

Article

Automated Diagnosis of Prostate Cancer Using mpMRI Images: A Deep Learning Approach for Clinical Decision Support

Anil B. Gavade ¹, Rajendra Nerli ², Neel Kanwal ³, Priyanka A. Gavade ⁴, Shridhar Sunilkumar Pol ¹ and Syed Tahir Hussain Rizvi ^{3,*}

¹ Department of E&C, KLS Gogte Institute of Technology, Belagavi 590008, India; abgavade@git.edu (A.B.G.)

² Department of Urology, JN Medical College, Belagavi 590008, India

³ Department of Electrical Engineering and Computer Science, University of Stavanger, 4036 Stavanger, Norway

⁴ Department of Computer Science and Engineering, KLE Society's Dr. M. S. Sheshgiri College of Engineering and Technology, Belagavi 590008, India

* Correspondence: tahir.rizvi@uis.no

Abstract: Prostate cancer (PCa) is a significant health concern for men worldwide, where early detection and effective diagnosis can be crucial for successful treatment. Multiparametric magnetic resonance imaging (mpMRI) has evolved into a significant imaging modality in this regard, which provides detailed images of the anatomy and tissue characteristics of the prostate gland. However, interpreting mpMRI images can be challenging for humans due to the wide range of appearances and features of PCa, which can be subtle and difficult to distinguish from normal prostate tissue. Deep learning (DL) approaches can be beneficial in this regard by automatically differentiating relevant features and providing an automated diagnosis of PCa. DL models can assist the existing clinical decision support system by saving a physician's time in localizing regions of interest (ROIs) and help in providing better patient care. In this paper, contemporary DL models are used to create a pipeline for the segmentation and classification of mpMRI images. Our DL approach follows two steps: a U-Net architecture for segmenting ROI in the first stage and a long short-term memory (LSTM) network for classifying the ROI as either cancerous or non-cancerous. We trained our DL models on the I2CVB (Initiative for Collaborative Computer Vision Benchmarking) dataset and conducted a thorough comparison with our experimental setup. Our proposed DL approach, with simpler architectures and training strategy using a single dataset, outperforms existing techniques in the literature. Results demonstrate that the proposed approach can detect PCa disease with high precision and also has a high potential to improve clinical assessment.

Keywords: computer-aided diagnosis; classification; deep learning; image segmentation; multiparametric MRI; prostate cancer



Citation: Gavade, A.B.; Nerli, R.; Kanwal, N.; Gavade, P.A.; Pol, S.S.; Rizvi, S.T.H. Automated Diagnosis of Prostate Cancer Using mpMRI Images: A Deep Learning Approach for Clinical Decision Support. *Computers* **2023**, *12*, 152. <https://doi.org/10.3390/computers12080152>

Academic Editor: Hersh Sagreiya
Sagreiya

Received: 25 June 2023

Revised: 24 July 2023

Accepted: 26 July 2023

Published: 28 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Prostate cancer (PCa) is the most commonly diagnosed male malignancy, with over 1.4 million new cases and 375,000 deaths in 2020 alone [1]. It is ranked as the fifth leading cause of death due to cancer in men and the most frequently diagnosed cancer in over 50% of countries, worldwide. PCa starts in the small walnut-shaped prostate gland below the bladder and in front of the rectum. If not diagnosed in the early stages, the fatality rate of PCa can be significant, with a 20-year actuarial cumulative likelihood of death from prostate cancer [2]. Within clinical settings, the diagnosis of prostate cancer primarily relies on prostate-specific antigen (PSA) testing, prostate tissue biopsies, and CT/MRI scans. These clinical procedures demand significant time and expertise from radiologists, pathologists, and physicians, as they carefully observe and assign a grade or stage. Later, the treatment options are considered based on the stage, severity of cancer, and other

factors. Unfortunately, the routine diagnostic process requires human intervention and results in variability in the outcomes, which may lead to delayed or wrong diagnosis.

Deep learning (DL) approaches are becoming popular in biomedical image analysis [3,4]. DL approaches are promising due to their potential to identify complex patterns and hidden representations in the data. DL models, such as deep neural networks (DNNs) have shown great potential in accurately detecting and classifying medical images, which can help improve the accuracy of diagnosis and treatment planning [4,5]. Computer-aided (CAD) systems using DL approaches can provide decision support to doctors by localizing regions of interest and saving their time in collecting second opinions [6,7]. However, developing effective DL models with high precision is complex, as it requires the availability of significant medical data and clinical labels. Among various imaging modalities, multiparametric magnetic resonance imaging (mpMRI) is widely utilized due to its high sensitivity in detecting PCa and its ability to offer superior anatomical imaging of the prostate gland [8,9]. This superiority stems from the advanced spatial and contrast resolution of mpMRI, surpassing that of other imaging techniques. This comprehensive imaging technique of mpMRI is a valuable source of data for DL models, as it provides detailed information about the size, location, and aggressiveness of tumors within the prostate gland and surrounding tissues. This information is helpful in guiding treatment decisions and improving outcomes for patients. DNNs have already been developed to segment and classify PCa using other imaging modalities (i.e., histopathology) [10,11]; unfortunately, preparing histological medical images is expensive and time-consuming. Moreover, the problems of stain variation and artifacts in histological images require additional steps in preprocessing [7]. DNN models can identify specific features in mpMRI images associated with PCa, such as tumor volume and location, and classify cancer with high precision. This will allow the development of CAD systems for the more precise localization and staging of PCa, as well as more personalized treatment plans based on the characteristics of cancer [12,13].

Figure 1 illustrates an overview of our proposed DL approach, which uses DL models to segment first and then classify the region-of-interest (ROI) for mpMRI images. In this work, we combine two DL models, (i) U-Net architecture for segmentation, and (ii) the long short-term memory (LSTM) model for classification to benefit from both models, providing a reliable CAD system with improved performance. The proposed pipeline leverages the power of convolutional and recurrent neural networks to capture intricate patterns and spatial dependencies, yielding promising results and improving clinical decision support. Our training strategy uses data augmentation and transfer learning techniques to overcome the limitation of labeled data and provides high precision. Furthermore, we compare the results against using only recurrent neural network (RNN), ResNet-50 DNN and existing methods in literature, where the proposed approach outperforms others, underscoring effectiveness in the classification and segmentation of PCa.

The remainder of the paper is organized as follows: Section 2 presented related work on DL models (U-Net and LSTM) for CAD systems on PCa diagnosis. Section 3 details the method of building the pipeline, training setup, and the dataset used for training. We provide results and discuss them in Section 4. We conclude in Section 5 and finally provide limitations and future direction for this work in Section 6.

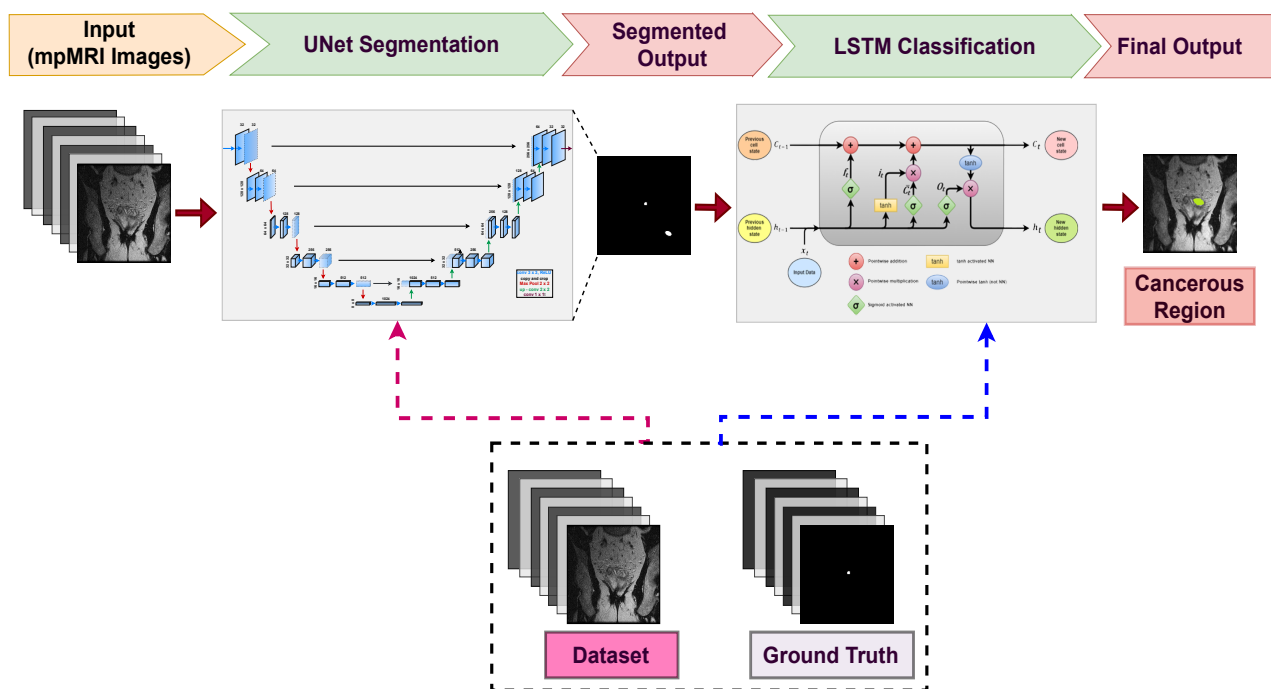


Figure 1. Overview of our proposed deep learning (DL) approach to classify mpMRI images. Trained U-NET and LSTM models are deployed in the proposed pipeline for segmentation first and then the classification of the segmented region for a binary prediction (cancerous vs. non-cancerous).

2. Related Work

The increasing use of artificial intelligence (AI) technologies in healthcare has led to the development of a predictive CAD system for PCa diagnosis. The mpMRI images are a popular modality for DL approaches due to their non-invasive collection method [9,14–17]. With the combination of mpMRI and AI technologies, CAD systems can provide clinicians with accurate and timely PCa diagnosis for improved patient care [3].

Brosch et al. [18] proposed a DL-based boundary detection and segmentation for prostate in MRI images. Their approach achieved a Dice similarity coefficient (DSC) and a mean absolute distance of 0.89 and 1.4 mm, respectively. Litjens et al. [19] developed a CAD system using MRI images for screening purposes. Their system achieved a sensitivity of 69% and a specificity of 83%. In a similar way, Zhang et al. [13] achieved an accuracy of 82.5% in detecting PCa in their CAD system. Through their methodology, they were able to achieve an average DSC of 0.76 in accurately segmenting prostate lesions. Aldoj et al. [20] used a DenseNet-like U-net architecture to perform prostate zonal segmentation from MRI images automatically and obtained an overall DSC of 0.92 in their limited dataset. Leveraging cost-sensitive support vector machines (and conditional random fields to refine segmentations, Artan et al. [21] proposed a CAD system for PCa localization on multi-spectral MRI. Their approach achieved a sensitivity of 88.9% and specificity of 92.2% for cancer detection and an average DSC of 0.82. Similarly, Peng et al. [14] designed a CAD system for PCa detection and differentiation from normal tissue using mpMRI, achieving an AUC of 0.85 for distinguishing cancer from normal tissue.

Karimi et al. [22] proposed a prostate segmentation method using a custom convolutional neural network (CNN) architecture and a training strategy based on statistical shape models achieving an average DSC of 0.895. In a similar CNN-based approach, Tian et al. [23] also proposed PSNet, for MRI prostate segmentation. Their proposed method achieved an average DSC of 0.94 and an average Hausdorff distance of 4.16 mm on the test dataset. Liu et al. [5] proposed a method for prostate cancer segmentation using MRI; their method achieved an average DSC of 0.815 and 0.666 for segmenting the prostate and cancerous regions, respectively. For PCa grading, Abraham et al. [24] used VGG16 CNN with an ordinal classifier to perform Gleason scoring (GS). Cao et al. [9] proposed a

multiclass CNN, FocalNet, for the simultaneous detection of prostate cancer lesions and prediction of their aggressiveness using GS. Benefiting from deep attention models, Duran et al. [25] proposed ProstAttention-Net for the segmentation of PCa by aggressiveness in MRI scans, achieving a DSC of 0.69. ProstAttention-Net achieved micro-precisions of 0.88 and 0.83 for high-risk and intermediate-risk cancer classification, respectively.

Leveraging the efficacy of transfer learning, Zhong et al. [15] proposed a model to classify PCa in mpMRI, achieving an accuracy of 87%, a sensitivity of 87%, and a specificity of 88%. Mehta et al. [16] developed a patient-level classification framework for PCa diagnosis using mpMRI and clinical features. Their system achieved an AUC of 0.89 for distinguishing cancer from benign tissue and 0.79 and 0.85 AUC for differentiating low-grade and high-grade cancer, respectively. Combining textural and morphological analysis, Zhang et al. [13] proposed a new approach for diagnosing PCa using MRI. Their approach achieved overall accuracy of 89.6%, sensitivity of 87.5%, and specificity of 90.8%.

Among works involving mpMRI, Mahapatra and Buhmann [26] developed an active learning-based method for prostate MRI segmentation using visual saliency cues. The method achieved an average DSC of 0.807. Liu and Yetik [27] proposed an iterative normalization method to improve PCa localization with multispectral MRI, achieving a detection rate of 89.2% on a small dataset. Sun et al. [28] developed DL models for detecting and localizing clinically significant PCa in mpMRI. Their models reached an average AUC-ROC of 0.91 for the detection of clinically significant PCa. Later, Hasan et al. [29] created a fully automated and efficient deep features extraction algorithm that uses T2W-TSE and STIR MRI sequences to discriminate between pathological and healthy breast MRI scans. The obtained features were classified using the LSTM classifier. Recently, Detectron2, developed by Facebook AI Research (FAIR), gained popularity for object detection and instance segmentation applications. There are some works in the literature using Detectron2 for other cancer types. To the best of our knowledge, the literature that attends to automated PCa diagnosis is nearly non-existent, specifically using mpMRI images. It shows that DL researchers have been focused on using architectures like U-Net, CNNs, and specialized variants to address unique challenges in processing mpMRI images. It opens room for us to experiment and harvest the power of combining contemporary U-Net and LSTM architectures to obtain higher DSC for PCa classification using mpMRI images.

3. Materials and Methods

This section provides details about the dataset used for developing DL models, the preprocessing dataset, and integrating the DL pipeline (as shown in Figure 1). U-net is an important first step to accurately segment the prostate gland and cancerous lesions, generating a reliable feature map for classification. Later, LSTM uses the power of sequential dependence to classify the features. A different combination of U-net segmentation and LSTM classification has already been used in other medical imaging applications and has shown to be effective [30]. A graphical overview of the method used to develop DL models is illustrated in Figure 2.

3.1. Dataset

The Initiative for Collaborative Computer Vision Benchmarking (I2CVB) dataset provides annotated mpMRI images for developing CAD systems. The dataset includes data from two commercial scanners: a 1.5 Tesla General Electric (GE) scanner and a 3.0 Tesla Siemens scanner [31]. The modalities available in the dataset are T2-Weighted (T2-W) MRI, dynamic contrast enhanced (DCE) MRI, diffusion weighted imaging (DWI) MRI, magnetic resonance spectroscopic imaging (MRSI), and apparent diffusion coefficient (ADC) maps for data acquired with the Siemens scanner. The T2-W MRI, DCE MRI, and DWI MRI ADC are in DICOM format. This dataset contains expertly annotated data and provides ground truth on every image. The I2CVB dataset is an essential resource for researchers and clinicians working on the development of CAD systems for PCa diagnosis.

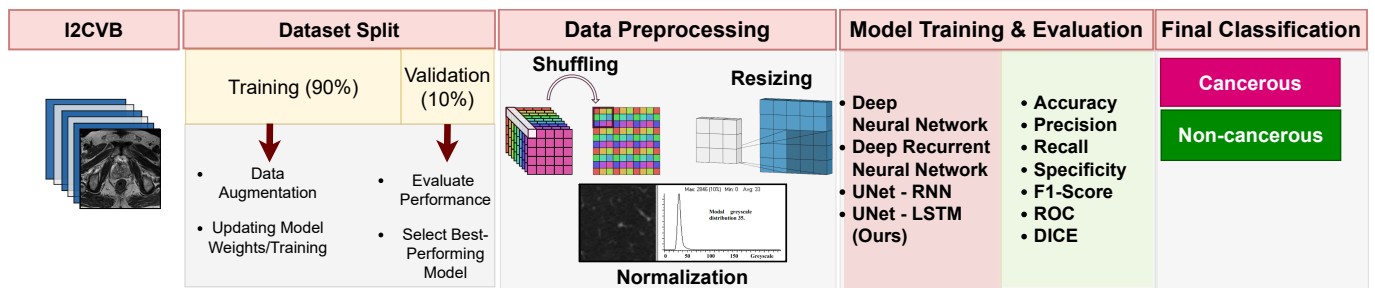


Figure 2. Overview of our proposed deep learning (DL) approach. The dataset is split into 90/10 for training and validation. Data preprocessing is applied to the training subset, and a set of chosen hyperparameters is used for the training setup. Later, the best-performing model on the validation set is used for developing the DL pipeline for inference.

3.2. Preprocessing

The preprocessing of mpMRI images is vital for improving the classification performance of DL models by reducing image acquisition artifacts, standardizing images across a data set, and isolating relevant areas [7]. To mitigate the challenge posed by a small number of training samples, we partitioned the dataset, allocating 90% for the training subset and 10% for the validation subset. This division recognizes the critical role of data augmentation in the network's development. We used data augmentation techniques, including generating new samples from existing ones by rotating, flipping, translating medical images, and adding noise to simulate imaging artifacts. Random elastic deformations were also used to train the segmentation network with a limited number of annotated images to enrich the model with acquisition invariance and robustness properties. By employing a coarse 3×3 grid, smooth deformations were generated through the utilization of random displacement vectors. To compute per-pixel displacements, bicubic interpolation was applied. These displacement vectors were sampled from a Gaussian distribution with a standard deviation of 10 pixels. Finally, the input images were resized to 256×256 .

3.3. Segmentation

We used U-Net architecture originally proposed by Ronneberger et al. [32] as shown in Figure 3. UNet proved its effectiveness and superiority in segmenting various organs and tissues in medical images [12]. U-Net is appropriate for mpMRI image segmentation since it can handle images of varying sizes and resolutions, making it flexible in handling the different image qualities and resolutions in mpMRI images. U-Net has already proven its ability to segment several structures in a single pass [32], particularly useful for complex structures in mpMRI images. It can also be trained with a relatively small amount of annotated data, making it suitable for medical imaging applications, where obtaining large amounts of annotated data is difficult.

The U-Net architecture consists of a contracting path and an expansive path, where the former comprises convolutional layers that reduce the input image's spatial resolution and extract high-level features. The latter contains up-convolutional layers that increase the spatial resolution and generate the segmentation map. To refine the segmentation in the expansive path, U-Net utilizes skip connections that enable the model to use low-level features from the contracting path. This feature overcomes the issue of spatial information loss in traditional CNNs. The contracting path of the network adheres to a standard convolutional architecture, where a series of two 3×3 convolutions, ReLU activation, and 2×2 max pooling operations are repetitively applied. On the other hand, the expansive path involves up-sampling the feature map, followed by a 2×2 convolution, concatenation with a cropped feature map from the contracting path, and two additional 3×3 convolutions. To map each 64-component feature vector to the desired number of classes, a final layer employs a 1×1 convolution. The entire network encompasses a total of 23 convolutional layers. It is essential to consider the input tile size in order to achieve seamless tiling of the

output segmentation map, ensuring that all 2×2 max-pooling operations are applied to layers with even x and y sizes.

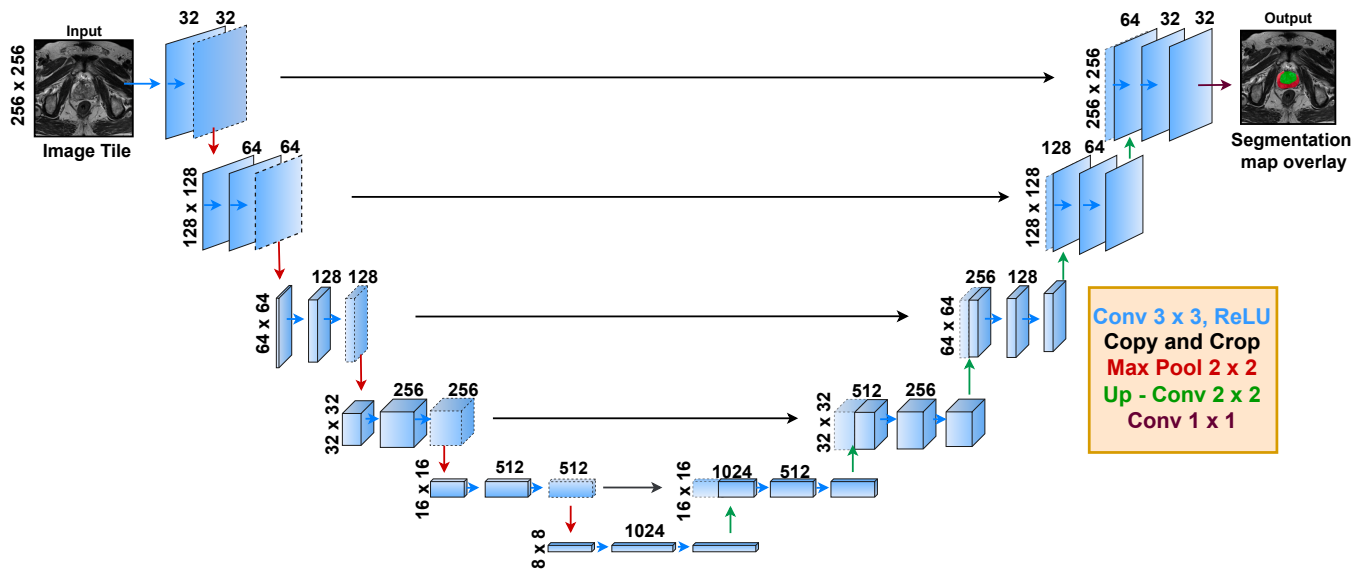


Figure 3. Schematic diagram of UNet architecture used in this work, originally proposed by [32].

Mathematically, the energy function in the U-Net architecture is computed by a pixel-wise softmax over the final feature map combined with the cross-entropy loss function. The energy function can be defined as

$$E = \sum_{x \in \Omega} \omega(x) \log(p_l(x)) \quad (1)$$

Let Ω represent the set of all pixel positions. Each pixel position x is assigned a weight $\omega(x)$ to ensure the contribution of each pixel is appropriately balanced in the loss function. Additionally, $p_l(x)$ denotes the predicted probability of pixel x belonging to the foreground class. The calculation of $p_l(x)$ involves applying a pixel-wise softmax operation over the final feature map, combined with the utilization of the cross-entropy loss function, with log referring to the natural logarithm.

The predicted probability $p_l(x)$ is calculated using a pixel-wise softmax over the final feature map combined with the cross-entropy loss function. The softmax is defined as

$$p_k(x) = \frac{\exp(a_k(x))}{\sum_{k=1}^K \exp(a_{k'}(x))} \quad (2)$$

Let $a_k(x)$ represent the activation in feature channel k at the pixel position $x \in \Omega$ with $\Omega \subset \mathbb{Z}^2$. The parameter K signifies the number of classes, and $p_k(x)$ denotes the approximated maximum function. In other words, $p_k(x) \approx 1$ for the value of K , which corresponds to the maximum activation $a_k(x)$ and $p_k(x) \approx 0$, while it is approximately equal to 0 for all other k values.

The weight map $\omega(x)$ is computed as

$$\omega(x) = \omega_c(x) + \omega_0 * \exp(-(d_1(x) + d_2(x))^2 / (2\sigma^2)) \quad (3)$$

ω_c represents the weight map utilized to balance class frequencies. The distance to the border of the nearest cell is denoted as d_1 , while d_2 refers to the distance to the border of the second nearest cell. The parameter σ controls the width of the Gaussian function. Prior to segmentation, the weight map is computed for each ground truth segmentation. This weight map aids in achieving balanced activation across the network, thereby improving

performance. It is crucial to appropriately set the initial weights to prevent certain parts of the network from exhibiting excessive activation while others remain dormant. To address this, the initial weights are drawn from a Gaussian distribution with a standard deviation of $\sqrt{2/N}$, where N represents the number of incoming nodes for a single neuron.

3.4. Classification

For classification, we used the LSTM network originally proposed by Staudemeyer et al. [33] (as shown in Figure 4). LSTM is a powerful DL technique for processing sequential data and has been successfully applied in medical image analysis. LSTM is an advanced type of RNN for sequence modeling and time series analysis. LSTM has several advantages, such as its ability to handle variable-length sequences, its capacity to capture long-term dependencies, and its potential for generalization on new sequences. mpMRI data can be organized as a sequence of measurements taken at various time points. The LSTM model learns to identify patterns and relationships between the data points over time and can be trained on a dataset of labeled examples.

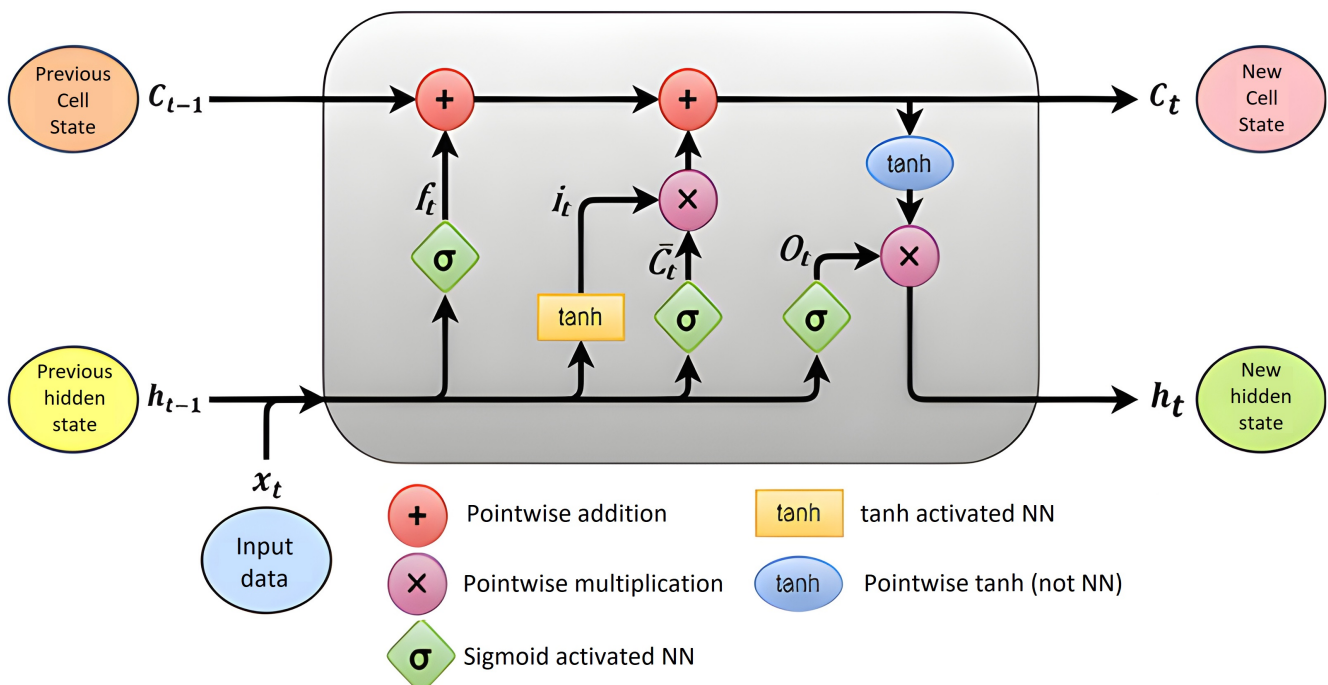


Figure 4. Schematic diagram of LSTM architecture used in this work, originally proposed by [33].

The strength of the LSTM architecture is the use of memory cells and gates to control the flow of information through the network, which allows them to better capture long-term dependencies in sequential data. The LSTM architecture comprises several components, including a cell state, an input gate, an output gate, and a forget gate. The cell state acts as the network's memory and is passed from one-time step to the next. The input gate regulates the influx of information into the cell state, while the output gate controls the outflow of information. The forget gate manages the removal of unnecessary information from the cell state. Let us denote x_t as the input at time step t , h_{t-1} as the hidden state at time step $t - 1$, i_t as the activation of the input gate, f_t as the activation of the forget gate, c'_t as the activation of the cell state, o_t as the activation of the output gate, c_t as the cell state at time step t , h_t as the hidden state at time step t , σ as the sigmoid function, \odot as the element-wise multiplication, and \tanh as the hyperbolic tangent function. With these definitions in mind, we can state that the input gate determines the extent to which new input should be added to the cell state:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (4)$$

The forget gate determines how much the previous cell state should be forgotten:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (5)$$

The cell activation updates create a candidate new cell state that can be added to the current cell state based on the input gate

$$c'_t = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \quad (6)$$

The cell state is updated by forgetting the previous cells based on the forget gate and adding the new candidate cell state based on the input gate:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (7)$$

The output gate determines how much of the current cell state should be output as the hidden state:

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (8)$$

The hidden state is updated based on the current cell and output gates:

$$h_t = o_t \odot \tanh(c_t) \quad (9)$$

During training, the LSTM model adjusts its internal parameters to minimize the difference between the predicted and actual labels. Once trained, the LSTM model can predict the probability of cancer recurrence for new patients.

3.5. Implementation and Experimental Setup

We implemented the DL models using the Caffe framework. The hyperparametric search was performed and later fixed to a learning rate of 0.0003, early stopping of 10 epochs, and stochastic gradient descent (SGD) optimizer. As a consequence of utilizing unpadded convolutions, the resulting output image exhibits a slight reduction in size compared to the input image. To make the most of the GPU memory and reduce the burden of computation, large input tiles were utilized over a large batch size, which led to using a batch size of a single image. We also used a high momentum value of 0.99, which allowed the network to take into account a larger number of previously seen training samples when updating the parameters during optimization. The proposed DL pipeline overcomes the limitations of each DL model to benefit from both in a comprehensive manner. The training algorithm can be summarized in Algorithm 1.

Algorithm 1 Training PCa detection from mpMRI.

Input: Input mpMRI images: $\rightarrow X = (x_1, x_2, \dots, x_N)$

Output: Segmented images $\rightarrow I_S = (I_{S1}, I_{S2}, \dots, I_{SN})$ and Classification output $\rightarrow C = (C_1, C_2)$

Initialize: Hyperparameters and set E epochs, data loader $\rightarrow D = (D_{train}, D_{val})$, LSTM model and U-Net model with random weights.

for $K = 1$ to E **do**

1: Sample the training data $I_K \in D_{train}$.

2: Obtain $I_{SK} \in UNET(x_k)$.

3: Calculate the DSC loss between I_{SK} and ground truth I_K .

4: Feed extracted features to the LSTM classifier.

5: Calculate binary cross-entropy loss.

6: Evaluate the model performance on D_{val} for PCa classification.

end for

return Optimal model weights

3.6. Evaluation Metrics

Evaluation criteria are essential in developing and deploying DL models to recognize their effectiveness in diagnosis. Let us assume TN as a true negative, TP as a true positive, FP as a false positive, and FN as a false negative. Then the metrics used to evaluate our proposed DL pipeline can be defined as follows:

- Precision: Precision is the ratio of TP predictions to the total number of positive predictions. In other words, it measures how many predicted positive cases are positive. A high precision value indicates that the model has a low rate of false positives.
- Recall: Recall, also referred to as sensitivity or true positive rate (TPR), quantifies the ratio of correctly predicted TP cases to the overall number of positive cases. Essentially, it evaluates the accuracy of identifying actual positive cases as positive. In the context of prostate cancer diagnosis, a high recall/sensitivity signifies the algorithm's capability to accurately detect cancerous tissue.
- F1 score: The harmonic means of precision and recall. In prostate cancer diagnosis, a high F1 score indicates that the algorithm is able to accurately identify cancerous tissue with few false positives and false negatives.
- Accuracy: The accuracy is the proportion of correct predictions made by the algorithm. In prostate cancer diagnosis, high accuracy indicates that the algorithm is able to identify both cancerous and healthy tissue accurately. Accuracy is $(TN + TP) / TP + FP + FN + TP$.
- Specificity: The specificity, also known as the false positive rate (FPR), is the proportion of actual negative cases correctly identified by the algorithm. In prostate cancer diagnosis, high specificity indicates that the algorithm is able to identify healthy tissue accurately.
- Receiver operating characteristic (ROC): The ROC plot illustrates the trade-off between sensitivity and specificity for varying threshold values. To assess the algorithm's overall performance, the area under the ROC curve, known as AUC, is commonly employed as a metric. The AUC captures the algorithm's ability to discriminate between positive and negative cases, providing a comprehensive evaluation of its performance.
- Dice similarity coefficient (DSC): The Dice index, also referred to as the Dice coefficient, serves as a commonly used metric for evaluating the performance of a segmentation model. It quantifies the degree of overlap between the predicted segmentation and the ground truth, with values ranging from 0 to 1. A value of 1 signifies a perfect agreement between the predicted and ground truth segmentation. A higher Dice coefficient indicates improved segmentation accuracy, which is particularly valuable when working with imbalanced data or when dealing with segmented objects of varying sizes.

4. Results and Discussion

This section provides a comparative analysis of the proposed DL approach. The implementation is compared with three other techniques: DNN (i.e., ResNet50), Deep RNN, and U-net RNN. Table 1 summarizes the performance of four models and comparison against other works in the literature.

The comparative analysis (in Table 1) shows that U-Net with LSTM outperforms other DL models. The U-Net with the LSTM model achieved a top accuracy of 90.69%, significantly improving a single ResNet-50 DNN by a huge margin. Our U-NET + LSTM pipeline surpasses other works in the literature by a significant margin. To evaluate the efficacy of models using a fraction of the training set, we evaluated all these models with evaluation metrics. Figure 5 shows classification performance as a function of the percentage of the total available training data. Each graph represents a different performance metric, with the x-axis representing the percentage of training data used and the y-axis representing the metric score. The results exhibit that all four different architectures improve their classification performance on the validation set when training data on more data. In particular, precision

improves significantly with an increased percentage of training data. It recapitulates that the availability of training data can enhance the performance of DL models, where the U-Net + LSTM pipeline shows superior performance at all fractions of the training set. Interestingly, our proposed DL approach can work well, even in the presence of smaller datasets. DSC is reported as a fractional number value for each model, where a higher value indicates better segmentation performance. The findings indicate that employing U-Net with LSTM achieved the highest DSC value of 0.670, followed by U-Net with RNN of 0.648, Deep RNN of 0.644, and DNN with 0.592. Our findings conclude that the U-Net architecture combined with an LSTM or RNN performs better than the traditional use of a single DNN or deep RNN for PCa diagnosis and segmentation task using the given training data.

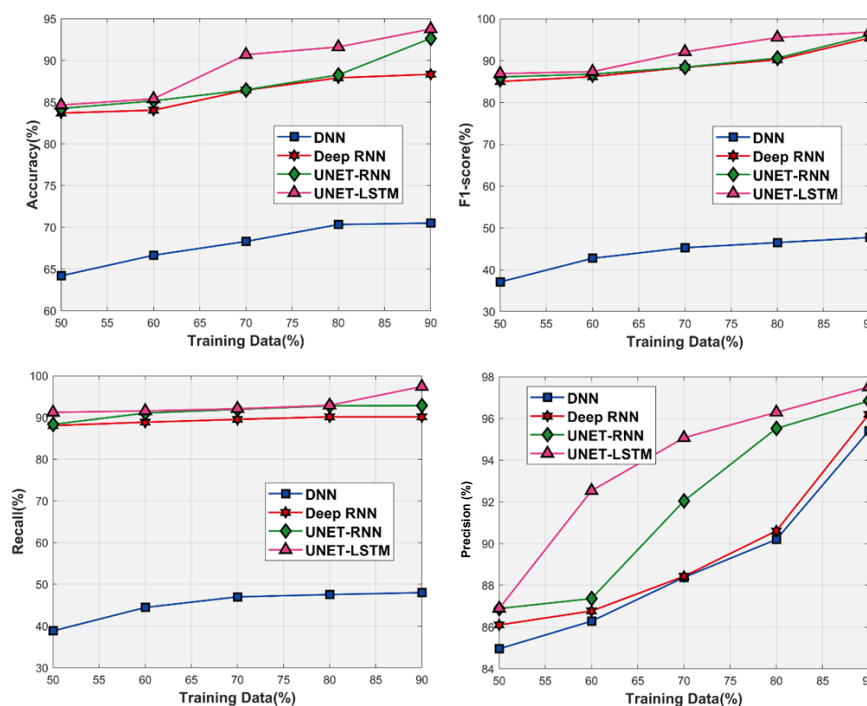


Figure 5. Analyzing the performance of our DL pipeline against other stand-alone networks and literature works. The x-axis represents the fraction of training data used in the development of models, while the y-axis shows the obtained scores on four evaluation metrics (in each sub-graph).

Table 1. Comparative analysis of mpMRI PCa segmentation and patient-level classification on the validation set. The best results are marked in bold, and results not reported in the literature are filled with a dash.

Architectures	Accuracy (%)	F1 (%)	Precision	Recall (Sens. (%))	RoC	Spec. (%)	Dice
Liu et al. [5]	89.38	-	-	87.5	-	89.5	0.62
Artan et al. [21]	-	-	-	85.0	-	50.0	0.34
Zhang et al. [13]	80.97	-	76.69	-	0.77	-	-
PCF-SEL-MR [16]	-	-	63.0	75.0	0.86	55.0	-
FocalNet [9]	-	-	-	89.7	-	-	-
DNN	68.31	45.28	88.38	46.98	0.719	89.53	0.59
Deep RNN	86.43	88.37	88.43	89.53	0.787	91.81	0.64
U-Net RNN	86.47	88.43	92.04	91.92	0.814	90.09	0.65
U-Net LSTM (Ours)	90.69	92.09	95.17	92.09	0.953	96.88	0.67

Finally, we used our best-performing model to find RoI examples for visualization. Four cancerous samples from the validation set are tested with outcomes depicted in Figure 6. It shows the original, pre-processed, ground truth, and predicted ROI images. In

the resulting mask using the U-Net and LSTM pipeline, the green-colored region typically refers to the area predicted to contain cancer. All predicted samples cover the annotated regions and show an extended cancerous area in some cases. The color scheme can be updated to aid in the visual interpretation of the segmentation results with different grades or other potential biomarkers.

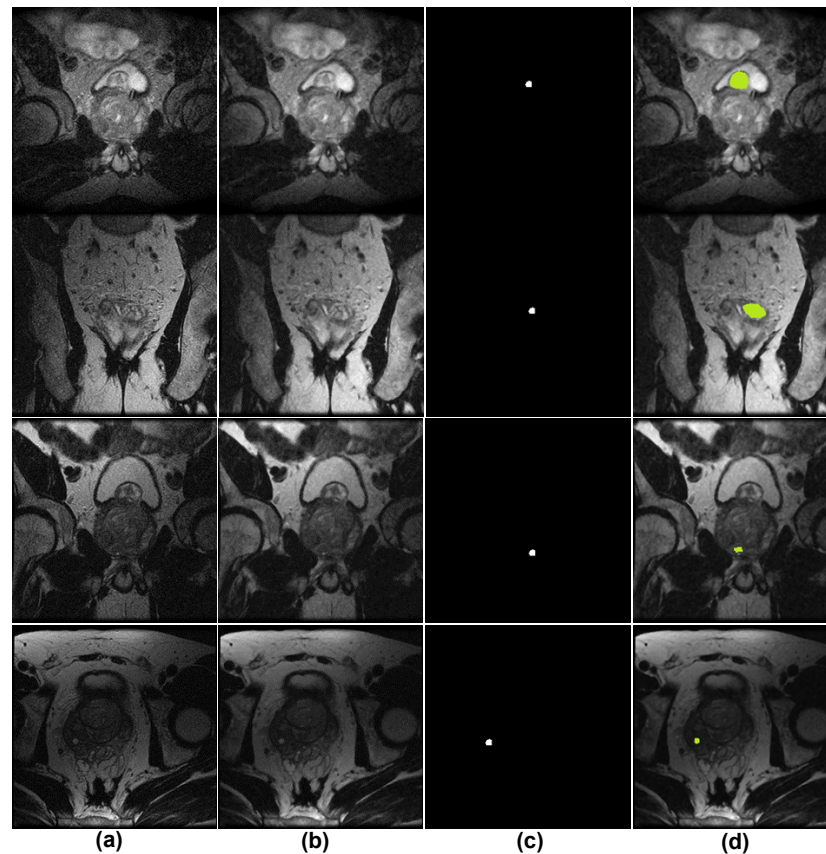


Figure 6. Visualization of the region of interest (RoI) using our proposed UNet+LSTM DL approach. (a) Example mpMRI images; (b) preprocessed version of the images; (c) Ground truth; (d) map of predicted cancerous RoI.

5. Conclusions

The detection of prostate cancer (PCa) at the early stage may reduce the mortality rate, and deep learning (DL) holds the potential to benefit the precise detection and support of clinical decisions. This paper introduces a DL-based approach for the detection of prostate cancer using multiparametric magnetic resonance imaging (mpMRI). The proposed DL pipeline employs U-Net for segmentation and LSTM for classification to identify cancerous patients. The effectiveness of the proposed approach was assessed by employing a range of standardized performance evaluation metrics. Our proposed DL approach yields a Dice coefficient of 0.67 for the segmentation task, along with an accuracy of 90.69% and an F1-score of 92.09% for the classification metrics. These results clearly articulate the superiority of our approach over existing state-of-the-art methods, such as DNN and Deep RNN. These results strongly indicate that the proposed approach holds significant potential in improving the precision and efficiency of PCa diagnosis. Consequently, the research findings would significantly improve future prostate cancer diagnosis and treatment.

6. Limitations and Future Work

The proposed DL pipeline is developed and tested on a single dataset. However, incremental learning can be applied to the developed robust models, as they can benefit from other existing datasets, such as PROSTATEx [34], PICTURE [35], etc. This extensive training

and validation will reduce bias and enhance the generalizability of DL models. Ultimately, overcoming these limitations will boost the reliability and foster clinical adoption of CAD systems for automated PCa diagnosis.

The future of automated PCa grading DL models will benefit from vision transformers (ViTs) [4] and their multi-attention [36] to the spatial correlation of mpMRI images. This work can be further extended by combining ViTs and CNNs for PCa grading tasks, possibly in combination with data from different modalities. By incorporating state-of-the-art ViT models and staying attuned to advancements in multi-model and representation learning, PCa diagnosis can be more precise (even with fewer labeled data) for strong clinical decision support.

Author Contributions: Conceptualization, A.B.G.; Methodology, A.B.G. and P.A.G.; Software, S.S.P.; Investigation, A.B.G., S.S.P. and R.N.; Data Curation, S.S.P. and R.N.; Validation, P.A.G.; Formal Analysis, A.B.G. and N.K.; Writing—original draft, A.B.G.; Writing—editing & review; N.K. and S.T.H.R.; Mentoring for present-ability, N.K.; Supervision; N.K., R.N. and S.T.H.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset is publicly available.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stephen, W.L.; Larry, E.S.; Hussain, S.; R, I.A.; Taylor, S.S. *Prostate Cancer*; U.S. National Library of Medicine: Bethesda, MD, USA, 2022.
2. *Survival Rates for Prostate Cancer*; American Cancer Society: Atlanta, GA, USA 2023.
3. Tabatabaei, Z.; Colomer, A.; Engan, K.; Oliver, J.; Naranjo, V. Residual block convolutional auto encoder in content-based medical image retrieval. In Proceedings of the 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Nafplio, Greece, 26–29 June 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–5.
4. Kanwal, N.; Eftestøl, T.; Khoraminia, F.; Zuiverloon, T.C.; Engan, K. Vision Transformers for Small Histological Datasets Learned Through Knowledge Distillation. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Osaka, Japan, 25–28 May 2023; Springer: Cham, Switzerland, 2023; pp. 167–179.
5. Liu, X.; Langer, D.L.; Haider, M.A.; Yang, Y.; Wernick, M.N.; Yetik, I.S. Prostate cancer segmentation with simultaneous estimation of Markov random field parameters and class. *IEEE Trans. Med. Imaging* **2009**, *28*, 906–915. [[CrossRef](#)] [[PubMed](#)]
6. Kanwal, N.; Amundsen, R.; Hardardottir, H.; Janssen, E.A.; Engan, K. Detection and Localization of Melanoma Skin Cancer in Histopathological Whole Slide Images. *arXiv* **2023**, arXiv:2302.03014.
7. Kanwal, N.; Pérez-Bueno, F.; Schmidt, A.; Engan, K.; Molina, R. The Devil is in the Details: Whole Slide Image Acquisition and Processing for Artifacts Detection, Color Variation, and Data Augmentation: A Review. *IEEE Access* **2022**, *10*, 58821–58844. [[CrossRef](#)]
8. Sunoqrot, M.R.; Selnaes, K.M.; Sandsmark, E.; Langørgen, S.; Bertilsson, H.; Bathen, T.F.; Elschot, M. The reproducibility of deep learning-based segmentation of the prostate gland and zones on T2-weighted MR images. *Diagnostics* **2021**, *11*, 1690. [[CrossRef](#)]
9. Cao, R.; Bajgiran, A.M.; Mirak, S.A.; Shakeri, S.; Zhong, X.; Enzmann, D.; Raman, S.; Sung, K. Joint prostate cancer detection and Gleason score prediction in mp-MRI via FocalNet. *IEEE Trans. Med. Imaging* **2019**, *38*, 2496–2506. [[CrossRef](#)]
10. Gavade, A.B.; Nerli, R.B.; Ghagane, S.; Gavade, P.A.; Bhagavatula, V.S.P. Cancer Cell Detection and Classification from Digital Whole Slide Image. In *Smart Technologies in Data Science and Communication: Proceedings of SMART-DSC 2022*; Springer: Singapore, 2023; pp. 289–299.
11. Tabatabaei, Z.; Engan, K.; Oliver, J.; Naranjo, V. Self-supervised learning of a tailored Convolutional Auto Encoder for histopathological prostate grading. *arXiv* **2023**, arXiv:2303.11837.
12. Li, H.; Lee, C.H.; Chia, D.; Lin, Z.; Huang, W.; Tan, C.H. Machine learning in prostate MRI for prostate cancer: Current status and future opportunities. *Diagnostics* **2022**, *12*, 289. [[CrossRef](#)]
13. Zhang, L.; Li, L.; Tang, M.; Huan, Y.; Zhang, X.; Zhe, X. A new approach to diagnosing prostate cancer through magnetic resonance imaging. *Alex. Eng. J.* **2021**, *60*, 897–904. [[CrossRef](#)]
14. Peng, Y.; Jiang, Y.; Yang, C.; Brown, J.B.; Antic, T.; Sethi, I.; Schmid-Tannwald, C.; Giger, M.L.; Eggen, S.E.; Oto, A. Quantitative analysis of multiparametric prostate MR images: Differentiation between prostate cancer and normal tissue and correlation with Gleason score—A computer-aided diagnosis development study. *Radiology* **2013**, *267*, 787–796. [[CrossRef](#)]
15. Zhong, X.; Cao, R.; Shakeri, S.; Scalzo, F.; Lee, Y.; Enzmann, D.R.; Wu, H.H.; Raman, S.S.; Sung, K. Deep transfer learning-based prostate cancer classification using 3 Tesla multi-parametric MRI. *Abdom. Radiol.* **2019**, *44*, 2030–2039. [[CrossRef](#)]

16. Mehta, P.; Antonelli, M.; Ahmed, H.U.; 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**, *73*, 102153. [[CrossRef](#)]
17. Mehta, P.; Antonelli, M.; Singh, S.; Grondecka, N.; Johnston, E.W.; Ahmed, H.U.; Emberton, M.; Punwani, S.; Ourselin, S. AutoProstate: Towards automated reporting of prostate MRI for prostate cancer assessment using deep learning. *Cancers* **2021**, *13*, 6138. [[CrossRef](#)]
18. Brosch, T.; Peters, J.; Groth, A.; Stehle, T.; Weese, J. Deep learning-based boundary detection for model-based segmentation with application to MR prostate segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2018: 21st International Conference, Granada, Spain, 16–20 September 2018; Proceedings, Part IV 11; Springer: Cham, Switzerland, 2018; pp. 515–522.
19. Litjens, G.; Debats, O.; Barentsz, J.; Karssemeijer, N.; Huisman, H. Computer-aided detection of prostate cancer in MRI. *IEEE Trans. Med. Imaging* **2014**, *33*, 1083–1092. [[CrossRef](#)]
20. Aldoj, N.; Biavati, F.; Michallek, F.; Stober, S.; Dewey, M. Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net. *Sci. Rep.* **2020**, *10*, 14315. [[CrossRef](#)]
21. Artan, Y.; Haider, M.A.; Langer, D.L.; Van der Kwast, T.H.; Evans, A.J.; Yang, Y.; Wernick, M.N.; Trachtenberg, J.; Yetik, I.S. Prostate cancer localization with multispectral MRI using cost-sensitive support vector machines and conditional random fields. *IEEE Trans. Image Process.* **2010**, *19*, 2444–2455. [[CrossRef](#)]
22. Karimi, D.; Samei, G.; Kesch, C.; Nir, G.; Salcudean, S.E. Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *13*, 1211–1219. [[CrossRef](#)]
23. Tian, Z.; Liu, L.; Zhang, Z.; Fei, B. PSNet: Prostate segmentation on MRI based on a convolutional neural network. *J. Med. Imaging* **2018**, *5*, 021208. [[CrossRef](#)] [[PubMed](#)]
24. Abraham, B.; Nair, M.S. Automated grading of prostate cancer using convolutional neural network and ordinal class classifier. *Inform. Med. Unlocked* **2019**, *17*, 100256. [[CrossRef](#)]
25. Duran, A.; Dussert, G.; Rouvière, O.; Jaouen, T.; Jodoin, P.M.; Lartizien, C. ProstAttention-Net: A deep attention model for prostate cancer segmentation by aggressiveness in MRI scans. *Med. Image Anal.* **2022**, *77*, 102347. [[CrossRef](#)] [[PubMed](#)]
26. Mahapatra, D.; Buhmann, J.M. Visual saliency-based active learning for prostate magnetic resonance imaging segmentation. *J. Med. Imaging* **2016**, *3*, 014003. [[CrossRef](#)]
27. Liu, X.; Samil Yetik, I. Iterative normalization method for improved prostate cancer localization with multispectral magnetic resonance imaging. *J. Electron. Imaging* **2012**, *21*, 023008. [[CrossRef](#)]
28. Sun, Z.; Wu, P.; Cui, Y.; Liu, X.; Wang, K.; Gao, G.; Wang, H.; Zhang, X.; Wang, X. Deep-Learning Models for Detection and Localization of Visible Clinically Significant Prostate Cancer on Multi-Parametric MRI. *J. Magn. Reson. Imaging* **2023**. [[CrossRef](#)]
29. Hasan, A.M.; Qasim, A.F.; Jalab, H.A.; Ibrahim, R.W. Breast Cancer MRI Classification Based on Fractional Entropy Image Enhancement and Deep Feature Extraction. *Baghdad Sci. J.* **2023**, *20*, 0221. [[CrossRef](#)]
30. Agnes, S.A.; Anitha, J.; Solomon, A.A. Two-stage lung nodule detection framework using enhanced UNet and convolutional LSTM networks in CT images. *Comput. Biol. Med.* **2022**, *149*, 106059. [[CrossRef](#)] [[PubMed](#)]
31. Lemaître, G.; Martí, R.; Freixenet, J.; Vilanova, J.C.; Walker, P.M.; Meriaudeau, F. Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: A review. *Comput. Biol. Med.* **2015**, *60*, 8–31. [[CrossRef](#)] [[PubMed](#)]
32. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; Springer: Cham, Switzerland, 2015; pp. 234–241.
33. Staudemeyer, R.C.; Morris, E.R. Understanding LSTM—A tutorial into long short-term memory recurrent neural networks. *arXiv* **2019**, arXiv:1909.09586.
34. Armato, S., III; Huisman, H.; Drukker, K.; Hadjiiski, L.; Kirby, J.; Petrick, N.; Redmond, G.; Giger, M.; Cha, K.; Mamonov, A.; et al. PROSTATEx Challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images. *J. Med. Imaging* **2018**, *5*, 044501. [[CrossRef](#)] [[PubMed](#)]
35. Simmons, L.A.; Kanthabalan, A.; Arya, M.; Briggs, T.; Barratt, D.; Charman, S.C.; Freeman, A.; Gelister, J.; Hawkes, D.; Hu, Y.; et al. The PICTURE study: Diagnostic accuracy of multiparametric MRI in men requiring a repeat prostate biopsy. *Br. J. Cancer* **2017**, *116*, 1159–1165. [[CrossRef](#)] [[PubMed](#)]
36. Kanwal, N.; Rizzo, G. Attention-based clinical note summarization. In Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing, Virtual, 25–29 April 2022; pp. 813–820.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.