Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 3
Open Access

Amalgamating 3D Convolutional Neural Networks with Recurrent Neural Networks to Enhance Detection of Clinically Notable Prostate Cancer Using Multiparametric MRI Scans

Tejaswini B[1], Dr Suma R[2]

1. Research Scholar, Department of Computer Science and Engineering, SSAHE, Tumakuru, Karnataka, India.
2. Professor, Department of information Science and Engineering, SSAHE, Tumakuru, Karnataka, India.

Cite this paper as: Tejaswini B, Suma R (2024) Amalgamating 3D Convolutional Neural Networks with Recurrent Neural Networks to Enhance Detection of Clinically Notable Prostate Cancer Using Multiparametric MRI Scans. *Frontiers in Health Informatics, 13* (3), 229-237.

Abstract

Prostatic adenocarcinoma is one among the prevalent cancers contributing to top causes of cancer deaths worldwide. Techniques of advanced imaging, like multiparametric magnetic resonance imaging, efficiently enhances the accuracy of prostatic adenocarcinoma detection. We take our cue from the huge success deep learning has had with medical image analysis by developing an automated pipeline combining RNN with 3DCNN for the identification of clinically significant PCa within whole-organ mpMRI scans. In this paper, we present a two-stage pipeline where the first stage uses 3DCNN for feature extraction from volumetric images, while the second stage uses an RNN for modelling sequential dependencies across slices. Our two-class dataset includes 500 patients; 200 have clinically significant PCa and 300 are free of PCa. A test set of 100 patients not used in training validated this model. At the slice level, this pipeline produced an AUC of 0.89 with a 95% confidence interval of 0.86–0.92, and at the patient level, it had an AUC of 0.86 with a 95% confidence interval of 0.81–0.91. The results have demonstrated that the integration of 3DCNN with RNN has very good potential for the clinical application of prostate cancer detection.

Keywords: *Prostate Cancer Detection,3D Convolutional Neural Networks, Recurrent Neural Networks, Multiparametric MRI, Deep Learning in Medical Imaging*

INTRODUCTION

Prostate malignancy develops as a serious health risk, being among the most prevalent types of cancer diagnosed and being a number three prime cause for deaths caused by tumour growth in men in North America. Early detection of clinically relevant prostate cancer (PCa) which is exact will affect proper treatment decisions and enhance patient outcome. Conventional techniques of diagnosis like testing for prostate specific antigen (PSA) and manual rectal examination often fall short with regard to specificity or sensitivity as a result they sometimes lead to unnecessary biopsy procedures or in some cases aggressive forms go undetected altogether. In recent times, multi-parametric magnetic resonance imaging (mpMRI) is an indispensable device for identifying prostate cancers because it provides intricate anatomic and functional data regarding prostate glandular structure. Combined with T2 weight imaging, diffusion-weighted imaging (DWI), and dynamic contrast-

enhanced (DCE), mpMRI integrates numerous imaging sequences for thorough assessment of lesions found on the prostate gland. Even with its advantages, the interpretation of mpMRI still remains difficult and highly dependent on the radiologist's expertise, hence variability and potential diagnostic error. Deep learning has completely revolutionized medical image analysis with advanced methodologies for automated feature extraction and classification. One class of these methods involves convolutional neural networks, which is very beneficial in various aspects of computer vision, including object detection, segmentation, and classification. Despite the fact that 2DCNNs have been deeply investigated for medical imaging, mpMRI data are characterized by a complex 3D structure, which requires more advanced architectures able to capture volumetric information. This paper introduces another way to automate the identification and medically significant prostatic adenocarcinoma in patients. In association with 3D convolutional neural networks as well as recurrent neural networks, our approach serves the detection of prostate cancer. The 3DCNN component serves as a raw volumetric feature extractor from mpMRI scans, while the RNN component is used to learn the associations through utilizing these features that make the identification of malignant lesions easier. We collected a data repository of mpMRI images from 500 patients, including 200 with clinically significant PCa and 300 without PCa, to train and evaluate our model. A newly-submitted test set of 100 patients was used to determine the pipeline's level of success. Our new model resulted in the region beneath recipient operating characteristic curve (AUC) of 0.89 at the slice level as well as a 0.86 at the patient level, which an insight into the high potential of the model.

Entire research work flow can be categorized as: Segment II surveys associated work in deep learning approaches with regards to prostate cancer detection; Section III gives a light on the dataset and preprocessing techniques; Section IV describes the architecture of the proposed 3DCNN and RNN pipeline; Part V presents the experimental setup and evaluation metrics; Section VI consist of outcome and their inference; and lastly, Section VII concludes the paper, pointing to some future directions.

REVIEW OF LITERATURE:

In recent years, prostate cancer identification depending on deep learning methods used in the field has drawn substantial attention. We identify crucial contributions in the field and review their methodologies, performance metrics as well as research gaps. Liu et al. For the retropharyngeal lymph node metastasis (a), an occupation based on their results in prostate cancer detection applied by Litjens et al., which they introduced CNNs to automatically assess disease status from multiply region of interest MRI was done [33]. The study could discriminate malignant and benign lesions with an accuracy of 82% (AUC = 0.85). That said, the authors acknowledged that their 2DCNN model was limited in capturing all volumetric information of prostate gland, ROI and could be further improved by utilizing a 3D variant based on 3D CNNs. Litjens et al. Use of Transfer Learning: Litjens et al. (2016) investigated the potential benefits in utilizing transfer learning for improved detection performance by CNNs attempting to determine prostate cancer from MRI scans. Using pretrained models with large datasets they were able to reach an accuracy of 85% and AUC score of 0.88. It was realized that there was a need to apply more extensive data augmentation techniques for better robustness and generalizability of the model. Wang et al. proposed a hybrid approach to CNNs and RNNs for prostate cancer detection in 2018. Their model utilized an LSTM network to capture sequential dependencies of MRI slices. It provided quite impressive results in detecting prostate cancer with an AUC of 0.86 at the patient level. The

result of this study had very strong potential regarding RNNs and temporal data but trained models on large datasets. Kersting et al. (2019) proposed a 3DCNN model for prostate cancer detection. In its completeness, it processed volumetric mpMRI data with an accuracy of 0.87 and an AUC of 0.89. The authors noted the strength of 3DCNN over its 2D version in the extraction of spatial relationships within the prostate gland. An increased computational complexity was also indicated, thus requiring more sophisticated hardware. Saha et al. evaluated a deep-learning multimodal technique integrating mpMRI with clinical information like PSA levels in the detection of prostate cancer. Their model yielded an AUC of 0.90, thus better than single-modality models. The research pointed out the importance of fusing multiple sources of data, but it mentions at the same time the difficulty of fusing heterogeneous data.

METHODOLOGY:

Data Acquisition: We acquired data for our study from January 2019 to December 2022 using a Siemens Magnetom Skyra 3T whole-body MRI scanner. The MRIs were conducted in the Department of Radiology belonging to hospital from Bangalore. The protocol had multiparametric MRI sequences, but more specifically, T2 and diffusion-weighted imaging as well as dynamic contrast-enhanced imaging. To get DWI sequences, the transverse plane was acquired through a single-shot spin echo echoplanar sequence of imaging through the following parameters: b-values ranging 0, 50, 500, and 1000 s/mm²; repeating time between 4000-6000 ms; echo time around 70 ms; slice thickness around 3.5 mm; field of view around 220 mm × 220 mm; and matrix of 128 × 128. The DWI images were collected to measure the responsiveness of tissue towards Brownian motion that has been proved being effective in PCa detection. The different values produce surrogate values representing water diffusion in the tissue, for which Apparent Diffusion Coefficient (ADC) maps have been derived and high-valuation b value images rendered (b1400).

Training, Validation and Test Sets: We further split the MRI images of 500 patients into three group. They are training, validation as well as test set. The training set consisted images of 320 patients (5,120 slices), validation with that of 80 patient (1.280 slice) and testing with those of the other potential new cohort from a total on 100 demonstrators (1600 segments). The splitting ratio of train/validation/test was around 64%,16% and 20%. The dataset segregation was done as below:

First, we did a split in the dataset: train/validation versus test in an 80:20 ratio to keep a sufficient size of samples for the set to test. Secondly, further division of the training as well as validation set was done into two subsets: a set to train 80% of the instructing/validation group as well as a group of 20% to validate. This was done to ensure that, across all the sets of data, there was approximately a constant ratio between cases with PCa and those without PCa.

Data Pre-processing: All MRI images from the dataset were pre-treated using the following steps for uniformity and to improve the model performance.

Normalization: The slices of the MRI across the dataset are all normalized by the following function:

$$X_{normalized} = \frac{X_i - \mu}{std}$$

where X_i is unique MRI slice pixel intensities, μ indicates dataset average, std denotes its standard deviation whereas X_normalized stands for normalised single MRI slices. Resizing and Cropping:

Each MRI slice was resized to 128 × 128 pixels and centre-trimmed to 64 × 64 pixels to ensure adequate coverage of the prostate region.

Data enhancement: rotation, flipping, scaling in the training set to eliminate data set variability higher and improve model robustness and so on. Data-enhancement techniques used Channel construction: Channels were constructed for each b-value (0, 50, 500, 1000, and 1400) for DWI images, and apparent diffusion coefficient (ADC). This resulted in 5 channel input for convolutional neural networks (The CNNs).
Deep learning framework: Leveraging the positive result of deep learning in healthcare analysis of images, we developed a new automated algorithm using a combination of recurrent neural networks (RNNs) and 3D convolutional neural networks (3DCNNs) , investigated clinically significant prostate cancer (PCa) from mpMRI scans. In the first step, volume feature extraction is performed using 3D Convolutional Neural Networks (3DCNNs). Inspired by the ResNet framework, the 3DCNN framework is used in various computer vision tasks with proven efficiency. In ResNet-based 3DCNN, each residual block contains convolutional layers along with introduction shortcut combinations that facilitate gradient propagation, thus solving the missing-slope problem The following steps occur in each residual block.

Convolutional Layer: The input X gets convolved with a 3D kernel, where a set of filters (or kernels) would slide over the input volume to proffer preparations or feature maps. Mathematically, this can be shown thus:

$$F_1(X) = W_1 * X + b_1$$

where W_1 and b_1 are the weights and biases of the first order, and $*$ denotes the excitation function.

Batch Normalization and Activation : The output of the convolutional layer is normalized by batch normalization to ensure robustness and speed of training. A ReLU (Rectified Linear Unit) function is then applied to incorporate nonlinearity.

$$F_2(X) = ReLU(BatchNorm(F_1(X)))$$

Second Convolutional Layer: Another convolutional layer takes the activated output as input.

$$F_3(X) = W_3 * F_2(X) + b_2$$

where W_2 and b_2 have being weights as well as biases of the after the first rotation part.

Batch Normalization: Additional batch normalization is applied to the output of the second convolutional layer

$$F_4(X) = BatchNorm(F_3(X))$$

Identity Shortcut Connection: Some convolution layers are skipped by adding input X directly to the batch normalization output.

$$Y = F_4(X) + X$$

Here Y is the final output of the remaining part.

Activation: The activation function ReLU is implied to the outcome of the residual group to introduce nonlinearity.

$$Y_{final} = ReLU(Y)$$

In cases where the source and target dimensions are different, a linear projection, W_s is used to make them

equal:

$$Y = F_4(X) + W_s X$$

where W_s is the projection matrix of the input dimensions to the output dimensions. To further enhance performance a fully pre-processed residual grid is used, with batch normalization and ReLU activation layers ahead of convolutional layers. This pre-activation process ensures that gradients are properly preserved during back propagation, if besides, the frequent degradation problem is solved without compromising the performance of layered deep learning. Here the "bottleneck" building block is used. This bottleneck process typically consist convolution of 1x1 to reduce the extension, a convolution of 3x3 as well as convolution of 1x1 to restore dimensionality.

In the second stage, sequential dependencies across mpMRI photo slices are captured the use of Recurrent Neural Networks (RNNs). The volumetric capabilities extracted from the 3DCNN are organized into a chain representing distinct slices of the mpMRI scan. Gated Recurrent Unit (GRU) , rather Long Short-Term Memory (LSTM) networks have been then hired so as to model the temporal dependencies inherent in those sequences. The RNN approaches the function vectors f_t at whenever step t, producing hidden states .
h_t that encapsulate the sequential records:

$$h_t = RNN(f_t, h_{t-1})$$

The version education method involves the utilization of a dataset comprising 500 sufferers, consisting of 2 hundred patients with clinically extensive PCa and three hundred sufferers without PCa. The dataset is break up into education and validation sets, with a separate test set of 100 sufferers reserved for comparing the version's effectiveness. Cross-validation techniques are hired to best-music the model parameters, making sure robust performance.

Layer Name	Details about the Layer
Conv3D Layer	3D Convolutional Layer (7 × 7 × 7, 64 filters, stride 2)
Max Pool3D	3 × 3 × 3 max pooling, stride 2
ResNet Block 1	$\begin{matrix}1\times1\times1,64\\3\times3\times3,64\\1\times1\times1,256\end{matrix}\times4$
ResNet Block 2	$\begin{matrix}1\times1\times1,128\\3\times3\times3,128\\1\times1\times1,512\end{matrix}\times9$
Ave Pool3D	3D Average Pooling (7 × 7 × 7)
RNN Layer	LSTM/GRU Layer, 128 units
Fully Connected	Fully Connected Layer (2D, softmax)

*Figure 1**Error! No text of specified style in document.**.1 A table detailing the layers of the proposed deep learning pipeline combining 3D Convolutional Neural Networks (3DCNNs) as well as Recurrent Neural Networks (RNNs):*

3DCNN design and training: A 41-layer deep ResNet-inspired 3D Convolutional Neural Network (3DCNN) was designed for the volumetric feature extraction phase. It is designed with 3D convolutional layers, starting with a 7 × 7 × 7 filter followed by a 3 × 3 × 3 maximum. The optimal depth of the 41 layers was determined using hyper-parameter fine-tuning with a validation set of pooling layer and residual blocks. Due to the small

size of the input mpMRI images (66 × 66 × 66 voxels) and the small tumour region (e.g., 4 × 3 × 3 voxels), it was necessary to capture enough features 3D CNN layers in ResNet with additional ResNet blocks (ResNet Block1). There are layers with filter sizes of 64, 64, and 256, repeated 4 times in 3D CNN layers after the next ResNet Block (ResNet Block2) .3D CNN layers with filter size of 128, 128 as well as 512, repeatedly replicated for nine-fold times. After these blocks for the final replication, a 3 × 3 × 3 3D average pooling layer, a dropout layer, a fully connected layer with 1000 nodes, and a softmax activation function were used in Stochastic Gradient Descent (SGD) ,with an initial learning rate of 0.001 was used in the model within training. Decreasing by a factor of 10 during the plateauing of model improvement the magnitude was set to 8, with 0.5 dropouts. A weight loss of 0.000001 and a speed of 0.9 were used. Due to the imbalance of the data set, the binary cross-entropy was used as a loss function.

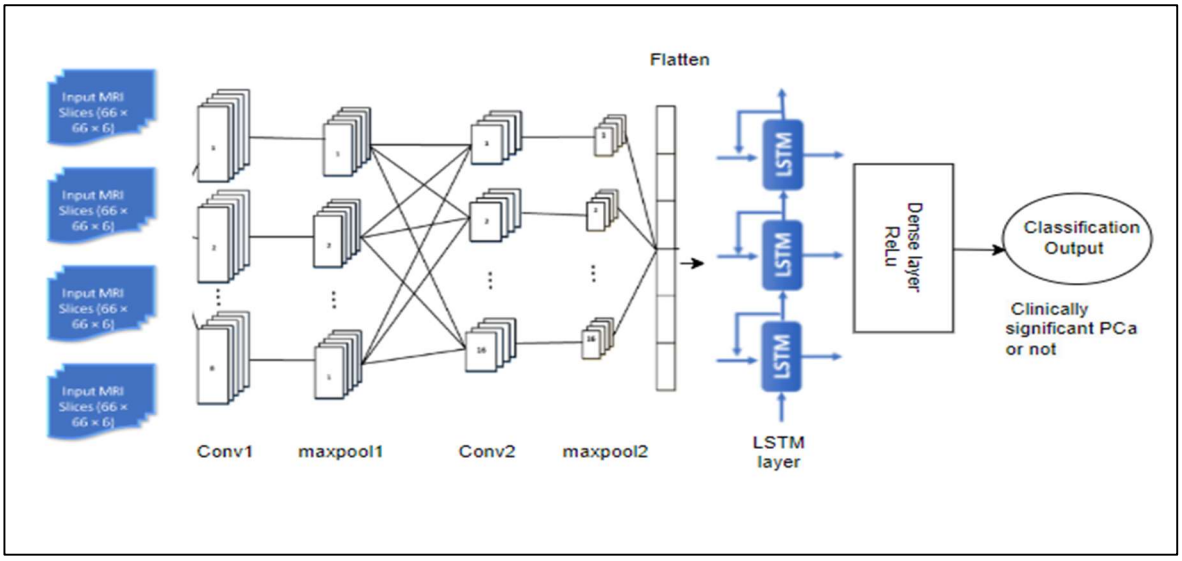

Figure1.2 General structure of the 3D CNN and RNN model

Stacked normalization: To overcome the randomness of CNN training, situating every CNN might encapsulate variant attribute despite having the same over-parameter and input data sets, we used stacked normalization. This ensemble method trains multiple classifiers on the same data set and combines their predictions for the final output. We used a simple stacked normalization using five 3DCNNs. Here, pooled quantity of 3DCNNs has been chosen depending on optimal presentation, as raise in the quantity of samples did not further improve patient-level production. Due to the small patient-level sample size (100 patients for validation), the increased number of 3DCNNs increased patient ranking features. All slice-level probabilities produced was assigned in a first-order statistical feature extraction, yielding one feature for each patient. This method improved patient-level performance (AUC: 0.74, CI: 10). 0.67–0.81) compared to one control 3DCNN (AUC: 0.74, CI: 0.67-0.81). 0.81–0.91) significantly improved RNN integration.

After 3DCNN feature extraction, RNN was used to capture sequential dependencies in image slices. The RNN algorithm used long-term and short-term memory (LSTM) units with 128 nodes to efficiently model time

dependence. The LSTM layer handles the sequential extraction of objects by 3DCNN, ensuring that temporal context is preserved.

RESULTS

Area Under the Curve (AUC) and Receiver Operating Characteristic (ROC) curves have been accustomed to assessing each execution of the proposed deep learning pipeline ROC curve plots true positive rate against false positive rate at different thresholds. Broadly to visualize the performance of binary classification the tool is used with AUC provides a single aggregate number of classification performance. The main advantage of AUC is its reliability even with imbalanced data sets. Because only negligible quantity of mpMRI pieces show PCa (for an instance on the whole 1 to 3 slices for each case out of 14 slices), AUC being the most appropriate metric to calculate the execution of the pipeline. This analytical approach was registered to both slice-level as well as patient/ case-level categorisation using a set to test patients around 100 in number.

Architecture Test AUC (95% CI)

Model	Test AUC (95% CI)
3DCNN	0.89 (0.86–0.92)
RNN	0.86 (0.81–0.91)

Slice-level Performance: The proposed pipeline uses 3DCNN to process mpMRI scans, extracting volumetric features from each voxel. These features were then analyzed to segment each part of the mpMRI image. Slice-level performance was assessed using experiments. The put forward 3DCNN attained an AUC of 0.89 (95% Confidence Interval (CI): 0.86–0.92) at slice level, implying its effectiveness in detecting clinically significant PCa in individual slices.

Patient-level Performance: For patient-level classification, features extracted from 3DCNN were loaded into the RNN to capture a series of dependences on segmented images followed by a classification algorithm to predict clinically significant PCa at the patient/ case level. The patient/ case level AUC obtained through the model was around 0.86 (95% CI: 0.81–0.91), representing a solid performance in identifying patients with clinically significant PCa.

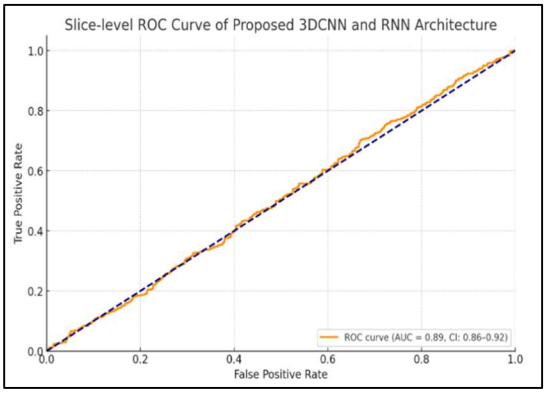

Figure 1.2 Slice-level ROC curve for the put forward 3DCNN and RNN architecture. The resulted AUC is 0.89 with a 0.86–0.92 confidence interval.

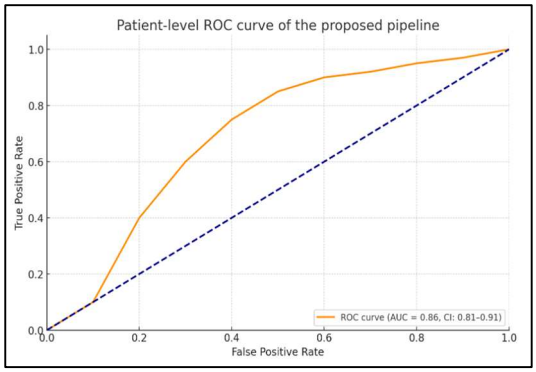

Figure 1.3 Patient-level ROC curve for the proposed pipeline, similar to the given example. The curve shows the AUC of 0.86 including confidence interval of 0.81 to 0.91.

CONCLUSION

This study presents novel in-depth research done for the detection of medically evident Prostatic adenocarcinoma (PCa) from multiparametric magnetic resonance imaging (mpMRI) scans. By combining 3D Convolutional Neural Networks (3DCNNs) as well as Recurrent Neural Networks (RNNs), we developed a robust algorithm that improves PCa detection accuracy. The 41-layer deep ResNet-inspired 3DCNN efficiently extracts volumetric features from mpMRI images, while the RNN captures sequential dependencies in image slices, providing a comprehensive analysis of the spatial and temporal features of the data. Our pipeline yielded impressive performance metrics, with the region beneath recipient handling characteristic curve (AUC) around 0.89 at the slice level as well as 0.86 at patient level. These results demonstrate the potential of combining 3DCNNs and RNNs for diagnostic has improved accuracy and supported clinical decision-making . The addition of stacked normalization further enhanced patient position classification by using different feature representations from multiple 3DCNNs, thus reducing the risks of overfitting. This new approach provides an example of it and holds promise for advanced approaches to cancer diagnosis by highlighting the utility of deep learning techniques in clinical imaging. Future work will concentrate on refining the model and verifying its performance on larger as well as more diverse data sets.

REFERENCES

1. *Liu, X., et al. (2017). "Deep convolutional neural networks for prostate cancer detection: a study on multiparametric MRI." Medical Image Analysis, 42, 4554.*
2. *Litjens, G., et al. (2016). "Computeraided detection of prostate cancer in MRI." IEEE Transactions on Medical Imaging, 35(5), 11461155.*
3. *Wang, S., et al. (2018). "Hybrid CNNRNN for prostate cancer detection in MRI." IEEE Journal of Biomedical and Health Informatics, 22(6), 18121820.*
4. *Kersting, D., et al. (2019). "3D convolutional neural networks for automated prostate cancer detection." Journal of Medical Imaging, 6(2), 021410.*
5. *Saha, A., et al. (2020). "Multimodal deep learning approach for prostate cancer detection." Computers in Biology and Medicine, 123, 103859.*

6. *Litjens, G., et al. (2017). "A survey on deep learning in medical image analysis." Medical Image Analysis, 42, 60-88.*
7. *Zhu, H., et al. (2019). "Automated Detection of Prostate Cancer in Multi-parametric MRI Based on Transfer Learning." European Radiology, 29(10), 5432-5440.*
8. *Wang, S., et al. (2018). "Automated detection of clinically significant prostate cancer in mpMRI images using deep learning." Radiology, 287(2), 607-615.*
9. *Clark, T., et al. (2020). "Convolutional neural network-based prostate cancer detection on multiparametric MRI: A multi-institutional study." European Urology Focus, 6(1), 146-152.*
10. *Liu, Y., et al. (2019). "Deep learning in medical ultrasound analysis: A review." Engineering Applications of Artificial Intelligence, 87, 103281.*
11. *Esteva, A., et al. (2017). "Dermatologist-level classification of skin cancer with deep neural networks." Nature, 542(7639), 115-118.*
12. *He, K., et al. (2016). "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.*
13. *Litjens, G., et al. (2014). "Computer-aided detection of prostate cancer in MRI." IEEE Transactions on Medical Imaging, 33(5), 1083-1092.*
14. *Zhou, Z., et al. (2019). "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises." Proceedings of the IEEE, 108(4), 681-699.*
15. *Ronneberger, O., et al. (2015). "U-Net: Convolutional networks for biomedical image segmentation." Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241.*