Deep Learning-based Detection, Segmentation of Prostate Cancer from mp-MRI Images

Yahya Bouslimi*, Takwa Ben Aïcha Gader* and Afef Kacem Echi*

* *University of Tunis, ENSIT-LaTICE, 05 Avenue Taha Hussein, Tunis, 1008, Tunisia*

Abstract

In this work, we proposed a fully automatic computer-aided diagnostic system to help radiologists detect and segment Prostate cancer (PCa) lesions from multi-parametric Magnetic resonance imaging (mp-MRIs) and predict whether the segmented lesions are clinically significant or not (benign vs. malignant lesion). Our proposed system is based on deep learning methods: residual networks (ResNet) models (ResNet50, ResNet101V2, ResNet152V2), inception networks (InceptionNetV3), U-Net, and MultiResU-Net models to classify clinically significant cancer, and properly segment the prostate lesion and determine its GGG. We tested our fully automatic detection solution on two publicly available datasets: Radboudumc and ProstateX. We achieved encouraging results where the PCa lesion classification model AUROC exceeded 98.4% accuracy. On the other hand, MultiResU-Net Model achieved 98.34% accuracy for the PCa lesion segmentation, which competes with current state-of-the-art approaches.

Key Words: Computer-aided Diagnosis, Convolutional Neural Network, Magnetic Resonance Imaging, MultiResU-Net, Prostate Cancer, U-Net.

1 Introduction

According to the latest statistics presented by the international agency for research on cancer data [1] announced in 2020, Prostate Cancer (PCa) was the fifth most common cancer in Tunisia, excluding non-melanoma skin cancer. It has registered 1186 (6.1%) of the new cancer cases total and 11.3% of new male patients. Prostate cancer is the sixth leading cause of death by cancer in Tunisian males, and it was responsible for 541 (or 4.6%) death cases. Current diagnostic techniques like digital rectal examination, PSA levels, and biopsies lack precision [2]. In most cases, those employed in screening for cancer approaches are insufficient to identify, locate, and describe it accurately, which returns in the low quality of ultrasound imaging. Additionally, the randomly selected biopsy samples may help misinterpret the prostate cancer lesion, resulting in an incorrect estimation of the tumor's aggressiveness.

In the last few years, Multi-parametric Magnetic Resonance Imaging (mp-MRI) has become a bright line of medical diagnosis research. It has become a PCa screening, diagnosis, and lesion localization benchmark. mp_MRI had demonstrated encouraging outcomes because its localization capabilities created new prospects

for essential PCa detection, diagnosis, localization, risk stratification, and clinically significant PCa staging. However, there is a considerable lack of consistency in the interpretation of mp-MRI images despite today's growing use of mp-MRI for prostate cancer screening. Given the difficulty and complexity of the process, the radiologist often relies only on visual interpretation to accomplish the work, which is not straightforward. That can also cause errors and may present crucial inter-reader variability in the diagnosis, particularly when the mp-MRI sequences contradict each other. As deep learning methods have proven their usefulness in biomedical image classification and segmentation, we propose a deep learning-based Computer-Automated Decision-making system (CAD) to aid the radiologist in diagnostics and decision-making.

CAD systems can lower the rate of incorrect diagnoses and give better results, especially when images contradict each other. Furthermore, in recent years, deep learning has achieved substantial advancements in medical image segmentation. U-Net is the most well-known Convolutional Neural network (CNN) architecture in the medical imaging industry in this approach. However, the conventional U-Net design is lacking in certain aspects. In terms of architecture, U-Net has a fixed shape and structure with a symmetric encoder-decoder architecture. This can make it less adaptable to different types of image segmentation tasks. In contrast, MultiRes-Unet has a more flexible architecture that can adapt to different scales and resolutions of images. Another area for improvement of U-Net is its relatively shallow architecture, with only a few layers in the encoder and decoder, which can make it less able. Additionally, U-Net typically uses transposed convolutional layers for upsampling in the decoder, which can introduce artifacts in the segmentation results. MultiRes-Unet uses a multi-resolution approach where the decoder path combines features from multiple scales, which can produce more accurate segmentation results. In summary, while U-Net is a powerful image segmentation model, MultiRes-Unet offers more versatility and improved accuracy by adapting to different scales and resolutions and having a deeper, more flexible architecture. MultiRes-Unet also tends to produce more accurate results by avoiding the artifacts introduced by transposed convolutional layers. MultiResU-Net was first created for segmenting skin cancer cases, and its design was updated and expanded to operate with fewer training photos and provide more accurate segmentation. It is based on the fully convolutional neural network (U-Net). This work used it to identify PCa lesions and segment mp_MRI images for precise lesion outlines.

This paper is organized as follows: In the next section, we will briefly present state-of-the-art diagnoses of PCa using deep learning techniques. In the third section, we will break down our proposed model architecture and take it to explain the data, what we did to prepare and augment it, and clarify the technique we used to achieve our objective. Afterward, we will present our results and compare them with related work in the last few years. Finally, we conclude and give some insight into our future studies.

2   State-of-the-Art

In recent years, considerable research has been done to help radiologists correctly segment, locate and identify clinically significant prostate cancer. These solutions take a new approach to detect prostate cancer and look at the issue from several perspectives. Starting with the earliest models and moving up to the most recent, this section will introduce the state-of-the-art use of deep neural networks in prostate cancer. From this works, we can mention [1] who developed XmasNet based on CNN to classify mp-MRI prostate cancer lesions on the PROSTATEx dataset. XmasNet outperformed all the traditional machine learning models for training and testing data with an AUC of 0.84%. Meanwhile, [3] used 2D U-Net with MRI slices as input, with lesion segmentation maps that encode the Gleason Grade Group (GGG). The model outperforms standard multi-class classification and multi-label ordinal regression by scoring a voxel-wise weighted kappa of 0.446% and a Dice similarity coefficient of 0.370%.On the other hand, [4] used their data, composed of a 49-patient mp-MRI database and SVM-L and logistic regression, to create a CADe to generate a cancer probability map for the radiologist. The limitation was that they used the dictionary learning method to construct their CADe, and the dictionaries were estimated from the characteristics of already extracted images. Even though their result was statistically better than all other CADe diagrams, this difference is not visible on the sample probability maps. Meanwhile, [5] proved that the mp-MRI-based segmentation scheme yields better performance than the

previous T2W-based schemes by adopting a state-of-the-art, Fully Convolutional Network (FCN) architecture with residual connections to segment prostate mp-MRI. In comparison, [6] applied a 3D deep, dense multi-path Convolutional Neural Network (CNN) based on the encoder-decoder architecture on two independent datasets to segment the prostate in MRI images. The encoder is based upon densely connected layers, and the decoder interprets the features and predicts the whole prostate volume. Also, the authors of [7] used the PROSTATEx dataset in their two-stage approach. They trained a Mask R-CNN model to segment the prostate structures automatically and then detected and classified lesions using a weakly supervised deep neural network. This work achieved an average AUC of 0.912% and 0.882% on their validation set. In the same context, [8] used Mask-RCNN for prostate segmentation and Dominant Intra-prostatic Lesion (DIL) segmentation and localization on Mp-MRI. Furthermore, [9] trained two separate CNNs to form a global CNN to create an automatic segmentation system of the prostate in T2W MRI images. Moreover, in [10], the authors fine-tuned a trained CNN for PCa detection using a Genetic Algorithm to achieve a higher AUC on their 6-channel diffusion-weighted prostate MRI dataset. This work resulted in an AUC of 0.707 on their test dataset. Also, [11] developed and implemented an automated CNN-based pipeline to detect clinically significant PCa for a given axial DWI image and each patient. Recently, [12] made a deep learning-based analysis of prostate cancer from mp-MRI. They tried various models, such as 3D U-net, 3D Res-Net, and XmasNet. The result was around 0.69% DICE score, whereas other researchers obtained a 0.83%. In the meantime, [13] proposed Pytorch V-Net deep learning framework architecture on volumetric CNN by borrowing the U-net network and dividing it into residual stages of learning to perform MRI prostate volume segmentation in a fast and accurate manner. While [14] applied nine different classification algorithms to a prostate cancer database to develop a novel prototype for PCa clinical management. Authors get a classification rate of 98.71%. Sensitivity was around 97.4%, while specificity was perfect and equal to 100% for the AdaBoost classifier. The other algorithms (Naïve Bayes, Multi-layer Perception, Simple Logistics, Nearest Neighbor, Random Committee, PART, LMT, and Random Forest) also presented great results. In [15], authors trained and tested 13 different CNN architectures (binary classification models) on Raudboundumc MRI scans of 204 patients suspected of having prostate cancer, the same one used in our research. The best model achieved an accuracy of 86.9% and an area under the ROC (AUROC) of 90.3%. In the same year (2021), for prostate segmentation, [16] proposed an upgraded 2D U-Net model with an incorporated Squeeze-and-Excitation (SE) layer using the public dataset PROMISE12. The model was based on an encoder stage that aims to extract features from the input using CONV blocks, SE layers, and max-pooling layers, and a decoder stage used to map the returned features to the original picture. Experiments revealed that the proposed model might surpass other existing approaches regarding segmentation accuracy and DSC, which reached 87% for mean DSC.

3   Materials and Methods

3.1   Data Description

In this work, we used two different datasets. For the development, testing, and validation of the proposed model, we first used the Radboudumc dataset to detect the presence of prostate cancer. After that, we used PRostateX for the segmentation and Gleason grade estimation.

3.1.1   Radboudumc Dataset

We used a data set from the Netherlands' Radboud University Medical Centre (Radboudumc) [17]. It represents a collection of studies on 204 clinically diagnosed patients with 326 prostate cancer lesions. As shown in Table 1, for each patient, there were three MRI parametric maps: the T2-weighted images (T2W), which represent a basic pulse MRI image and highlight differences in the T2 relaxation time of tissues, the amount of water molecule diffusion within the tissue is measured using the Apparent Diffusion Coefficient (ADC) and K-trans maps (MR perfusion), which is used to measure capillary permeability obtained using Dynamic

Contrast-Enhanced (DCE). All the images are in the NIfTI-1 data format. They were accompanied by descriptive Comma-Separated Values (CSV) containing the patient ID, the finding number, the 3-D coordinates of the finding, and a boolean value describing whether or not it is evidence of clinically significant prostate cancer. The only setback was that the masks were missing (not given in the dataset).

Table 1: Detailed Radboudumc Prostate Cancer Dataset Description.

Detailed Radboudumc Dataset Description.	
Modalities	MRI
Number of Patients	204
Number of findings	326
Number of images/masks before data augmentation	326
Number of images/masks after data augmentation	1956

3.1.2 ProstateX Dataset

For the development and validation of the model, we employed the ProstateX dataset, which is part of an ongoing online challenge and is freely available for download and used with no constraints. The ProstateX dataset consists of 345 mpMRI studies acquired by two different Siemens 3-Tesla scanners. The studies were divided into two groups: training and testing datasets. The training part was composed of 204 mp-MRI studies, where each study concerned one patient. Each study contains a T2-weighted image sequence (T2w), diffusion-weighted (DW) with b-values of b50, b400, and b800 s/mm2, an apparent diffusion coefficient (ADC) map (calculated from the b-values), and Ktrans (computed from dynamic contrast-enhanced-DCE-T1-weighted series). The only problem with this dataset is that it needs to contain information about lesions or GGG score information. Hopefully, the ProstateX2 challenge, which shares the same data with ProstateX, provides one to four lesion locations (points marking their position on the mp-MRI sequences) per patient; the average is 1.62 per patient. The same can be mentioned for the provided GGG scores. The lesion location had been reported under the supervision of a professional radiologist. The test dataset comprises 140 mpMRIs, which includes all the previously mentioned information except for the GGG score.

3.2 Experimentation

3.3 Execution Environment

We have trained, tested, and validated our models online using the Kaggle platform. This cloud-based solution owned by Google offers more computational power with its 15GB NVIDIA TESLA Graphics Processing Unit and 25.3 GB of SSD RAM, which helped us a lot in accelerating the training process. It has replaced Jupyter Notebook, not entirely, but in many good ways.

3.3.1 Radboudumc Dataset Preprocessing

The Radboudumc dataset was well presented. The only setback was that it was missing lesion masks and needed some cleaning and cropping. For that, some image cropping was needed to remove unnecessary data (muscle, bones, tissues) and leave only the prostate zone and its surroundings. Initially, we set as our image center the provided lesion coordinate; we took as input a $348 \times 348 \times 3$ image; we cropped it from all sides by 102 pixels, leaving only an image of the prostate zone with a shape of $144 \times 144 \times 3$. The same operation has been done on ADC and Ktrans. We needed to generate the missing masks automatically. Therefore, we took the ROI (lesion) coordinates where each lesion was present for each finding ID number from the attached CSV data, and we grew a region of 5×5 white pixels in both height and width, where the white box represents

the mask. Then, we generated a $144 \times 144 \times 3$ image, putting that box in its center, corresponding with the mp-MRI images. The final result was as follows in Figure 1, where (a) represents an image of a cropped T2W image and (b) represents a cropped mask.
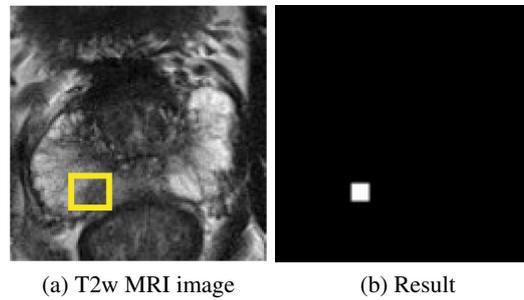

(a) T2w MRI image   (b) Result

Figure 1: *Mask Creation Resul.*

After mask generation, we had to do some data augmentation, but before that, we needed to merge our mp-MRI sequences to visualize the lesion better. Therefore, we joined the T2W images with the corresponding ADC and Ktrans MRI images. An example of the final result is shown in the figure (2), where (a) represents an image of a cropped T2W, (b) and (c) are the corresponding ADC and Ktrans, and (d) represents the merged image.
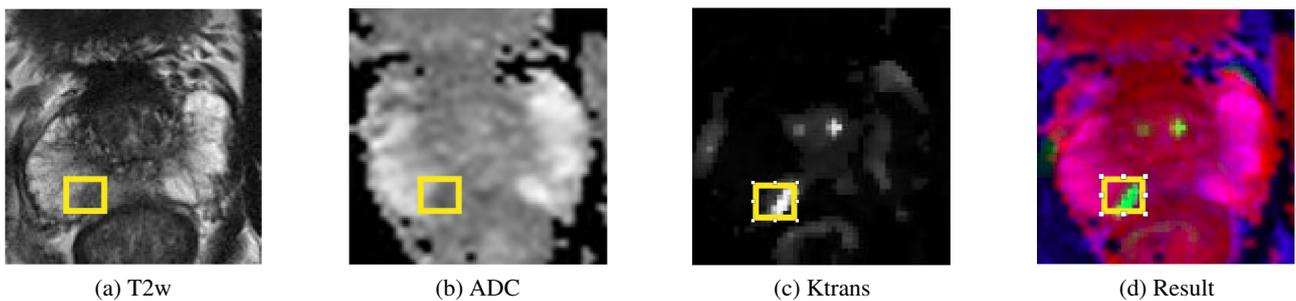

(a) T2w   (b) ADC   (c) Ktrans   (d) Result

Figure 2: *mp-MRI images merging results.*

3.3.2 ProstateX Dataset Pre-processing

For the ProstateX dataset, we needed to preprocess the mp-MRI images to clean the data and enrich the images with extra information that might be useful for model training and testing. Therefore, we cropped all T2w images around the prostate area to have dimensions of $160 \times 160 \times 24$ with a spacing of (0.5, 0.5, 3) mm. Afterward, we applied some B-Spline interpolation of third order for all image interpolation tasks, while Gaussian label interpolation was used for the segmentation masks. The ProstateX dataset comes with only the location of the lesion center. So we had to grow them from the provided coordinates; the process is the same as previously used in the Radboudumc dataset, except that it was built using a threshold level set method from the Python library SimpleITK72, as it gives better results and more apparent boundary contours. the result are shown in figure 3.

3.4 Data Augmentation and Splitting

To increase the image number, enhance our model results, and avoid over-fitting, we have done some data augmentation on the cropped and merged mp-MRI images. Therefore, the data pool was augmented by eleven-folds, and the number of images and masks increased from 326 to 3586. After we applied some image distor-
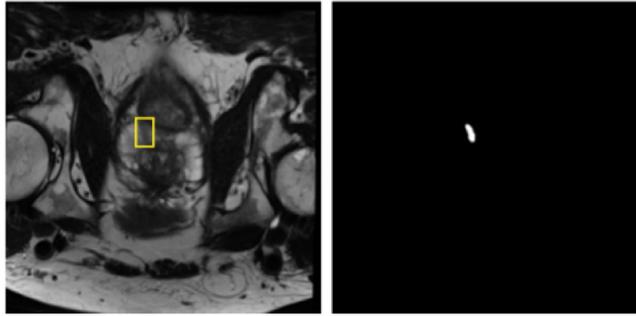
Figure 3: *PROSTATEx T2w Mask Generated Images.*

tions, liquefying, and inward zooming techniques, the zoom was random for each image and mask (less or equal to 50%). We also applied horizontal and vertical mirroring and a random rotation up to 20 degrees from their original axes. An example of the data augmentation process on an image is shown in figure (4). Figure (4)(a) and (b) represent original images of a merged T2W, ADC, and Ktrans sequence and (c) and (d) are examples of augmented images. When the dataset was ready, we randomly split it into training and testing datasets in 1 to 7 ratios. As a result, we got 3075 images or 85.72% for training and 513 or 14.28% for testing.
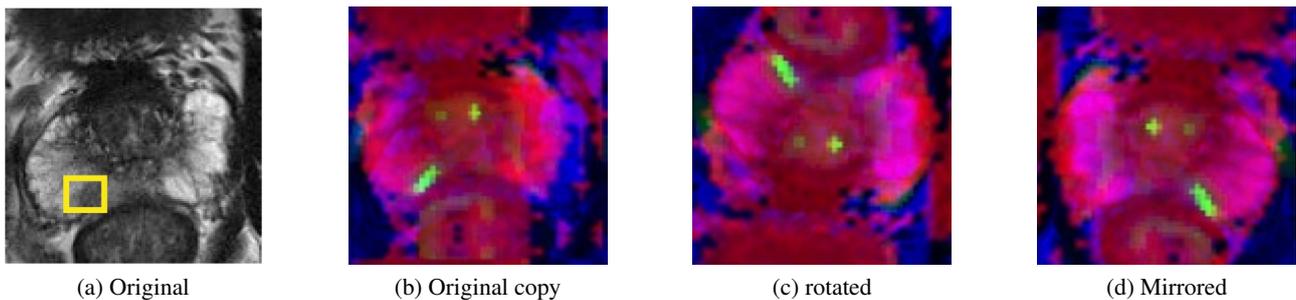

| (a) Original | (b) Original copy | (c) rotated | (d) Mirrored |

Figure 4: *Example of augmented images.*

black

As for ProstateX, we used a different approach. We employed the Python library "Batchgenerators" 14 to apply some substantial 3D data augmentation properly. This library can easily apply rigid and non-rigid transformations from scaling, rotations, and elastic deformations. As for data splitting, the latter was already divided by the ProstateX creators into train and test datasets. black

black

3.5 Proposed Models

In this study, we used two different model groups. The first group involves ResNet50, the ResNet101_V2, and the ResNet152_V2 models. We used it for the PCa classification problem to clinically determine significant lesions from the benign ones. We used the second model group, which contains the U-Net and MultyResUnet models for the semantic segmentation of PCa lesions.

3.6 Proposed Classification Pipeline

The classifier algorithm aims to tell us if there is a malignant lesion in the prostate mp-MRI given image or not. We turned to deep learning techniques to create a suitable classifier for PCa. In our case, we used three different models to see which would give the best result; we used the ResNet50, the ResNet101_V2, and the
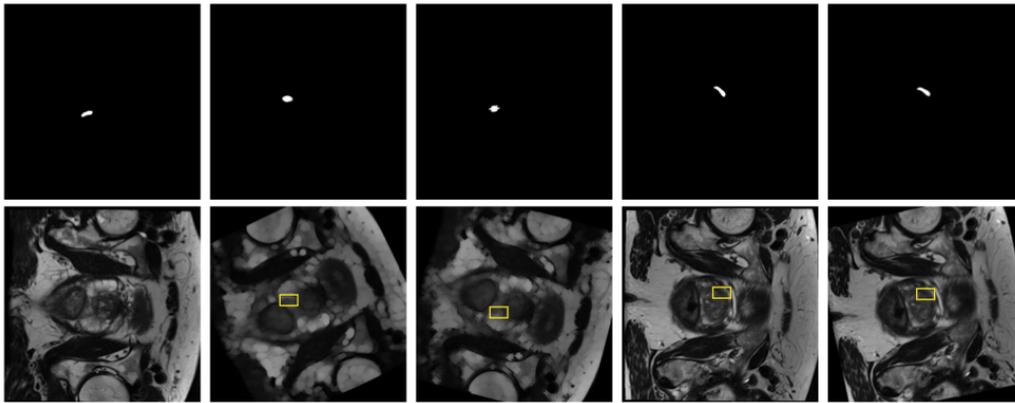
Figure 5: *Example of PROSTATEx augmented images.*

ResNet152_V2 models. The ResNet50 provided the most promising results compared to other models. We will present its architecture in detail in the next paragraph.

3.6.1   ResNET50

Currently, there are many variations of residual networks (ResNet); each has a specified number of layers; as examples, we site ResNet16, ResNet18, ResNet34, ResNet50, ResNet101_V2, ResNet110, ResNet152_V2, $ResNet$164, etc. In our work, we used pre-trained ResNet50, ResNet101_V2, and ResNet152_V2 to classify PCa lesions. ResNet50 gave the best result, a deep convolutional neural network (CNN) of 50 deep layers. Its architecture enables large numbers of convolutional layers to function efficiently. Still, adding multiple deep layers in ResNet50, ResNet101_V2, and ResNet152_V2 has some drawbacks; thus, they cause vanishing gradient problems, which results in a degradation of the output and the performance of training. The idea behind ResNet is to use shortcut connections (also known as jumping connections or skip connections) to jump over one or multiple layers, forming shortcuts between them. To solve the previously mentioned vanishing gradient problem, we use the activations of the previous layer. This process helps compress the network and expand layers by adding more feature space, which helps accelerate the learning process. Figure 6 shows the layered architecture of the ResNet50 CNN model.

3.6.2   Classification Models Hyper-parameters

As shown in Table 2, we trained a group of three ResNet CNNs to get a suitable classifier for our data; we used ResNet50, ResNet101_V2, and ResNet152_V2 as the backbones for our models. To train them, we used the Radboudumc dataset, which is adequate for the classification process and has been made and used for this purpose before. Our proposed method takes the augmented and merged mp-MRI images collected from the radboudumc dataset after preprocessing them and setting them to shapes ($143 \times 143 \times 3$). We trained the three models on 3073 images and tested on 513 images. We set their loss function to binary cross-entropy; we optimized them using the Adam gradient descent algorithms. We achieved the training by fitting each model to the augmented merged partitioned Radboudumc dataset for up to 150 epochs, and the ResNet50 gave better results and stopped earlier at epoch 52. The batch size was 52 images for all models (both for train and test), and their learning rate was set to 0.001. The only difference between them was the number of total parameters (trainable and non-trainable), as it seems to pass from one model to another.
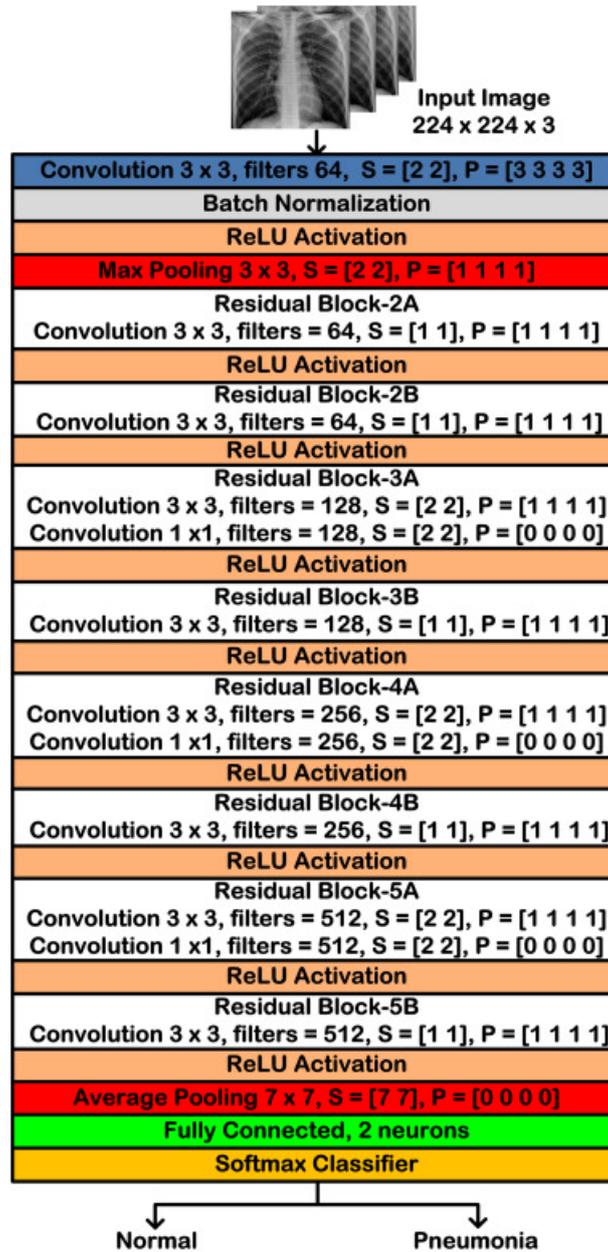
Figure 6: Used ResNet50 architecture.

3.7   Proposed Segmentation Pipeline

Our proposed method takes an mp-MRI image (DICOM format) as input and processes an automatic seamless segmentation of the prostate lesion. It reduces computational resources by decompressing and converting Dicom files into a more readable form: a sequence of 2D images in PNG format. Then, it reduces their size by half (setting all images to the shape of (128×128)). Finally, the proposed system applies our trained segmentation model (MultiResU-Net) to the input images to segment and locate the prostate lesion.

Table 2: Classifier Models Hyperparameter.

Hyperparameter	ResNet50	ResNet101_V2	ResNet152_V2
Train images	3073	3073	3073
Validation images	513	513	513
Epochs	52	150	150
Batch_size	52	52	52
Test batch_size	52	52	52
Optimizer	adam	adam	adam
Alpha	1.36	1.36	1.36
Learning_rate	0.001	0.001	0.001
Loss	binary_crossentropy	binary_crossentropy	binary_crossentropy
Total params	23,866,754	42,905,602	58,610,690
Trainable params	23,813,634	42,807,938	58,466,946
Non-trainable params	53,120	97,664	143,744

3.7.1 Modified MultiResU-Net

As shown in figure 7, the MultiResU-Net architecture is inspired by the U-Net architecture. Therefore, it takes the symmetrical shape of a 'U'; likewise, it is composed of an encoder, a decoder, and a MultiResPath that relies on them. The encoder processes the input images to extract the main features and then passes them to the decoder, which constructs the corresponding segmentation map from the extracted features. As for MultiResPaths, they have upgraded skip connections that rely on the encoder and decoder together. In MultiResU-Net, all these components have been collected and put together into MultiResBloc, which replaces the skip connection.

MultiRes Block:

The MultiRes block not only replaces the three convolutions of the U-Net but also surpasses them by learning additional spatial information. It was built based on the architecture of ResNet, and this skip connection has proven its effectiveness in biomedical picture segmentation. Jump between convolutional layers to reduce, if not eliminate, the vanishing gradient problem caused by multiple deep CNN layers, as demonstrated in figure 7. Furthermore, the MultiRes block gradually increases the number of filters to avoid memory loss, whereas in the classical equal number of filters on each layer in the basic U-Net architecture. MultiResU-Net gradually increased the number of filters from 3×3, 5×5 to 7×7 and concatenated them by the residual connection of 1×1 filter to conserve dimensions while learning and avoiding the vanishing gradient problem.
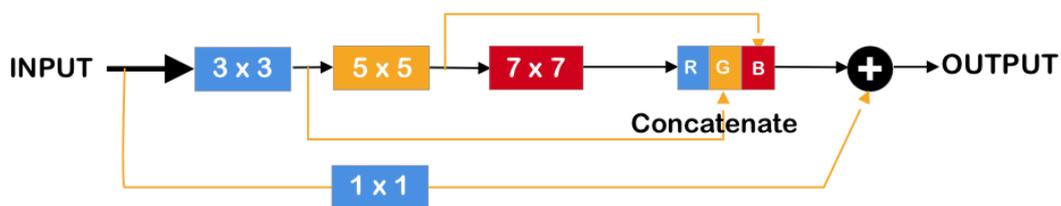

Figure 7: *MultiResU-Net block architecture.*

Skip Connection:

MultiResU-Net revolutionized U-Net by adding shortcut connections connecting convolutional layers before and after the max-pooling (skip connections). The latter permits the network to communicate spatial informa-

tion from the encoder to the decoder during the pooling operation. Various non-linear alterations are made to the skip connections. These transformations are convolutional layers combined with residual connections to make learning easier. As a result, a convolutional layer sequence is applied to the feature. Through residual linkages, these two are related. Afterward, the output will be concatenated with the decoder features. We modified our ResPath (skip connection) to use upscaling filters growing gradually from 3 x 3 filters to 7 x 7, then going back to 3 x 3 filters (see figure 8); we also implemented some random drop values going up 30% to conserve spacial information and reduce calculation time.
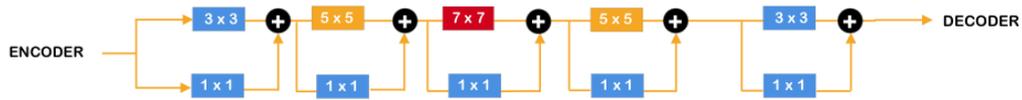

Figure 8: *Respath architecture.*

black ***Modified MultiResU-Net Architecture:***

Figure 9 shows our model architecture. It is very similar in overall architectural shape to U-Net, consisting of 13 Multires-blocks layers, each replacing the classic U-Net sequences with two convolutional layers. To make the model more robust and avoid overfitting, we added a dropout function of 30% to each Multires-block. We also removed the U-Net's casual skip connections and implied ResPaths instead. The same goes for ResPaths; each has a dropout function of 30%.
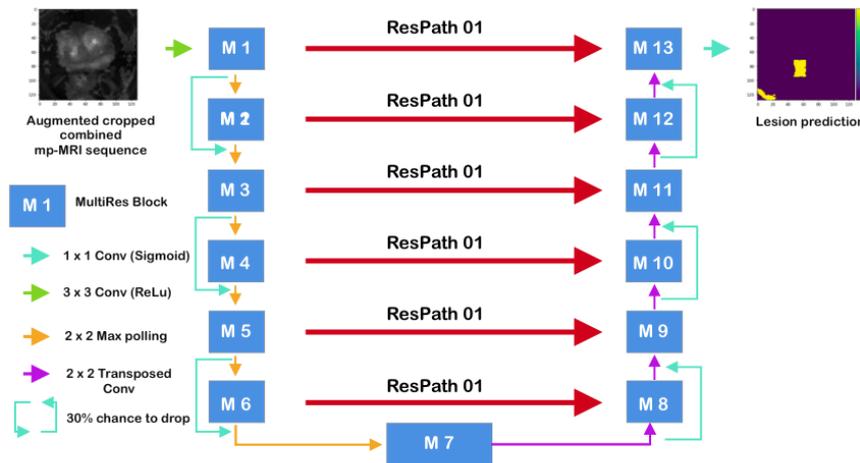

Figure 9: Modified MultiResU-Net model architecture.

3.7.2 Modified MultiResU-Net Hyperparameters

We preprocessed, cleaned, merged, and divided our mp-MRI images into train and test datasets. Then we trained our MultiResU-Net model; table 3 reveals a detailed description of the used parameters in the training process. It shows how we trained the model for 50 epochs with batch sizes equal to 30 for training and 15 for testing. Our model was optimized using the adamax optimizer with a learning rate equal to 0.0005, beta_1 and

beta_2 are equivalent to 0.8 and 0.888 successively, and epsilon of 1e-07. To effectively adjust the weights, we assigned the value 1.36 to alpha as it was the one that gave us the most promising results. Concerning the loss function, we used the binary_crossentropy one.

.

Table 3: MultiResU-Net Models Hyperparameter.

Hyperparameter	MultiResU-Net
Train images	1564
Validation images	392
Epochs	50
Batch_size	30
Test-evaluation-batch_size	15
Alpha	1.36
Optimizer	'adamax'
Learning_rate	0.0005
Beta_1	0.8
Beta_2	0.888
Epsilon	1e-07
Loss	binary_crossentropy

3.7.3   Used Metrics

This section will define the metrics we employed to assess our results and compare them to state-of-the-art ones. Usually, those metrics are decimal digit ranges from 0 to 1 or 0% to 100%, with 1 or 100% representing the best possible score. The used metrics are:

– **Accuracy**: we used this metric for both classification and segmentation. For classification, it represented the percentage of the predicted true positive from the total number of predictions. As for segmentation, it calculates the successfully classified pixel accuracy for each image in every specified class across all classes.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

black

– **Sørensen-Dice Similarity Coefficient (DSC)**: It is used to evaluate the overlap between two segmentation areas. Typically, it is calculated patient-by-patient and then averaged for all patients using the mean method.

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \tag{2}$$

– **Sensitivity and Specificity** : Sensitivity tests the capacity of the model to appropriately identify patients with Benign PCa Lesions or clinically significant lesions. As for the specificity, it tests the model's ability to appropriately identify those who do not have the illness or non-clinically significant lesions.

$$Sensetivity = \frac{TP}{TP + FN} \tag{3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

black

– **The Area Under the Receiver Operating characteristic Curve (AUC-ROC, or AUC)**: is an indicator of performance for classification issues at different threshold levels. It reveals how well the model can differ across classes.

$$P(X_1 > X_0) = P(X_1 - X_0 > 0) \tag{5}$$

where :

– X_1 is a continuous random variable giving the "score" output by our binary classifier for a randomly chosen positive sample.
– X_0 is a continuous random variable giving our binary classifier's "score" output for a randomly chosen negative sample.

4   Obtained Results

4.1   Prostate Cancer Classification

Our goal is to determine for each mp-MRI whether it is a clinically significant cancer or whether it presents a benign lesion. It is a simple binary classification; initially, we considered U-Net as it is one of the best models for processing biomedical images, from mp-MRI to ultrasound and microscopic images. However, it still suffers from many gaps in mp-MRI image classification as it was invented for semantic segmentation. Therefore, we tried using a different variation of pre-trained ResNet architectures and Inception-V3 CNNs to try and predict prostate cancer from mp-MRI. We trained our models on 150 epochs. The ResNet50 gave us the best result at epoch 62, and InceptionV3 gave us an excellent result at epoch 149. According to our results that ResNet50 outperforms both ResNet101-V2 and ResNet152-V2 in all fields. It achieved an excellent area under the receiver operating characteristic (AUROC) score (98%) compared to its siblings ResNet101V2 and ResNet152-V2, who reached successively 92.2% and 90.5%. The same goes for accuracy, precision, and sensitivity, where ResNET50 has scored them consecutively (94.7%, 83.4%, 95%, 93.5%), which surpasses ResNet101V2 results. We should also mention that this latter exceeded ResNet152-V2 by far in precision and recall (80% to 68% for precision and 92.3% to 84.61% for recall) and slightly top it for the other metrics. On the other hand, Inception-V3 exceeded ResNet50 in all metrics except the recall, which was less than expected from Inception-V3 high performance. Our best model scored successively for accuracy, precision, and AUROC, 96.3%, 98%, 98.4%, significantly better than most achieved results in the field.

In Figure 10 the images a, c, and e show the results of the AUROC curve; the three models give excellent results, above 90%, and the ResNet50 lead a nearly perfect result of 99%. We can also see the loss and accuracy curves from the images b, d, e, and f. ResNet50 presents an excellent learning rate compared to his siblings. It increases gradually with wavy movements to reach high accuracy in the learning and testing phases; the training and validation loss also present a good fit. Likewise, both other models offer a good fit; only the curves were a little noisier compared to ResNet50's curve.
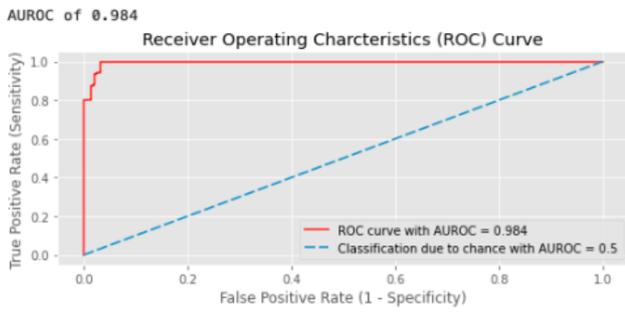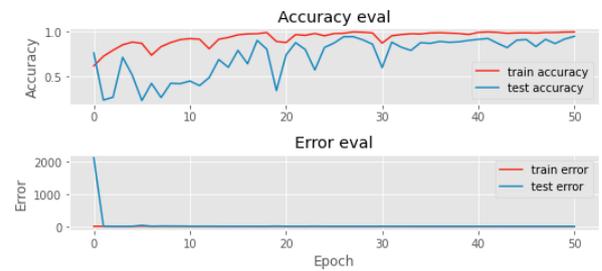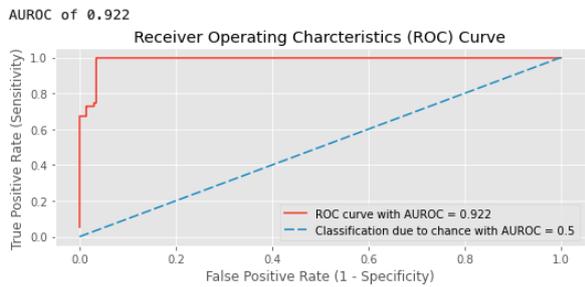
(a) InceptionV3 AUROC curve.
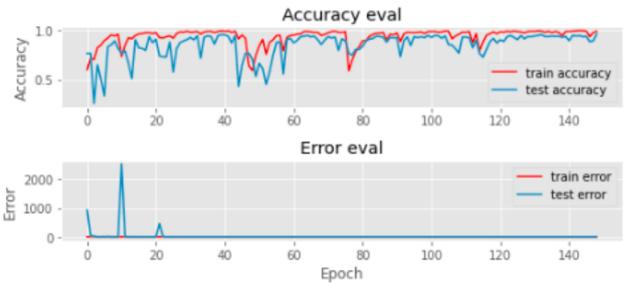
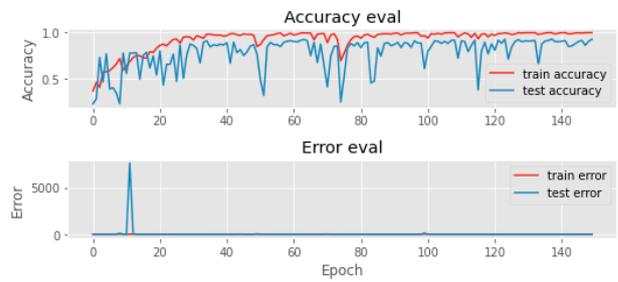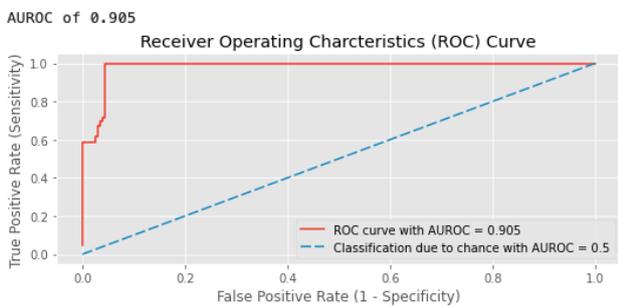(b) InceptionV3 accuracy and loss curve.
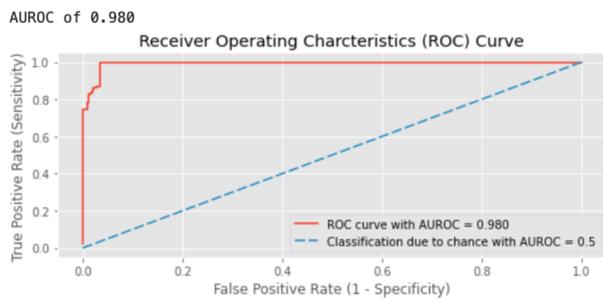
(c) ResNet50 AUROC curve.

(d) ResNet50 accuracy and loss curve.

(e) ResNet101V2 AUROC curve.

(f) ResNet101V2 accuracy and loss curve.

(g) ResNet152V2 AUROC curve.

(h) ResNet152V2 accuracy and loss curve.

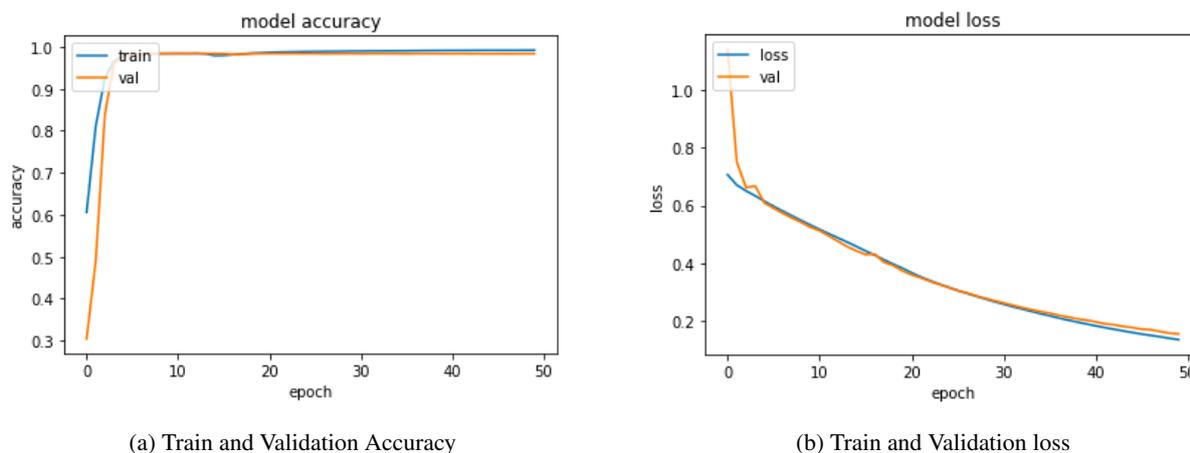Figure 10: *MultResU-Net Model Results.*

(a) Train and Validation Accuracy   (b) Train and Validation loss

Figure 11: *MultResU-Net Model Results.*

4.2 Prostate Cancer Segmentation

MultiResUNet has an excellent accuracy of 98.34%. It is followed by a nearly perfect validation loss score of 0.074%. As we mentioned, we automatically generated the masks, which made them very generic and not that specific, which gravely affected our result values. For that, it has achieved 0.74% AUC and only 59% DSK.

In Figure11 we can see the training and validation accuracy where (a) presents a nearly perfect learning rate. They increase gradually with nearly steady movements to reach high accuracy in the learning and testing phase. Likewise, in (b), the training and validation loss show a good fit with a slightly noisy decreasing curve for both the training and validation phases.

Figure (12) represents an example of a prediction of our trained MultiResU-net model, where the first image represents a gray-scale augmented T2w test image, the second image illustrates the ground truth and the third shows the predicted lesion mask (or segmentation result.)
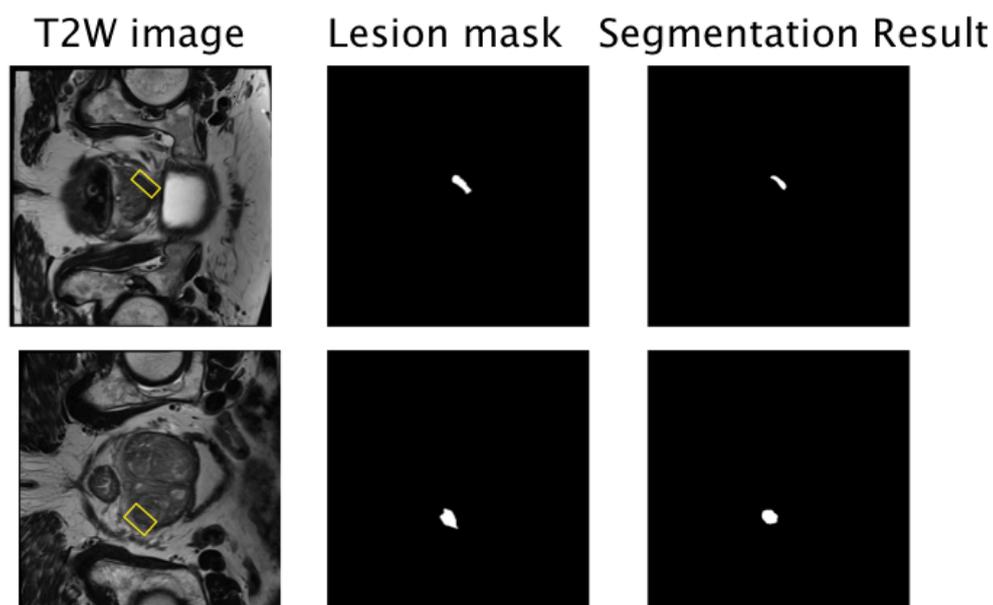

Figure 12: *Segmentation Results.*

5   Discussion

5.1   Prostate Cancer Detection

We developed the prostate cancer detection model to automate the detection of benign and clinically significant prostate lesions. Table 5 compares our model results to the state-of-the-art ones. As mentioned, we trained three models on our database and achieved remarkable results.

We trained our models (ResNet50, ResNet101V2, ResNet152V2, and Inception-V3) on 150 epochs; only ResNet50 and Inception-V3 gave us outstanding results at epochs 62 and 149 successively. They presented excellent results. From the previously created works, we can mention the one in [27], where the authors presented a perfect model that surpasses ours and used various models to detect PCa from MRI. They applied deep learning Long Short-Term Memory (LSTM) and Residual Net (ResNet101) on a publically available dataset and obtained an AUROC of 99% and 100%, which exceeds our 98.4% produced by InceptionNet-V3. However, we can see clearly from our results that other researchers like [26] tried to classify PCa lesions using Retina-U-net on two different datasets, the IVO dataset, which is a private one, and the ProstateX publicly available dataset. They obtained successively for both datasets an AUROC of 95% and 78%. The decadence of the efficiency returns to the heterogeneity of datasets and the missing slices that ProstateX presented.

The ResNet50 outperforms both the ResNet101-V2 and the ResNet152-V2 in all fields. It achieved an excellent area under the receiver operating characteristic (AUROC) score (98%) compared to its siblings ResNet101-V2 and ResNet152-V2, who reached successively 92.2% and 90.5%. The same goes for accuracy, precision, specificity, and sensitivity, where ResNET50 had scored them consecutively (94.7%, 83.4%, 95%, 93.5%), which surpasses ResNet101-V2 results. We can also mention that this latter exceeded ResNet152-V2 by far in precision and recall (80% to 68% in precision and 92.3% to 84.61% for recall) and slightly topped it for the other metrics. On the other hand, Inception-V3 exceeded ResNet50 in all metrics except the recall, which was less than expected from Inception-V3's high performance. Our best model scored successively for accuracy, precision, and AUROC, 96.3%, 98%, 98.4%, significantly better than most achieved results in the field. Meanwhile, in [21], authors created a computer-aided diagnosis of prostate cancer using multi-parametric MRI and clinical features based on PCF, which predicts a probability value that indicates whether a patient's MRI is presenting a clinically significant PCa lesion. This method achieved an AUROC of 79% and 86% at best for two different datasets. Our four models also exceeded [19], which used RMANet to classify multi-modal prostate cancer from mp-MRI via feature auto-encoder; their model achieved an AUROC of 84%. As for [18], they used various 3D CNN techniques to classify PCa lesions and achieved an accuracy of 82%. To conclude, our selected models have achieved excellent results compared to the state-of-the-art ones, especially with the InceptionNet-V3, which presented superior results to most CNN models.

5.2   Prostate Cancer Segmentation

As for PCa lesions, Many attempts have been made to auto-segment Prostate cancer using deep learning-based methods in the last few years. The research varies from segmenting the prostate normal zonal anatomy to PCa Legions, but no consensus on the combination of input sequences of mp-MRI images has been obtained.

In this work, we used MultiResU-Net to examine the prostate segmentation performance of models trained with various mp-MRI images paired with automatically generated masks. We did not segment the zonal prostate anatomy but were interested in PCa lesion segmentation. Therefore, we discovered that the three accepted models performed similarly. In segmenting the PCa lesions from the PROSTATEx dataset, the mentioned models in Table 4 produced comparable findings (DSC of 52.73, 41, 37.46 percent successively for [20], [23], [22]). The findings were typically better with the Modified-SegNet model. However, this latter was exceeded by our model performance with 59% Dsc. Previous research that employed simple T2W pictures for training also came to the same conclusion. We fused our T2W, ADC, and Ktrans MRI images as the three-channel of an RGB image to achieve this result. As for the accuracy results of the test phase, which are presented in Table 4, our used model outperforms the modified SegNet model presented in [20]. The latter achieves an accuracy

Table 4: Pca lesion classification models comparison.

Author	Year	Model	Acc*	Precision	Spec†	Rec/Sen‡	AUROC
our Work	2023	ResNet50	94.7%	83.4%	95%	93.5%	98%
	2023	ResNet101V2	92.6%	80%	92.65%	92.3%	92.2%
	2023	ResNet152V2	90.3%	68%	91.46%	84.61%	90.5%
[19]	2022	RMANet	-	-	-	-	84%
[25]	2021	3D CNN	82%	86%	-	78%	
[26]	2021	Retina-U-Net	-	-	79%	100%	95%
[26]	2021	Retina-U-Net	-	-	37.5%	100%	87%
[27]	2021	ResNet101V2	100%	-	-	-	100%
[27]	2021	LSTM	99.48%	-	100%	98.33%	99%

[a]Accuracy

[b]Specificity

[c]Recall / Sensitivity

of 96.97%, whereas our model exceeds it with an accuracy of 98.14% and a loss of 0.074%. Likewise, for the AUC metric, our model outperformed other models. We achieved a score of 74% compared to the 90%, 94%, 84% successively achieved by [20], [23], [22]). Furthermore, the ground truths are automatically generated and not presented by an expert. Therefore, the training masks were very general compared with the generated ones; they were different in shape and form and specific to the lesion, not the automatically generated mask.

Table 5: MultResU-Net Performance Comparison.

Works	Model	Accuracy	Loss	DSK	AUROC
Our Work	Modified MultiResU-Net	98.34%	0.074%	0.59%	0.74%
[26]	RetinaUNet		79%	100%	95%
[26]	RetinaUNet		37.5%	100%	87%
[20]	ModifiedSegNet	96.97%		52.73%	0.9%
[22]	UNet			37.46%	
[23]	Fully C. network			41%	-
[26]	RetinaUNet		37.5%	100%	87%
[28]	VGGNet				0.94%
[24]	XmasNet				0.84%

6   Conclusion and Future Work

This paper presented a CAD for fully automatic Pca detection, segmentation, and characterization of mp-MRI images. We used the MultiResUNet architecture to improve image segmentation tasks' performance by effectively integrating high-resolution and low-resolution information. This is particularly useful in medical imaging, where fine details are important. For that, we adjusted the Multires-Unet by adding two more MultiResblocks to have more features and make the model deeper to achieve more accurate segmentation results. We tested it on publicly available benchmark databases: the radboudumc dataset and the ProstateX dataset. Our proposed system has shown excellent efficiency in all three main processes it invented. It achieved a classification accuracy of 94.7% and an AUROC of 89%; meanwhile, for segmentation, it achieved 98.34%

accuracy with 74% AUC, which can still be intensified with a few more enhancements. The proposed models may learn additional global and local features and give better results. Currently, we are working towards the characterization of PCa Gleason grade from mp-MRI. In future work, we aim to improve the performance of the proposed model. We aim to test other benchmark databases, and we also plan to try the model for detecting and segmenting different cancer types to prove the approach's genericity.

References

[1] Liu, Saifeng and Zheng, Huaixiu and Feng, Yesu and Li, Wei. Prostate Cancer Diagnosis using Deep Learning with 3D Multiparametric MRI, 1013428, 03, 2017.

    doi: 10.1117/12.2277121. https://doi.org/10.1117/12.2277121

[2] Ibtehaz Nabil and Rahman, M Sohel,"MultiResUNet: Rethinking the U-Net architecture for multi-modal biomedical image segmentation", Elsevier: Neural Networks, vol. 121, pp.74–87, February 2020. https://doi.org/10.1016/j.neunet.2019.08.025

[3] Vente, Coen and Vos, Pieter and Hosseinzadeh,et all. Deep Learning Regression for Prostate Cancer Detection and Grading in Bi-parametric MRI, IEEE Transactions on Biomedical Engineering, Vol.PP, pages.1-1, Mai.2020, doi:10.1109/TBME.2020.2993528 https://doi.org/10.1109/TBME.2020.2993528

[4] Lehaire, Jerome,Détection et caractérisation du cancer de la prostate par images IRM 1.5T multi-paramétriquesOctober, October 2016.

[5] Yunzhi Wang, Bin Zheng, Dashan Gao, and Jiao Wang. Fully convolutional neural networks for

    prostate cancer detection using multi-parametric magnetic resonance images: an initial investigation. pages 3814-3819, 08-2018.

[6] Minh To, Quoc Vu, Baris Turkbey, Peter Choyke, and Jin Tae Kwak. Deep dense multi-path neural network for prostate segmentation in magnetic resonance imaging. International Journal of Computer Assisted Radiology and Surgery, 13, 08-2018. https://doi.org/10.1007/s11548-018-1841-4

[7] Zhiyu Liu, Wenhao Jiang, Kit-Hang Lee, Yat Long Lo, Yui Lun Ng, Qi Dou, Vince Vardhanabhuti, and Ka-Wai Kwok. A Two-Stage Approach for Automated Prostate Lesion Detection and Classification with Mask R-CNN and Weakly Supervised Deep Neural Network, pages 43-51. 10 2019. https://doi.org/10.1007/978-3-030-32486-5_6

[8] Zhenzhen Dai, Eric Carver, Chang Liu, Joon Lee, Aharon Feldman, Weiwei Zong, Milan Pantelic, Mohamed Elshaikh, Ning Wen, Segmentation of the Prostatic Gland and the Intraprostatic Lesions on Multiparametic MRI Using Mask-RCNN, Computer Vision and Pattern Recognition. 4 Apr 2019, https://doi.org/10.48550/arXiv.1904.02575

[9] Davood Karimi, Golnoosh Samei, Yanan Shao, and Septimiu Salcudean. A deep learning-based method for prostate segmentation in t2-weighted magnetic resonance imaging, 2019.

[10] Khashayar Namdar, Isha Gujrathi, Masoom Haider, and Farzad Khalvati. Evolution-based fine- tuning of cnns for prostate cancer detection, 11 2019.

[11] Yoo, Sunghwan and Gujrathi, Isha and Haider, Masoom and Khalvati, Farzad,Prostate Cancer Detection using Deep Convolutional Neural Networks, Scientific Reports, Vol.9, P.19518, December-2019. doi: 10.1038/s41598-019-55972-4. https://doi.org/10.1038/s41598-019-55972-4

[12] Pedro Neto. Deep Learning Based Analysis of Prostate Cancer from MP-MRI. PhD thesis, 05 2020.

[13] Yizhou Liu. 3d image segmentation of mri prostate based on a pytorch implementation of v-net. Journal of Physics: Conference Series, 1549:042074, 06 2020. https://doi.org/10.1088/1742-6596/1549/4/042074

[14] Jasna Nuhič and Jasmin Kevric. Prostate Cancer Detection Using Different Classification Tech- niques, pages 67-73. 01 2020. https://doi.org/10.1007/978-3-030-17971-7_10

[15] Andrew Hwang, Emlyn Evans, Ernest Mo, Jiatong Li, Sasha Jenner & Tom Linstrom. Research Report on Deep Learning Methods in Prostate Cancer, 2021

[16] B. Liu, J. Zheng, P. Chen, S. Li, C. Lai, An Improved 2D U-Net Model Integrated Squeeze-and-Excitation Layer for Prostate Cancer Segmentation. 2021 https://doi.org/10.1155/2021/8666693

[17] A. Hwang, E. Evans, E. Mo, Jiatong Li, Sasha Jenner,"Research Report on Deep Learning Methods in Prostate Cancer", Elsevier: Neural Networks, vol. 121, pp.74–87, February 2020.

[18] Patel, A., Singh, S. & Khamparia, A. Detection of prostate cancer using deep learning framework. *IOP Conference Series: Materials Science And Engineering*. 1022 (2021,1) https://doi.org/10.1088/1757-899X/1022/1/012073

[19] Bochong Li, Ryo Oka, Ping Xuan, Yuichiro Yoshimura, Toshiya Nakaguchi, Robust multi-modal prostate cancer classification via feature autoencoder and dual attention, Informatics in Medicine Unlocked, Volume 30, 2022, 100923, ISSN 2352-9148, https://doi.org/10.1016/j.imu.2022.100923. https://doi.org/10.1016/j.imu.2022.100923

[20] Lai,C.-C.;Wang,H.-K.; Wang, F.-N.; Peng, Y.-C.; Lin, T.-P.; Peng, H.-H.; Shen, S.-H. Autosegmentation of Prostate Zones and Cancer Regions from Biparametric Magnetic Resonance Images by Using Deep-Learning- Based Neural Networks. Sensors 2021, 21,2709. https://doi.org/10.3390/ s21082709 https://doi.org/10.3390/s21082709

[21] Mehta P, Antonelli M, Ahmed HU, Emberton M, Punwani S, Ourselin S. Computer-aided diagnosis of prostate cancer using multiparametric MRI and clinical features: A patient-level classification framework. Med Image Anal. 2021 Oct;73:102153. doi: 10.1016/j.media.2021.102153. Epub 2021 Jun 29. PMID: 34246848. https://doi.org/10.1016/j.media.2021.102153

[22] Vente, C.; Vos, P. Deep learning regression for prostate cancer detection and grading in bi-parametric mri. IEEE Trans. Biomed. Eng. 2021, 68, 374-383. https://doi.org/10.1109/TBME.2020.2993528

[23] Kohl, S.; Bonekamp, D.; Schlemmer, H.; Yaqubi, K.; Hohenfellner, M.; Hadaschik, B.; Radtke, J.; Maier-Hein, K. Adversarial networks for the detection of aggressive prostate cancer. arXiv 2017, arXiv:1702.08014.

[24] Liu, S.; Zheng, H.; Feng, Y.; Li, W. Prostate cancer diagnosis using deep learning with 3d multiparametric mri. In International Society for Optics and Photonics; Armato, S.G., Petrick, N.A., Eds.; SPIE: Bellingham, WA, USA, 2017; Volume 10134.

[25] Abhishek Patel and Sanjay Kumar Singh and Aditya Khamparia, Detection of Prostate Cancer Using Deep Learning Framework, IOP Conference Series: Materials Science and Engineering, 2021, 1022, 1, 012073, https://doi.org/10.1088/1757-899x/1022/1/012073.

[26] Pellicer-Valero, O. J., Jiménez, J. L. M., Gonzalez-Perez, et al. Deep Learning for fully automatic detection, segmentation, and Gleason Grade estimation of prostate cancer in multiparametric Magnetic Resonance Images. http://arxiv.org/abs/2103.12650.

[27] S. Iqbal et al., "Prostate Cancer Detection Using Deep Learning and Traditional Techniques," in IEEE Access, vol. 9, pp. 27085-27100, 2021, doi: 10.1109/ACCESS.2021.3057654.

[28] Song, Y.; Zhang, Y.D. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multipara- metric mri. J. Magn. Reson. Imaging 2018, 48, 1570–1577.