Private	116	Bayesian method combines a deformable prostate model, learned from examples, and a geometrical likelihood strategy that maps actual observation into the space of model organs and selects the best shape to superimpose it in the original CT image	91.00
[50]	2015	CT	3D	Private	70	Boundary regression method	84.00
[48]	2017	CT	3D	Private	92	Learning-based segmentation scheme that combines the population learning and patient-specific learning together	87.18
[53]	2016	CT	3D	Private	313	Estimate a displacement field for guiding deformable segmentation	87.00
[46]	2014	CT	3D	Private	73	Interactive segmentation method	86.50
[51]	2018	CT	3D	Private	70	Semiautomatic, learning-based technique for full 3D segmentation of the prostate	88.00
[42]	2019	CT	3D	Private	92	Algorithm based on a customized U-Net architecture	83.00

3.2.1 Previous Work

Image Segmentation is the process that decomposes an image into regions with similar properties in particular gray level, color, texture, brightness and contrast [55]. The function of segmentation is to subdivide the objects in an image.

The segmentation process can be complex, as is the case with the segmentation of male pelvic regions. The anatomy of this region is composed of organs with similar density (soft tissue), such as the prostate, bladder and rectum [28]. In addition, the boundaries between these organs are not easy to distinguish, as it is possible to see on Figure 2.1, especially on CT images (Figure 3.2), because they have small contrast variations between them. It's also important to note that, due to variations in the contents of the bladder and rectum, this is one of the areas with the most variance not just across individuals, but also in the same individual. The following images of the Figure 3.2 its possible to see three slices of CT images of the same patient. In these images it is possible to identify the bladder in yellow, the prostate in blue and the rectum in orange.



(a) Coronal section through a male pelvis.

(b) Transverse section through a male pelvis.

(c) Sagittal section through a male pelvis.

Figure 3.2: CT slices from a male pelvic cavity. In these slices the bladder is segmented in yellow, the prostate in blue and the rectum in orange.

In past works it was possible to see that the simplest methods, using little processing, were not enough to segment the prostate correctly, showing how difficult it is to segment this gland.

In the previous work was used a CT image of the pelvis of a male, which was previously put into .vtk format. The size (number of pixels) of the 3D image is 512 x 512 x 165, with pixel spacing of 0.98 x 0.98 x 2.5. The methodology used in this work is represented in the pipeline of Figure 3.3.

Once the image was read, through ITK, the histogram equalization was performed. This was the method used to improve the contrast quality of the image by changing the range of pixel intensity values. After that the region of interest (ROI) was removed from the image. This step was important to reduce the size of the image. In addition to making the image processing faster, it will minimize possible segmentation errors, because the size of the image will be smaller. Segmentation was then performed by using an ITK region-growth based filter. Many filters were tested but the one that obtained the best result was itkNeighborhoodConnectedImageFilter, even so, the results were not as good as expected. Several parameters of this filter have been tested, such as the

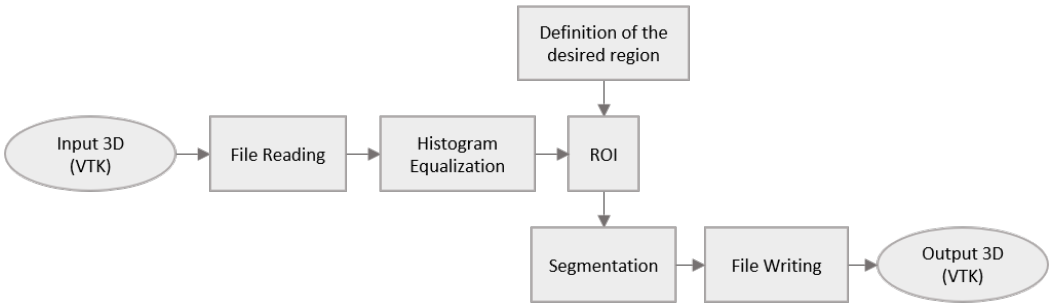


Figure 3.3: Pipeline of the taken approach.

indices of the pixels that are spatially connected, and the results show the best possible outcomes. Finally, the three-dimensional model obtained is written in .vtk format.

Figure 3.4 shows the histograms obtained after histogram equalization was performed. The first histogram shows a large peak in the -1000 HU region, which corresponds to air. Peaks are also shown around the 0 HU value, which corresponds to fat and soft tissue. The second histogram demonstrates the pixel density distribution when the image intensity values are equalised to values between -100 and +210. These values proved to be good enough to create more contrast in the image.

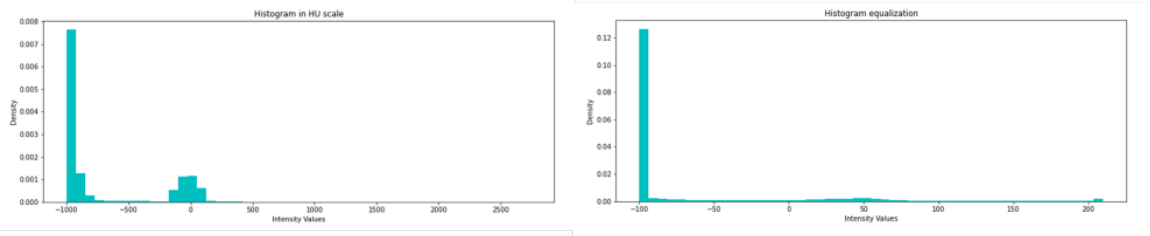


Figure 3.4: Before and after histograms.

Figure 3.5 despicts demonstrate the difference in contrast of the initial image before and after histogram equalisation.

As previously written, the segmentation did not go as expected (Figures 3.6 and 3.7). The segmentation of the prostate is a very challenging process since this organ is surrounded by others with similar densities.

Changing the contrast helped to obtain a segmentation closer to what was expected, but still not enough to guarantee a good three-dimensional reconstruction of the prostate (as shown in the red model of Figure 3.6)

Figure 3.7 shows clearly how the result obtained is far from being a perfect model of the prostate. It is possible to see in both views that there are actually points that have been well segmented as "prostate", but also unimportant pixels have been included.

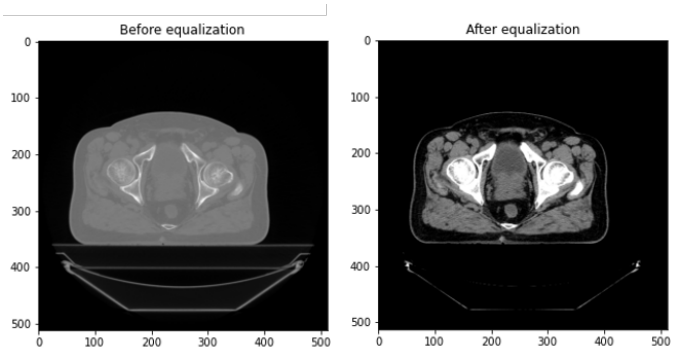
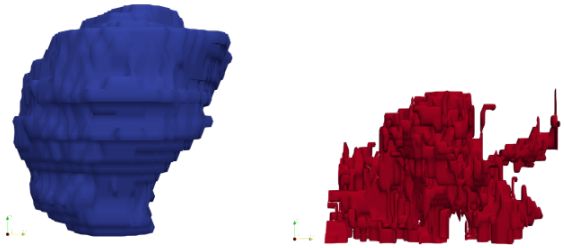


Figure 3.5: Before and after equalization of the histogram in slice 81.



(a) (b)

Figure 3.6: (a) Ground Truth obtained from the original image for the prostate. (b) Segmentation result of the applied algorithms. These models were visualized using ParaView.

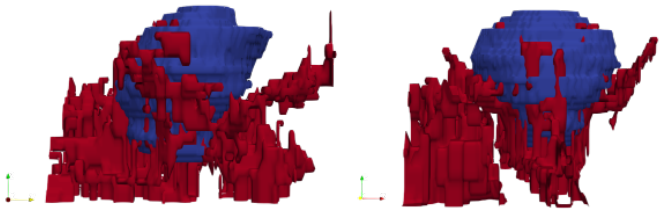


Figure 3.7: Overlap of the model obtained with the algorithms with the model developed manually by the medical technician. It's possible to see the overlap in two different views. These models were visualized using ParaView.

3.2.2 Datasets

From the systematic review only in [43], the authors used a public dataset (PROMISE12[1]). However this dataset is of MRI images, and no public dataset with CT images was found.

The dataset that will be used are images corresponding to radiotherapy planning CTs for prostate cancer of patients who underwent treatment between 2011 and 2019 in the IPO of Porto. These were acquired using two similar, wide diameter (80cm) machines - GE LightSpeed RT16 and GE Optima 580 - both 16-slice. The peak voltage was always 120 kV, with mAs modulation. Both are calibrated in the same way and the curves of Hounsfield units vs e-density are identical.

[1] MICCAI Prostate MR Image Segmentation 2012 Grant Challenge dataset

All the images have in common the presence of the segmentation of the prostate, however in some it is possible to observe, besides the prostate, the bladder, the rectum, the bulb and the femurs and in some it is even possible to observe the lymphatic glanglia. Other images have only the prostate with the bladder and rectum, or even only the prostate. The image dataset is composed by 2164 images. The size (number of pixels) of the 3D images is variable, with pixel spacing of 0.98 x 0.98 x 2.5. This dataset was divided in this work into two parts:

- **DatasetP**: this dataset consists of 1500 images. These are images that have all the prostate segmented and may or may not have yet the rectum, the bulb and the femurs;
- **DatasetP_GG**: composed of 324 images, this dataset consists of the images where not only the prostate has been segmented but also the lymphatic glanglia have been segmented.

3.3 Statistical Modeling Fundaments-Statistical Shape Models

Image analysing can be performed using a wide range of techniques to extract information about the original dataset. Statistical Shape Models (SSMs[2]) are a method of extracting information and are a tool to describe the shape variation of a dataset. SSMs allow both the qualitative and quantitative interpretation of image data by describing the change in shape of an object class over a group of images [56].

The Statistical Shape Models was developed by [57], it is a statistical model that consider the shape of objects/structures.

Based on the principle that each shape is a deformed version of a reference shape, Statistical Shape Models may be used to examine differences in a dataset as well as synthesis new related forms [56]. As a result, it represents the object's mean form as well as permissible variations in relation to the same mean shape [58]. In medical imaging research, statistical shape models are commonly used in a similar way to organ modeling [59].

SSM is predicated on the idea that every object is made up of a collection of points, and that a dataset of samples must be generated. To enable comparison, each sample from the dataset must contain the same number of landmark points and be dispersed in the same order. The average model and any potential variations of these models can be created by calculating the distance between the respective landmark locations in each sample of the training dataset, allowing for variability. A number of significant changes must be made before creating the model, including alignment, scaling, translation, and rotation, in order to acquire only the variations of the shape, keeping in mind that this set of statistical data defines the model [60].

To create a statistical shape model or a point distribution model for a set of n_s surfaces$\{\mathscr{M}_i\}$ correspondence needs to be established and the surfaces have to be aligned in a common reference coordinate system [61, 62]. Assuming that x_i are the parameterizations in the common reference coordinate system that express this relationship. Thus, the goal of SSMs is to obtain the shape present in the set of surfaces using a probability density function. A set of annotated images is

[2]SSMs are also known as Point Distribution Models in the literature.

required for the creation of an SSM. The training dataset is the dataset on which the model is developed.

Principal Component Analysis (PCA), an algorithm that discovers the directions with the largest variation in the training data, is the most frequent approach for this analysis [63, 64]. The mean shape ($\bar{F}$) may be defined as indicated in equation 3.2 using a collection of **aligned** forms (N) with point to point correspondence [64].

$$\bar{F} = \frac{1}{N}\sum_{i=1}^{N} F_i \tag{3.2}$$

Singular Value Decomposition is applied to the shapes $\bar{F} - F$ to conduct PCA on that set. The output of this procedure is a P matrix of eigenvectors and associated eigenvalues λ. A Primary Component (PC) or principal mode of variation is each eigenvector. PC can be as many as $min(N, 3n)$, where N is the number of samples in the training set and n refers the number of points.

Thus, the shape set can be defined as:

$$F = \bar{F} + \sum_{m=1}^{M} PC_m b_m \tag{3.3}$$

where M is the number of eigenvectors in the model, b_m are vectors of scalar weights, and $-3\lambda_m < b_m < 3\lambda_m$

The Statistical Shape Model can be ilustrative and predictable. Descriptive models allow for the discovery of unexpected patterns like, trends, clusters, and outliers by examining the features form. Predictive models use regression or classification approaches to investigate correlations between form and continuous or discrete (clinical or functional) characteristics [65]. In [65] it's possible to see an example of the creation of an SSM, Figure 3.8. This figure shows a process to obtain the model and take some features from the dataset.

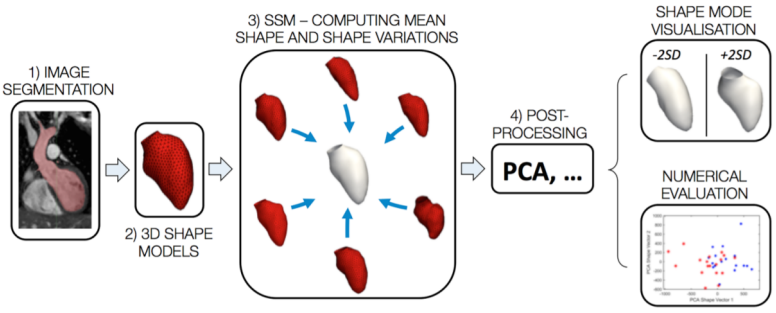


Figure 3.8: Resume of a statistical shape analysis framework (SSM). Segmentation of medical imaging data is the first step. The anatomical segment of interest, for example, the left ventricle, is reconstructed and these two steps are repeated for all the patients in the dataset, generating the inputs for computing the mean shape. Further down the line is made a post processing involves methods such as principle component analysis (PCA), allowing to compare variations in shape (eg, ±2 SD from the mean shape) and perform quantitative assessments, from [65].

It is unquestionable how these models change and continue to change the medical paradigms, making processes such as pelvis reconstruction after surgery more efficient and safer. In the Figure 3.9 it is possible to see the approach taken from [66] for the reconstruction of male pelvis. With SSM it was possible reconstruct the anatomically optimal bone shape of the pelvis after tumour resection.

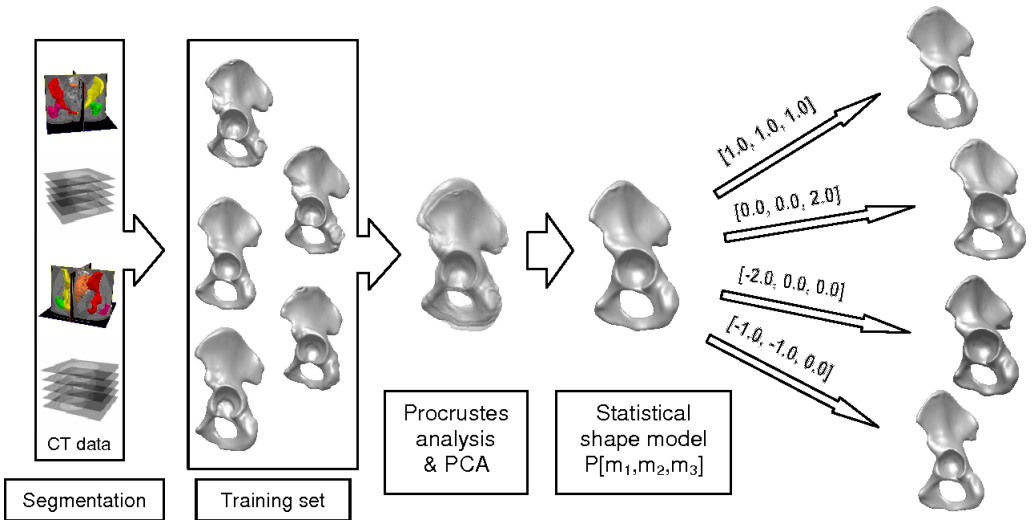


Figure 3.9: Different stages of the data processing pipeline for generation of the statistical pelvis shape model, from [66].

3.4 Summary

With the work carried out it was possible to prove how complex the segmentation of the prostatic gland is. In this sense, a literature review was conducted where it was found that most of the works developed in the area use more complex techniques, mostly of Machine and Deep Learning.

In the work to be developed, an approach involving Statistical Shape Models will be carried out. Thus, in this way, in this chapter the fundamentals were carried out.

This chapter will be important for the realisation of the thesis, as it will be based on the foundations set out here.

Chapter 4

Methodology, Results and Discussion

This chapter aims to cover the approached methodology to get the statistical shape models of the prostate from CT images using surface principal components. It will also address the results obtained for the proposed objectives.

4.1 Methodology and results for the Statistical Shape Model

As already explained in Chapter 3 the project was developed in two different datasets. The first one containing 1500 CT images from patients with prostate cancer who were underwent treatment (DatasetP), where was segmented besides the prostate, the rectum, the bulb and the femurs or, in some cases, just the prostate. The other dataset, with 324 CT images (DatasetP_GG) which was images from patients underwent treatment in the prostate and in the lymphatic glanglia, that where segmented.

All the files where in the DICOM format to be processed were all passed to the .npz format for handling in python.

4.1.1 DatasetP

To produce the statistical shape model we started by centring the images, through their centroids, and cutting the images in $100 * 100 * 100$. Once the images were all centred with the same size we proceeded to build the shape model.

Initially, the dataset was made an average. If the value of this average, in each point, was less than 0.5 (<0.5) then the returned value would become 0, while if it was equal to or greater than 0.5 (=>0.5) this point would become 1. Thus, in a binary way, it would be possible to perform a threshold for the visualization of the shape model of the prostate.

In this way the shape model of the used prostate images was obtained, in other words, the geometric model that describe the collection of similar objects in a very compact way was computed, as shown in Figure 4.1. The model was visualized using from *mpl_toolkits.mplot3d.art3d* the *Poly3DCollection* and from *matplotlib* the *pyplot*, libraries for python.

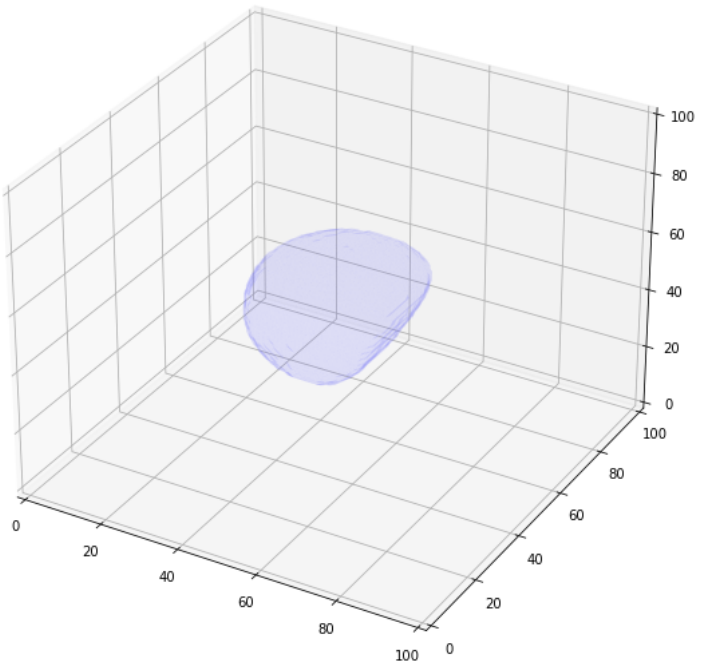


Figure 4.1: Shape Model of DatasetP.

Principal component analysis is applied to decompose the enormous variety in the natural shapes of prostate into a set of characteristics that explain the majority of the variation. To obtain the results of this part was notorious a slight difficulty, so it will be demonstrated part of the code performed for better understanding.

The following code demonstrates the creation of a function that will allow to calculate the distance map of the input image as a pixel-accurate approximation to the Euclidean distance, using the *SignedDanielssonDistanceMap* filter from ITK[1].

```
def makeDistanceMask(mask):
    mask = sitk.GetImageFromArray(np.moveaxis(mask, 0, -1))
    dist_mask = sitk.SignedDanielssonDistanceMap(mask)
    dist_mask = np.moveaxis(sitk.GetArrayFromImage(dist_mask),
    -1, 0)
    return dist_mask
```

The function takes in a binary image of a structure (e.g. prostate segmentation where 1 is prostate, 0 is background) as a numpy array. Then converts to ITK image. Then an ITK function *SignedDanielssonDistanceMap* is applied to create a distance map where each voxel contains the distance to the edge of the structure. Negative values inside the structure, positive values outside, zero on the edge/boundary. In the end converts back to a numpy array, whith same shape as input.

[1] Insight Toolkit (open-source, cross-platform package that offers programmers a wide range of methods for image analysis.)

To visualise the variance, the *PCA* library, from *sklearn.decomposition*, was used in the algorithm. This library allowed linear dimensionality reduction using Singular Value Decomposition of the data to project it to a lower dimensional space. The input data is centered but not scaled for each feature before applying the singular value decomposition. From these extracted characteristics it was taken the variance.

In a feature set, the number of principal components to keep varies on a range of factors, including storage space, training time, performance, etc. All the primary components are essential to the predictions in some datasets because all the features contribute equally to the overall variance. A basic rule is to disregard principle components with diminishing variance returns and focus on those principal components that contribute significantly to variance [67]. Plotting the variance against the principal components it's possible to see where the variance starts to be negligible. In this way it's possible to see, on the Figure 4.2, that at the 5 principal component, the variance is minimum. However, for the study up to 9 components were considered.

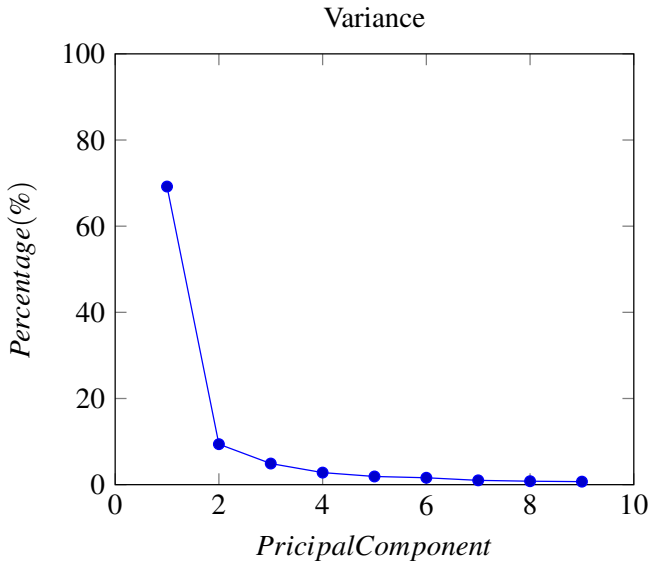


Figure 4.2: Variance in percentage (%) for each principal component.

In detail it's possible to see the variance of the first nine principal components on Table 4.1. The first 4 principal components can describe the majority of the variation at about 86%. In this way it's possible to affirm that with these four principal components we can describe 86% of the prostates of the dataset used. The impact of the other components is evident despite being harder to describe the variation.

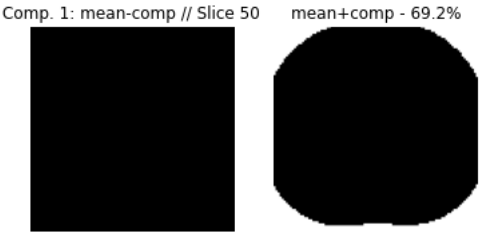
Subsequently, some mathematical manipulation was carried out in order to verify how the components influenced the shape model, later the visualisation of that influences was performed. Figure 4.3 demonstrates the subtraction and addition of components to the resulted shape model.

For viewing the influence on the mean shape of the first nine principle components it was chosen the slice 50, once is probably the middle slice. Another slice could have been chosen, however the central slice will have a larger area and will allow a better view of the influence of the

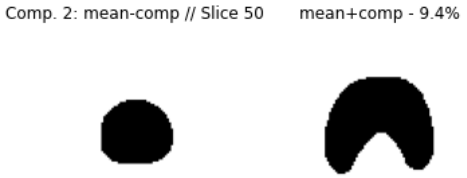
Table 4.1: Principal component and their variance on the shape model.

Component	**Variance**
1	69.2%
2	9.4%
3	4.9%
4	2.8%
5	1.9%
6	1.6%
7	1.0%
8	0.8%
9	0.7%

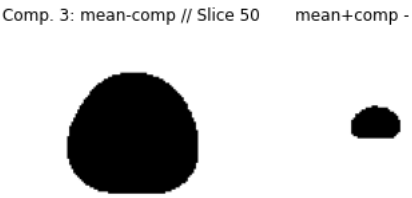
components. Figure 4.3 shows the first nine principle components influence on the mean shape. The influence of the components is striking with the images obtained. The components influence width, size, and, in other ways, that always influences the shape.



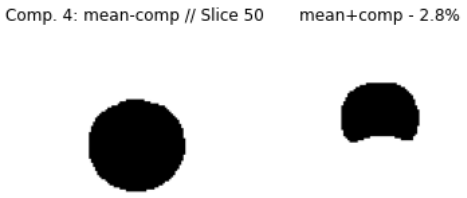
(a) Slice 50 of the model minus component 1 and plus the same.



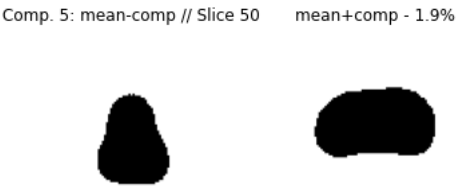
(b) Slice 50 of the model minus component 2 and plus the same.



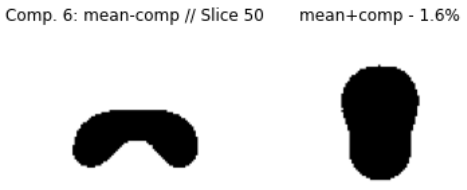
(c) Slice 50 of the model minus component 3 and plus the same.



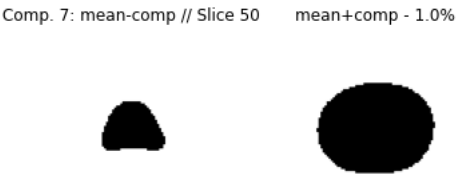
(d) Slice 50 of the model minus component 4 and plus the same.



(e) Slice 50 of the model minus component 5 and plus the same.



(f) Slice 50 of the model plus component 6 and plus the same.



(g) Slice 50 of the model minus component 7 and plus the same.

(h) Slice 50 of the model minus component 8 and plus the same.

(i) Slice 50 of the model minus component 9 and plus the same.

Figure 4.3: Influence on the mean shape of the first nine principle components for the DatasetP.

4.1.2 DatasetP_GG

Equivalent to what was done for DatasetP, the CT images of the DatasetP_GG were centered and cutted. With the same produced algorithms it was possible to obtain a new statistical shape model, Figure 4.4.

At first look this produced model is similar/equal to the one represented in Figure 4.1 however both will be analyzed in more detail later.

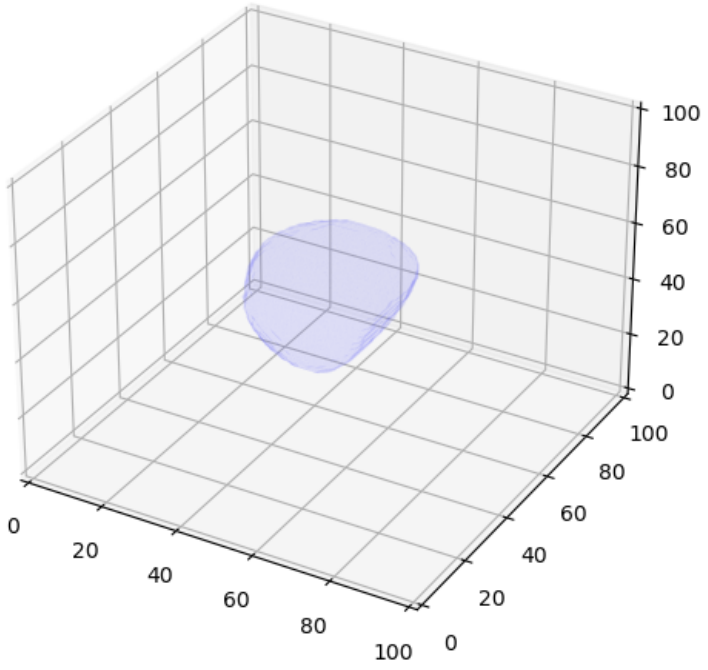


Figure 4.4: Shape Model of DatasetP_GG.

Both figures resemble a walnut, which is to be expected as this is the normal shape of the prostate. Although the patients in both datasets have prostate cancer, these images confirm that prostate cancer does not affect the shape of the prostate, so further medical tests such as MRI and PET are essential for the confirmation of prostate cancer. Nevertheless, it should be noted that CT images are essential in the treatment and planning of radiotherapy, and therefore this type of image should not be discarded.

For the visualization of the influence of the principal components, as it was done in DatasetP a chart was created (Figure 4.5). Figure 4.5 allows stating that only with the first. component the variation of DatasetP_GG could be explained, being then the other eight components dispensable for the explanation of the same

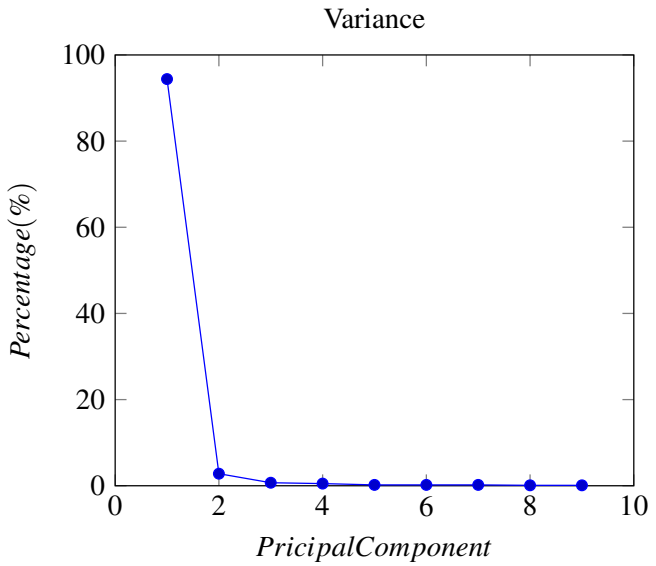


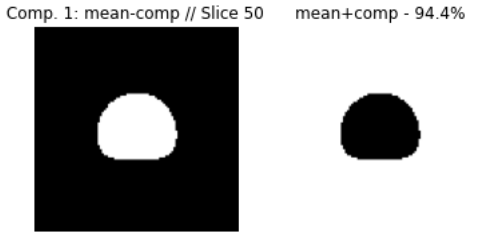
Figure 4.5: Variance in percentage (%) for each principal component.

Table 4.2 allows us to validate the above statement. By analysing the table we verify that the first principal component corresponds to 94.4% of the variance. This high number can be explained by the size of the dataset. The datasets under analysis do not have the same order of magnitude, so it was observed that the first principal component for DatasetP is 69.2% and for Datasetp_GG is 94.4%. Thus, based on the results obtained, in this case, it is possible to state that the smaller the dataset, the lower the amount of principal components needed to explain the dataset variance.

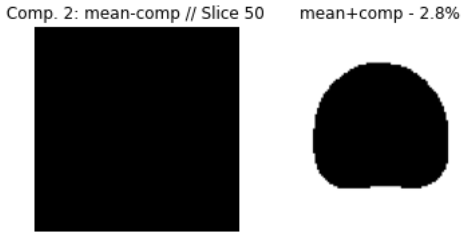
Table 4.2: Principal component and their variance on the shape model.

Component	Variance
1	94.4%
2	2.8%
3	0.7%
4	0.5%
5	0.2%
6	0.2%
7	0.2%
8	0.1%
9	0.1%

Figure 4.6 shows the influence of the principal components on the statistical shape model of DatasetP_GG. This influence, on the shape of the prostate, is quite noticeable to the point where the first image (Figure 4.6a) the variance makes the shape of a walnut appear (prostate shape).The remaining images in the Figure show the influence of the other principal components on the average statistical model.

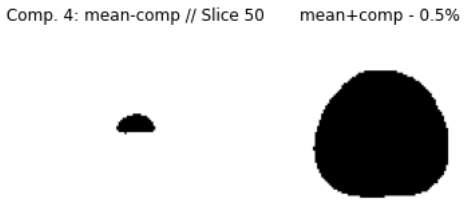


(a) Slice 50 of the model minus component 1 and plus the same.

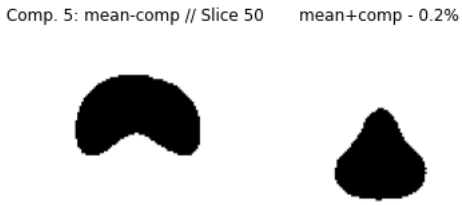


(b) Slice 50 of the model minus component 2 and plus the same.

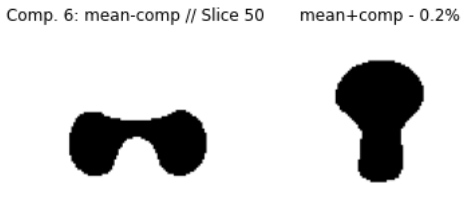
(c) Slice 50 of the model minus component 3 and plus the same.



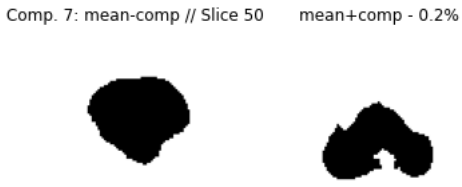
(d) Slice 50 of the model minus component 4 and plus the same.



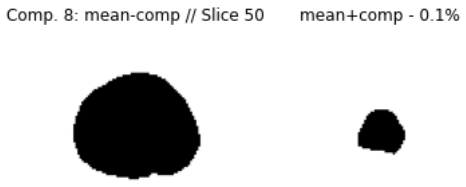
(e) Slice 50 of the model minus component 5 and plus the same.



(f) Slice 50 of the model plus component 6 and plus the same.



(g) Slice 50 of the model minus component 7 and plus the same.



(h) Slice 50 of the model minus component 8 and plus the same.

(i) Slice 50 of the model minus component 9 and plus the same.

Figure 4.6: Influence on the mean shape of the first nine principle (DatasetP_GG).

4.1.3 Segmentation Analysis

In order to analyse the produced segmentations, the volumes of the two statiscal shape models produced were calculated.

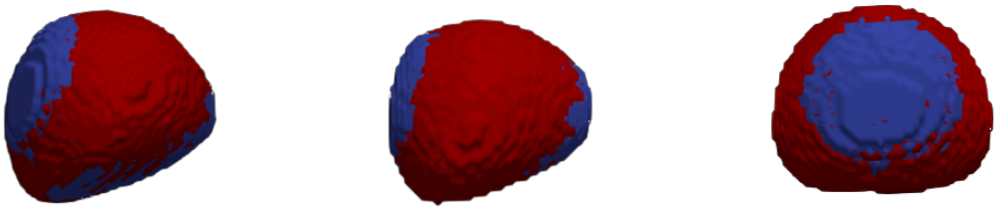
- Volume DatasetP: 34008 mm^3
- Volume DatasetP_GG: 31879 mm^3

Therefore, it can be stated that the difference in volumes is 2129 mm^3, which is equivalent to a minimal difference.

It should also be noted, as described in Chapter 2, that the average age of prostate cancer diagnosis is 66 years, which is assumed to be the target age for the datasets under analysis. By analysing the graph in Figure 2.2, the prostate at this age has an average volume of 30 ml (30000 mm^3). Thus, it is possible to state that by the volume of the prostates, it could be concluded that they were from patients without pathology, since the calculated volumes are close to the average value. This proves the need to use more images from other medical modalities for a correct diagnosis.

Although the volumes are very similar, the shape could be different. So, with the aid of Paraview[2], the images were overlapped to understand in what way there was a change in their shape (Figure 4.7). In red it's possible to see the statistical shape model from DatasetP, and in blue from DatasetP_GG. As can be seen, there are parts where one medium prostate is more evident and in others the other. However it is possible to state that they are not significant parts which alter its format.

Despite all the above, segmentation varies for the most diverse reasons, even if they are minimal variations, so it is important to create computational tools to assist physicians.



(a) Lateral view. (b) Coronal plane. (c) Axial plane.

Figure 4.7: Statiscal shape models overlaps. In red, statistical shape model from DatasetP, and in blue from DatasetP_GG.

[2]https://www.paraview.org/

4.2 Summary

This chapter covers the methodologies used to obtain the results of this dissertation. It is possible to refer that the Statistical Shape Model was successfully performed, having been possible to obtain the shape model and the values of the features of the PCA for both datasets. The difference in order of magnitude influences the variance, so the larger the dataset, the greater the variance between images, so more components will be needed to explain their changes.

The segmentation analysis allowed us to verify that there are no changes in the segmentations, despite the fact that they are made for different treatments.

Still to be remembered from this chapter is the importance of using different medical modalities when making a diagnosis, since, as observed in this case, all the data indicated that we were dealing with patients without pathologies.

Chapter 5

Conclusion

In this final chapter the main ideas of this report are shown, providing an analysis of the work developed and the conclusions taken from the project. This chapter will include discussion of the results.

5.1 Discussion and Final Considerations

The prostate is an interesting organ, from a biomedical engineering point of view, as few works exist in the area that aid medical diagnosis and it is the spot of the fifth leading cause of death from cancer among men. The prostate is a gland that is surrounded by other organs with similar densities (bladder and rectum), thus the segmentation of the prostate becomes a difficult task.

The most commonly used medical modality to support prostate imaging is Computed Tomography. However, MRI and PET are often used in order to complement the results obtained with CT.

The importance of segmentation for the planning of radiotherapy treatments is so important that in past work were developed aproaches to try the segmentation os the prostate. In this report was presented the past work, that shows to be a simplistic way to try to reconstruct the prostate and as such was unsuccessful. The conducted literature review allows us to state that most of the work developed in the area of 3D reconstruction of the prostate is in the Machine and Deep Learning area.

The segmentation made by each observer differ greatly. Correct segmentation is crucial to the radiotherapy treatment process since the more accurate the segmentations are, the less healthy tissue is exposed to radiation.

The main goal of this document was to outcome statistical shape models of the prostate from CT images using surface principal components to investigate whether there is any difference in the accuracy of how the segmentations are performed when a patient receives treatment in the prostate, for those who, in addition to the prostate, also receive treatment in the lymphatic glanglia.

The goals were achived, and turnned possible to conclude that the accuracy of how the segmentations are performed when a patient receives treatment in the prostate, for those who besides receiving treatment in the prostate also receive in the Lymphatic ganglia is extremely precise.

In addition to what has already been mentioned throughout the report, this research proved the importance of using various medical modalities for medical diagnosis.

5.2 Future Work

As future works we can refer the search for new algorithms that can make the 3D segmentation of the prostatic gland. It would be interesting, once having the 3D reconstruction, to calculate the DSC values, in order to be able to compare the results obtained with the segmentations performed manually, in order to see if the segmentations performed by an algorithm allow results close to reality.

It would also be interesting to carry out Data Augmentation. This is the process of expanding the amount of data utilized to train a model. In this way it would be possible to extend the dataset by e.g. flipping, rotation, translation the initial dataset.

Data Augmentation is important for reliable predictions. Deep learning generally requires the use of a lot of training data, which is not always available, and with this technique can be augmented to obtain better model results.

References

[1] Susan Standring. *Gray's anatomy : the anatomical basis of clinical practice*. Elsevier, New York, 2021.

[2] Michael Schuenke, Erik Schulte, and Udo Schumacher. *THIEME Atlas of Anatomy - Internal Organs*, volume 2. Thieme Medical Publishers, Inc, 2nd edition, 2016.

[3] Michael Mitterberger, Wolfgang Horninger, Friedrich Aigner, Germar M. Pinggera, Ilona Steppan, Peter Rehder, and Ferdinand Frauscher. Ultrasound of the prostate. *Cancer Imaging*, 10(1):30–48, 2010.

[4] Richard Berges and Matthias Oelke. Age-stratified normal values for prostate volume, PSA, maximum urinary flow rate, IPSS, and other LUTS/BPH indicators in the german male community-dwelling population aged 50 years or older. *World Journal of Urology*, 29(2):171–178, January 2011. URL: https://doi.org/10.1007/s00345-010-0638-z, doi:10.1007/s00345-010-0638-z.

[5] M. Mendes Silva and Rui C. Sousa. Hiperplasia benigna da próstata. *Acta Médica Portuguesa*, 12:103–11, 1999.

[6] Omesh Singh and Srinivasa Rao Bolla. *Anatomy, Abdomen and Pelvis, Prostate*. StatPearls Publishing, Treasure Island (FL), 2020.

[7] Bruno Mendes, Inês Domingues, Augusto Silva, and João Santos. Prostate cancer aggressiveness prediction using CT images. *Life*, 11(11):1164, October 2021. URL: https://doi.org/10.3390/life11111164, doi:10.3390/life11111164.

[8] Anil Bhavsar and Sadhna Verma. Anatomic imaging of the prostate. *BioMed Research International*, 2014:1–9, 2014. URL: https://doi.org/10.1155/2014/728539, doi:10.1155/2014/728539.

[9] Muhammad Amin, A Khalid, Nadia Tazeen, and Muhammad Yasoob. Zonal anatomy of prostate. 16:138, 1 1. URL: https://annalskemu.org/journal/index.php/annals/article/view/212, doi:10.21649/akemu.v16i3.212.

[10] LaTayia Aaron, Omar E Franco, and Simon W Hayward. Review of prostate anatomy and embryology and the etiology of benign prostatic hyperplasia. *Urol Clin North Am.*, 43:279–288, 2016.

[11] Shu-Jie Xia, Di Cui, and Qi Jiang. An overview of prostate diseases and their characteristics specific to asian men. *Asian J. Androl.*, 14(3):458–464, May 2012.

[12] Soum D Lokeshwar, Benjamin T Harper, Eric Webb, Andre Jordan, Thomas A Dykes, Durwood E Neal, Jr, Martha K Terris, and Zachary Klaassen. Epidemiology and treatment modalities for the management of benign prostatic hyperplasia. *Transl. Androl. Urol.*, 8(5):529–539, October 2019.

[13] Claus G. Roehrborn. Benign prostatic hyperplasia: an overview. *Reviews in urology*, 7 Suppl 9(Suppl 9):S3–S14, 2005.

[14] Neelima Dhingra and Deepak Bhagwat. Benign prostatic hyperplasia: An overview of existing treatment. *Indian J. Pharmacol.*, 43(1):6–12, February 2011.

[15] Farhad Fakhrudin Vasanwala, Michael Yuet Chen Wong, Henry Sun Sien Ho, and Keong Tatt Foo. Benign prostatic hyperplasia and male lower urinary symptoms: A guide for family physicians. *Asian J. Urol.*, 4(3):181–184, July 2017.

[16] Arkadiusz Miernik and Christian Gratzke. Current treatment for benign prostatic hyperplasia. *Dtsch. Arztebl. Int.*, 117(49):843–854, December 2020.

[17] Farhan Ullah Khan, Awais Ullah Ihsan, Hidayat Ullah Khan, Ruby Jana, Junaid Wazir, Puregmaa Khongorzul, Muhammad Waqar, and Xiaohui Zhou. Comprehensive overview of prostatitis. *Biomed. Pharmacother.*, 94:1064–1076, October 2017.

[18] J C Nickel. Classification and diagnosis of prostatitis: a gold standard? *Andrologia*, 35(3):160–167, June 2003.

[19] Richard J Rebello, Christoph Oing, Karen E Knudsen, Stacy Loeb, David C Johnson, Robert E Reiter, Silke Gillessen, Theodorus Van der Kwast, and Robert G Bristow. Prostate cancer. *Nat. Rev. Dis. Primers*, 7(1):9, February 2021.

[20] Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, February 2021.

[21] Shahneen Sandhu, Caroline M Moore, Edmund Chiong, Himisha Beltran, Robert G Bristow, and Scott G Williams. Prostate cancer. *Lancet*, 398(10305):1075–1090, September 2021.

[22] D Mazhar and J Waxman. Prostate cancer. *Postgrad. Med. J.*, 78(924):590–595, October 2002.

[23] Prashanth Rawla. Epidemiology of prostate cancer. *World Journal of Oncology*, 10(2):63–89, 2019. URL: https://doi.org/10.14740/wjon1191, doi:10.14740/wjon1191.

[24] David G Bostwick, Harry B Burke, Daniel Djakiew, Susan Euling, Shuk-Mei Ho, Joseph Landolph, Howard Morrison, Babasaheb Sonawane, Tiffany Shifflett, David J Waters, and Barry Timms. Human prostate cancer risk factors. *Cancer*, 101(10 Suppl):2371–2490, November 2004.

[25] Samuel W D Merriel, Garth Funston, and Willie Hamilton. Prostate cancer in primary care. *Adv. Ther.*, 35(9):1285–1294, September 2018.

[26] Catarina Pinto Oliveira. O cancro da próstata no idoso. Master's thesis, Faculdade de Medicina da Universidade de Coimbra, 2015.

[27] Nigel Borley and Mark R Feneley. Prostate cancer: diagnosis and staging. *Asian J. Androl.*, 11(1):74–80, January 2009.

[28] Gonçalo Almeida and João Manuel R.S. Tavares. Deep learning in radiation oncology treatment planning for prostate cancer: A systematic review. *Journal of Medical Systems*, 44(10), August 2020.

[29] Kunio Doi. Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology. *Physics in Medicine and Biology*, 51(13):R5–R27, jun 2006. doi:10.1088/0031-9155/51/13/r02.

[30] Paulo Alexandre da Costa Ferreira Brandão Pereira. *Erro em Imagiologia Médica*. PhD thesis, Universidade do Minho-Escola de Engenharia, 2011.

[31] Sean L Kitson, Vincenzo Cuccurullo, Andrea Ciarmiello, Diana Salvo, and Luigi Mansi. Clinical Applications of Positron Emission Tomography (PET) Imaging in Medicine: Oncology, Brain Diseases and Cardiology. *Current Radiopharmaceuticals*, 2:224–253, 2009.

[32] Harjit Singh and Janet A Neutze. *Radiology Fundamentals: Introduction to Imaging & Technology*. Springer, 4th edition, 2012.

[33] S A Trip and D H Miller. Imaging in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry*, 76 Suppl 3(suppl_3):iii11–iii18, September 2005.

[34] Edson Amaro Júnior and Helio Yamashita. Aspectos básicos de tomografia computadorizada e ressonância magnética. *Rev. Bras. Psiquiatr.*, 23(suppl 1):2–3, May 2001.

[35] Manoj Diwakar and Manoj Kumar. A review on CT image noise and its denoising. *Biomedical Signal Processing and Control*, 42:73–88, April 2018.

[36] Harjit Singh and Janet Neutze. *Radiology Fundamentals: Introduction to Imaging & Technology.* Springer, 4 edition, 2012.

[37] Daniel Thomas Ginat and Rajiv Gupta. Advances in computed tomography imaging technology. *Annual Review of Biomedical Engineering*, 16(1):431–453, July 2014.

[38] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *J. Big Data*, 6(1), December 2019.

[39] Eliott Brion, Jean Léger, A M Barragán-Montero, Nicolas Meert, John A Lee, and Benoit Macq. Domain adversarial networks and intensity-based data augmentation for male pelvic organ segmentation in cone beam CT. *Comput. Biol. Med.*, 131(104269):104269, April 2021.

[40] Zhuangzhuang Zhang, Tianyu Zhao, Hiram Gay, Weixiong Zhang, and Baozhou Sun. Semi-supervised semantic segmentation of prostate and organs-at-risk on 3d pelvic CT images. *Biomedical Physics & Engineering Express*, 7(6):065023, oct 2021. URL: https://doi.org/10.1088/2057-1976/ac26e8, doi:10.1088/2057-1976/ac26e8.

[41] Christian Jamtheim Gustafsson, Johan Swärd, Stefan Ingi Adalbjörnsson, Andreas Jakobsson, and Lars E Olsson. Development and evaluation of a deep learning based artificial intelligence for automatic identification of gold fiducial markers in an MRI-only prostate radiotherapy workflow. *Phys. Med. Biol.*, 65(22):225011, November 2020.

[42] Maysam Shahedi, Martin Halicek, James D Dormer, David M Schuster, and Baowei Fei. Deep learning-based three-dimensional segmentation of the prostate on computed tomography images. *J. Med. Imaging (Bellingham)*, 6(2):025003, April 2019.

[43] Qikui Zhu, Liang Li, Jiangnan Hao, Yunfei Zha, Yan Zhang, Yanxiang Cheng, Fei Liao, and Pingxiang Li. Selective information passing for MR/CT image segmentation. *Neural Comput. Appl.*, October 2020.

[44] Anjali Balagopal, Samaneh Kazemifar, Dan Nguyen, Mu-Han Lin, Raquibul Hannan, Amir Owrangi, and Steve Jiang. Fully automated organ segmentation in male pelvic CT images. *Phys. Med. Biol.*, 63(24):245015, December 2018.

[45] Shuai Wang, Qian Wang, Yeqin Shao, Liangqiong Qu, Chunfeng Lian, Jun Lian, and Dinggang Shen. Iterative label denoising network: Segmenting male pelvic organs in ct from 3d bounding box annotations. *IEEE Transactions on Biomedical Engineering*, 67(10):2710–2720, October 2020. doi:10.1109/TBME.2020.2969608.

[46] Sang Hyun Park, Yaozong Gao, Yinghuan Shi, and Dinggang Shen. Interactive prostate segmentation using atlas-guided semi-supervised learning and adaptive feature selection. *Med. Phys.*, 41(11):111715, November 2014.

[47] Wei Li, Shu Liao, Qianjin Feng, Wufan Chen, and Dinggang Shen. Learning image context for segmentation of the prostate in CT-guided radiotherapy. *Phys. Med. Biol.*, 57(5):1283–1308, March 2012.

[48] Ling Ma, Rongrong Guo, Guoyi Zhang, David M Schuster, and Baowei Fei. A combined learning algorithm for prostate segmentation on 3D CT images. *Med. Phys.*, 44(11):5768–5781, November 2017.

[49] Shuai Wang, Kelei He, Dong Nie, Sihang Zhou, Yaozong Gao, and Dinggang Shen. CT male pelvic organ segmentation using fully convolutional networks with boundary sensitive representation. *Med. Image Anal.*, 54:168–178, May 2019.

[50] Yeqin Shao, Yaozong Gao, Qian Wang, Xin Yang, and Dinggang Shen. Locally-constrained boundary regression for segmentation of prostate and rectum in the planning CT images. *Med. Image Anal.*, 26(1):345–356, December 2015.

[51] Maysam Shahedi, Martin Halicek, Rongrong Guo, Guoyi Zhang, David M Schuster, and Baowei Fei. A semiautomatic segmentation method for prostate in CT images using local texture classification and statistical shape modeling. *Med. Phys.*, 45(6):2527–2541, June 2018.

[52] Najeeb Chowdhury, Robert Toth, Jonathan Chappelow, Sung Kim, Sabin Motwani, Salman Punekar, Haibo Lin, Stefan Both, Neha Vapiwala, Stephen Hahn, and Anant Madabhushi. Concurrent segmentation of the prostate on MRI and CT via linked statistical shape models for radiotherapy planning. *Med. Phys.*, 39(4):2214–2228, April 2012.

[53] Yaozong Gao, Yeqin Shao, Jun Lian, Andrew Z Wang, Ronald C Chen, and Dinggang Shen. Accurate segmentation of CT male pelvic organs via regression-based deformable models and multi-task random forests. *IEEE Trans. Med. Imaging*, 35(6):1532–1543, June 2016.

[54] Fabio Martínez, Eduardo Romero, Gaël Dréan, Antoine Simon, Pascal Haigron, Renaud de Crevoisier, and Oscar Acosta. Segmentation of pelvic structures for planning CT using a geometrical shape model tuned by a multi-scale edge detector. *Phys. Med. Biol.*, 59(6):1471–1484, March 2014.

[55] Neeraj Sharma, AmitK Ray, KK Shukla, Shiru Sharma, Satyajit Pradhan, Arvind Srivastva, and LalitM Aggarwal. Automated medical image segmentation techniques. *Journal of Medical Physics*, 35(1):3, 2010.

[56] Claudia Lindner. Automated image interpretation using statistical shape models. In *Statistical Shape and Deformation Analysis*, pages 3–32. Elsevier, 2017.

[57] T F Cootes, C J Taylor, D H Cooper, and J Graham. Active Shape Models-Their Training and Application. *COMPUTER VISION AND IMAGE UNDERSTANDING*, 61(1):38–59, 1995.

[58] Maria João Medeiros de Vasconcelos. *Computational algorithms for image analysis: Applications on human vocal tract and silhouette*. PhD thesis, Faculdade de Engenharia da Universidade do Porto, 2015.

[59] Femke Danckaers, Toon Huysmans, Mattias Van Dael, Pieter Verboven, Bart Nicolaï, and Jan Sijbers. Building 3D statistical shape models of horticultural products. *Food Bioproc. Tech.*, 10(11):2100–2112, November 2017.

[60] Ana Bandarrinha Brandão. Predicting thumb osteoarthrosis using morphology: a 3d statistical shape analysis. Master's thesis, Instituto Superior Técnico de Lisboa, 8 2019.

[61] Toon Huysmans, Jan Sijbers, and Brigitte Verdonk. Automatic construction of correspondences for tubular surfaces.

[62] Toon Huysmans. Parameterization and correspondence for improved modeling, analysis, and visualization of tubular surfaces. Phd thesis, University of Antwerp, Antwerp, 2009.

[63] Alan Brunton, Augusto Salazar, Timo Bolkart, and Stefanie Wuhrer. Review of statistical shape spaces for 3D data with comparative analysis for human faces. *Comput. Vis. Image Underst.*, 128:1–17, November 2014.

[64] Ana Luísa Pereira da Costa. 3d generative models of lower limb sockets. Master thesis, Faculdade de Engenharia da Universidade do Porto, Porto, 2020.

[65] Giovanni Biglino, Claudio Capelli, Jan Bruse, Giorgia M Bosi, Andrew M Taylor, and Silvia Schievano. Computational modelling for congenital heart disease: how far are we from clinical translation? *Heart*, 103(2):98–103, January 2017.

[66] Pawel Skadlubowicz, Zdzislaw Król, Zygmunt Wróbel, Fritz Hefti, and Andreas Krieg. Using of statistical shape models for pelvis reconstruction in the oncology surgery. *Journal of Medical Informatics and Technologies*, 13:151–156, 2009.

[67] Principal Component Analysis (PCA) in Python with Scikit-Learn. Accessed at: 16 june 2022. URL: https://stackabuse.com/implementing-pca-in-python-with-scikit-learn/.