Prostate Cancer Detection, Segmentation, and Classification using Deep Neural Networks

Yahia Bouslimi, Takwa Ben Aicha Gader\textsuperscript{a} and Aref Kacem Echi\textsuperscript{b}

National Superior School of Engineering, University of Tunis, LR: LATICE, Tunisia

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

Abstract: This paper provides a fully automated computer-aided medical diagnostic system that assists radiologists in segmenting Prostate Cancer (PCa) Lesions from multi-parametric Magnetic Resonance Imaging (mp-MRIs) and predicting whether those lesions are benign or malignant. For that, our proposed approach used deep learning neural networks models such as residual networks (ResNet) and inception networks to classify clinically relevant cancer. It also used U-Net and MultiResU-Net to automatically segment the prostate lesion from mp-MRI's. We used two publicly available benchmark datasets: the Radboudumc and ProstateX. We tested our fully automatic system and obtained positive findings, with the AUROC of the PCa lesion classification model exceeding 98.4% accuracy. On the other hand, the MultiResU-Net model achieved an accuracy of 98.34% for PCa lesion segmentation.

1 INTRODUCTION

Prostate Cancer (PCa) is a common type of cancer that affects the prostate gland, a small walnut-shaped gland in the male reproductive system. It is the second most common type of cancer in men worldwide, with an estimated 1.4 million new cases diagnosed each year, which is nearly (7.3%) of the total cancer cases worldwide.\textsuperscript{(Sung, 2021)} The majority of prostate cancer cases occur in men over the age of 65, and the risk of developing the disease increases with age. The highest incidence rates of prostate cancer are found in North America, Europe, and Eastern Asia, while it is less common in the other parts of Asia and Africa.\textsuperscript{(global cancer observatory, )} Despite of the mortality rate for prostate cancer has been decreasing in many countries due to improvements in early detection and treatment. Current diagnostic methods, such as digital rectal examination, PSA levels, and biopsies, are imprecise \textsuperscript{(Nabil and Rahman, 2020).} In most situations, those used in cancer screening procedures are insufficient to effectively identify, locate, and describe it, resulting in poor ultrasound imaging quality. Furthermore, the biopsy samples still just a group of random specimen that can be inaccurate by missing their target. Multi-parametric Magnetic Resonance Imaging (mp-MRI) has emerged as a promising area of medical diagnosis research in recent years. It has shown promising results since its localization skills opened new avenues for critical PCa detection, diagnosis, localization, risk stratification, and clinically significant PCa staging. Despite the increasing use of mp-MRI for PCa screening, there remains a significant need for uniformity in the interpretation of mp-MRI images. Given the process’s difficulty and intricacy, the radiologist frequently depends solely on visual interpretation to complete the work, which is complex. This could lead to inaccuracies and significant inter-reader variability in diagnosis, especially when the mp-MRI sequences contradict one other. When images contradict each other, a Computer-Aided Decision-making System (CAD) can reduce the rate of inaccurate diagnoses and provide better results. Given the use of deep learning approaches in biomedical image classification and segmentation, we propose a deep learning-based CAD to assist radiologists in diagnoses and decision-making. It is about a MultiResU-Net, an enhanced version of the basic U-Net that allows it to run with fewer training sets while still providing accurate segmentation. We will discuss the state-of-the-art PCa diagnostics-based deep learning approaches in the upcoming section. In the third section, we will describe and pre-

\textsuperscript{a} https://orcid.org/0000-0002-3786-3649  
\textsuperscript{b} https://orcid.org/0000-0001-9219-5228

 DOI: 10.5220/0011795100003411

Copyright © 2023 by SCITEPRESS – Science and Technology Publications, Lda. Under CC license (CC BY -NC-ND 4.0)
process the used data. We will present our proposed deep learning model architectures in the fourth section. Afterward, we will report our findings and compare them to previous research. Finally, we conclude by providing some insight into our future research.

2 RELATED WORKS

In the last few recent years, there has been many research aiming to assist radiologists in appropriately segmenting, locating, and identifying clinically significant prostate cancer. We can mention (Liu et al., 2017) used XmasNet, a CNN-based classifier to categorize mp-MRI prostate cancer lesions on the PROSTATEx dataset. With an AUC of 0.84, XmasNet surpassed all typical machine learning models for training and testing data. (Vente and Vos, 2021), on the other hand, used 2D U-Net with MRI slices as input and lesion segmentation maps that encode the GGG By achieving a voxel-wise weighted kappa of 0.446% and a Dice similarity coefficient of 0.37%, the model beats traditional multi-class classification and multi-label ordinal regression. (Lehaire, 2016), on the other hand, used their data, which included a 49-patient mp-MRI database and SVM-L and logistic regression, to construct a CADe that generated a cancer probability map for the radiologist. The drawback was that they built using the dictionary learning method, and the dictionaries were approximated from the features of previously retrieved images. Although their results were statistically superior to the other CADe diagrams, this difference needed to be more evident on the sample probability maps. Meanwhile, (Yunzhi Wang and Wang, 2018) demonstrated in 2018, the mp-MRI-based segmentation scheme outperforms earlier T2W-based schemes by using a state-of-the-art Fully Convolutional Network (FCN) architecture with residual connections to segment prostate mp-MRI. (To et al., 2018), on the other hand, used a 3D deep, dense multi-path Convolutional Neural Network (CNN) based on the encoder-decoder architecture to segment the prostate in MRI images on two distinct datasets. The encoder is composed of densely connected layers, and the decoder interprets the features and predicts the total prostate volume.

The developers of (Dai Z, 2020) used the PROSTATEx dataset in their two-stage approach in 2019. They used a poorly supervised deep neural network to detect and classify lesions after training a Mask R-CNN model to segment the prostate structures automatically. On their validation set, this work attained an average AUC of 0.912% and 0.882%. (Zhenzhen Dai, 2019) employed Mask-RCNN for prostate segmentation and Dominant Intra-prostatic Lesion (DIL) segmentation and localization using Mp-MRI in the same scenario. Furthermore, (Karimi et al., 2019) trained two different CNNs to produce a global CNN to create an automatic prostate segmentation method in T2-weighted images (T2w) in prostate MRI sequence.

Furthermore, in (Stahl, 2020), the authors used a Genetic Algorithm to fine-tune a trained CNN for PCa detection to get a higher AUC on their 6-channel diffusion-weighted prostate MRI dataset. On their test dataset, this work yielded an AUC of 0.707. (Yoo et al., 2019) also created and implemented an automated CNN-based process to detect clinically significant PCa in axial DWI images for each patient.

The authors of (Neto, 2020) published a deep learning-based study of PCs via mp-MRI in 2020. They experimented with numerous models, including 3D U-net, 3D Res-Net, and XmasNet. The outcome was a DICE value of roughly 0.69, whereas other researchers received a DICE score of 0.83. Meanwhile, (Liu, 2020) suggested a 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 conduct rapid and accurate MRI prostate volume segmentation. (Nutić and Kevric, 2020) developed a novel PCa clinical management prototype using nine classification algorithms in a PCa database. The classification rate for authors is 98.71%. For the AdaBoost classifier, sensitivity was around 97.4%, while specificity was perfect and equivalent to 100%. Other algorithms (Nave Bayes, Multi-layer Perception, Simple Logistic, Nearest Neighbor, Random Committee, PART, LMT, and Random Forest) also produced excellent results.

The authors of (Andrew Hwang, 2021) trained and evaluated 13 different CNN architectures (binary classification models) on Raudboundumc MRI scans of 204 patients suspected of having prostate cancer, the same one we utilized in our study. The best model has an accuracy of 86.9% and an area under the ROC (AUROC) of 90.3% percent. (B. Liu, 2021) proposed an improved 2D U-Net model with an included Squeeze-and-Excitation (SE) layer for prostate segmentation using the public dataset PROMISE12 in the same year (2021). The model is built around an encoder stage that uses CONV blocks, SE layers, and max-pooling layers to extract features from the input and a decoder stage that maps the returned features to the original image. Experiments demonstrated that the suggested model might outperform other approaches in terms of segmentation accuracy and DSC, with a mean DSC of 87%.
We used two separate datasets in this study. We initially used the Radboudumc dataset (see Table 1) for PCa classification. Following that, we used PRostateX for PCa segmentation. The ProstateX dataset is made up of 345 mp-MRI studies, and these studies were split into training and test datasets. The training set consists of 204 mp-MRI studies, each involving one patient. Each study includes a T2w image sequence, DW with b-values of b50, b400, and b800 s/mm2, an apparent diffusion coefficient (ADC) map (derived from the b-values), and Ktrans (computed from dynamic contrast-enhanced-DCE-T1-weighted series). The ProstateX2 challenge, which uses the same data as ProstateX, should produce one to four lesion sites (points indicating their position on the mp-MRI sequences) per patient; the average is 1.62. The lesion’s location was documented under a competent radiologist’s supervision. The test set consists of 140 mp-MRIs containing all previously specified information.

Table 1: Detailed Radboudumc PCa Dataset Description.

<table>
<thead>
<tr>
<th>Modality</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>204</td>
</tr>
<tr>
<td>Findings</td>
<td>326</td>
</tr>
<tr>
<td>Images before data augmentation</td>
<td>326</td>
</tr>
<tr>
<td>Images after data augmentation</td>
<td>1956</td>
</tr>
</tbody>
</table>

The only drawback with the Radboudumc dataset was that it lacked lesion masks and needed cleaning and cropping. Image cropping was required to remove unnecessary data (muscle, bones, tissues) and leave only the prostate zone and its environs. Initially, we set the supplied lesion coordinate as our image center; we took a $348 \times 348 \times 3$ image as input; we cropped it from all sides by 102 pixels, leaving just an image of the prostate zone with a shape of $144 \times 144 \times 3$. A similar procedure was performed on ADC and Ktrans. The missing masks have to be generated automatically. As a result, we extracted the ROI (lesion) coordinates for each discovery ID number from the attached CSV data and grew a region of $5 \times 5$ white pixels in both height and width, where the white box represents the mask. Then, to coincide with the mp-MRI images, we constructed a $144 \times 144 \times 3$ image with that box in the center. The result is shown in figure 1, where (a) is an image of a cropped T2W image and (b) is a cropped mask.

We also needed to undertake some data augmentation, but first, we needed to integrate our mp-MRI sequences to visualize the lesion better. As a result, we combined the T2W images with the ADC and Ktrans.

MRI images. Figure 2 shows an example of the result, where (a) represents a cropped T2W image, (b) and (c) are the matching ADC and Ktrans, and (d) represents the merged image.

We needed to preprocess the mp-MRI images for the ProstateX dataset to clean the data and enrich the images with extra information relevant to model training and testing. As a result, we cropped all T2w photos surrounding the prostate to have dimensions of $160 \times 160 \times 24$ with a spacing of $(0.5,0.5,3)$ mm. Following that, we used third-order B-Spline interpolation for all image interpolation tasks, whereas Gaussian label interpolation was used for segmentation masks. Only the location of the lesion center is included in the ProstateX dataset. So we had to grow them from the specified coordinates; the technique is the same as in the Radboudumc dataset, except that we utilized a threshold level set method from the Python module SimpleITK72, which produces better results and more visible boundary contours. We thought about using prostate zone masks because the radiologist could need such information.
We also did some data augmentation on the cropped and merged mp-MRI pictures to increase the image quantity, improve our model performance, and avoid over-fitting. As a result, the data pool was expanded by eleven-folds, and the number of images and masks grew from 326 to 3586. After applying image distortions, liquefying, and inward zooming techniques, the zoom for each image and mask was random (less or equal to 50%). We also used horizontal and vertical mirroring and random rotations up to 20 degrees off their original axes. Figure 3 depicts an example of the data augmentation procedure on an image. 3(a) and (b) are original images of a merged T2W, ADC, and Ktrans sequence, whereas (c) and (d), are examples of augmented images. We randomly divided the dataset into training and testing datasets in 1 to 7 ratios once it was complete. Consequently, we obtained 3075 photos (or 85.72%) for training and 513 images (or 14.28%) for testing.

We took a different strategy with ProstateX. To correctly apply some significant 3D data augmentation, we applied rigid and non-rigid transformations derived from scaling, rotations, and elastic deformations. In terms of data splitting, the ProstateX designers have already divided it into train and test datasets.

4 PROPOSED DEEP MODELS

We used two separate model groups. The first includes the ResNet50, ResNet101 V2, and ResNet152 V2 models. We used it to differentiate between severe and benign lesions in the PCa classification challenge. For the PCa segmentation or semantic segmentation of PCa lesions, we used the second model group, which includes the U-Net and MultiResUnet models.

4.1 PCa Classification

For PCa classification, we attempt to determine whether or not there is a malignant lesion in the supplied prostate mp-MRI image. To construct a good classifier, we used deep learning models. We tested four models to see which would produce the most significant results: we used InceptionV3, the ResNet50, the ResNet101 V2, and the ResNet152 V2. Compared to other models, Both InceptionV3 and ResNet50 (see figure 4) produced the most promising results. The ResNet50 architecture enables large numbers of convolutional layers to function efficiently. Adding many deep layers in ResNet50, ResNet101 V2, and ResNet152 V2 has significant drawbacks; for example, they generate vanishing gradient difficulties, resulting in a degradation of output and training performance.

4.2 PCa Segmentation

For PCa segmentation, we used an mp-MRI image (DICOM format) as input and performed automatic seamless prostate lesion segmentation. We saved computer resources by decompressing and transforming DICOM data into a more legible format: a series of two-dimensional images in PNG format. Then we cut their size by half (making all photos $128 \times 128$ in size). Finally, the proposed system uses the trained segmentation model (MultiResU-Net) to segment and locate the prostate lesion in the input images.

The MultiResU-Net architecture is inspired by the U-Net architecture, as seen in figure 5. As a result, it has the symmetrical shape of a ‘U’, and it is made up of an encoder, a decoder, and a MultiResPath that relies on them. The encoder extracts the main features from the input images before passing them to the decoder, which creates the corresponding segmentation map from the extracted features. As for MultiResPaths, they have upgraded skip connections that rely on the encoder and decoder together. All these components have been collected and combined in MultiResU-Net to form MultiResBloc, which substitutes the skip connection.

The MultiRes block not only substitutes the U-Net’s convolutions but also outperforms them by learning additional spatial information. It was constructed on the ResNet architecture. This skip link is helpful in biomedical image segmentation. Jump between convolutional layers to mitigate, if not elimi-
nate, the vanishing gradient problem induced by numerous deep CNN layers, as seen in figure 5. Furthermore, the MultiRes block gradually increases the number of filters to reduce memory loss, whereas the standard U-Net architecture uses an equal number of filters on each layer. MultiResU-Net gradually raised the number of filters from $3 \times 3$, $5 \times 5$ to $7 \times 7$ and concatenated them using the residual connection of $1 \times 1$ filter to conserve dimensions while learning and avoiding the vanishing gradient problem.

MultiResU-Net transformed U-Net by introducing shortcut connections between convolutional layers before and after max-pooling (skip connections). During the pooling procedure, the latter allows the network to send spatial information from the encoder to the decoder. The skip connections are subjected to a number of non-linear modifications. To facilitate learning, these transformations are subjected to a number of non-linear modifications. To facilitate learning, these transformations are convolutional layers paired with residual connections.

As a result, the feature is subjected to a convolutional layer sequence. These two are linked via residual links. The output will then be concatenated with the decoder features. Our model architecture is depicted in 6. Its overall architectural shape is highly similar to that of U-Net. It comprises 13 Multires-Blocks layers, replacing the traditional U-Net sequences with two convolutional layers. To make the model more resilient and reduce over-fitting, we added a 30% dropout function to each multires-block. Additionally, we eliminated the skip connections of the U-casual Net in favor of ResPaths. The same is true for ResPaths, which have a 30% dropout function. Table 2 displays a detailed description of the parameters used during the training process.

### 5 EXPERIMENTATION AND RESULTS DISCUSSION

#### 5.1 PCa Classification

We aim to assess whether each mp-MRI shows a clinically significant cancer or a benign lesion. It is a straightforward binary classification; initially, we

![Figure 5: MultiResUNet block architecture.](image_url)
Table 3: Used classification models results.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Specificity</th>
<th>Recall/Sensitivity</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>InceptionV3</td>
<td>96.3%</td>
<td>90.8%</td>
<td>98%</td>
<td>92.5%</td>
<td>98.4%</td>
</tr>
<tr>
<td>ResNet50</td>
<td>94.7%</td>
<td>83.4%</td>
<td>95%</td>
<td>93.5%</td>
<td>98%</td>
</tr>
<tr>
<td>ResNet101V2</td>
<td>92.6%</td>
<td>80%</td>
<td>92.65%</td>
<td>92.3%</td>
<td>92.2%</td>
</tr>
<tr>
<td>ResNet152V2</td>
<td>90.3%</td>
<td>68%</td>
<td>91.46%</td>
<td>84.61%</td>
<td>90.5%</td>
</tr>
</tbody>
</table>

Figure 6: MultiResU-Net model architecture (Nabil and Rahman, 2020).

picked U-Net because it is one of the finest models for processing biomedical images ranging from mp-MRI to ultrasound and microscopic images. However, because it was designed for semantic segmentation, it still has numerous limitations in mp-MRI image categorization. As a result, we attempted to predict PCa from mp-MRI using a distinct variety of pre-trained ResNet architectures and Inception-V3 CNNs. Our models were trained across 150 epochs.

The ResNet50 performed best at epoch 62, while InceptionV3 performed admirably at epoch 149. ResNet50 beats both ResNet101-V2 and ResNet152-V2 in all fields, as shown in Table 3. It has a higher area under the receiver operating characteristic (AUROC) score (98%) than its siblings ResNet101V2 and ResNet152-V2, which had 92.2% and 90.5%, respectively. The same is true for accuracy, precision, and sensitivity, which ResNet50 has scored sequentially (94.7%, 83.44%, 95.4%, and 93.5%), outperforming ResNet101V2 scores. It is also worth noting that this latter outperformed ResNet152-V2 in precision and recall (80% vs. 68% for precision and 92.3 vs. 84.61%) and slightly beat it in the other metrics. Inception-V3, on the other hand, surpassed ResNet50 in all metrics except recall, which was lower than expected from Inception-outstanding V3’s performance. Our top model got 96.3%, 98%, and 98.4% for accuracy, precision, and AUROC, which is much higher than most achieved outcomes in the field.

The table above show the AUROC curves results. The four models produce good results, above 90%, and the InceptionV3 gives near perfect result of 98.4%. They also gave a noisy loss and accuracy curves in the beginning that leads to be stedy over time. ResNet50 outperforms his siblings in the learning rate, gradually increasing with wavy movements to achieve high accuracy in the training and testing phases. In our case, InceptionV3 surpassed ResNet50 by a small percentage and shows a good fit of training and validation loss curves. Similarly, the other models provide a good fit; however, the curves were slightly noisier than ResNet50’s.

5.2 PCa Segmentation

Recently, many attempts have been undertaken to auto-segment Prostate cancer utilizing deep learning-based approaches. The research ranges from segmenting the prostate’s normal zonal structure to PCa lesions, but an agreement has yet to be reached on the combination of input sequences of mp-MRI images. In this study, we used MultiResU-Net to automatically generated Lesion masks from mp-MRI images. We found that the three recognized models behaved similarly. The indicated models in Table 5 showed equivalent results when segmenting the PCa lesions from the PROSTATEx dataset (DSC of 52.73, 41, 37.46 percent consecutively for (Lai and Shen., 2021), (Vente and Vos, 2021)). The Modified-SegNet model produced consistently improved results. Our model performance, however, outperformed this with a Dsc of 59%. Previous studies that used simple T2W images for training came to the same conclusion. To reach this result, we combined our T2W, ADC, and Ktrans MRI images as three channels of an RGB image. Regarding test accuracy findings, shown in Table 5, our employed model beats the modified SegNet model presented in (Lai and Shen., 2021). The latter has an accuracy of 96.97%, whereas our model has an accuracy of 98.14% and a loss of 0.074%. Similarly, our model beat other models in terms of AUC. We received a score of 74%, compared to the 90%, 94%, and 84% obtained by (Lai and Shen., 2021) ((Song, 2018), (Vente and Vos, 2021)). Furthermore, the ground truths are generated automatically rather than delivered by an expert. As a result, the training masks were more general than the generated ones; they differed in shape and form and were particular to the lesion, unlike the automatically created mask.
Table 4: Prostate cancer classification curve results.

<table>
<thead>
<tr>
<th>Model</th>
<th>AUROC Curve</th>
<th>ACCURACY/LOSS Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>InceptionV3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUROC of 0.930</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: MultResU-Net Performance Comparison.

<table>
<thead>
<tr>
<th>Author</th>
<th>Model</th>
<th>Accuracy</th>
<th>Loss</th>
<th>DSK</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our work</td>
<td>MultiResU-Net</td>
<td>98.34%</td>
<td>0.074%</td>
<td>59%</td>
<td>74%</td>
</tr>
<tr>
<td>(P.Valero and et al., 2021)</td>
<td>Retina-U-Net</td>
<td>100%</td>
<td>-</td>
<td>79%</td>
<td>95%</td>
</tr>
<tr>
<td>(P.Valero and et al., 2021)</td>
<td>Retina-U-Net</td>
<td>100%</td>
<td>-</td>
<td>37.5%</td>
<td>87%</td>
</tr>
<tr>
<td>(Lai and Shen., 2021)</td>
<td>Modified-SegNet</td>
<td>96.97%</td>
<td>-</td>
<td>52.73%</td>
<td>90%</td>
</tr>
<tr>
<td>(Vente and Vos, 2021)</td>
<td>U-Net</td>
<td>-</td>
<td>-</td>
<td>37.46%</td>
<td>-</td>
</tr>
<tr>
<td>(P.Valero and et al., 2021)</td>
<td>Retina-U-Net</td>
<td>-</td>
<td>-</td>
<td>100%</td>
<td>87%</td>
</tr>
<tr>
<td>(Song, 2018)</td>
<td>VGG-Net</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94%</td>
</tr>
<tr>
<td>(Liu et al., 2017)</td>
<td>XmasNet</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>84%</td>
</tr>
</tbody>
</table>
6 CONCLUSION

A CAD for fully automatic Pca segmentation, and classification of mp-MRI images was reported in this study. We tested it using two publicly available benchmark databases: the RadboudUMC dataset and the ProstateX dataset. Our proposed system has demonstrated high efficiency in all the key operations for which it was designed. It achieved a classification accuracy of 96.3% and an AUROC of 90% with Inceptionv3 model. Likewise, for segmentation using MultiResUnet, it achieved 98.34% percent accuracy with a 74% AUC, which can still be improved with a few more tweaks. The proposed models could learn more global and local features and produce better results. We are now working on determining the PCa Gleason grade using mp-MRI. In future work, we aim to improve the performance of the proposed model. We also intend to test the model on other benchmark databases and use it to detect and segment different cancer kinds to demonstrate the approach’s generalizability.

REFERENCES

global cancer observatory. Cancer today.
Lehaire, J. (2016). Détecton et caractérisation du cancer de la prostate par images irm 1.5t multiparamétriques.