République Algerienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la Recherche scientifique
Université Abderrahmane Mira de Béjaia
Faculté des Sciences Exactes
Département d'informatique

# Mémoire de Master

Domaine : Mathématiques-informatique
Filière : Informatique
Spécialité : Intélligence Artificielle

Par :

Mr AZEGGAGH Bilal
Mr METMATI Mohammed Sadek

## Prostate cancer Detection and segmentation in MRI Images using advanced ML and DL techniques

Soutenu le 04 Juillet 2024 devant le jury :

| | | |
|---|---|---|
| Pr. | Rebouh Nadjette | Encadrant |
| Pr. | Belkhiri Louiza | Examinateur |
| Pr. | Ouyahia Samira | Président du jury |

Année Universitaire : 2023 - 2024

## ⁕ Remerciements ⁕

*Avant tout nous remercions dieu le tout puissant, qui nous a donné la force et la patience pour l'accomplissement de ce travail qui est pour nous le point de départ d'une merveilleuse aventure, celle de la recherche, source de remise en cause permanente et de perfectionnement perpétuel.*

*Nous vifs remerciements vont en premier lieu à notre encadrante Mme REBOUH Nadjette pour avoir accepté de nous guider tout au long de ce travail, pour sa disponibilité et son implication pour l'aboutissement de ce travail.*

*Aux membres du jury Mme BELKHIRI Louiza et Mme OUYAHIA Samira pour l'intérêt qu'elles ont porté à notre travail en acceptant de l'examiner et de l'enrichir par leurs propositions.*

*À tous les enseignants qui ont cultivé en nous l'amour de l'Informatique. Qu'ils trouvent dans ce modeste travail un petit geste de reconnaissance qui ne sera jamais à la hauteur de la grandeur de leurs âmes.*

*Enfin, nous remercions, de tout cœur, tous ceux qui ont contribué de près ou de loin à la réalisation de ce mémoire.*

## ⁕ *Dédicaces* ⁕

*Rien n'est aussi beau á offrir que le fruit d'une labeur qu'on dédie du fond du coeur á ceux qu'on aime et qu'on remercie en exprimant la gratitude et la reconnaissance durant toute notre existence.*
*Nous dédions ce mémoire : À nos très chers parents qui ont toujours été là pour nous, qui ont sacrifié leur vie pour notre réussite et nous ont éclairé le chemin par leurs conseils judicieux. Nous espérons qu'un jour, nous pourrions leur rendre un peu de ce qu'ils ont fait pour nous. Que Dieu leur prête bonheur et longue vie,*
*À nos frères et sœurs et, à nos familles,*
*À nos amis et tous ceux qui nous sont chers ,*
*À tous les bons professeurs qui nous ont enseignés, Et à notre encadrante qui nous a suivi jusqu'au bout de notre projet.*

*M. AZEGGAGH Bilal*
M. METMATI Mohamed Sadek

# Table of Contents

# List of figures

# List of tables

# Notations

| | |
|---|---|
| ADC | Apparent Diffusion Coefficient |
| AAPM | American Association of Physicists in Medicine |
| AFU | Association Française d'Urologie |
| AI | Artificial Intelligence |
| ANN | Artificial Neural Networks |
| ANOVA | Analysis of Variance |
| AUC | Area Under the Curve |
| ASD | Average Surface Distance |
| ASSD | Average Symmetric Surface Distance |
| b-value | Diffusion Sensitizing Gradient Value |
| BC | Breast Cancer |
| BPH | Benign Prostatic Hyperplasia |
| Bow | bag-of-words |
| CNN | Convolutional Neural Network |
| CS | Clinically Significant |
| CZ | Central Zone |
| DCE | Dynamic Contrast-Enhanced |
| DCE+ | Dynamic Contrast-Enhanced Positive |
| DCNN | Deep convolutional Neural Networks |
| DICOM | Digital Imaging and Communications in Medicine |
| DL | Deep Learning |
| DRE | Digital Rectal Examination |
| DSC | Dice Similarity Coefficient |
| DT | Decision Tree |
| DWI | Diffusion-Weighted Imaging |
| EBRT | External Beam Radiotherapy |
| ENet | Efficient Neural Network |
| ERFNet | Efficient Residual Factorized Network |
| FC | Fully Connected |
| FCN | Fully Convolutional Networks |
| GS | Gleason Score |
| GPU | Graphics Processing Unit |
| HAS | Haute Autorité de Santé |
| HD | Hausdorff Distance |

| | |
|---|---|
| HIFU | High-Intensity Focused Ultrasound |
| HPC | High Performance Computing |
| IoU | Intersection over Union |
| ISUP | International Society of Urological Pathology |
| k-fold | Cross-validation |
| KNN | K-Nearest Neighbors |
| LASSO | Least Absolute Shrinkage and Selection Operator |
| LFPR | Labor Force Participation Rate |
| LR | Logistic Regression |
| ML | Machine Learning |
| MLP | Multilayer Perceptron |
| MP-MRI | Multi-Parametric Magnetic Resonance Imaging |
| MRI | Magnetic Resonance Imaging |
| $ng/ml$ | nanograms per milliliter |
| NCI | National Cancer Institute |
| NIFTI | Neuroimaging Informatics Technology Initiative |
| NVIDIA | Nvidia Corporation |
| PCa | Prostate Cancer |
| PI-RADS | Prostate Imaging Reporting and Data System |
| PPV | Positive Predictive Value |
| PSA | Prostate-Specific Antigen |
| PZ | Peripheral Zone |
| RCNN | Region-based Convolutional Neural Network |
| ReLU | Rectified Linear Unit |
| ResNet-v2 | Residual Network version 2 |
| RF | Random Forests |
| ROC | Receiver Operating Characteristic |
| ROI | Region of Interest |
| RT | Radiotherapy |
| SIFT | Scale-Invariant Feature Transform |
| SPIE | the international society for optics and photonics |
| SVM | Support Vector Machine |
| T2-W | T2-Weighted |
| T2-WI | T2-Weighted Imaging |
| TP | Total Prostatectomy |
| TTP | Time to Peak |
| TZ | Transition Zone |
| VD | Volume Difference |
| VOE | Volume Overlap Error |

# General Introduction

In over half of the world's countries, prostate cancer (PCa) is the most frequently diagnosed cancer in men. In 2020, PCa was the second most prevalent cancer and the fifth deadliest cancer among men, with more than 1.4 million new cases and 375,000 deaths worldwide.

Magnetic resonance imaging (MRI) is now included in the traditional diagnostic process for PCa. This includes multiparametric MRI (mpMRI), which comprises a three-dimensional (3D) T2-weighted anatomical sequence, a three-dimensional diffusion-weighted imaging (DWI) sequence, and a four-dimensional dynamic contrast-enhanced (DCE) sequence. Consequently, the volume of data to be examined simultaneously is considerable, making the task challenging, especially when different sequences lead to different conclusions. This results in significant variability between observers, with less experienced radiologists performing worse than those trained specifically for this task.

For all these reasons, research into diagnostic aid methods to help radiologists analyze mp-MRI has been prolific for over a decade. We can distinguish approaches based on learning, in particular deep learning. In this work, we have chosen to study deep learning methods for the analysis of prostate MRI imaging data, more specifically for the detection and segmentation by aggressiveness of cancers in the peripheral and transition zones.

The study begins with an overview of the anatomy and function of the prostate, followed by an in-depth discussion on the epidemiology, risk factors, and characteristics of prostate cancer. The research highlights the importance of early detection and accurate diagnosis, which are crucial for effective treatment and management. It also emphasizes the role of machine learning (ML), particularly convolutional neural networks (CNNs), which are widely used for medical image recognition.

At the core of our work, we developed and evaluated various ML and deep learning (DL) models for the detection and segmentation of PCa. Different imaging techniques, including T2-weighted imaging, diffusion-weighted imaging (DWI), and dynamic contrast-enhanced imaging (DCE), were utilized to enhance the accuracy of these models. The study also addresses the limitations of MRI in PCa diagnosis and explores the potential of machine learning to overcome these challenges.

In the end, we combined two approaches to propose a better method for tumor segmentation. The proposed methods were tested on a public MRI dataset, and the results demonstrated significant improvements in accuracy.

# Generality

1

Sommaire

## 1.1 Introduction

In this chapter, we present some elements of the anatomy and function of the prostate within the male reproductive system. We then focus on prostate cancer by providing some epidemiological data before describing the diagnostic protocol and therapeutic management of this pathology. As we will also review the fundamental concepts of machine learning, which are the subject of our study. The diagnostic and treatment sections are based on the latest French recommendations from the Cancer Committee of the AFU for 2020-2022 described in [2]

## 1.2 Epidemiology of Prostate Cancer

Prostate cancer (PCa) is the most commonly diagnosed cancer in men in more than half of the countries worldwide [3]. With over 1.4 million new cases and 375,000 deaths globally, it was the second most frequent cancer and the fifth deadliest cancer among men in 2020.

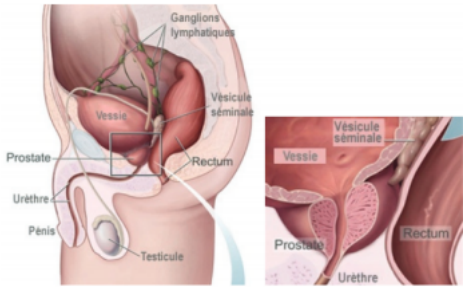



Figure 1.1 – *Pelvis and Male Reproductive System. via Wikimedia Commons*

### 1.2.1 Risk factors

The risk is particularly high among elderly individuals : nearly 79% of deaths involve men aged 75 and older, with a median age at diagnosis of 68 in 2018 [4]. In addition to age, family history represents a significant risk factor ; PCa is the cancer for which the weight of heredity is predominant [2]. Lastly, ethnicity is also a risk factor, with African or Afro-Caribbean descent increasing the risk of PCa. In the French West Indies - where 90% of the population is of African ethnic ancestry - the incidence and mortality of PCa are twice as high compared to metropolitan France (respectively 173/100000 for incidence and 23/100000 for mortality according to the report published by Public Health France in 2019).

## 1.3 Anatomy and Function of the Prostate

The prostate (see Figure 1.1) is a gland of the male genital system. It is located in the pelvis, beneath the bladder in front of the rectum, and surrounds the beginning of the urethra, the canal through which urine is eliminated from the bladder. A healthy prostate has the shape of a chestnut, approximately 3

centimeters in height and 4 centimeters in width, and it weighs no more than 20 grams in adulthood. The prostate is surrounded by a capsule and consists of several lobes : an anterior prostatic lobe, two lateral lobes, and a median lobe, also known as the lobe of Home. It is divided into 3 main glandular zones :

- The peripheral zone (PZ) : this is the region of the prostate closest to the rectum. It constitutes the largest area of the prostate and covers its lateral and posterior surfaces.
- The transition zone (TZ) : this is the area comprising 2 lobes located in the middle of the prostate in front of the peripheral and central zones. It surrounds the urethra and represents about 5% of the prostate until the age of 40. With aging, this zone increases in size to become the largest part of the prostate. This is what is called a prostate adenoma (also known as benign prostatic hyperplasia, noted BPH), which occurs in almost all men over 70 years old. The increase in size of the transition zone results in pushing the peripheral zone towards the rectum.
- Central zone (CZ) : this is the part of the prostate located at the base surrounding the ejaculatory ducts. This cone-shaped zone is often associated with the TZ.

## 1.4 Characteristics and Typology of Prostate Cancer

The vast majority of prostate cancers are located in the peripheral zone (PZ), they represent approximately 75% of cancers, with the remaining 25% found in the transition zone (TZ).

### 1.4.1 Gleason Score

The Gleason Score (GS) allows for the assessment of the aggressiveness of PCa. It is a major prognostic factor : it measures the extent and aggressiveness of the disease and will be considered for the choice of treatment to be offered.
It is defined based on biopsy analysis : depending on the level of cell differentiation, an architectural grade ranging from 1 to 5 is first assigned to the cells composing the tumor (see Figure 1.2).
Cancer is often heterogeneous, with tumor foci of different evolution stages, and varying degrees of differentiation can coexist within the same tissue sampled during biopsy. The Gleason score is then obtained by adding together the 2 most represented histological grades within the cancer.
The classification defined by Gleason in 1966 consisted of 5 architectural grades ranging from 1 to 5, with the sum defining 9 scores from 2 (1+1) to 10 (5+5). The predominant grade is the first term of the addition : a Gleason 7 tumor (3+4) is predominantly composed of grade 3 and to a lesser extent of grade 4, unlike a Gleason 7 tumor (4+3), which will therefore be more aggressive.

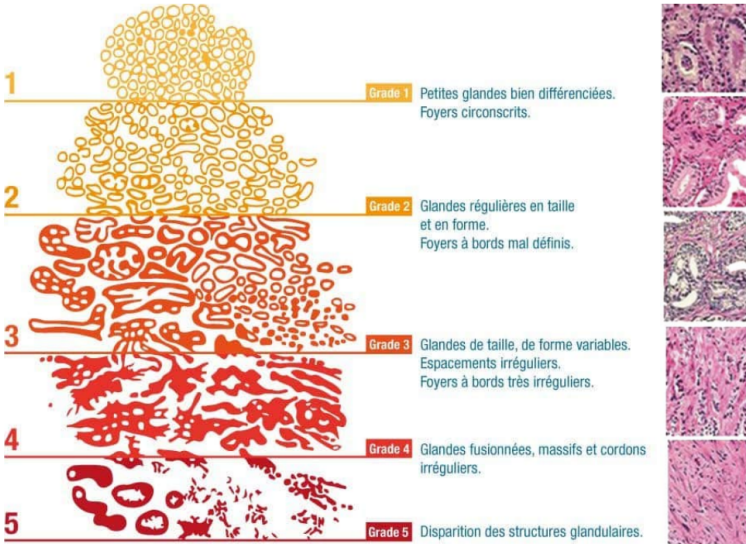



FIGURE 1.2 – *Representation of the Gleason grade based on cell differentiation (diagram and histological section)*

| | |
|---|---|
| Group 1 | Gleason Score 6 (3 + 3) |
| Group 2 | Gleason Score 7 (3 + 4) |
| Group 3 | Gleason Score 7 (4 + 3) |
| Group 4 | Gleason Score 8 (3 + 5, 5 + 3 or 4 + 4) |
| Group 5 | Gleason Score 9 or 10 |

TABLE 1.1 – *Prognostic Groups of the ISUP 2016 Classification*

Indeed, this classification had several shortcomings. Firstly, almost all currently diagnosed PCa have a minimum score of 6, corresponding to very well-differentiated cancers. It is therefore difficult for patients to understand that they have an indolent cancer when their score is in the median range of the Gleason scale. Furthermore, the Gleason score per se does not differentiate between scores of 7 (3 + 4) and 7 (4 + 3).
A new classification was therefore proposed by the ISUP [5], with 5 prognostic groups presented in Tableau 1.1. This classification serves as the reference nowadays.

## 1.5 Detection and Diagnostic Methods

In France, systematic screening for PCa in asymptomatic men is not recommended by the Haute Autorité de Santé (HAS). Despite this non-recommendation, the number of individual screenings is high and leads to the detection of early forms of PCa. Nearly 40 to 50% of identified cancers are slow-growing and would never have manifested during the individuals' lifetimes.

### 1.5.1 Early Detection

1. **PSA :**
   Prostate-Specific Antigen (PSA) is a molecule produced only by the prostate. PSA testing in the blood is an indicator for diagnosis since the risk of PCa increases with the total PSA value.
   A value higher than 4 ng/ml is generally considered abnormal, but this value must be interpreted by the physician based on the clinical context. There is no threshold value of PSA below which there is no risk of cancer.
   However, PSA is characteristic of prostatic epithelium and not specific to PCa, so it can be elevated in other situations (prostate infection, adenoma, benign hypertrophy, etc.). PSA kinetics (doubling time, velocity) are particularly useful for patient monitoring after treatment.
   Early detection of PCa also relies on the search for family history and ethnicity, in addition to digital rectal examination and total PSA testing.
2. **Digital Rectal Exam :**
   Digital Rectal Examination (DRE) is recommended prior to prescribing total PSA testing. A suspicious DRE is an indication for prostate biopsies regardless of the total PSA value.

### 1.5.2 Diagnosis

1. **Imagery :**
   In case of suspected PCa, the latest version of urological *guidelines* recommends performing an MRI acquisition before any initial series of prostate biopsies, contrary to previous practices [6]. Indeed, MRI has proven to increase the identification of significant PCa and guide prostate biopsies on these lesions. The MRI consists of 3 sequences : T2-weighted (T2-w), diffusion-weighted (DWI), and dynamic contrast-enhanced (DCE).

2. **PI-RADS Score :**
   The Prostate Imaging Reporting and Data System (PI-RADS) score was established to reduce inter-observer variability and improve the analysis of prostate MRI exams to reliably diagnose significant tumors (GS$\geqslant$7). The score is obtained by combining the analysis of T2-weighted, DWI, and DCE sequences for each suspicious area. The criteria differ depending on the zones (PZ or TZ), and a sequence will be favored in case of discordant sequences : DWI for the PZ and T2-w for the TZ. (see Figure 1.3)
   The PI-RADS ranges from 1 to 5, corresponding to the probability that the cancer is clinically significant (CS) :
   — PI-RADS 1 : Very low risk of CS cancer.
   — PI-RADS 2 : Low risk of CS cancer.
   — PI-RADS 3 : Equivocal risk of CS cancer.
   — PI-RADS 4 : High risk of CS cancer.
   — PI-RADS 5 : Very high risk of CS cancer.

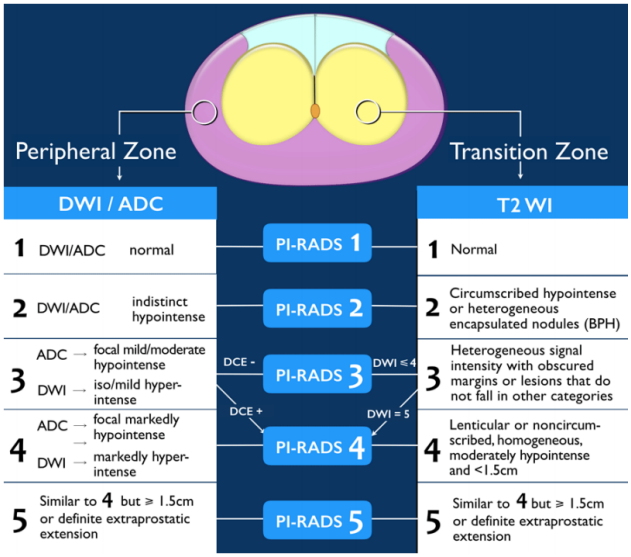


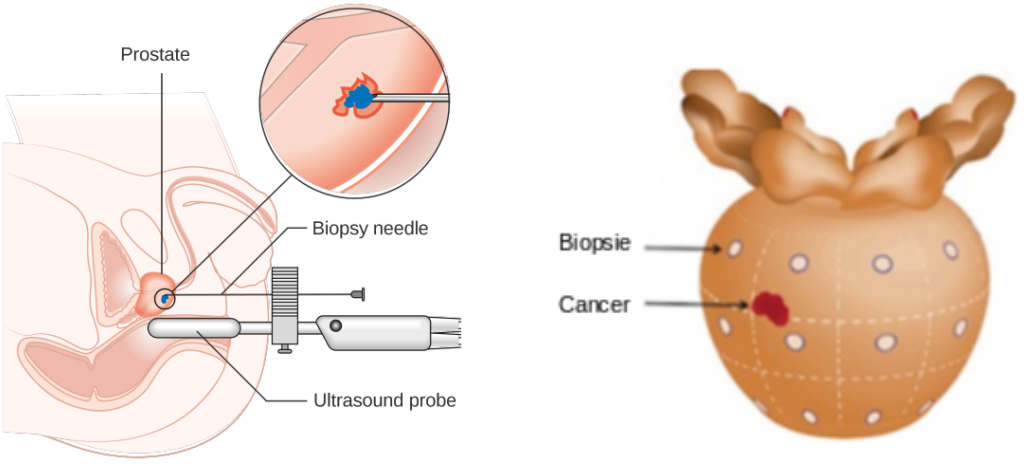

FIGURE 1.3 – *Criteria for scoring T2-w and DWI/ADC sequences based on the considered zone (PZ or TZ) for establishing the final PI-RADS score. via Radiology Assistant.*

Examples of cases corresponding to the different PI-RADS are illustrated in Figure 1.5.
The tumor risk stratification on MRI is based on the PI-RADS score, which partly determines the biopsy strategy : either simple biopsies not guided by imaging (PI-RADS score $<3$), or addition of MRI-guided biopsies if the MRI is positive (PI-RADS score $\geqslant 3$).

3. **Biopsy :**
A biopsy is a sampling of a part of tissue (here, from the prostate gland) for analysis. It is an invasive procedure that can be painful and have complications. The recommended standard scheme for systematic biopsies consists of 12 samples (Figure 1.4B). For each lobe, samples are taken at the medial and lateral levels : at the base, middle, and apex. It is important to note that a biopsy is a localized tissue sample that reflects only one area of the prostate.
In case of a positive MRI, targeted biopsies are combined with systematic biopsies. Ultrasound is the reference examination for performing targeted biopsies on suspicious lesions detected on MRI, either through visual guidance (cognitive targeting) or through MRI-ultrasound image fusion techniques (Figure 1.4A). MRI-guided biopsies are technically more challenging and expensive.
The histopathological examination (microscopic analysis of tissue samples) allows for defining the Gleason score and the prognostic group of the corresponding ISUP 2016 classification.

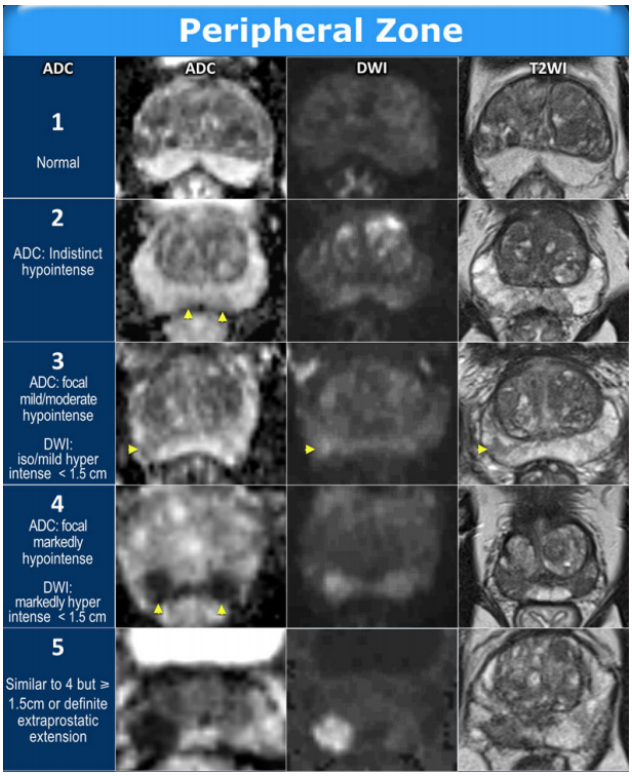
In case of persistent suspicion of PCa after a first series of negative biopsies, a second series of prostate biopsies may be indicated. There is no

consensus on the optimal interval between biopsy series. Prior to a second series of biopsies, MRI-guided biopsies significantly increase the rate of clinically significant cancers (ISUP grade $\geqslant 2$).

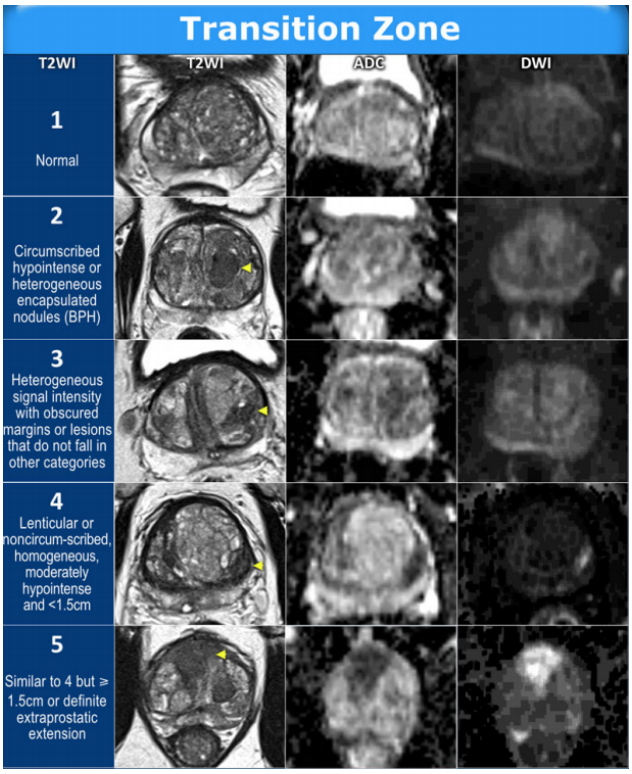



**(A)** ***Source : Cancer Research UK, via Wikimedia Commons*** **(B)** ***Adapted to CCM urology***

FIGURE 1.4 – *(A) Prostate Biopsy Under Ultrasound Guidance. (B) The 12 biopsy points performed during random biopsies, which may miss a cancer.*




**(A)** *for PZ*




**(B)** *for TZ*

FIGURE 1.5 – *Example of cases corresponding to the different PI-RADS scores for the two main zones of the prostate. via Radiology Assistant.*

## 1.6 Therapeutic Approaches

1. **Total Prostatectomy (TP) :**
   TP is one of the standard treatments for localized PCa, offering the greatest long-term oncological guarantee. The objective of TP is the removal of the entire prostate and seminal vesicles. However, its inevitable side effects on urinary continence and erectile function - despite sphincter and neurovascular bundle preservation - have a significant impact on the quality of life of treated patients.

2. **Radiotherapy (RT) :**
   External beam radiotherapy and brachytherapy are among the potentially curative therapeutic modalities offered in the treatment of non-metastatic prostate cancer.

3. **Brachytherapy or Internal Radiotherapy :**
   Brachytherapy is a highly localized treatment for cancer that involves placing radioactive implants (iodine 125 seeds or iridium 192 sources) inside the prostate. These implants emit radiation that destroys prostate cancer cells. Brachytherapy is a possible therapeutic modality for certain low-risk localized prostate cancers.

4. **External Beam Radiotherapy (EBRT) :**
   External beam radiotherapy is a localized treatment for cancer that uses ionizing radiation to destroy cancer cells by preventing them from multiplying. It involves precisely directing these radiation beams onto the area to be treated, while preserving healthy tissues and neighboring organs, known as organs at risk (including the bladder and the last segment of the digestive system : rectum and anal canal) as much as possible. These radiations are produced by a device called a particle accelerator. They are directed in beams towards the prostate to reach, through the skin, the tumor as well as neighboring lymph nodes. It is also a standard treatment, in combination with hormone therapy, for high-risk and locally advanced PCa. It is also associated with a risk of impotence and potentially urinary incontinence, as well as a risk of rectal inflammation (radiation proctitis).

5. **Hormone Therapy :**
   PCa is a hormonally sensitive cancer, meaning its development is stimulated by male hormones : androgens, particularly testosterone. Hormone therapy aims to prevent the stimulating action of testosterone on cancer cells to halt the cancer's development. Its effect is only temporary. Hormone therapy, combined with EBRT, is the standard treatment for locally advanced prostate cancers and one of the possible treatments for high-risk localized forms. Hormone therapy is usually initiated before radiotherapy and continued after radiotherapy for up to three years.

6. **Chemotherapy :**
   Chemotherapy is a treatment that targets the mechanisms of cell division. It is a general, or systemic, treatment that works throughout the entire body. This allows it to reach cancer cells regardless of their location. It is indicated for the treatment of metastatic hormone-resistant cancers with the aim of relieving pain or controlling the symptoms of the disease.

7. **HIFU :**
   High-Intensity Focused Ultrasound is a non-surgical therapy developed over the past thirty years for selected patients with localized prostate cancer. It is a minimally invasive treatment that targets prostate cancer by concentrating high-intensity focused ultrasound waves to destroy cancer cells through heat without damaging surrounding tissues. This highly precise local treatment reduces the risk of side effects. Notably, the first clinical trials began in 1993 at Edouard Herriot Hospital (Lyon, France), using a prototype developed at the LabTAU laboratory (Lyon, France). The RHU PERFUSE project, which funds this thesis, continues this work and aims to evaluate the oncological outcomes of focal HIFU treatment.

8. **Focal Treatments :**
   Other types of alternative focal treatments are being studied but are still under evaluation : cryotherapy and laser therapy. In general, focal therapy should be considered an evolving technique. Focal treatment requires precise knowledge of the location of tumor foci within the gland.

9. **Active Surveillance :**
   Active surveillance allows for delaying the initiation of curative treatment (and its associated side effects). It is a standard therapeutic option for low-risk tumors with slow progression. The principle of active surveillance is based on regular examinations : clinical exams, repeated PSA tests to monitor its progression, prostate biopsies, and possibly MRI. If disease progression is detected, a treatment aimed at curing the disease can be scheduled.

## 1.7 Imaging and Technology in Diagnosis

Recently, MRI has become a key imaging technique for prostate cancer. It guides biopsies, non-invasively assesses the cancer stage, and locates the cancer for focal treatment.
Prostate MRI is a standardized examination that must follow specific criteria in its implementation and analysis.

### 1.7.1 Acquisition

The MRI must be multiparametric and composed of the following sequences :
— Anatomical.
— Morphological (T2-w).

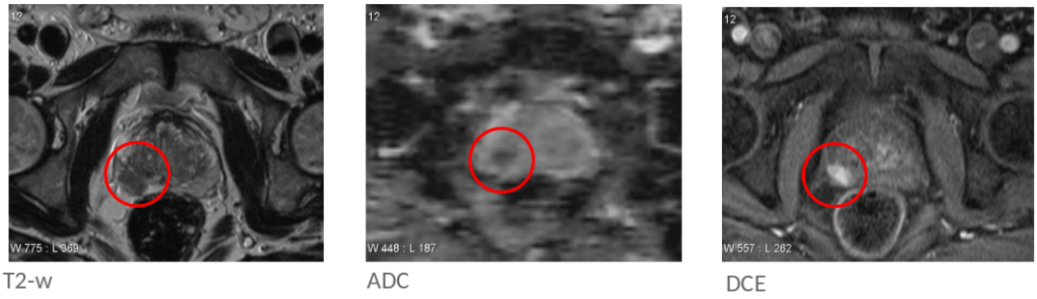



FIGURE 1.6 – *Multiparametric MRI of the prostate. Here, the slices are axial in the median plane.*

— Functional.
— Diffusion with a high b-value $\geqslant$ 1400 (DWI).
— T1 perfusion with contrast agent injection (DCE).

The reference plane is the oblique axial plane perpendicular to the rectal wall. Prostate MRI can be performed at 1.5T or 3T. The examination can be conducted with only an endorectal coil, with an external coil (see Figure 1.7), or by combining these two coils. A full examination takes 20 to 30 minutes.

**(A)** ***Pelvic coil*** **(B)** ***Endorectal coil*** **(C)** ***Siemens 1.5T MRI***

FIGURE 1.7 – *Used for some of the MRI-MP acquisitions. Adapted from Niaf [1].*

### 1.7.2 T2-weighted imaging

The T2-weighted sequence is also referred to as the "morphological" sequence : it allows visualization of the different zones of the prostate (PZ, TZ, and CZ). It is the reference sequence for visualizing prostatic tissues. On T2-weighted images, cancerous lesions appear as areas of low signal intensity.

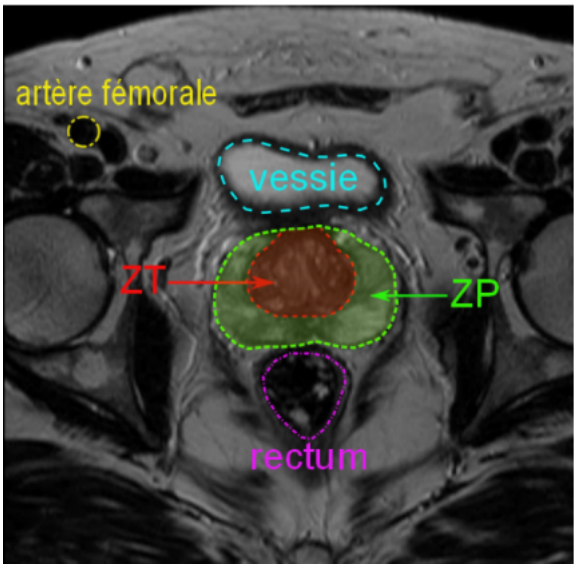



FIGURE 1.8 – *MRI of the male pelvis : T2-weighted sequence. Adapted from Niaf [1]*

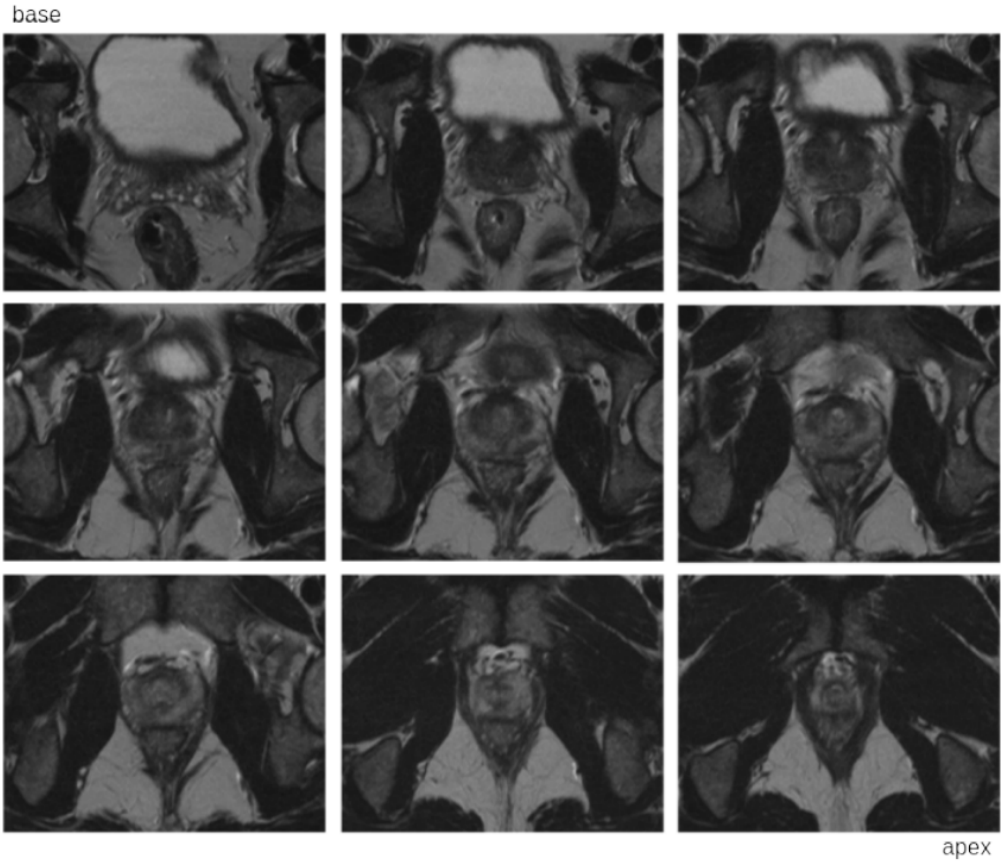



FIGURE 1.9 – *Axial slices of the prostate on T2-weighted MRI acquired from a 53-year-old patient. Source : Niaf [1]*

### 1.7.3 Diffusion-weighted imaging (DWI)

Diffusion sequences allow the study of molecular water movements. The acquisition involves applying several diffusion gradients *b* to a rapid MRI sequence. The higher the *b-value*, the more the signal decreases. In abnormal tissues, the signal persists despite the increase in *b*, and tumors appear as hyperintense on diffusion. The slope of the signal curve as a function of *b-value* allows the extraction of a feature map called the *Apparent Diffusion Coefficient (ADC)*, on which cancers appear as hypointense.
Diffusion imaging also provides information about tumor aggressiveness. Studies do not all report the same conclusions, but there appears to be an inversely proportional correlation between the ADC value and the Gleason score : the lower the ADC, the higher the Gleason score would be. A recent literature review of 39 studies concluded a moderate correlation in the PZ and a weak correlation in the TZ [7].

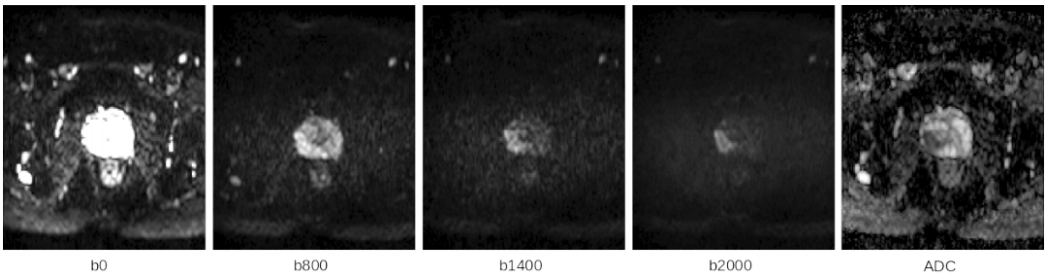



FIGURE 1.10 – *Axial diffusion-weighted MRI slices for different b values and corresponding ADC map.*

### 1.7.4 Dynamic contrast-enhanced (DCE) imaging

Dynamic contrast-enhanced (DCE) imaging allows the study of microvascularization (tissues affected are generally rich in micro blood vessels). The acquisition relies on the intravenous injection of a contrast agent (gadolinium) in a bolus, the signal of which can be tracked over time. For this purpose, short T1-weighted gradient echo sequences of 10-15 seconds are repeated over 2-3 minutes. This sequence is also referred to as a "dynamic" sequence.

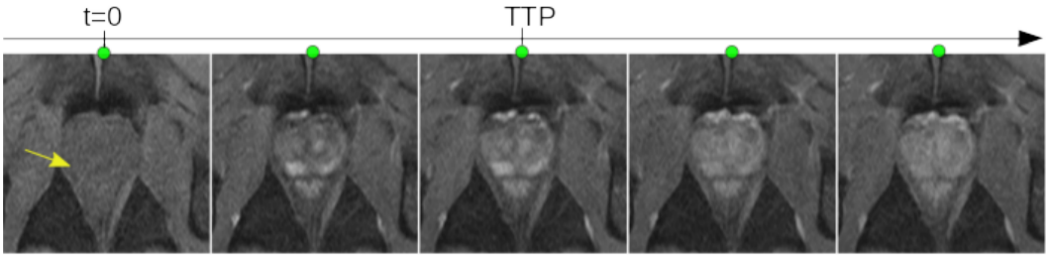



FIGURE 1.11 – *Images corresponding to different times of a perfusion MRI sequence. Adapted from Niaf [1]*

Analysis of the intensity curve over time allows for the extraction of discriminative parameters such as :

- Maximum enhancement : the maximum intensity of the signal during the time course.
- Time to peak (TTP) : the time between the onset of enhancement and the peak enhancement.
- Wash-in rate : the slope of the line between the onset of enhancement and the peak enhancement.
- Wash-out rate : the slope of the linear regression line to the enhancement curve between the peak enhancement and the end of the sequence.
- Area under the enhancement curve (AUC) : the integral of the curve between the onset of enhancement and the end of the sequence.

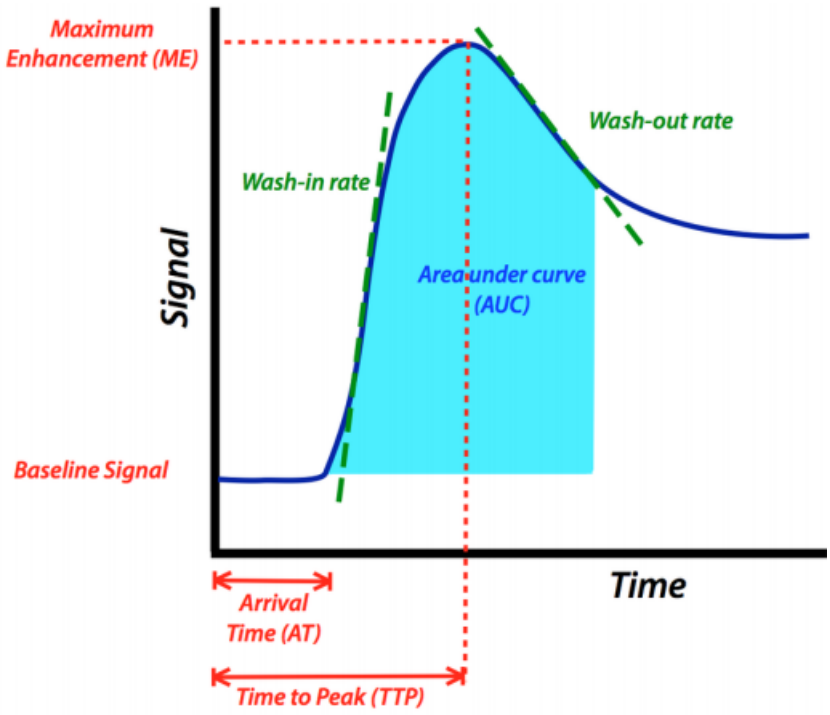



FIGURE 1.12 – *Extraction of quantitative parameters from the curve showing the mean signal over time of the DCE sequence. Source : MRIquestions*

A cancer is characterized by early, intense contrast enhancement and rapid *wash-out*.
A pharmacokinetic model (Tofts' two-compartment model) can also be established to calculate parameters such as the transfer constant $K^{trans}$, which reflects the transfer of the contrast agent into the interstitium *wash-in*.

### 1.7.5 Limitations of MRI for PCa Diagnosis

MRI has proven itself in recent years and is now at the heart of PCa diagnosis and patient monitoring. The combined analysis of morphological and functional sequences enhances cancer detection.

However, the combined analysis of MRI sequences is time-consuming and can be particularly challenging in the following cases :

- Presence of artifacts due to post-biopsy hemorrhagic foci.
- Prostatitis (inflammation of the prostate).
- Presence of benign nodules (abnormal round formations) that may be mistaken for lesions.
- Discordant sequences : the lesion may only be visible on one or two MRI sequences.
- Cancers not visible on MRI (mostly low-grade).

For all these reasons, intra- and inter-observer variability is high, experienced radiologists achieve higher sensitivity when analyzing mp-MRI sequences.
Despite all these efforts, the analysis remains a difficult and laborious task. Furthermore, certain information is currently not detectable on the images, such as the aggressiveness of the lesion (or Gleason score). This information would be valuable in clinical practice, to avoid or limit biopsies by targeting suspicious areas, or to allow non-invasive active surveillance of patients with low-grade cancers.
In the following section, we present the fundamentals of machine and deep learning for medical image segmentation.

## 1.8 Machine Learning

### 1.8.1 Generalities

Machine learning (ML) is a subclass of artificial intelligence (AI) that enables a system to learn from data rather than from a predetermined sequence of operations.

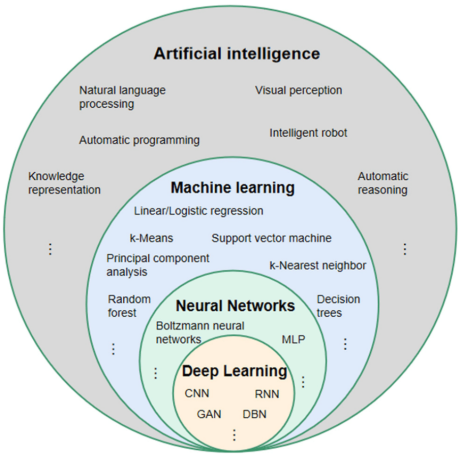



Figure 1.13 – *Relationship between AI, machine learning, neural networks, and deep learning. Adapted from Wikipedia*

### 1.8.2 Types of Machine Learning

There are several types of ML, each with its own advantages and limitations. Supervised learning involves learning from labeled examples, where each example is associated with a known response. Unsupervised learning involves finding structures and patterns in unlabeled data. Finally, reinforcement learning is a technique where the system learns through trial and error by interacting with a dynamic environment [8].

### 1.8.3 Applications of Machine Learning

ML is used in many fields, such as image recognition, natural language processing, product recommendation, financial fraud detection, and market trend prediction. ML algorithms are, also, used in broader applications such as malware recognition and natural disaster prediction [9].

### 1.8.4 Machine learning models

ML models are algorithms enabling computer systems to learn from data without explicit programming. In the following, we will present some used ML models in the studied domain.

#### Transfer Learning

Transfer learning (TL) is a ML technique that involves leveraging the knowledge acquired to solve a given task and applying it to a different but related task. Instead of starting the learning process from scratch, TL allows models to benefit from pre-trained models or knowledge gained from previous tasks.
In TL, a model is initially trained on a large dataset and a complex task, such as image classification or natural language processing. The knowledge and representations learned during this initial training are, then, transferred to a different but related task. This approach is particularly useful when the target task has a limited amount of labeled data, as the pre-trained model can provide an initial advantage and improve generalization.

1. **Decision Trees :**
   Decision trees are a ML architecture used for classification and prediction tasks. Decision trees consist of nodes and leaves. Nodes represent decisions to be made based on input data, while leaves represent the final classification outcomes[8]. Decision trees have been used for tasks such as fraud detection, predicting machine lifespan, and forecasting election results.
2. **Ensemble Learning :**
   Ensemble learning (EL) is a ML architecture that involves combining multiple ML models to improve prediction accuracy[10]. Individual models are often decision trees, neural networks, or other types of models. EL can be used for classification, regression, or clustering tasks.

The ML architectures described above are not the only ones that exist, but they are the most important and commonly used in the industry. The advantages of these architectures include their flexibility, their ability to work with large amounts of data, and their capability to discover complex relationships between data. However, each architecture has its advantages and disadvantages, and selecting the appropriate architecture depends on the task at hand and the available data.

Additionally, there are hybrid architectures that combine multiple types of ML models to achieve better results. For example, Convolutional Neural Networks can be combined with recurrent neural networks to handle both spatial and temporal data.

It's important to note that the choice of ML architecture doesn't always guarantee prediction accuracy. Results depend on many factors such as data quality, dataset size, task complexity, computational capacity, and the quality of the learning algorithm[10].

### 1.8.5 Challenges and limitations of Machine Learning.

Despite its numerous applications and benefits, ML also presents challenges and limitations. Firstly, it requires a significant amount of high-quality training data to produce accurate results. The data must also be representative of the population to which the model will be applied ; otherwise, the model may be biased[11].
Moreover, ML models can be difficult to interpret as they can be highly complex and may not provide a clear explanation of their predictions. This can be a problem in critical applications such as medicine, where predictions need to be understandable and justifiable[12].
Finally, ML is often used in sensitive applications such as surveillance and facial recognition, raising concerns about privacy and data security. It is important to consider these challenges and limitations when using ML in real-world applications[13].

### 1.8.6 Deep learning

Deep learning (DL) is a subfield of Machine Learning that focuses on creating relationships between input feature variables and output variables [14]. The architecture of DL was proposed for the first time in 2006 as an Artificial Neural Network (ANN) with an important learning capacity [15]. Nowadays, DL is one of the most important fields where it dominates and provides the most powerful solutions for many tasks such as image processing [16]. However, there are multiple extensions of DL models.

In the following section, we will present the main model for DL.

### Convoluional Neural Network (CNN)

In classical neural networks, the most common layers are fully connected layers, where neurons between two adjacent layers are connected pairwise, but neurons within the same layer do not share connections. CNNs are very similar to classical neural networks. Their particularity comes from convolution operations instead of fully connected layers, operations that apply to spatially structured signals such as text, audio signals, and images, exploiting their properties to the fullest. A CNN consists of three main types of layers : Convolutional layers, pooling layers, and fully connected layers. The concatenation of all these layers forms a CNN. These layers are further detailed in the following sections.

### Architecture of CNN

The architecture of a Convolutional Neural Network (CNN) is a key component of deep learning, specifically designed to efficiently process data such as images.
These layers are further detailed in the following section :

1. **Convolutional Layer :**
   Two Dimensions (2D) convolution involves sliding a filter over an image to extract features. In DL, the convolutional layer corresponds in practice to the mathematical operation of cross-correlation. The term 'convolution' will be used, henceforth, in the context of DL. The result of a convolution is called a feature map.
   The use of convolutional layers significantly reduces the number of parameters compared to fully connected layers, since the same filter is applied across the entire image (see Figure 1.14). Thus, the weights are shared among different neurons.

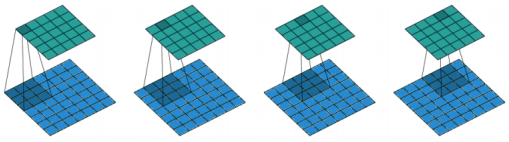


FIGURE 1.14 – *Convolution operation on a 7×7 image.*

2. **Pooling Layer :**
   Pooling layers are, often, inserted between convolutional layers in a CNN. This layer reduces the spatial size of the representation to decrease the number of network parameters, thus limiting the overfitting. The pooling layer operates, independently, along the z-dimension on each input feature map and spatially resizes it in the (x, y) plane using an operation such as maximum or average pooling. The most common form is max pooling with 2×2 filters applied with a stride of 2, discarding 75% of activations without altering the depth dimension (see Figure 1.15).
3. **Fully Connected Layer (FC) :**
   The Fully Connected (FC) layer is applied to an input where each neuron

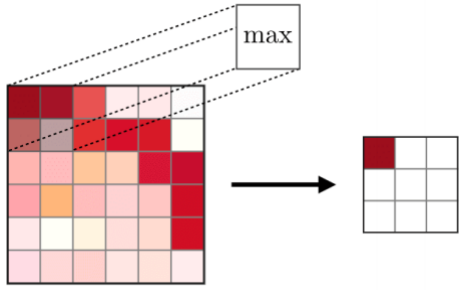



FIGURE 1.15 – *Pooling layers allow for downsampling the volume and reducing the number of parameters.*

is connected to all neurons. FC layers are, typically, present at the end of CNN architectures, before the classifier.

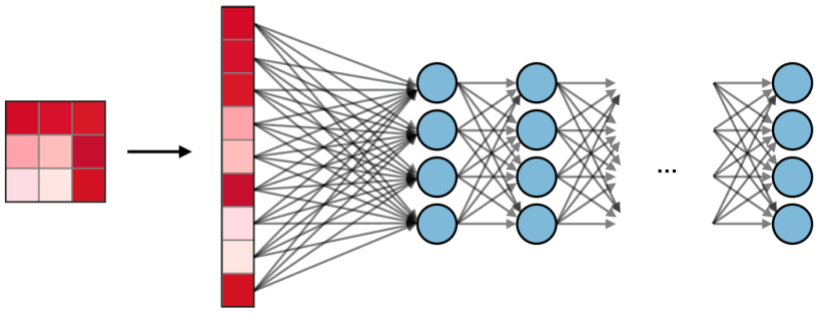


FIGURE 1.16 – *Fully Connected (FC) Layer, where each neuron in a layer is connected to all neurons in the following layer.*

4. **Normalization Layer :**
   The Normalization layers, when used, are inserted after a convolutional layer and before the activation function (see Figure 1.17). Normalization allows the use of a larger learning rate, limits gradient disappearance, and reduces excessive dependence on initialization.
5. **Classification Layer :**
   The classification layer corresponds to the last layer of the network. Generally, the **Sigmoid** function is used for a binary problem, and the **Softmax** classifier is used for a multi-class problem. The softmax classifier is based on linear regression and is also known as multinomial logistic regression. Figure 1.17 illustrates an example of a CNN architecture.

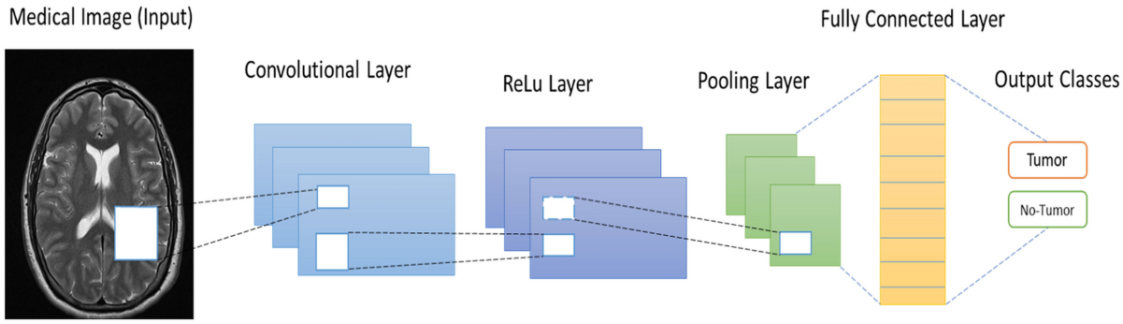



FIGURE 1.17 – *CNN Architecture*

## 1.9 CONCLUSION

The PCa is a major public health issue. Its diagnosis is continuously evolving, with the recent inclusion of MRI in the diagnostic process even before performing biopsies, following the demonstrated improvement in performance by targeting biopsies. However, the joint analysis of these multimodal images is tedious and time-consuming, especially when the sequences lead to different conclusions. Furthermore, MRI sensitivity remains low for cancers with a low Gleason score, and there is high inter-observer variability.

In the following chapter, we present a state-of-the-art that encompasses different approaches for the segmentation and detection of the PCa.

# Literature Review 2

Sommaire

## 2.1 Introduction

The literature proposes various methods for the segmentation and detection of PCa in MRI images, varying based on the type of images to be processed and the objectives of the analysis. This chapter presents a state-of-the-art review, defining the key concepts of image segmentation and highlighting its importance as well as its specific challenges. Then, it provides an overview of some recent methods proposed in the literature, followed by a synthesis and a comparative table of the presented methods according to some selected criteria.

## 2.2 Image Segmentation

Image processing is an important function of any processing and analysis process found in many computer programs. It consists of dividing an image into several parts or sections, each representing a specific part of the image. The segmentation phase includes image acquisition, enhancement, interpretation, and final decision-making. Segments are often defined by properties such as color, texture, brightness, and shape of objects in the image. Various classification methods have been developed to extract the characteristic features of interest. The effectiveness of each method depends largely on the characteristics of the image viewing.

### 2.2.1 Preliminaries and Definition

Image processing is an image analysis process that aims to simplify or transform an image into an image that is simple and easy to analyze. Due to its importance, it has been incorporated into many important applications such as automatic interpretation of biological and satellite images, tracking and comparison of motion capture images, and more. In our case, we are interested in a practical application of image segmentation in the medical field, which is the segmentation of the prostate in MRI images. As it is defined, image segmentation is the procedure of separating a digital image into numerous regions where each of the pixels in a region is comparable to particular features or calculated properties, such as color, intensity, or texture. If the image is signified by R, then, the segmentation is a procedure that partitions R into k sub-regions, $R_1$, $R_2 \ldots$, $R_k$.
Where $R_k \cap R_j = \emptyset \; \forall$ k $\neq$ j, and each $R_k$ is a associated region [17]. Segmentation aims to simplify and/or deviate the design of an image into an alternative form that is further sensitive and easier to observe. Image segmentation is typically used to find objects and boundaries (lines, curves, etc.) in images.

### 2.2.2 Importance of the PCa segmentation

MRI (Magnetic Resonance Imaging) segmentation plays a crucial role in the detection and treatment of the PCa. Segmentation techniques allow for precise delineation of the prostate and any tumor lesions, facilitating a more accurate assessment of the disease's extent. Deep learning-based methods, such as

convolutional neural networks (CNNs), have demonstrated remarkable efficacy in analyzing MRI images, offering better resolution and increased contrast of soft tissues[18]. These technological advances not only enhance diagnostic accuracy but also enable personalized treatment plans based on each patient's specific characteristics. Furthermore, automatic MRI image segmentation reduces the time and effort required for manual analysis while minimizing the risk of human errors [19]. Thus, integrating MRI image segmentation into clinical practice represents a significant advancement in the PCa management, contributing to more targeted interventions and improved patient outcomes.

## 2.3 Methods for the segmentation and detection of PCa

In this section, we present some recent methods that have been highlighted in the literature within the framework of segmentation and detection of the PCa.

### 2.3.1 The PCa grading framework based on deep transfer learning and Aquila optimizer

in this paper, Hossam et al. [20] proposed a hybrid framework for the diagnosis and segmentation of the PCa. The proposed framework is divided into two stages namely classification and segmentation, where they utilized 3 different datasets : PANDA (resized train data (512 * 512)) is used in the segmentation phase. Transverse Plane Prostate Dataset and ISUP grade-wise the PCa are used in the classification phase. Different data preprocessing methods were applied to make these datasets more efficient in performing the classification and segmentation processes. In the classification part, eight different CNN architectures via transfer learning were finetuned using Aquila Optimizer (AO). The AO was applied to tune the hyper-parameters of these models and then applied to diagnose patients with the PCa from normal ones. For the segmentation phase, the authors used U-net which is a convolutional network architecture and is applied for each masked layer on the PANDA dataset.

#### Method validation

The obtained results show the effectiveness of the proposed framework. In the classification phase, the best Accuracy, F1, Precision, Recall, Sensitivity, Specificity, AUC, IoU, Dice, and Cosine Similarity are obtained using the MobileNet pre-trained model for the two datasets. In the second stage U-Net achieved promising results with 98.46% of accuracy, the average loss is 0.0368, AUC is 0.9778, IoU is 0.9865, and Dice is 0.9873.

#### Review

The methodology used in this study is appropriate because the authors used a wide range of useful literature to identify the most effective constructs for the classification purposes. The data set used is not very diverse and may not be a true representation of the population. The results showed that this was true only

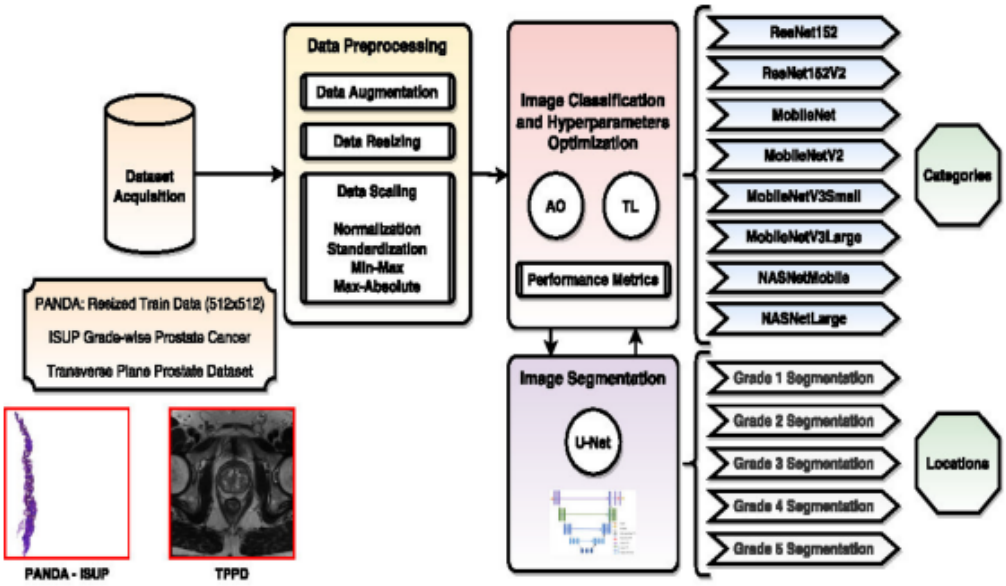



FIGURE 2.1 – *The suggested approach for the PCa classification and segmentation by Hossam et al.*

for performance criteria. The authors provide no evidence of how the model was tested on real data.

### 2.3.2 A new approach to diagnosing the PCa through Magnetic Resonance Imaging

The authors in [21] proposed a novel methodology that integrates advanced algorithms and machine learning techniques to enhance the accuracy and efficiency of the PCa detection and lesion segmentation in MRI images. The approach combines advanced techniques such as the improved GrowCut algorithm, Zernik feature extraction, and ensemble learning with machine learning algorithms like KNN, SVM, and MLP (Multilayer Perceptron), which is used in machine learning and pattern recognition tasks. Firstly, Li Zhang et al. used the GrowCut algorithm, a segmentation process that starts by initializing seed points within the MRI image. These seed points serve as markers to indicate the regions that are likely to belong to the prostate lesions. Then, the algorithm grows regions based on similarity criteria. It propagates the segmentation labels from the seed points to neighboring pixels that exhibit similar characteristics. This region-growing process helps delineate the boundaries of the prostate lesions. Following segmentation, authors applied the Zernik feature extraction algorithm to extract segment properties, providing valuable information for further analysis and diagnosis. The Zernik moments are computed to capture the geometric and textural properties of the segmented regions using a series of Zernik polynomials, which are orthogonal functions defined over a circular domain. These polynomials are used to represent the intensity distribution of the segmented regions in the MRI images. After the segmentation of the prostate lesion was

performed, all data should be divided into two sections. The training images are 80% of the total images and the Testing images are 20% of the sample. Then, the study incorporates ensemble learning, a technique that involves training multiple base models, such as SVM, KNN, and MLP algorithms, on the same dataset. Each base model learns to make predictions based on different aspects of the data by a voting mechanism to combine the predictions of each model. The authors use MRI and prostate imaging datasets for men (they used 271). A set of 217 samples is used in the ensemble learning system for training and 54 samples are for testing. The Figure 2.2 displays the key steps of the algorithm.

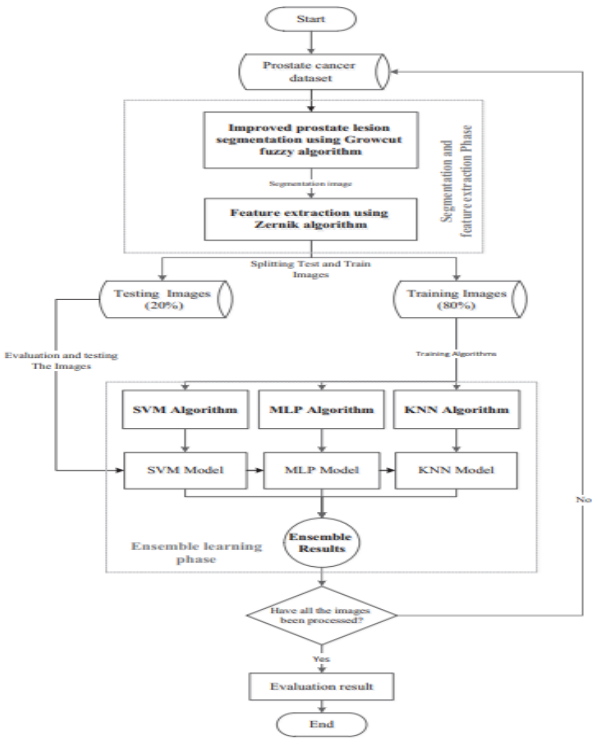



FIGURE 2.2 – *The structure and steps of the proposed approach.*

**Method validation**

The proposed method shows a significant improvement compared to other the PCa diagnostic methods. The approach achieved an accuracy of 80.97%, a precision of 76.69%, a recall of 77.32%, and an error rate of 19.02%. Compared to the KNN, SVM, and MLP methods, the proposed method demonstrated an improvement in accuracy by 4.19% and 24.23%, respectively. The Dice, LFPR, PPV, and VD criteria were also improved, with respective values of 0.79%, 0.07%, 1.24%, and 0.16%, thus surpassing the other evaluated methods.

#### Review

Combining these methods could improve diagnostic accuracy, but the work lacks details on their application and evaluation. A detailed explanation of each technique's steps and a well-structured approach would make the method more consistent and reproducible. The paper should, also, provide more information about the dataset used, as it's crucial for assessing the study's quality and generality. Basic information like dataset size, data entry method, and potential biases in data collection are important for understanding the reliability of the results. Regarding the results, there is a lack of in-depth analysis and comparison with existing methods or standards.

### 2.3.3 Prediction of the PCa aggressiveness with a combination of radiomics and machine learning-based analysis of dynamic contrast-enhanced MRI

In [22], authors explore a novel approach to predict the PCa aggressiveness using radiomics and machine learning techniques applied to dynamic contrast-enhanced MRI (DCE-MRI) images. The primary aim of the study was to investigate if this combination can predict PCa aggressiveness before biopsy. It included forty consecutive biopsy-confirmed PCa patients who underwent DCE-MRI examinations. Lesion segmentation was performed on the first and strongest phase of enhancement on the original DCE-MRI images, quantitative radiomics features were automatically calculated from each lesion, resulting in three datasets (Dataset-F, Dataset-S, and Dataset-FS). Techniques such as the variance threshold method, the select k-best method, and the LASSO algorithm were employed to reduce feature dimensions and select optimal subsets for analysis. Many machine learning approaches were utilized, including logistic regression (LR), support vector machine (SVM), random forests (RF), decision tree (DT), and k-nearest neighbor (KNN). The Cross-validation was used to evaluate the models, and the area under the receiver operating characteristic curve (AUC) was calculated to assess performance.

#### Method validation

Features were selected as optimal subsets in Dataset-F, Dataset-S, and Dataset-FS with eight, four, and 16 respectively, with LR-based analysis on Dataset-FS showing the highest prediction efficacy (AUC=0.93). In Dataset-FS, ten features showed a significant positive correlation with the Gleason score (GS) of the PCa lesion, indicating their potential as predictive markers for aggressiveness. The study concluded that the combination of radiomics and machine learning analysis of DCE-MRI images can predict PCa aggressiveness non-invasively, accurately, and automatically. The LR model based on Dataset-FS demonstrated the best classification efficacy, offering a valuable tool for clinicians in determining appropriate treatment strategies for PCa patients.

**Review**

Although the study demonstrated strengths in the use of multiple machine learning methods and inference techniques, a clear weakness was the lack of discussion on specific reproducibility criteria and algorithm selection. The quality of the data provided by biopsy-confirmed PCa and multiple DCE-MRI images was good, but the small sample size and lack of inter-observer variability raised concerns. The results showed that the study could provide additional clinical information and address biases.

### 2.3.4 A novel deep learning-based technique for detecting the PCa in MRI images

The proposed method in [23] involves a deep learning framework based on a 3D convolutional neural network (CNN) for detecting the PCa in MRI images that integrate advanced object detection and classification techniques. the authors used the SPIE-AAPM-NCI Prostate dataset. This dataset consists of MRI images of the prostate gland, including T2-weighted images, Dynamic Contrast-Enhanced (DCE) images, and Apparent Diffusion Coefficient (ADC) images from 346 patients with PI-RADS scores of 3 or higher. Different data pre-processing techniques and data augmentation are applied to enhance the dataset's diversity. The model is trained on the pre-processed dataset, with specific attention to lesion locations, and Gleason scores. The researchers apply transfer learning using the Faster RCNN with Inception ResNet-v2 architecture during the training process of the 3D convolutional neural network.

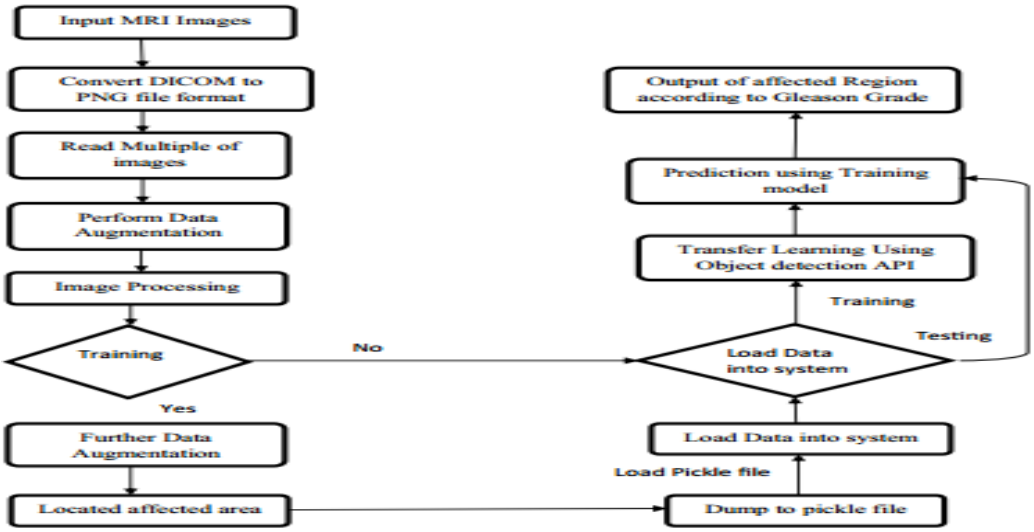



FIGURE 2.3 – *The structure and steps of the proposed approach.*

**Method validation**

The results obtained with the proposed model showed 87% overall accuracy, 85% specificity, and 89% sensitivity. This method provides more reliable detection of the PCa compared to other models in this paper. Regardless of the results obtained, the authors recommend the use of "ROI Align" for further improvement.

**Review**

The paper explains the process in detail, including various steps and procedures. However, it does not provide enough information to adequately know about the dataset and represent the problem at hand. The authors do not compare the performance of the proposed method in detail with other existing methods, which limits the ability to provide a comprehensive evaluation of the effectiveness of the method used.

### 2.3.5 Searching for the PCa by fully automated magnetic resonance imaging classification : Deep learning versus non-deep learning

This paper [24] presents a comprehensive analysis of the application of deep learning techniques, specifically Deep convolutional Neural Networks (DCNN), in comparison to non-Deep Learning methods for the automated classification of the PCa patients based on MRI imaging data. The study involved 172 patients with 2,602 morphological images of the prostate obtained through axial 2D T2-weighted imaging. The deep learning method used DCNN to automatically learn image features from prostate morphological images. The model consisted of five convolution layers, two inner product layers, max-pooling layers, and non-linear ReLU layers. The input of the model was a 288x288x3 MR image, and the output layer provided the probability of a patient having PCA. It was trained using high-performance computational resources, including two Nvidia Titan X GPUs, and for 1000 iterations it took only 10 minutes, and testing a patient's image required less than 0.5 seconds. On the other hand, non-deep learning methods used SIFT image extraction and Bag-of-words (Bow) models for recognition and analysis.

**Method validation**

To compare the performance of the two models, ROC curves were generated. To quantify the classification, the authors used the area under the ROC curve (AUC). The study concluded that deep learning with DCNN exhibited superior differentiation performance compared to the non-deep learning method in distinguishing PCa patients from the Breast Cancer (BCs) patients based on MRI imaging data.

**Review**

In terms of methodology, the lack of information regarding hyper-parameter tuning and the lack of comparison with other deep learning methods or en-

semble methods raise concerns about the effectiveness and generalizability of the method. Due to the quality of the data, the study is limited to describing data augmentation measures and quality control measures that may affect the representativeness of the dataset and introduce bias. In terms of results, the failure to discuss positive and negative falses and the model descriptions limits the understanding of the error evaluation and the decision-making

### 2.3.6 A novel solution of using deep learning for the PCa segmentation enhanced batch normalization

The proposed work [25] aimed at improving the accuracy of prostate segmentation on MRI images. This method consists of four main stages. In the first stage, the noise has been eliminated using a de-speckle filter. For the segmentation phase, the U-Nets, based on the FCN 8S standard, are implemented to delineate the prostate boundary and segment the prostate without dividing the network with the addition of pooling and up-sampling layers to enhance the segmentation model for more performing segmentation. A The authors used ReLU as an activation function. Moreover, for the Feature extraction, they used the batch normalization and dropout layer to optimize the performance.

#### Method validation

The proposed approach for PCa segmentation using deep learning demonstrates significant advancements regardless of the accuracy and the efficiency. By combining multi-level features and employing a deep neural network, the model effectively extracts high-level and low-level features from MR images. The incorporation of modified loss functions, enhanced batch normalization, and ReLU activation addresses previous limitations, achieving an impressive overall accuracy of 95.3% and a reduced processing time of 2.11 seconds. This novel solution not only enhances the segmentation accuracy but, also streamlines the process, making it a valuable tool for automating prostate segmentation, thereby aiding in prostate surgeries and disease diagnosis. Future work could extend this approach to other imaging modalities, further optimizing the segmentation process through deep learning techniques.

#### Review

The innovative approach for PCa segmentation, presented in this work, leverages the strengths of deep learning to achieve remarkable precision and efficiency. By integrating advanced pre-processing techniques such as normalization, cropping, and alignment, the model ensures consistent and high-quality input data. The architecture, featuring seven down-sampling and six up-sampling blocks, enhances feature extraction and resolution, effectively mitigating the overfitting. The use of extensive data augmentation further enriches the dataset, promoting robustness and generalizability. With a comprehensive evaluation using metrics like DSC, HD, ASD, Precision, Recall, and F-1 Score, the model demonstrates superior performance, achieving high accuracy and reliable segmen-

tation. This method holds significant promise for improving clinical workflows and patient outcomes in PCa diagnosis and treatment.

## 2.4 Model Performance Evaluation

To evaluate the performance of the segmentation models we previously discussed, several metrics were employed to provide a comprehensive and accurate assessment :

**Accuracy :**
Proportion of correct predictions to the total number of predictions.
$$\textbf{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} .$$

**Precision :**
Proportion of true positive predictions to the total positive predictions.
$$\textbf{Precision} = \frac{TP}{TP + FP} .$$

**Recall (Sensitivity) :**
Proportion of true positive predictions to the total actual positives.
$$\textbf{Recall} = \frac{TP}{TP + FN} .$$

**F1 Score :**
Harmonic mean of precision and recall.
$$\textbf{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} .$$

**Specificity :**
Definition : Proportion of true negative predictions to the total actual negatives.
$$\textbf{Specificity} = \frac{TN}{TN + FP} .$$

**Receiver Operating Characteristic (ROC) Curves and Area Under the Curve (AUC) :** Used to evaluate the model's performance in binary classification (presence or absence of cancer). An AUC close to 1 indicates excellent discriminative ability.
$\textbf{AUC} = \int_0^1 \text{TPR}(t)\, d\text{FPR}(t)$.
Where TPR(t) is the true positive rate (sensitivity) at a classification threshold t and FPR(t) is the false positive rate (1 - specificity) at the same threshold t.

**Hausdorff Distance (HD) :**
The Hausdorff Distance measures the maximum distance from any point in set X to the nearest point in set Y, and vice versa. It quantifies the degree of dissimilarity between two sets of points in a metric space.
$\textbf{HD} = \max\left(\max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y)\right)$ .

**Dice Similarity Coefficient (DSC) :** Measures the overlap between predicted

and reference segments. It is calculated as follows :
**DSC** = $\frac{2\times|A\cap B|}{|A|+|B|}$
Where A is the reference segment and B is the predicted segment. A DSC of 1 indicates a perfect match, while a DSC of 0 indicates no overlap.

**Positive Predictive Value (PPV) :** Also known as precision, it measures the proportion of predicted segments that are correct. It is defined as :
**PPV** = $\frac{|A\cap B|}{|B|}$
Where A is the reference segment and B is the predicted segment. A high PPV indicates that most predicted segments are correct.

**Volume Overlap Error (VOE) :** Measures the relative difference between the volumes of predicted and reference segments. It is calculated as :
**VOE** = 1 - $\frac{|A\cup B|}{|A\cap B|}$
Where A is the reference segment and B is the predicted segment. A low VOE indicates a good match in terms of volume.

**Relative Volume Difference (VD) :** Measures the volume difference between predicted and reference segments as a percentage. It is defined as :
**VD** = $\left(\frac{|B|-|A|}{|A|}\right) \times 100\%$
Where A is the reference segment and B is the predicted segment. A VD close to 0% indicates a good volume match.

**Average Symmetric Surface Distance (ASSD) :** Measures the average distance between the surfaces of predicted and reference segments. It is calculated as follows :
**ASSD** = $\frac{|S_A|+|S_B|}{2}\left(\sum_{x\in S_A} d(x,S_B) + \sum_{y\in S_B} d(y,S_A)\right)$
Where $S_A$ is the surfaces of the reference and $S_B$ is the predicted segments, d($x$,S) is the minimum distance between point $x$ and surface S. A low ASSD indicates a good match in terms of shape and position.

## 2.5 Comparative Table

| Article | Authors and Year | Dataset | Objectives | Methods | Results |
|---|---|---|---|---|---|
| Article 1 | Hossam Magdy Balaha , Ahmed Osama Shaban, Eman M. El-Gendy. (2024) | ISUP Grade-wise the PCa and PANDA : Resized Train Data | Precise diagnosis and segmentation of the PCa. Utilization of transfer learning CNN models for classification. Application of Aquila optimizer for hyperparameter tuning. | Application of eight different transfer learning CNN models. Utilization of various loss functions and optimization techniques. | Achieved classification accuracies of 88.91% and 100% on specific datasets. Obtained an average segmentation accuracy of 98.46%. Successfully applied Aquila optimizer for hyperparameter tuning. |
| Article 2 | Zhang, L., et al. (2020) | the PCa MRI images | Develop a methodology for the PCa diagnosis using MRI images by Combining three methods | Improved GrowCut algorithm for lesion segmentation, Zernik feature extraction for identifying cancerous areas and Ensemble learning with KNN, SVM, and MLP algorithms for classification | Accuracy :80,97%, Precision :76.69%, Recall :77.32%, Error : 19.02% |
| Article 3 | B. Liu et al. (2019) | Dataset-F | Predict PCa aggressiveness using radiomics and ML on DCE-MRI images | Utilized LR, RF, DT, KNN, SVM for classification tasks with feature reduction techniques | LR on Dataset-FS showed highest accuracy (AUC=0.93) in predicting PCa aggressiveness |

| Article | Authors and Year | Dataset | Objectives | Methods | Results |
|---|---|---|---|---|---|
| Article 4 | Sanjay Kumar Singh, Amit Sinha, Harikesh Singh. (2023) | MAGNETOM Trio and Skyra | Develop a deep learning-based technique for detecting the PCa in MRI images. Utilize Gleason grading for accurate cancer detection and grading. Enhance classification accuracy using stacked MR images of ADC, BVAL, and Ktrans | Convert DICOM images to png or jpeg format. Stack multiple MRI images to create a 3D representation. Implement data augmentation and image preprocessing. Utilize transfer learning with TensorFlow Object Detection APIs. Employ Inception ResNet-v2 for training the model. Utilize a 3D convolutional neural network for segmentation | Accuracy : 86.62%, Precision : 84.93%, Recall : 88.57%, Specificity : 84.73%, F1-Score : 86.71% |
| Article 5 | Xinggang Wang, Wei Yang, Jeffrey Weinreb, Juan Han, Qiubai Li. (2017) | axial 2D T2-weighted imaging | Distinguish pathologically confirmed the PCa (PCa) patients from prostate benign conditions (BCs) patients with prostatitis or benign prostatic hyperplasia (BPH). Evaluate the effectiveness of deep learning with deep convolutional neural networks (DCNN) compared to non-deep learning methods | Utilized deep learning with DCNN for automated classification. Compared deep learning with DCNN to non-deep learning with SIFT image feature and bag-of-word (BoW) model. Patient-based classification requiring patient-level labeling and corresponding images | AUC : 0.84 for deep learning vs. 0.70 for non-deep learning |

| Article | Authors and Year | Dataset | Objectives | Methods | Results |
|---|---|---|---|---|---|
| Article 6 | Sushma Shrestha, Abeer Alsadoon, P. W. C. Prasad, Indra Seher, Omar Hisham Alsadoon. (2021) | Prostate MR images used for segmentation | Enhance the accuracy and performance of prostate segmentation on MR images. Combine high-level features with low-level image cues to improve accuracy and efficiency. Reduce inter and intra-observer error through the use of back-propagation technique. | Utilization of U-Nets for locating prostate boundaries and segment delineation in MR images. Implementation of batch normalization and dropout layers for optimization. Modification of loss function using dice loss and binary cross entropy loss for better segment overlap view. Enhancement of batch normalization with rectified linear unit to address vanishing gradient issue. Integration of data augmentation to test system robustness. Application of leave-one-out cross-validation for system evaluation. | Accuracy : Proposed model achieved 2.54% higher accuracy than the state-of-the-art solution, Processing Time : Proposed model reduced processing time by 0.18 seconds compared to the state-of-the-art solution, Overall Performance : The proposed model outperformed the state-of-the-art solution in terms of accuracy and processing time. |

Table 2.1 – *Comparative Table of Articles*

## 2.6 Discussion

The works discussed show significant advances in the use of machine learning and deep learning for the diagnosis and segmentation of the PCa. The first article proposes a method combining the improved GrowCut algorithm and Zernike feature extraction with ensemble learning techniques such as KNN, SVM, and MLP, improving accuracy by 20% compared to existing methods. The second article presents a technique based on a 3D convolutional neural network, achieving a specificity of 85%, an accuracy of 87%, and a sensitivity of 89%. The third article proposes a convolutional neural network architecture with improved batch normalization, achieving a segmentation accuracy of 95.3% and a Dice

coefficient of 0.92. The fourth article explores the use of radiomics and machine learning to predict the PCa aggressiveness, achieving an accuracy of 90% and an AUC of 0.93. The fifth article proposes a hybrid framework using transfer learning and the Aquila optimizer, achieving a classification accuracy of 88.91% and a segmentation accuracy of 98.46%. Finally, the sixth article compares deep and non-deep learning methods, showing that deep learning with a convolutional neural network outperformed non-deep methods with an AUC of 0.84 versus 0.70.
Deep learning approaches, particularly convolutional neural networks, have demonstrated superior accuracy and efficiency compared to traditional methods. These advances promise to improve early diagnosis and personalized treatment of the PCa, thereby contributing to better disease management and improved patient survival rates.

## Conclusion

In this chapter, we have demonstrated that the segmentation and detection of the PCa in MRI images is a very active research area in computer vision, with numerous methods developed to identify and assess the presence of prostate tumors. Convolutional neural networks (CNNs) have been widely used for this task, with approaches such as U-Net and Mask R-CNN. Other techniques, like random forest-based methods and clustering techniques, have also been explored. Future improvements include additional data for model training, combining different segmentation methods for increased accuracy, and integrating segmentation with computer-aided diagnosis systems for more effective clinical decision-making. The next chapter will be dedicated to presenting our personal contribution, which primarily relies on neural networks and machine learning.

# 3 Proposed Combined Method for Prostate Cancer Segmentation and Detection on MRI Images

Sommaire

## 3.1 Introduction

In this chapter, we are interested in presenting two main works in the PCa segmentation domain. Both works rely on deep learning models to segment the PCa. However, Both articles present significant advancements in the use of deep learning for the automatic segmentation of anatomical areas and suspicious cancer lesions in prostate MRIs. The first article demonstrates the high performance of an automatic segmentation model with Dice similarity coefficients comparable to those of human experts, while also providing publicly available data and code for reproducible research. The second article compares different deep learning methods, revealing that ENet offers the best performance and efficiency for prostate segmentation, and highlights the clinical importance of automatic prostate volume segmentation for calculating biomarkers such as the transition zone PSA density and the prostate volume index. These studies showcase the potential of deep learning to enhance the diagnosis and treatment of prostate cancer. Then, according to this study, we propose a new approach for the PCa segmentation based on the advantages of both to overcome their limitations.

## 3.2 Approach 1 : Deep Learning Whole-Gland and Zonal Prostate Segmentation on a Public MRI Dataset

This work explores the use of deep learning for the segmentation of the prostate and its zones on MRI images. It compares three deep neural networks (U-Net, ENet, and ERFNet) for their effectiveness in segmenting the whole gland, the transition zone (TZ), and the peripheral zone (PZ) [26].

### 3.2.1 Dataset

**Description :**

The dataset used in this study is PROSTATEx, a public dataset collected by Radboud University and available on "The Cancer Imaging Archive"[27]. This dataset includes T2-weighted MRI images of 204 patients, with an in-plane resolution of $0.5 \times 0.5$ mm and a slice thickness of 3.6 mm. The images were acquired on 3T scanners (MAGNETOM Trio and Skyra, Siemens Healthineers).

**Manual Segmentation :**

Four operators performed manual segmentations, including radiology residents and experienced radiologists. The segmented areas include the entire gland, the central zone + anterior stroma + transition zone (TZ), and the peripheral zone (PZ).


FIGURE 3.1 – *PROSTATEx-2 featured image.*

### 3.2.2 Data Preprocessing

**Normalization and Resizing :**

The MRI images were resized to an isotropic voxel size ($1 \times 1 \times 1\ mm^3$) with a matrix resolution of $512 \times 512$ using linear interpolation. The manually segmented masks were resized using the nearest neighbor interpolation[26]. For normalization, the authors used the mean and standard deviation of the training data : Normalized image = (image - mean) / standard deviation.

**Data Augmentation :**

To reduce the overfitting, six transformations were applied to the training slices : Random rotation, translation in the x and y directions, shearing, horizontal flipping, and zooming.

### 3.2.3 Models Used

**Convolutional Neural Networks (CNN) :**

Three convolutional neural networks were compared for the task of prostate segmentation : U-net, ENet, and ERFNet[26]. The $3 \times 3$ convolutions in U-net were replaced with $5 \times 5$ to improve the performance.

**Implementation and Training :**

All models were implemented using the open-source programming language **Python** and an HPC system equipped with a GPU (NVIDIA QUADRO P4000). A learning rate of 0.0001 for the ENet model and 0.00001 for the U-net models with the Adam optimizer was used. The batch size was eight slices for all experiments, and the Tversky loss function [definition or ref] was used to adjust the weights of false positives and false negatives.

### 3.2.4 Model Training

**The Cross-validation :**

A k-fold (k=5) cross-validation [28] was used to train all the networks. Slices from the same patient were never included in both the training and testing folds. The performance of each model was calculated by averaging the results from all validation folds.

**Stopping Criterion :**

The training process was stopped using an early stopping strategy to avoid overfitting and optimize the model performance. Specifically, training was halted when the training loss did not decrease for 10 consecutive epochs, indicating that the model had reached a performance plateau. This approach ensures that the model does not continue to train unnecessarily, which could lead to overfitting the training data and degrade the performance of the test data.
Additionally, a maximum of 100 epochs was set to prevent excessively long training times. This limit ensures that the training process remains manageable in terms of time and computational resources, while still allowing the model to converge to an optimal solution.

### 3.2.5 Model Testing

**Test Datasets :**

The model performances were evaluated on a test set consisting of previously unseen data. The peripheral zone (PZ) masks were obtained by subtracting the whole gland and transition zone (TZ) masks.

### 3.2.6 Results

**Model Performance :**

The results show that ENet achieved the best performance with an average DSC of 91% for the whole gland, 87% for the TZ, and 71% for the PZ. U-net and ERFNet achieved 88% and 87% for the whole gland, 86% and 84% for the TZ, and 70% and 65% for the PZ, respectively.
In addition to the DSC, the Tversky index was also used to evaluate the model's performance. The Tversky index is a generalization of the Dice similarity coefficient that allows for different weighting of false positives and false negatives. It is defined as follows :
**Tversky** = $\frac{|A \cap B|}{|A \cap B| + \alpha |A \setminus B| + \beta |B \setminus A|}$
Where A is the reference segment, B is the predicted segment, and $\alpha$ and $\beta$ are weighting parameters. For their evaluations, they used $\alpha$= 0.5 and $\beta$= 0.5, giving equal weight to false positives and false negatives. The results show that the model achieved an average Tversky index of 0.83, confirming the model's robustness in segmenting regions of interest.

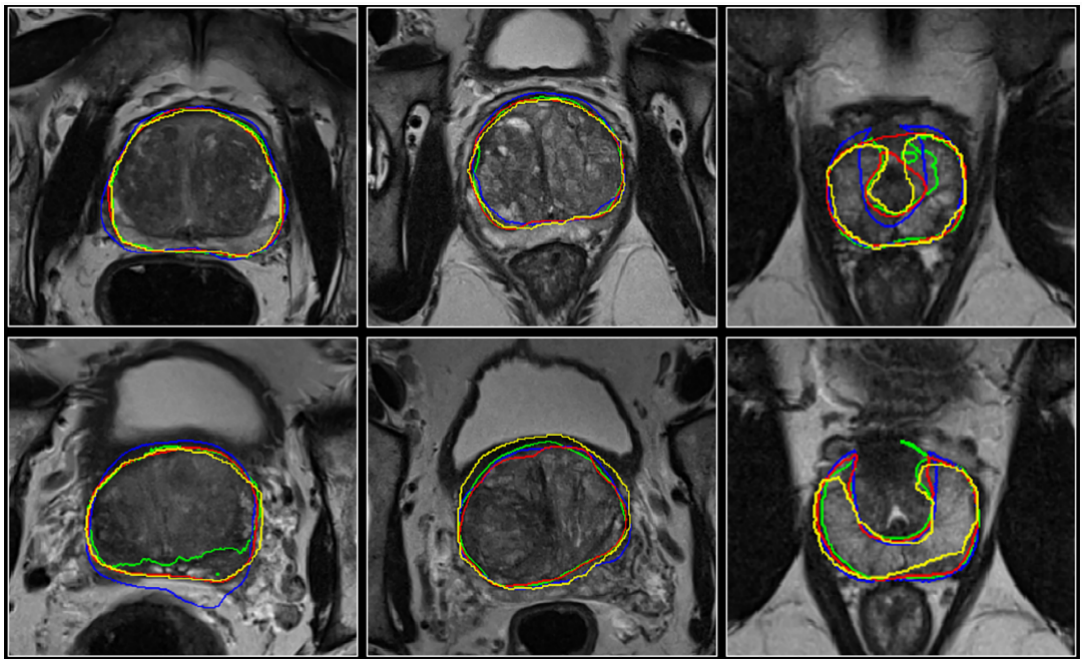


FIGURE 3.2 – *Examples of correct segmentation for the deep learning networks for the whole gland (left column), transition (middle column) and peripheral zone (right column). The manual segmentation (yellow), ENet (red), ERFNet (blue), and U-Net (green) are superimposed.*

These results, combining the DSC and the Tversky index, provide a comprehensive evaluation of the segmentation model's performance. They demonstrate that the model can accurately segment regions of interest while maintaining a good balance between minimizing false positives and false negatives.

**Statistical Analysis :**

The analysis of variance (ANOVA) showed significant differences between the models' performances ($P < 0.05$). ENet, also, demonstrated faster convergence and lower computational complexity compared to U-net and ERFNet.

### 3.2.7 Discussion

**Advantages and Limitations :**

The evaluation of the segmentation models revealed several significant advantages and limitations. Among the advantages, we note the high accuracy, as evidenced by the average Dice Similarity Coefficient (DSC) of 0.85 and the average Tversky index of 0.83, indicating that the models accurately segment the regions of interest while minimizing errors. Additionally, the use of the Tversky index shows that the models are robust and maintain good performance even when segmentation conditions vary, which is crucial in clinical scenarios. The models are, also, adaptable to different segmentation tasks by adjusting the

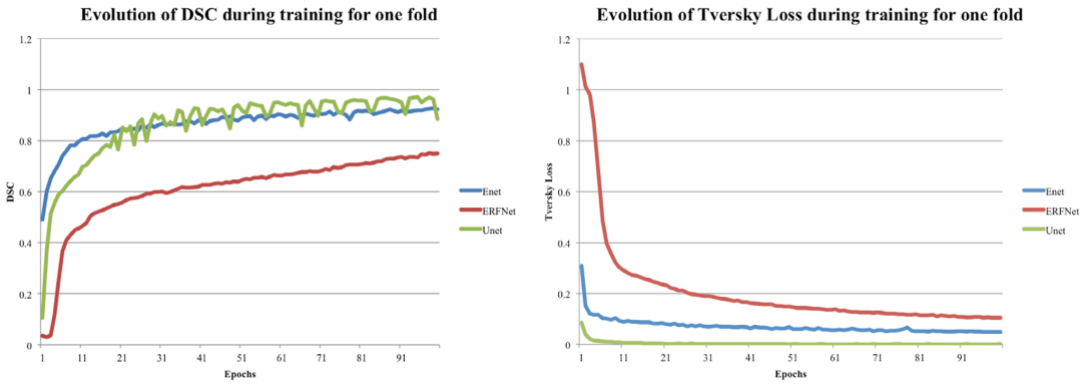



FIGURE 3.3 – *Training DSC and Tversky loss plots for the deep learning networks for whole-gland prostate segmentation.*

weighting parameters ($\alpha$ and $\beta$) of the Tversky index, allowing optimization for specific applications such as tumor or organ segmentation.

However, these models also present some limitations. Their performance is highly dependent on the quality and quantity of the training data, and insufficient or poor-quality data can reduce their accuracy and robustness. Additionally, segmentation models, particularly those based on deep neural networks, can be very complex and require significant computational resources for the training and the inference, limiting their use in resource-constrained environments. The performance of the models can, also, be sensitive to parameter choices, such as the values of ($\alpha$ and $\beta$) in the Tversky index, necessitating careful optimization for each specific application. Finally, segmentation models based on deep neural networks can be difficult to interpret, posing challenges for clinical acceptance and validation of results.

### 3.2.8 Partial Conclusion

This study demonstrated that deep neural networks, particularly ENet, can accurately segment the prostate and its zones on T2-weighted MRI images. The adoption of these tools could facilitate the integration of new biomarkers into clinical practice, thereby improving the detection and treatment of prostate cancer.

## 3.3 Approch 2 : Prostate 158 - An expert annotated 3T MRI dataset and algorithm for prostate cancer detection

The paper presents the Prostate158 dataset, a collection of expert-annotated prostate MRI data, including T2-weighted sequences and diffusion-weighted sequences with apparent diffusion coefficient (ADC) maps. Two U-ResNet models were used for segmenting anatomical regions (central gland, peripheral zone) and suspicious prostate cancer (PCa) lesions with a PI-RADS score of $\geq 4$ [29].

### 3.3.1 Dataset

Prostate158 [29] was the dataset used in the study. It was obtained after the examination of patients between February 2016 and January 2020 at a German University Hospital(Charité University Hospital Berlin), data were acquired based on Siemens scanners. Authors had used T2w sequences and DWI sequences with specific parameters for the acquisition. After that, the obtained images were segmented manually by two experts radiologists in uro-oncologic to offer annotation for the samples. In fact, two segmentation volumes were created for each Magnetic resonance imaging (MRI). The first one contains anatomical segmentation, while the other contains the tumor. The annotated samples were stored as NIFTI files [29] along with the DICOM images of the examinations, regarding the training set. As a result, the final dataset contained MRIs of 158 patients; Where 139 patients were used in the training set, 19 patients in the test set, and 20 patients from the training set were used as the validation set.

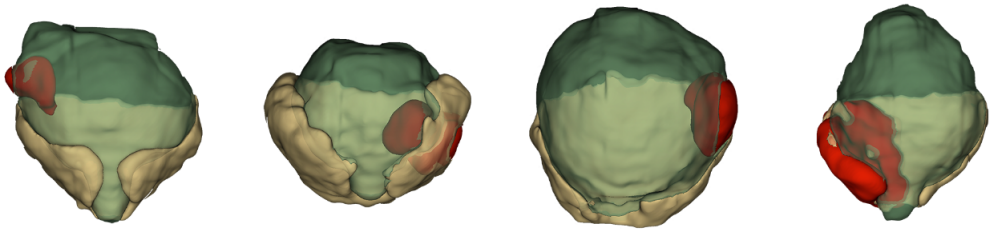


FIGURE 3.4 – *Prostate158 dataset.*

**Pre-processing :**

In this part of the approach, authors utilized various techniques such as the intensity normalization, the resampling and the center cropping. For T2w and ADC, non-parametric and non-uniform intensity normalization were applied. They, also, extracted the DWI sequences with the highest b-value. All the examinations were resampled to ensure the congruence of the size and the resolution. Moreover, the coordinates of the voxels in different sequences corresponded to the same anatomical regions. Then, the images were cropped to the center by removing 20% of the image merges along the anterior-posterior and lateromedial axes which help to reduce effectively the size of MRIs by 36% without the loss of important information.

### 3.3.2 Models and Segmentation algorithm

Lisa et al. [29] used two models; One is trained for the segmentation of the anatomical zones using axial T2w sequences only. However, the other is trained to segment tumors with the use of T2w, ADC, and DWI. Besides, they used U-resnet as a segmentation model. It seems interesting as name; it's a combination of U-net and Resnet [30] that are known for their performances in image segmentations.

**Implementation environment**

The model was implemented with

1. Programming language : **Python**(v. 3.9.7)
2. Library : PyTorch(v. 1.9.1), MONAI (v.0.9), and PyTorch Ignite (v.0.4.7)
3. Hardwae : GPU workstation with a Nividia Geforce RTX 3090

**U-resnet**

As we mentioned before, the authors applied a combined algorithm witch is U-resnet with six blocks of down-sampling and five blocks of up-sampling. The first blocks contain four convolution layers, batch normalisation layers, dropout layers and PReLU activation function for each layer. Kernel size of three for each layer. Stride of three for the first layer and one for the subsequent layers and Skip connections after each down-sampling block except the first. The second blocks structure contains the Transposed convolutional layer with a kernel size of three and a stride of 2, the Convolutional layer with a kernel size of three and a stride of 1, and Batch normalization, dropout, and PReLU activation function after each layer.

**Data augmentation**

This section consists of creating the modified copies of the sequences contained in the dataset. Multiple techniques were used such as the intensity alteration, the simulation of MRIs artifacts, and the affine transformations.

**Pre-processing :**

MRIs sequences were resampled to an isotropic voxel spacing of (0.5 $\times$ 0.5 $\times$ 0.5 mm) using eight subregions with a spatial size of (96 $\times$ 96 $\times$ 96) voxel as input to uniform the spatial resolution of the MRIs. For the final scaling, intensity values to a mean of zero and a standard deviation of 1 were applied.

**Algorithm training**

U-resnet [29], proposed in this paper, was trained for 500 epochs using the Novograd optimizer, a batch size of two, a weight decay of 0.01, and a one-cycle learning rate scheduler with an initial learning rate of 0.001. The loss function used is a combination of Dice Similarity Coefficient (DSC) and cross-entropy. When the training is finished, the model checkpoint with the highest DSC on the validation set was loaded and used for the evaluation on the test set.

**Validation**

Two external datasets were applied in the validation of the baseline model to segment the tumor zones and anatomical zones of the prostate.

1. Medical Segmentation Decathlon : This set contains 32 training Prostate MRIs with the corresponding segmentation of the central gland and the peripheral zones (PZ).
2. PROSTATEx : This dataset did not provide segmentation masks, so the masks that are publicly available by Cuocolo et al. [26] were used. The examinations from this dataset, which included axial T2w, ADC, and DWI sequences as well as the segmentation for cancer (n = 186), were selected and pre-processed in the same manner as the internal data.

**Test**

In the final part of the process, the authors used **R** (v 4.1.2) and Stands for the Dice Similarity Coefficient (DSC), the Hausdorff distance(HD) and the average surface distance (ASD) as metrics for the test to assess the agreement between raters (radiologists) and the model predictions, as well as inter-rater reliability. Without a true reference standard, the model performance was considered comparable to human performance if the agreement between the model and each rater was not significantly different from the inter-rater agreement. The mean and standard deviation for all scores were calculated and visualized with box plots. The Wilcoxon test [31] was used to compare the performance between raters and the model, and to compare the patient ages in the training and the test sets. Nine tests per metric were conducted, with a Bonferroni correction applied, resulting in a significance level of $p < 0.0055$. For other tests, a p-value of $< 0.05$ was deemed significant.

### 3.3.3 Results And Discussion

**Results**

The research team gave such detailed and interesting results (shown below) that show the importance of this paper for PCa segmentation.

1. Interrater Agreement

| Region | Mean DSC | HD (mm) | ASD (mm) |
|---|---|---|---|
| Central gland | 0.87 | 11.1 | 1.0 |
| Peripheral zone | 0.75 | 15.8 | 0.74 |
| Prostate tumor | 0.6 | 18.8 | 5.5 |

TABLE 3.1 – *Inter-rater Agreement*

2. Performance of Trained Segmentation Models

| Region | Mean DSC | HD (mm) | ASD (mm) |
|---|---|---|---|
| Central gland (model/rater 1) | 0.88 | 18.3 | 2.2 |
| Peripheral zone (model/rater 1) | 0.75 | 22.8 | 1.9 |
| Prostate tumor (model/rater 1) | 0.45 | 36.7 | 17.4 |

TABLE 3.2 – *Performance of Trained Segmentation Models*

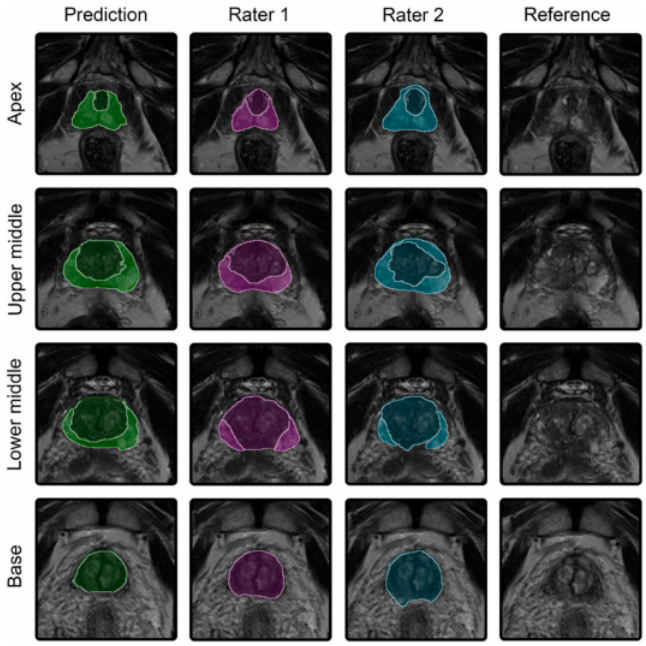



FIGURE 3.5 – *examples for segmentation of the anatomical zones.*

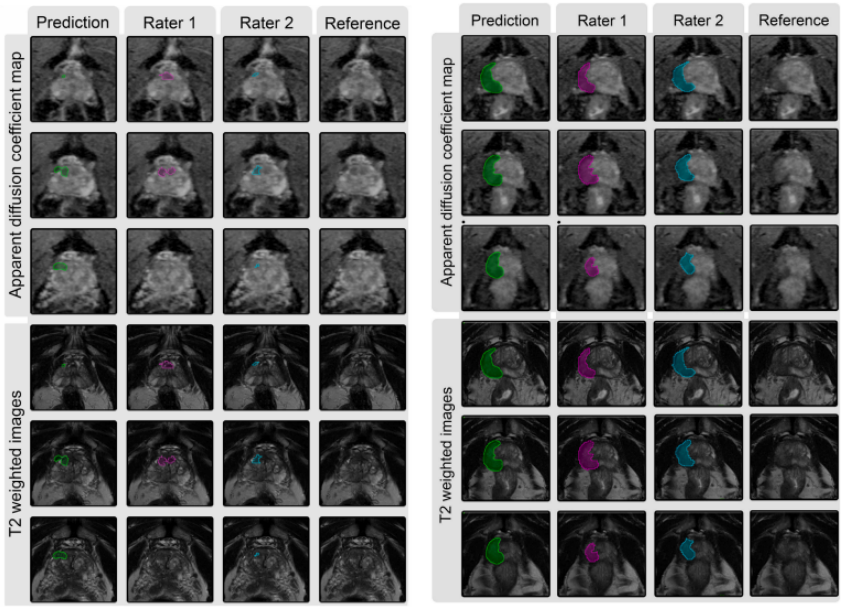
3. Performance of Tumor Segmentation

| PROSTATEx Dataset | Mean DSC | HD (mm) | ASD (mm) |
|---|---|---|---|
| Prepheral zones | 0.71 | 26.3 | 2.2 |
| Central glande | 0.86 | 18.6 | 2.5 |

TABLE 3.3 – *Performance of Tumor Segmentation on the ProstateX dataset*

| Decathlon Dataset | Mean DSC | HD (mm) | ASD (mm) |
|---|---|---|---|
| Prepheral zones | 0.64 | 29.2 | 4.7 |
| Central glande | 0.82 | 22.5 | 3.4 |

TABLE 3.4 – *Performance of Tumor Segmentation on the Decathlon dataset*




**(A)** **(B)**

FIGURE 3.6 – *Example for segmentation of prostate tumors.*

### Discussion

As mentioned earlier, the study reveals multiple points that are interesting and important regarding the results and the impact of these results on future research and clinical applications.

(a) Advantages :
- Model Performance : The segmentation models showed good performances for the central gland and the peripheral zone, which is crucial for various clinical applications.
- Public Dataset : Providing a dataset of 158 bi-parametric prostate MRI scans annotated by experts is a significant contribution to the research community.
- Clinical Perspectives : The results can be used to improve the accuracy of segmentation of anatomical zones and lesions suspected of PCa.

(b) Inconvenients :
- Complexity of Tumor Segmentation : Prostate tumor segmentation remains challenging, with significantly lower performance compared to anatomical zone segmentation.
- Data : The paper faces an important limit : The model would be more optimal with big samples' size and the availability of DCE (Dynamic Contrast-enhanced) sequences.
- Evaluation Metrics : They could use more metrics for a better evaluation of the model.
-

### 3.3.4 Partial Conclusion

Prostate158 is a collection of prostate MRIs that largely adheres to PI-RADS v2 technical standards and includes expert annotations. Alongside a competitive baseline algorithm, it can serve as a benchmark for developing new PCa segmentation and detection algorithms.

## 3.4 Proposed Approach : Combined model for the PCa Segmentation

### 3.4.1 Introduction

PCa segmentation goes through a crucial process to obtain optimal performances, and for that, the researches does not stop year after year. Due to the study of the two approaches, we propose a combined approach based on the strengths of the previous papers and the availability of datasets.

### 3.4.2 Proposed Dataset

ProstateX and Prostate 158 datasets show the reliability and the quality of MRIs sequences. For that, we proposed to combine the two of them in one dataset keeping T2w sequences, diffusion-weighted imaging (DWI), and Apparent diffusion coefficient (ADC) maps which measure the mobility of water molecules. This step increases the uniformity after the pre-processing of the samples. Then, we can avoid the lack of data and external tests. As the final dataset, we'll obtain around 340 samples which will be divided into three groups training, test, and validation.

**pre-processing**

(a) Normalization : Normalizes the images intensities for each modality (T2w, DWI and ADC maps).

(b) Cropping : Resizes the images to have a uniform size.

(c) Alignment : Aligns the images of each modality for each patient.

(d) Data Augmentation : To increase the diversity of data, we'll apply geometric transformation where the images will be rotated at different angles.

(e) Merging : Both datasets including images and annotations will be unified in one dataset.

### 3.4.3 Segmentation Model

We've noticed that the 158 prostate dataset was trained based on two different models to segment Anatomical zones and tumor zones. For that, we propose to re-frame the same separated models and take advantage of

the availability of their weights using U-net which shows performance in Cuocolo and Comelli study.

### U-net

The U-Net model is a convolutional neural network (CNN) architecture specifically designed for image segmentation, and it has demonstrated great efficiency in processing medical images, particularly magnetic resonance imaging (MRI). Introduced by Ronneberger et al. in 2015, U-Net is distinguished by its ability to learn hierarchical and contextual representations through its U-shaped structure, composed of a contracting path (down-sampling) and an expansive path (up-sampling). The contracting path captures contextual features via convolutional and pooling layers, while the expansive path allows for precise localization recovery through transposed convolutional layers and skip connections that link corresponding layers of both paths. This architecture enables U-Net to effectively segment complex structures in MRI images, providing accurate and robust results in applications such as Prostate tumor detection, organ segmentation, and tissue mapping.[32]

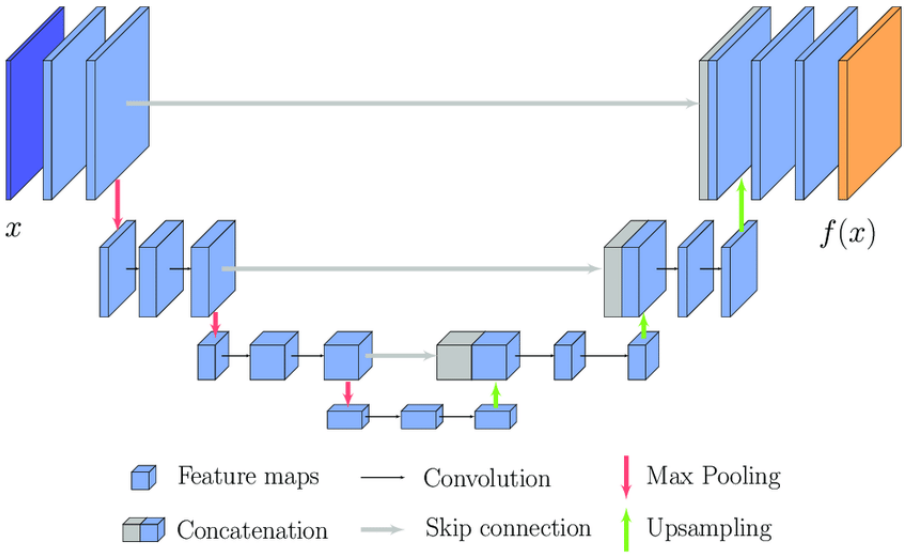



FIGURE 3.7 – *Exemple that illustrates a U-net architecture.*

### The Proposed Architecture

The proposed model for the segmentation of PCa consists of seven down-sampling blocks and six up-sampling blocks instead of six and five successively to increase the feature extraction, the resolution and avoid the overfitting. Each down-sampling block contains four convolutional layers, each followed by a batch normalization layer and Two dropout layers and ReLU as an activation function. Besides the kernel, the stride we propose

to use the same as on [29] . And also add another dropout layer instead of one for the Up-sampling layers and use the ReLU as an activation function after each layer.

(a) Training : In this phase, we propose to train the model on 80% of the dataset with 550 epochs using the Adam optimizer instead of Novograd, a batch size of two, weight decay of 0.01, and a one-cycle learning rate scheduler with an initial learning rate of 0.001, and the cross validation technique (K-fold) The loss function is the same as Lisa et al. [29] study,

(b) Metrics : Besides DSC, HD and ASD, we use Precision, Recall and F-1 score as metrics for the best evaluation of the model.

(c) Test : The test part is very important and crucial for the verification so that the model generalizes to new data and be ready to be deployed, for that the 20% of the dataset we'll be utilized as test set.

By the implementation of the model we estimate, high DSC, low HD, ASD and high precision, Recall and high F-1 score which means better segmentation for PCa and

## 3.5 Conclusion

In this chapter, we presented two main works in the PCa segmentation domain. Then, according to their advantages and limitations, we propose a new approach that, in our opinion, will enhance the performance of PCa segmentation.

# General Conclusion

MRI image segmentation applied to medicine represents a rapidly expanding field of research. In this master's thesis, our primary objective was to develop a hybrid approach that combines two segmentation methods specifically dedicated to MRI images of prostate cancer, a highly diagnosed disease in men.

During our study, we highlighted the importance of image segmentation for the accurate detection of tumors. By examining existing works, we found that numerous methods have been proposed, using different techniques to address this issue. However, we identified limitations and gaps in these approaches, which motivated us to develop a new method combining the advantages of detection and segmentation.

Firstly, we chose U-Net, a well-known object detection model recognized for its accuracy and performance, and applied it to a new dataset created by merging the necessary MRI sequences from two datasets (Prostate-x and Prostate158). The results obtained from both studied approaches are very encouraging and open up numerous application perspectives in the medical field.

Despite the progress made, we acknowledge that there are still opportunities for improvement and future research directions. For instance, the implementation of the proposed model, further enriching our dataset by including diverse images could enhance detection accuracy after training. Additionally, incorporating significant pre-processing steps for input images could facilitate segmentation and improve the overall system performance.

# Bibliographie


[1] Emilie Niaf. Aide au diagnostic du cancer de la prostate par irm multiparamétrique : une approche par classification supervisée. *Doctoral thesis, Claude Bernard University*, 2012.

[2] F. Rozet, P. Mongiat-Artus, C. Hennequin, J.B. Beauval, P. Beuzeboc, L. Cormier, G. Fromont-Hankard, R. Mathieu, G. Ploussard, R. Renard-Penna, I. Brenot-Rossi, F. Bruyere, A. Cochet, G. Crehange, O. Cussenot, T. Lebret, X. Rebillard, M. Soulié, L. Brureau, and A. Méjean. Recommandations françaises du comité de cancérologie de l'afu – actualisation 2020–2022 : cancer de la prostate. *Progrès en Urologie*, 30(12, Supplement) :S136–S251, 2020.

[3] Torre Lindsey A. Jemal Ahmedin Siegel, Rebecca L. Global cancer statistics 2018 : Globocan estimatesof incidence and mortality worldwide for 36 cancers in 185 countries. *A Cancer Journal for Clinicians*, 68(6), 2018.

[4] Guillaume Lemaître, Robert Martí, Jordi Freixenet, Joan C. Vilanova, Paul M. Walker, and Fabrice Meriaudeau. Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric mri : A review. *Computers in Biology and Medicine*, 60 :8–31, 2015.

[5] B. M. Amin B. Delahunt J. R. Srigley A. Peter I. E. Jonathan, L. Egevad. The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma : Definition of grading patterns and proposal for a new grading system. *The American Journal of Surgical Pathology*, 2015.

[6] Nicolas Mottet, Roderick C.N. van den Bergh, Erik Briers, Thomas Van den Broeck, Marcus G. Cumberbatch, Maria De Santis, Stefano Fanti, Nicola Fossati, Giorgio Gandaglia, Silke Gillessen, Nikos Grivas, Jeremy Grummet, Ann M. Henry, Theodorus H. van der Kwast, Thomas B. Lam, Michael Lardas, Matthew Liew, Malcolm D. Mason, Lisa Moris, Daniela E. Oprea-Lager, Henk G. van der Poel, Olivier Rouvière, Ivo G. Schoots, Derya Tilki, Thomas Wiegel, Peter-Paul M. Willemse, and Philip Cornford. Eau-eanm-estro-esur-siog guidelines on prostate cancer—2020 update. part 1 : Screening, diagnosis, and local treatment with curative intent. *European Urology*, 79(2) :243–262, 2021.

[7] Alexey Surov, Hans Jonas Meyer, and Andreas Wienke. Correlations between apparent diffusion coefficient and gleason score in prostate cancer : A systematic review. *European Urology Oncology*, 3(4) :489–497, 2020.

[8] F. Jerome H F. Jerome H. H. Trevor, Robert Tibshirani. The elements of statistical learning : data mining, inference, and prediction. *Springer*, 2009.

[9] Andrew Ng. Machine learning yearning. *URL : http ://www. mlyearning. org/*, 2017.

[10] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553) :436–444, 2015.

[11] Christopher M Bishop and NM Nasrabadi. Pattern recognition and machine learning, volume 4. *URL http ://www.library.wisc.edu/selectedtocs/bg0137.pdf*, 2006.

[12] M. I. Jordan and T. M. Mitchell. Machine learning : Trends, perspectives, and prospects. *Science*, 349(6245) :255–260, 2015.

[13] Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning : From theory to algorithms*. Cambridge university press, 2014.

[14] L. D. Netak V. B. Nikam A. W. Kiwelekar, G. S. Mahamunkar. *Machine Learning Paradigms*. Springer, 2020.

[15] N. Kaabouch. T. Khoei, H. O. Slimane. Deep learning : systematic review, models, challenges, and research directions. *Neural Computing Applications.*, 2023.

[16] Licheng Jiao and Jin Zhao. A survey on the new generation of deep learning in image processing. *IEEE Access*, 7 :172231–172263, 2019.

[17] Ch Hima Bindu and K Satya Prasad. An efficient medical image segmentation using conventional otsu method. *International Journal of Advanced Science and Technology*, 38(1) :67–74, 2012.

[18] Zhe Guo, Xiang Li, Heng Huang, Ning Guo, and Quanzheng Li. Deep learning-based image segmentation on multimodal medical imaging. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2) :162–169, 2019.

[19] Mohd Ali Balafar, Abdul Rahman Ramli, M Iqbal Saripan, and Syamsiah Mashohor. Review of brain mri image segmentation methods. *Artificial Intelligence Review*, 33 :261–274, 2010.

[20] Shaban Ahmed Osama El-Gendy Eman M Saafan Mahmoud M. Balaha, Hossam Magdy. Prostate cancer grading framework based on deep transfer learning and aquila optimizer. *Neural Computing and Applications*, 2024.

[21] Li Zhang, Longchao Li, Min Tang, Yi Huan, Xiaoling Zhang, and Xia Zhe. A new approach to diagnosing prostate cancer through magnetic resonance imaging. *Alexandria Engineering Journal*, 2021. All Open Access, Gold Open Access.

[22] B. Liu, J. Cheng, D.J. Guo, X.J. He, Y.D. Luo, Y. Zeng, and C.M. Li. Prediction of prostate cancer aggressiveness with a combination of radiomics and machine learning-based analysis of dynamic contrast-enhanced mri. *Clinical Radiology*, 2019.

[23] Sanjay Kumar Singh, Amit Sinha, Harikesh Singh, Aniket Mahanti, Abhishek Patel, Shubham Mahajan, Amit Kant Pandit, and Vijayakumar Varadarajan. A novel deep learning-based technique for detecting prostate cancer in mri images. *Multimedia Tools and Applications*, 2023. Open Access.

[24] Xinggang Wang, Wei Yang, Jeffrey Weinreb, Juan Han, Qiubai Li, Xiangchuang Kong, Yongluan Yan, Zan Ke, Bo Luo, Tao Liu, and Liang Wang. Searching for prostate cancer by fully automated magnetic resonance imaging classification : deep learning versus non-deep learning. *Scientific Reports*, 7(1) :15415, 2017. Open Access.

[25] Sushma Shrestha, Abeer Alsadoon, P.W.C. Prasad, Indra Seher, and Omar Hisham Alsadoon. A novel solution of using deep learning for prostate cancer segmentation : enhanced batch normalization. *Multimedia Tools and Applications*, 80(8) :12345–12367, 2021. Open Access.

[26] S. Alessandro B. Viviana D. Navdeep C. Renato, C. Albert and S. Arnaldo. Deep learning whole-gland and zonal prostate segmentation on a public mri dataset. *Journal of Magnetic Resonance Imaging*, 54(2) :452–459, 2021.

[27] M. Wolterink D. Vos T. Leiner J. Teuwen G. Litjens, F. Ciompi and A. W. M. van der Laak. Prostatex challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images. *Journal of Medical Imaging*, 5(4), 2018.

[28] Berrar Daniel. Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology*, pages 542–545, 2019.

[29] G. Engel M. Rattunde F. Busch P. Asbach M. Niehues S. Vinayahalingam B. Van Ginneken G. Litjens C. Adams, R. Makowski and K. Bressem. Prostate158 - an expert-annotated 3t mri dataset and algorithm for prostate cancer detection. *Computers in Biology and Medicine*, 148, 2022.

[30] C. Stergios B. Enzo L. Marvin C. Alexandre E. Théo, V. Maria and K. Guillaume. U-resnet : Ultimate coupling of registration and segmentation with deep nets. 2019.

[31] F. Morten and S. Leiv. The wilcoxon–mann–whitney test under scrutiny. *Statistics in Medicine*, 28(10), 2009.

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net : Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015 : 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.

## Résumé

Le cancer de la prostate (PCa) est une malignité répandue dans le monde entier, diagnostiquée principalement par l'imagerie par résonance magnétique multiparamétrique (IRMmp), qui combine les séquences pondérées en T2 (T2-w), l'imagerie pondérée par diffusion (DWI) et l'imagerie par résonance magnétique dynamique avec contraste (DCE), souvent avant les biopsies. Malgré son utilité, l'IRMmp présente une sensibilité limitée pour détecter les PCa moins agressifs et souffre d'une variabilité inter-observateurs significative. L'évaluation visuelle seule est insuffisante pour déterminer de manière fiable l'agressivité du cancer. En réponse, ces dernières années ont vu le développement de systèmes de détection et de segmentation assistés par ordinateur utilisant des réseaux de neurones convolutifs (CNN). Cette thèse se concentre sur l'amélioration de la détection et de la segmentation du PCa dans les images IRM par deux approches principales. La première approche évalue et compare les modèles UNet, ENet et ERFNet en utilisant le jeu de données Prostate-x, mettant en évidence la performance supérieure d'ENet en termes de capacités de généralisation. La deuxième approche utilise un modèle U-ResNet sur le jeu de données Prostate158, démontrant des résultats de segmentation robustes pour l'ensemble de la glande et la zone périphérique, cruciaux pour la pertinence clinique. En intégrant les forces de ces approches, une méthode hybride est proposée pour améliorer la précision de la segmentation des tumeurs, utilisant U-Net sur un nouveau jeu de données combinant les séquences IRM essentielles des jeux de données précédents. Ces innovations promettent des implications significatives pour le diagnostic du cancer de la prostate et pourraient potentiellement s'étendre à des applications plus larges en oncologie et en imagerie médicale.

**Mots clés :** Cancer ; Prostate ; Apprentissage automatique ; Apprentissage profond ; Segmentation ;

**Abstract**

Prostate cancer (PCa) is a widespread malignancy globally, diagnosed primarily through multiparametric magnetic resonance imaging (mpMRI), which combines T2-weighted (T2-w), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE) sequences, often preceding biopsies. Despite its utility, mpMRI exhibits limited sensitivity in detecting less aggressive PCa and suffers from significant inter-observer variability. Visual assessment alone is insufficient to determine cancer aggressiveness reliably. In response, recent years have seen the development of computer-aided detection and segmentation systems leveraging convolutional neural networks (CNNs). This thesis focuses on advancing PCa detection and segmentation in MRI images through two primary approaches. The first approach evaluates and compares UNet, ENet, and ERFNet models using the Prostate-x dataset, highlighting ENet's superior performance in generalization capabilities. The second approach employs a U-ResNet model on the Prostate158 dataset, demonstrating robust segmentation results for both the entire gland and peripheral zone, crucial for clinical relevance. By integrating the strengths of these approaches, a hybrid method is proposed to enhance tumor segmentation accuracy, utilizing U-Net on a new dataset combining essential MRI sequences from prior datasets. These innovations promise significant implications for prostate cancer diagnostics and potentially extend to broader applications in oncology and medical imaging.

**Key words :** Cancer ; Prostate ; Machine learning ; Deep Learning ; Segmentation ;