International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I12P224

Volume 70 Issue 12, 252-271, December 2022

Original Article

Prostate Gland Segmentation using Semantic Segmentation Models U-Net and LinkNet

M. N. Rajesh[1], B. S. Chandrasekar[2]

[1,2]*Department of Electronics and Communication Engineering, Faculty of Engineering & Technology, Jain (Deemed-to-be University), Bangalore, Karnataka, India.*

[1]*Corresponding Author : rajeshmn.mn@gmail.com*

Received: 17 August 2022 Revised: 12 November 2022 Accepted: 16 December 2022 Published: 24 December 2022

Abstract *- The segmentation and classification of the prostate lesion or malignant growth through manual observation are highly challenging. Machine learning-based semantic segmentation architecture was used to segment the diseases based on lesion appearance and characteristics automatically. Still, those models consume more energy and processing time and will lead to reduced scalability and reliability. To tackle these limitations, deep learning (DL) based semantic segmentation architecture can be implemented, which has more advantages in discriminating the features of the lesions efficiently and accurately. This paper proposes DL-based semantic segmentation models such as LinkNet, U-Net, PSPNet and FPN. These proposed segmentation models are integrated with convolutional neural network (CNN) based backbone architectures like ResNet-34 and SE-ResNet-34. Initially, the nearest neighbour interpolation technique is employed as preprocessing technique for scaling the image. Next, normalization of intensity is employed to minimize the variations in the intensity distributions of the image. Normalized image is used for processing various settings of LinkNet and U-Net architectures. Furthermore, the proposed model uses hyper-parameter optimization with optimizers such as Adam, Adamax, Stochastic Gradient Descent, RMSProp, and Nadam for both U-Net and LinkNet to minimize the complexity of the network and enhance computing efficiency. Experimental results have been evaluated using Python on Google Colab with NCI-ISBI 2013 dataset. Performance analysis of the proposed model is assessed in terms of the Intersection of the Union (IoU) score. The LinkNet with SE-ResNet-34 model optimized using Adamax generated the best result with 0.7454453 IoU score, following U-Net with SE-ResNet-34 optimized using Adamax generated 0.738271933 IoU.*

Keywords *- Deep learning, FPN, Prostate cancer, PSP-Net, LinkNet, ResNet-34, Semantic segmentation, SE-ResNet-34, U-Net.*

1. Introduction

The developing biomedical image acquisition system has shifted the research community's focus toward the convention of disease diagnosis that does not involve invasive procedures. Every diagnostic test requires a comprehensive and analytical review of the patient's medical scans, which depict the intricate internal structure of the body and show how its various organs function. The medical field has witnessed an exponential expansion in the number of diagnostic practices due to the availability of a wide range of medical imaging techniques, including X-ray, magnetic resonance image (MRI), computed tomography (CT), etc.

The imaging process, types of applications, and typical time consumption for diagnosis are all different for each of these scans. Deep learning (DL) methodologies are analyzed well to consider the automatic support in diagnosis procedures, which results in a quick and best way to cure, monitor, and treat the disease. Because analysing such complex scans is laborious and time-consuming for any radiologist, DL methodologies are being explored to fill this void of complexity. Segmentation is one of the automation tasks that help identify and discover the required items or regions of interest for the issue. Segmentation can be broken down into two stages, known as instance and semantic, depending on the level of specificity with which the classes of objects are identified. Semantic segmentation separates the items into their respective classes; however, instance segmentation goes one step further and separates the objects that belong to the same class [1].

Prostate cancer (PCa) is a complex prevalent form of tumor caused in males due to instability and accumulation of multiple molecular alterations. Molecular alteration of normal cells becomes malignant in the form of a lesion with irregular appearance and boundaries [2]. Diagnosis of PCa can be made using non-invasive techniques such as MRI and CT. However, accurate diagnoses of the prostate lesion are difficult, error-prone and time-consuming due to the

heterogeneous appearance, boundaries and irregular shapes of lesions [3]. Manual lesion segmentation and classification are highly intensive and challenging on features with high intra-class variation and low interclass variations [4].

1.1. Problem Statement & Solution

Machine learning (ML) based semantic segmentation algorithm such as conditional random field [5] and gray level segmentation [6] has been employed to automatically segment the diseases based on the lesion appearance and its characteristics on shape, size and border into benign and malignant. The ML model is incapable of representing the complex structures of the malignant features, and these models are time-consuming and lead to reduced scalability and reliability. In addition, this model produces limited discriminative capability and is less adaptable to lesion boundary changes in the various categories of lesion features. In order to tackle those limitations, DL-based semantic segmentation architecture [7] has been exploited as it is more advantageous in discriminating the features of the lesions efficiently and accurately [8].

In this paper, various optimization techniques have been employed for the DL-based semantic segmentation on convolution network architectures such as U-Net, LinkNet, PSP-Net and FPN on the backbone of CNN such as ResNet-34 and SE-ResNet-34 for PCa lesion segmentation. Initially, image interpolation and intensity normalization were employed as preprocessing techniques for image resizing and image normalization for reducing variations in the intensity distributions of the image. Further, those processed images are applied for processing in the various setting of U-Net [9] [34], LinkNet [10] [22], PSP-Net and FPN architecture which is a full convolution decoder and encoder network with the skip connections among the encoder and decoder blocks using ResNet-34 [11] and SE-ResNet-34 [12]. Furthermore, the proposed model uses hyperparameter optimization with various optimization algorithms such as Adam, Adamax, SGD, RMSProp, and Nadam for both U-Net, LinkNet, PSP-Net and FPN to enhance prostate boundary detection.

The remaining article has been sectioned as follows; section 2 details the problem statement and literature review for lesion classification. Section 3 provides the proposed semantic segmentation architectures using CNN models for segmenting the lesions. Experimental analysis of the proposed methodology on the ImageNet dataset has been carried out in Section 4, along with performance analysis based on IoU Score in various settings. Finally, Section 5 concludes the research with future recommendations.

2. Related Works

In this part, numerous conventional models have been employed for lesion semantic segmentation as an automated system for the analysis of MRI by utilizing ML. The DL model has been detailed as follows.

A conditional random field is most effective in semantic segmenting for lesion classification. The classification process has been processed with preprocessing of the MR images using interpolation techniques [13]. Next, preprocessed image is segmented using a finite set of possible states using the cost function. The prior relationship among pixels [14] was to extricate the multiple attributes of the images. Those computed attributes were utilized for training the classifiers. The classifier model classifies the feature into a malignant class or benign class.

Gray-level segmentation is employed to segment the lesions of the prostate. The segmentation process has been carried out after preprocessing the images using interpolation and normalization [15]. Those processes generate the contrast-enhanced lesion for effective feature segmentation with reduced variation between dataset images. Segmented images are applied to KNN classification [16] on extracted features to categorize the normal and malignant lesions with higher accuracy and efficiency in the neighbour context.

The segmentation of the PCa on an MRI is an essential step in the process of diagnosing and determining how well treatment is working. However, automatic segmentation is difficult since there is no distinct prostate boundary, heterogeneity inside the prostate tissue, a wide range of prostate shapes, and a shortage of training data has been annotated. To segment the prostate automatically in an accurate and reproducible way, [41] presented a two-stage segmentation methodology using the combination of U-Net and residual blocks. This method included a two-stage neural network. The first was a 2D U-Net, used to identify an approximation of the prostate's position, and the second was an integration of U-Net and Res-Net, which was used to segment the prostate accurately.

An analysis of encoder–decoder CNN was proposed in [18] for the prostate glands of segmentation using T2W MRI images. SegNet, FCN, U-Net, and DeepLabV3+ were chosen to be the four selected CNNs. DeepLabV3+ was initially offered for segmenting of natural and biomedical images. In [19], a deep neural network, watershed, multi-atlas, and SSM-based reconstruction model was presented to segment the PCa in MRIs. DNN determines the ROI for the prostate and eliminates the least relevant pixels. It reduces the amount of time it takes for the code to run while simultaneously improving the accuracy of the registration. The watershed technique helped enhance the outcome in the base and apex regions, and the dynamic probabilistic atlas was able to produce more exact initial segmentation as a result. The findings that were achieved were improved with the application of the dynamic SSM technique.

In [20], an automated MRI segmentation technique was proposed that uses DL architectures. CASE-Net is an architecture developed to improve segmentation performance by combining the benefits of adjustable feature weights with attention gating. Using a sequential model, the CASE-Net, U-Net, USE-Net, and Link-Net were each manually generated in Keras. It was proposed in [21] that a multisite network, also known as an MS-Net, could improve prostate segmentations by learning more robust representation and using numerous different data sources. In order to account for inter-site variability present in the various MRI datasets, the domain-specific batch normalization layers were added to the network's backbone. It allowed the networks to evaluate statistics and independently normalise site features. A multi-site-guided knowledge transfer was presented to enhance kernels for extracting more generic descriptions from multisite information. This was done in light of the challenge associated with capturing the shared information contained within various datasets.

The prostate gland and prostate lesions could be segmented using a radiomics-based, highly supervised U-Net model proposed in [22]. The most widely used SGD approach, along with the Adam optimization strategy as an additional hyperparameter, was utilized in the training of this model. In the various stages of prostate cancer, this pipeline model has been utilized to segment the prostate capsule and prostate lesion.

An automatic deep learning system called prostate segmentation (PSNet) was developed in [23] to segment the PCa on an MRI. This deep CNN approach was trained from endwise in a single learning step, and the inputs that it was given were prostate MRIs and the ground facts that corresponded to them. An inference for pixel-wise segmentation is something that can be accomplished by using the learnt CNN model. In [24], a CNN architecture called the USE-Net model was proposed for the zonal prostate segmentation of MRI. Squeeze-and-Excitation (SE) blocks have been incorporated into U-Net in this architecture. In particular, the blocks of SE were included after each encoder and encoder-decoder block.

2.1. Research Gap and Analysis

The segmentation of the PCa using MRI is a crucial step in the process of adaptive radiotherapy and radiomics research. The goal is to determine which imaging characteristics are associated with which patient outcomes. In this research, DL approaches such as LinkNet, U-Net, PSPNet, and FPN are proposed. These DL approaches aim to tackle the prostate gland's real-time, fully automated delineation process on T2-w-MRI. Delineation manually is a time-consuming operation, which is why this research proposes these approaches. While U-Net is used in various applications for delineating medical images, LinkNet, PSPNet, and FPN are employed in applications for object detection.

3. Proposed Methodology

In this section, deep semantic segmentation architectures named U-Net, LinkNet, PSP-Net and FPN, and ResNet-34 and SE-ResNet-34 as encoder backbones for prostate gland segmentation are adopted for the prostate gland MR image. This architecture is modelled to segment the prostate gland based on its appearance.

3.1. Image Preprocessing

Images may contain various sizes; the following image preprocessing technique must be employed to achieve common image sizes.

3.1.1. Nearest Neighbor Interpolation Technique

It has been employed to resize an image by computing an average value to the data points in the image vector by weighting criteria. Initially, NCI-ISBI 2013 data set [25] was represented in different sizes (256×256), (320×320), and (384×384). It can be resized by performing row-wise and column-wise interpolating the image matrix by normalizing the row and column-wise pixel positions using the ceil function.

3.2. Semantic Segmentation

It is referred to as pixel-wise classifications in the process of assigning labels to every individual pixel into object class in the image. A network of semantic segmentation is composed of two symmetrical paths convolution parts and deconvolution parts. The convolution part down-samples the input image, and the deconvolution part will up-sample the down-sampled features with a fully connected layer. Finally, the Softmax layer was utilized to label the segmented features of the input image. The novelty of the work includes the semantic segmentation of prostate MRI with various deep models.

3.2.1. U-Net-ResNet-34

U-Net is a pretrained network model consisting of an encoder, bottleneck convolution layer and decoder. The encoder network has 34 convolution blocks that encrypt the input images into the representation of features at various levels with the features map. Feature maps will be increased in each convolution block to efficiently learn the complicated shape of the lesions. The block of bottleneck convolution transforms the features learned out of the encoder network into decoder networks. Decoder networks semantically transform the features of discrimination to pixel space to obtain the dense network. Fig. 1 represents the proposed architecture of the U-Net for prostate classification applied in this work [1].

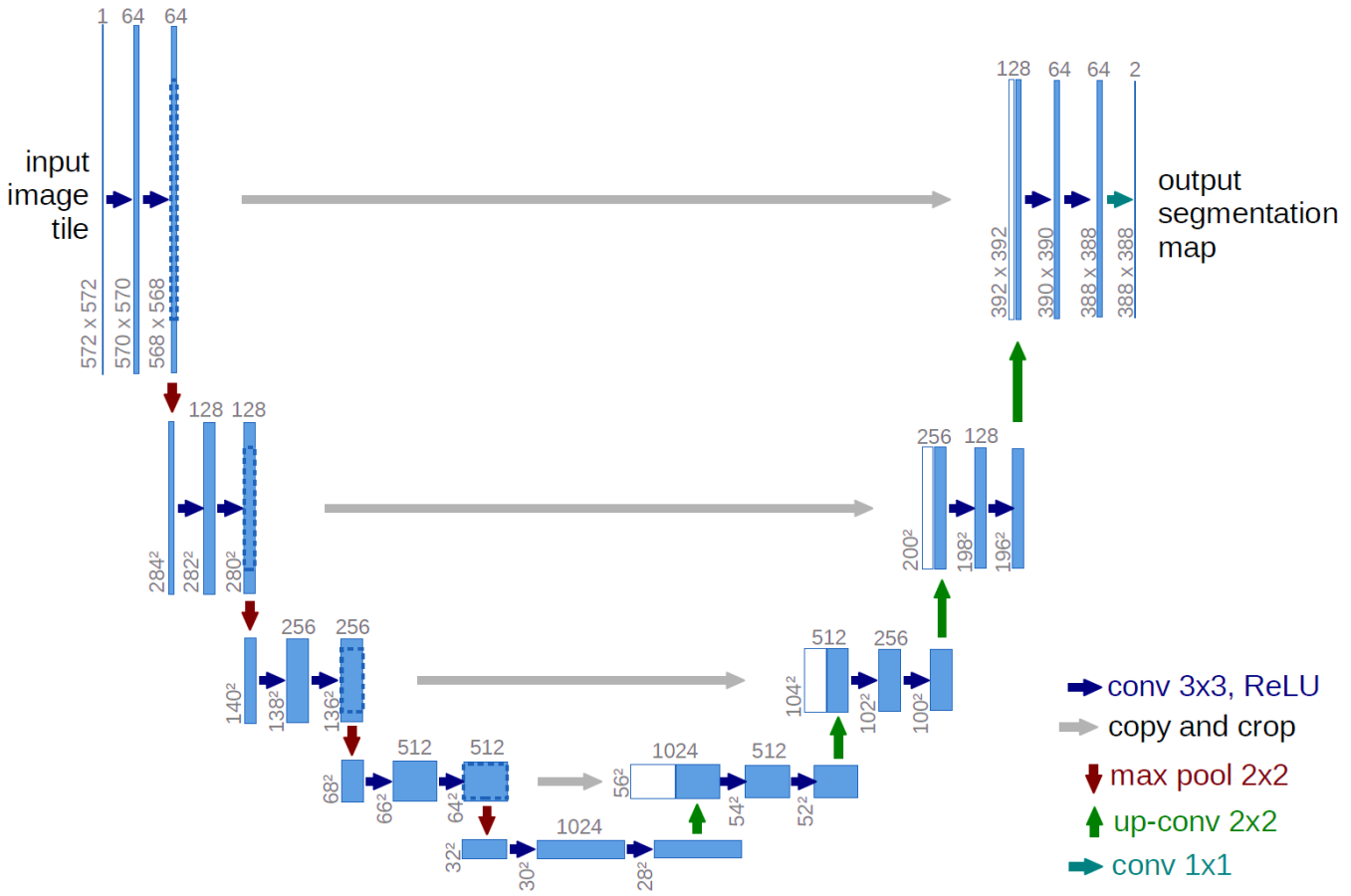


Fig. 1 U-Net Architecture for ResNet-34 Layer (Ronneberger et al, 2015).

3.2.2. Residual Neural Network

An increase of layers in the neural network will be transformed into a dense network, which leads to a vanishing gradient problem. The network becomes deeper due to a change of weights in the layers to the batch of training inputs. The ResNet network was employed to reduce the various layers' tuning, which can skip one or more layers to decrease the depth of the network. In this work, ResNet-34 was employed as an encoder network to U-Net architecture to enable feature propagation gradient flow and solve the vanishing gradient issues [26].

3.2.3. Encoder Network of U-Net

In this, ResNet-34 is a CNN composed of 34 layers deep. It has 16 residual blocks. Each block-down samples the images in each layer. It is considered a contraction path of the U-Net.

3.2.4. Decoder Network of the U Net

Decoder Network is considered the expansive path with a decoder block as residual connections are incorporated as it up-samples the down-sampled features. The up sample of the features is carried out in transposed convolutions for expanding the convolution size. The images were concatenated with the appropriate images from the contracting paths for combining the data of the past layers to obtain the accurate prediction of lesion boundaries and good spatial locations.

Block 1: The transposed convolutions (TC) were utilized with the 3*3 kernels and the stride of two residual blocks to hold 256 feature maps. The block's output was concatenated with the output of the contraction block. Dropout was included with the 0.5 drop rate. A 2D Convolution layer was utilized with the 3*3 kernel sizes in addition to padding to keep input and output sizes as same. At last, the function of sigmoid activation was utilized.

Block 2: TC was utilized with the 3*3 kernels and the stride of two residual blocks to hold a reduced 128 feature maps. The block's output was concatenated with the output of the contraction block. Dropout was included with the 0.5 drop rate. The 2D Convolution layer was utilized with the 3*3 kernel sizes and padding to keep input and output sizes. At last, the function of sigmoid activation was utilized.

Block 3: TC was utilized with the 3*3 kernels and the stride of two residual blocks to hold a reduced 64 feature

maps. The block's output was concatenated with the output of the contraction block. Dropout was included with the 0.5 drop rate. A 2D Convolution layer was utilized with the 3*3 kernel sizes in addition to padding to keep input and output sizes as same. Then, the function of sigmoid activation was utilized.

Block 4: TC was utilized with the 3*3 kernels and the stride of two residual blocks to hold a reduced 32 feature maps. The block's output was concatenated with the output of the contraction block. Dropout was included with the 0.5 drop rate. A 2D Convolution layer was utilized with the 3*3 kernel sizes in addition to padding to keep input and output sizes as same. Then, the function of sigmoid activation was utilized.

Block 5: TC was utilized with the 3*3 kernels and the stride of two residual blocks to hold a reduced 8 feature maps. The block's output was concatenated with the output of the contraction block. Dropout was included with the 0.5 drop rate. A 2D Convolution layer was utilized with the 3*3 kernel sizes in addition to padding to keep input and output sizes as same. Then, the function of sigmoid activation was utilized [27]. Thus, the output feature map was reduced to a single value using 1*1 convolution at the output layer, and the sigmoid activation function was applied to the output.

3.2.5. LinkNet-ResNet-34 Architecture

LinkNet is a pretrained network consisting of the encoder and decoder networks. The encoder block transforms the image into a feature map as it contains the convolution block. Feature maps learn the complex structures of the lesion. In this work, ResNet-34 and SE-ResNet-34 have been employed as the backbone [28].

3.2.6. Decoder Network of the LinkNet

The decoder network is considered with the decoder block as a residual connection connected to the related encoder blocks. The transferred blocks from the encoder were included in the related decoder block, which holds the feature map. Dropout was included with the 0.5 drop rate. All the decoder blocks included a 1 × 1 convolution function which minimizes the feature map using 4 filters, subsequently transposed convolution and batch normalization to upsample the feature maps. Finally, the sigmoid activation function precisely predicts lesion boundaries and better spatial locations [31].

3.2.7. LinkNet-SE-ResNet-34 Architecture

LinkNet is a pretrained network consisting of the encoder and decoder networks for semantic segmentation of the lesion. In this architecture [28], the encoder network containing convolution transforms the image into a feature map by down-sampling and the decoder network up-samples the feature map to yield better-predicted results.

3.2.8. SE-ResNet-34

SE-ResNet network has been employed as it can identify the interdependencies among the feature in the decoder block using the feature recalibration approach. Notably, the model automatically acquired the significance level of all the feature blocks through learning approaches for decreasing the depth of the network, and then according to the degree of importance using the sequence operation and excitation operation on the residual block of the decoder block.

3.2.9. Encoder Network of LinkNet

In this, ResNet-34 is a CNN composed of 34 layers deep. It has 16 residual blocks. Each block-down samples the images in each layer. Initially, kernel size of 7*7 and stride 2 performed convolution, and it was sequenced using max-pooling with a 2-stride. Finally, the network is composed of repetitive residual blocks. The initial convolutional operation was applied with stride two to provide the down-sampling since the convolution operations utilize stride one.

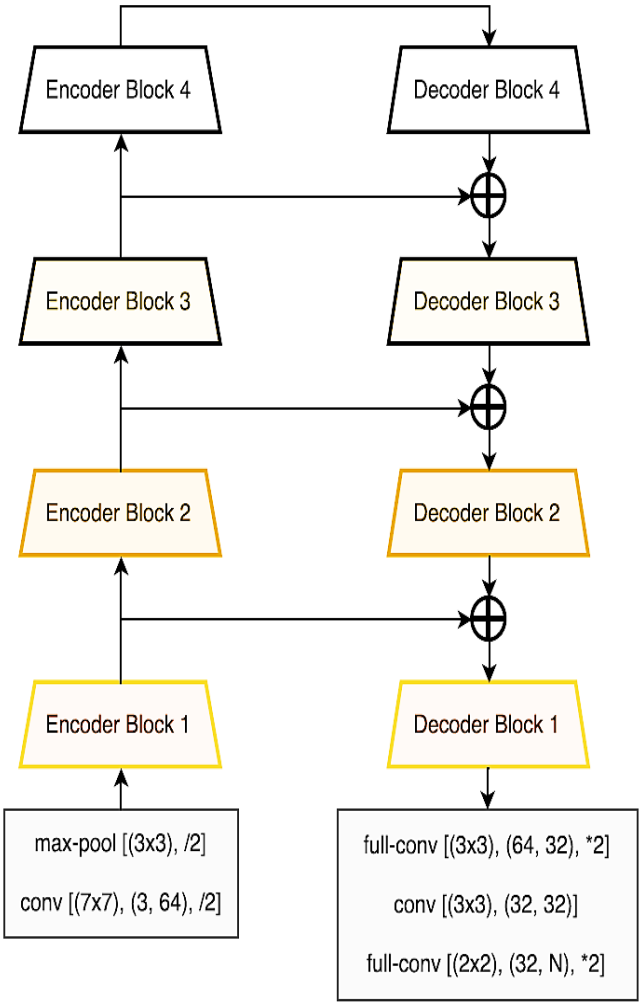


Fig. 2 LinkNet Architecture

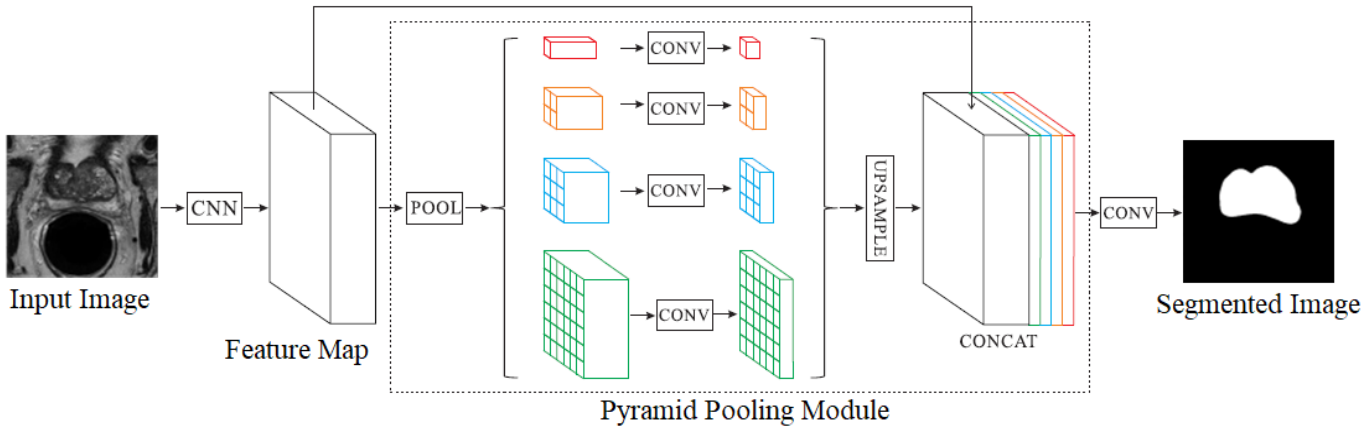


Fig. 3 PSP-Net Architecture

3.2.10. Decoder Network of the LinkNet

Decoder Network is considered with the decoder block as a residual connection connected to the corresponding encoder block. The transferred blocks from the encoder were included in the related decoder block, which holds the feature map. Dropout was included with the 0.5 drop rate. All the decoder blocks included the 1 × 1 convolution operations, which minimizes the feature map using 4 filters, sequenced by transposed convolution and batch normalization to up-sample the feature maps. Finally, the sigmoid activation function precisely predicts lesion boundaries and better spatial locations.

3.3. PSP-Net

Pyramid Scene Parsing Network, also known as PSP-Net, is a semantic segmentation model that uses a pyramid parsing module (PPM). This module makes use of global context information using different region-based context aggregation. To acquire the feature maps from the input images, PSP-Net employed a CNN that was pretrained and utilized the dilated networks methodology. The final feature map size was equal to one-eighth of the input image. The PPM was used at the very top of the map to collect information about the map context. When utilizing the four-level pyramids, the pooling kernels covered the overall image and half and tiny regions. They were combined to form global priors. After, the earlier initial feature map was combined into the subsequent section of the final part. After this layer, a convolution layer will be applied to produce the final prediction map.

Fig. 3 illustrates the architecture of the PSP-Net, which is created based on the PPM [29]. A CNN model that had been pre-trained with the dilated network strategy was employed to extract the feature map from an image provided as input. For the purpose of elucidating this structure, PSP-Net offers an efficient global contextual prior that may be used for scene parsing at the pixel level. The PPM can collect information on multiple levels, which is more representative than the global pooling module. Compared to the initially dilated FCN network, the PSP-Net does not significantly increase the computing effort required. When doing end-to-end learning, it is possible to optimize both the global PPM and the local FCN features simultaneously [29].

3.4. FPN

FPN, known as a Feature Pyramid Network, takes single-scale images of random size as input and produces a correspondingly sized feature map at many layers in a completely convolutional way. This type of feature extractor is also known as an FPN. This technique is not dependent on the convolutional backbone architectures in any way. Consequently, it acts as a general-purpose model for developing feature pyramids within the deep CNNs, which may be applied to problems such as object detection. The pyramid's building requires both a top-down and a bottom-up pathway to be completed. The feedforward computation of the backbone ConvNet is the bottom-up pathway. This pathway computes a feature hierarchy that consists of feature mappings at multiple scales with a scaling step of 2. Each step of the feature pyramid corresponds to a single level of the pyramid.

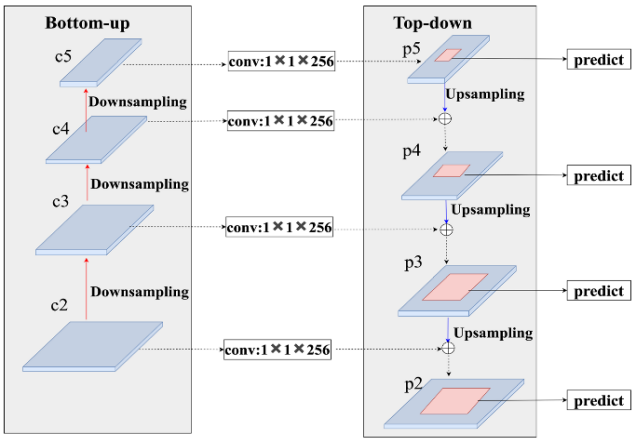


Fig. 4 FPN Architecture

Each stage's final layer output is used as a reference set of feature maps. Regarding ResNet, the feature activations output was used by the final residual block of each stage. Through upsampling spatially coarser but semantically stronger feature maps from higher pyramid levels, the top-down pathway can perceive features with a better resolution than other pathways. After that, features from the bottom-up pathway are added to these characteristics through lateral connections to improve them. Each lateral link combines feature maps from the bottom-up and the top-down pathways that are the same size in terms of spatial extent. The bottom-up feature maps may have low-level semantics, but their activations have been localized with greater precision due to the reduced number of times it was subsampled [30].

3.5. Optimization of Models

Optimization of the U-Net, LinkNet, PSP-Net and FPN is carried out using Adam, Adamax, SGD, RMSProp and Nadam to reduce the error rates for the training process and parameter tuning. Adam Optimization techniques estimate the learning rates for every parameter included in training gradients to reach an exact result with high dice similarity coefficient. Stochastic Gradient Descent (SGD) optimization with strategic relapse produces outputs by getting inputs for producing actual results with the weight parameters of the model. Nesterov adaptive momentum uses the correction factor to reach the exact result. Root Mean Square Propagation (RMSProp) uses the squared gradient's moving average, which utilizes the recent gradient descent's magnitudes to normalize the gradients. Thus, with the increased learning rate, the algorithm could shift in the horizontal direction with major steps converging faster [33].

3.6. Loss Function

The loss function of weighted cross-entropy was selected for the architecture of CNN as it can differentiate between the prostate pixels and penalizes the predicted outcome of the results on gradient descent from the exact value. It is formulated as,

$$LF = -\frac{1}{n}\sum_{i=1}^{n} W_{c,i}[T_i \log p_i + (1 - T_i)\log(1 - p_i)] \quad (1)$$

Where, p_i was predicted, segmentation class, $W_{c,i}$ is the encoder weight and T_i is the target segmentation label.

3.7. Backbone

The proposed models use pre-trained models as the backbone, such as ResNet-34 and SE-ResNet-34, for PCa gland segmentation. Backbone refers to a feature-extracting network usually used within architecture. This feature extractor encodes the input of the network input in the encoder part, from which the decoder section will be built up based on the program as part of the transfer learning process and into a certain feature representation. Hence based on the semantic segmentation model proposed herewith, for example, if U-Net is used, then it first loads the backbone with pre-trained ImageNet weights from the applications of Keras and thus, a function creates the decoder section by concatenating the prior outputs on the decoder section with the outputs from the related layers named in the list of skip connections and integrating more convolutions. ResNet-34 is a type of CNN that consists of a collection of residual blocks with skip connections. The ResNet-34 consists of a 34-layer ResNet architecture. The model is faster to train and uses less memory. It is used for the encoder/down-sampling section of the U-Net model. The ResNet model has to skip connections, allowing large layers to be skipped when needed. It results in the model being trained with optimal weights, thereby reducing the loss during downsampling in the segmentation model. SE block is a form of CNN that explicitly models the interdependencies across channels to recalibrate channel-wise feature responses adaptively. These blocks, when stacked together, can generate SE-Net designs that generalize exceedingly effectively across a wide variety of datasets. These architectures can be created by stacking the blocks. Even in the current results, it was observed that the SE blocks deliver significant performance improvements for existing semantic segmentation [34].

4. Experimental Results

Experimental results of the proposed model have been evaluated with Python program using Google Colab Notebook and with NCI-ISBI 2013 dataset containing 2276 images of various sizes as (384×384), (320×320) and (256×256) [27]. In processing the image, the data set is portioned into a training set containing 1744, a testing set containing 261 and a validation set containing 271. The performance considering IoU measure of semantic segmentation with U-Net, LinkNet, PSP-Net and FPN architectures in combination with backbones like ResNet-34 and SE-ResNet-34 has been evaluated while fixing hyper-parameter such as Batch Size, Validation Steps and variation in epoch values and the model was tuned with encoder weights set to "ImageNet" and activation parameter set to sigmoid with classes set to binary and varying optimizers.

4.1. Results Obtained with U-Net

This section describes various approaches for Prostate Semantic Segmentation using U-Net with ResNet34 and SE-ResNet34 as the backbone and the performance results obtained. Fig. 5 shows the Original Prostate Gland Image along with the mask that was provided as against the predicted mask by for

i. ResNet34 as backbone and Optimizer being Adam as in Fig. 5
ii. ResNet34 as backbone and Optimizer being Adamax as in Fig. 5
iii. ResNet34 as backbone and Optimizer being SGD as in Fig. 5
iv. ResNet34 as backbone and Optimizer being RMSProp as in Fig. 5
v. ResNet34 as backbone and Optimizer being Nadam as in

Fig. 5

vi. SEResNet34 as backbone and Optimizer being Adam as in Fig. 5

vii. (vi), SEResNet34 as backbone and Optimizer being Adamax as in Fig. 5

viii. SEResNet34 as backbone and Optimizer being SGD as in Fig. 5

ix. SEResNet34 as backbone and Optimizer being RMSProp as in Fig. 5 and

x. SEResNet34 as backbone and Optimizer being Nadam as in Fig. 5

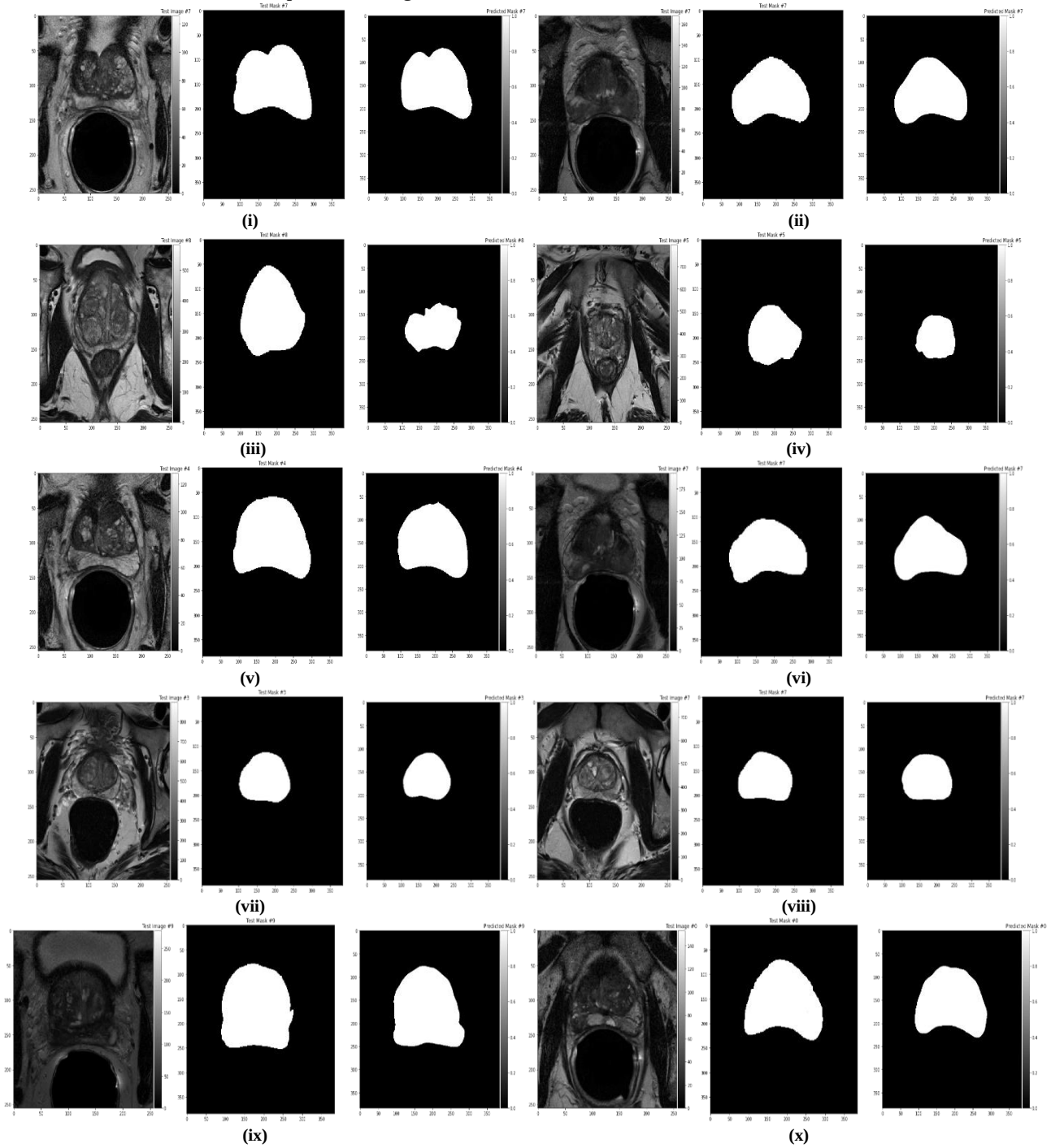


Fig. 5 Prostate MRI Image along with the ground truth or provided mask and the predicted mask are listed for U-Net model considering (i) ResNet34 as backbone and Optimizer being Adam, (ii) ResNet34 as backbone and Optimizer being Adamax, (iii) ResNet34 as backbone and Optimizer being SGD, (iv) ResNet34 as backbone and Optimizer being RMSProp, (v) ResNet34 as backbone and Optimizer being Nadam, (vi) SEResNet34 as backbone and Optimizer being Adam, (vii) SEResNet34 as backbone and Optimizer being Adamax, (viii) SEResNet34 as backbone and Optimizer being SGD, (ix) SEResNet34 as backbone and Optimizer being RMSProp, (x) SEResNet34 as backbone and Optimizer being Nadam

Table 1. The table indicates IoU Score obtained with various Optimizers used with U-Net Model with ResNet3-4 and SE-ResNet-34 as Backbone and in combination with various Optimizer

Architecture and Backbones	Adam	Adamax	SGD	RMSProp	Nadam
U-Net – ResNet-34	0.716165932	0.732374634	0.704223167	0.663843288	0.708662823
U-Net – SE-ResNet-34	0.737513949	0.742655936	0.703767031	0.694015578	0.697213093

Table 1 shows IoU Score obtained when testing the U-Net model with ResNet-34 and SE-ResNet34 as the backbone for untrained MRI images. Based on the trials, it can be inferred that the best performance - IoU Score were obtained with SE-ResNet-34 as the backbone with Adamax Optimizer giving 0.742655936, followed by Adam Optimizer giving 0.737513949, followed by Resnet-34 as the backbone with Adamax Optimizer giving 0.732374634, followed by Adam Optimizer giving 0.716165932. The bar chart in figure 6 shows the performance - IoU Score of various combinations of U-Net with backbones and Optimizers. Fig. 6 indicates a Bar Chart with IoU Score obtained with UNet Model in the combination of ResNet34 and SEResNet34 as Backbones plotted against various Optimizers.

Each model was set to 200 epochs for training, 16 as batch size, a scheduler of the iterative learning rate, 1e-3 initial learning rate, and various optimizers and backbones. The performance of the U-Net model-IoU score obtained for training and validation as against the epochs for the best 6 results is shown in Fig. 7. It was observed that for U-Net with ResNet-34 as the backbone and with adamax Optimizer, the training and validation IoU score went up to 0.95995 but IoU score obtained while testing is 0.732374634. For U-Net with SE-ResNet-34 as the backbone and with adam optimizer, the training and validation IoU score went up to 0.94412, but the IoU score obtained while testing is 0.737513949. For U-Net with SE-ResNet-34as backbone and with adamax Optimizer, the training and validation IoU score went up to 0.97588, but the IoU score obtained while testing is 0.742655936.

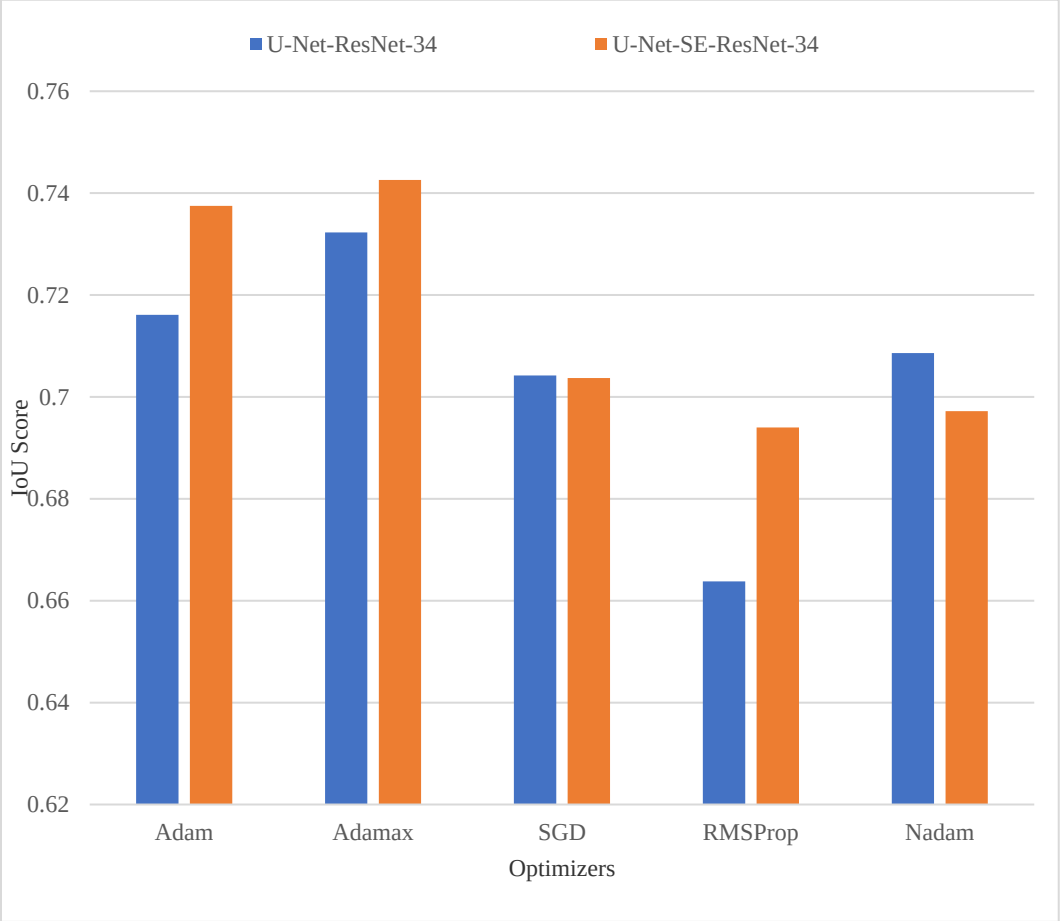


Fig. 6 Bar Chart indicates the IoU score obtained plotted against various optimizers used with the U-Net model in the combination of ResNet-34 and SE-ResNet-34 as Backbones

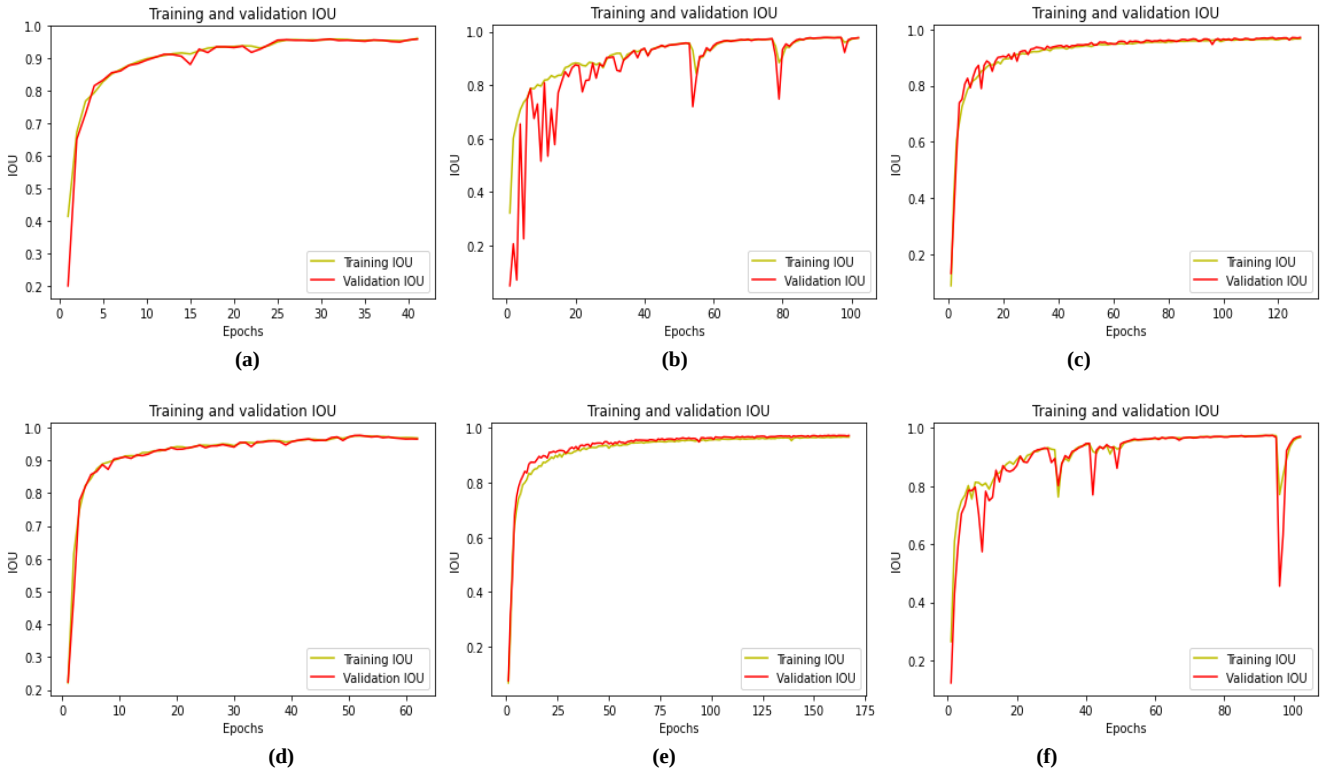


Fig. 7 Training and Validation Performance - IOU Vs epochs for U-Net with a) Resnet-34 as the backbone and Adamax Optimizer, b) Resnet-34 as the backbone and SGD Optimizer, c) Resnet-34 as the backbone and Nadam Optimizer d) SE-Resnet-34 as the backbone and Adamax Optimizer e) SE-Resnet-34 as the backbone and SGD Optimizer f) SE-Resnet-34 as the backbone and Nadam Optimizer

4.2. Results Obtained with LinkNet

This section describes various approaches for Prostate Semantic Segmentation using LinkNet Architecture with ResNet34 and SEResNet34 backbone, along with the performance results obtained.

Fig. 8 shows the Original Prostate Gland Image along with the mask that was provided as against the predicted mask by for ResNet-34 as backbone and Optimizer being Adamax as in Fig. 8 (i), ResNet-34 as backbone and Optimizer being Adamax as in Fig. 8 (ii), ResNet-34 as backbone and Optimizer being SGD as in Fig. 8 (iii), ResNet-34 as backbone and Optimizer being RMSProp as in Fig. 8 (iv), ResNet-34 as backbone and Optimizer being Nadam as in Fig. 8 (v), SE-ResNet-34 as backbone and Optimizer being Adam as in Fig. 8 (vi), SE-ResNet-34 as backbone and Optimizer being Adamax as in Fig. 8 (vii), SE-ResNet-34 as backbone and Optimizer being SGD as in Fig. 8 (viii), SE-ResNet-34 as backbone and Optimizer being RMSProp as in Fig. 8 (ix) and SE-ResNet-34 as backbone and Optimizer being Nadam as in Fig. 8 (x). Table 2 shows IoU Score obtained when testing the LinkNet model with ResNet-34 and SE-ResNet-34 as the backbone for untrained MRI images.

Based on the trials, it can be inferred that the best performance - IoU Score were obtained with SE-ResNet-34 as the backbone and Adamax Optimizer 0.7454453. This result was followed by ResNet-34 as the backbone, with Adamax Optimizer giving 0.738271933, followed by Nadam Optimizer giving 0.712770765, followed by SGD Optimizer with 0.701902515.

The bar chart in Fig. 9 shows the performance - IoU Score of various combinations of LinkNet with ResNet-34 and SE-ResNet-34 backbones and Optimizers.

Table 2. The table indicates IoU Score obtained with various optimizers used with the LinkNet model in the combination of ResNet-34 and SE-ResNet-34 as Backbone

Architecture and Backbones	Adam	Adamax	SGD	RMSProp	Nadam
LinkNet – ResNet-34	0.698139869	0.738271933	0.701902515	0.661088051	0.712770765
LinkNet – SE-ResNet-34	0.687389449	0.7454453	0.697475474	0.686124772	0.693827738

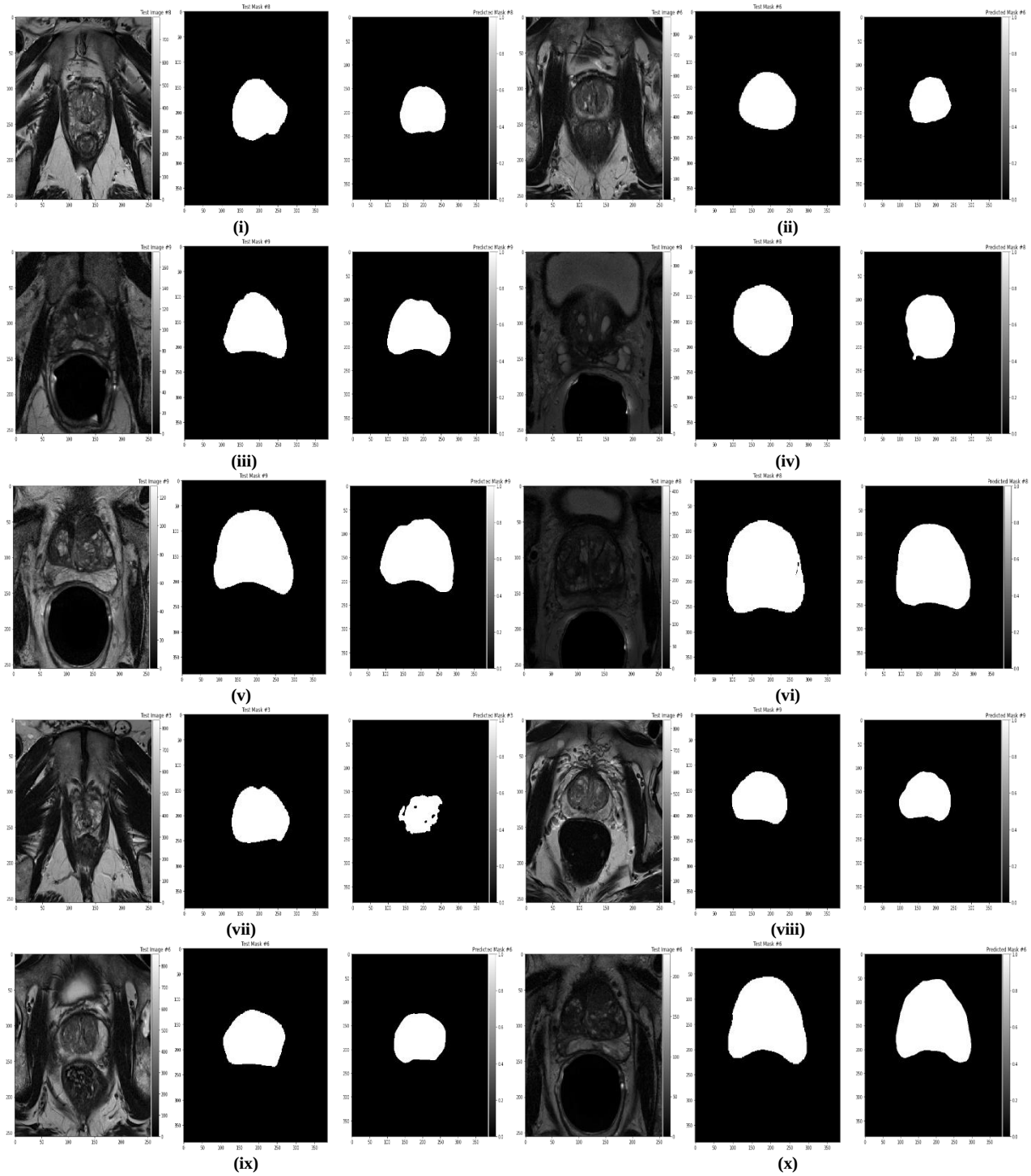


Fig. 8 Prostate MRI along with the ground truth or provided mask and the predicted mask are listed for the LinkNet model considering (i) ResNet-34 as backbone and Optimizer being Adam, (ii) ResNet-34 as backbone and Optimizer being Adamax, (iii) ResNet-34 as backbone and Optimizer being SGD, (iv) ResNet-34 as backbone and Optimizer being RMSProp, (v) ResNet-34 as backbone and Optimizer being Nadam, (vi) SE-ResNet-34 as backbone and Optimizer being Adam, (vii) SE-ResNet-34 as backbone and Optimizer being Adamax, (viii) SE-ResNet-34 as backbone and Optimizer being SGD, (ix) SE-ResNet-34 as backbone and Optimizer being RMSProp, (x) SE-ResNet-34 as backbone and Optimizer being Nadam

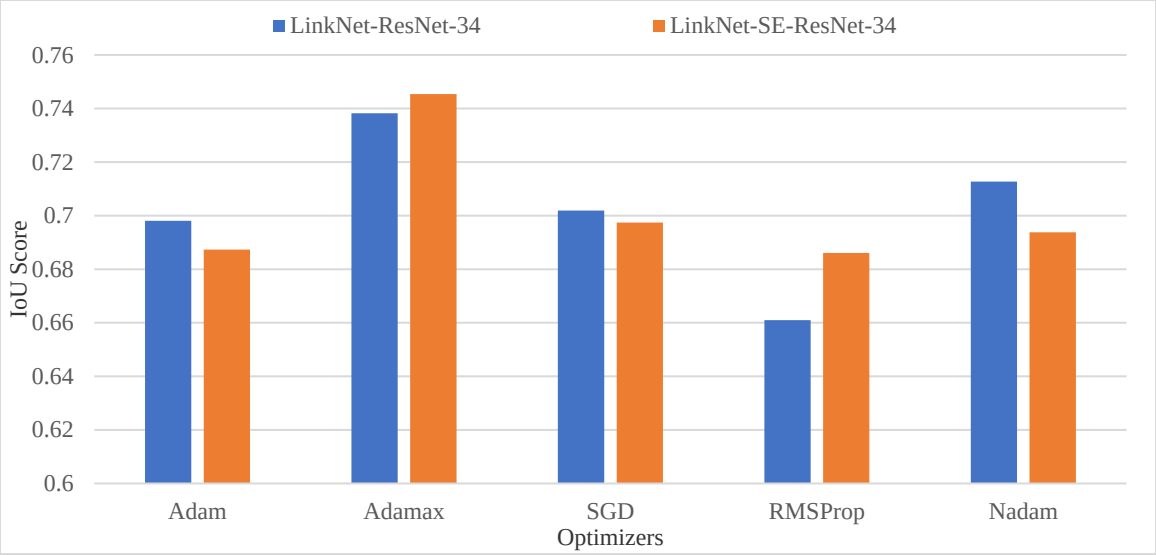
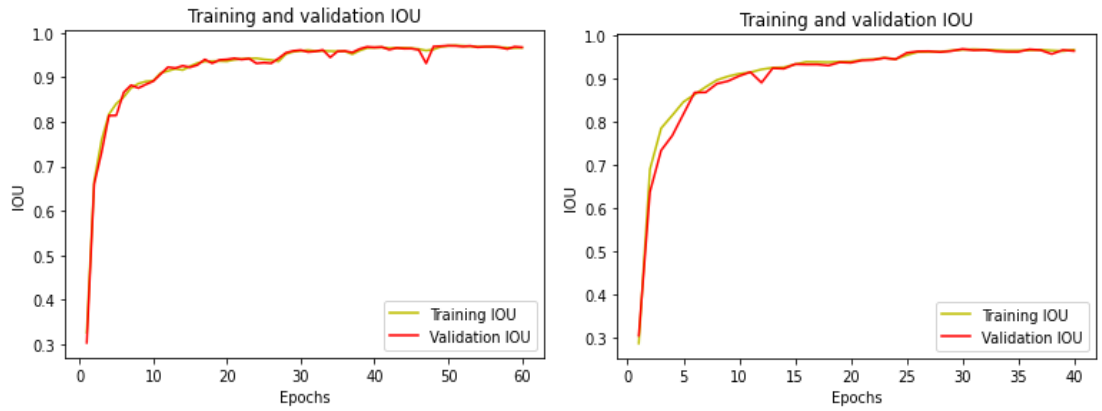


Fig. 9 Bar chart indicates the IoU score obtained plotted against various optimizers used with the LinkNet Model in the combination of ResNet-34 and SE-ResNet-34 as Backbone



(a) **(b)**

Fig. 10 Training and Validation, Performance-IoU Vs Epochs for LinkNet with Adamax optimizer and with backbone such as a) ResNet-34 and b) SE-ResNet-34

Each model was set to 200 epochs for training, 16 as batch size, a scheduler of the iterative learning rate, 1e-3 initial learning rate, and various optimizers and backbones. The performance of the LinkNet model-IoU score obtained for training and validation against the epochs for the best 3 results is represented in Fig. 10.

It was observed with LinkNet while using adamax Optimizer and with ResNet-34 as backbone obtained an IoU score during training and validation of 0.97074, but while testing resulted from 0.738271933. For LinkNet with SE-ResNet-34 as the backbone and with adamax Optimizer, the training and validation IoU score went up to 0.96763, but the IoU score obtained while testing is 0.7454453.

4.3. Results Obtained with PSP-Net

This section describes various approaches for Prostate Semantic Segmentation using PSP-Net architecture with ResNet-34 and SE-ResNet-34 backbone, along with the performance results obtained. Fig. 11 shows the Original Prostate Gland Image along with the mask that was provided as against the predicted mask by for ResNet34 as backbone and Optimizer being Adamax as in Fig. 11 (i), ResNet34 as backbone and Optimizer being Adamax as in Fig. 11 (ii), ResNet34 as backbone and Optimizer being SGD as in Fig. 11 (iii), ResNet34 as backbone and Optimizer being RMSProp as in Fig. 11 (iv), ResNet34 as backbone and Optimizer being Nadam as in Fig. 11 (v), SE-ResNet34 as backbone and Optimizer being Adam as in Fig. 11 (vi), SE-ResNet34 as backbone and Optimizer being Adamax as in Fig. 11 (vii), SE-ResNet34 as backbone and Optimizer being SGD as in Fig. 11 (viii), SE-ResNet34 as backbone and Optimizer being RMSProp as in Fig. 11 (ix) and SE-ResNet34 as backbone and Optimizer being Nadam as in Figure 11 (x).

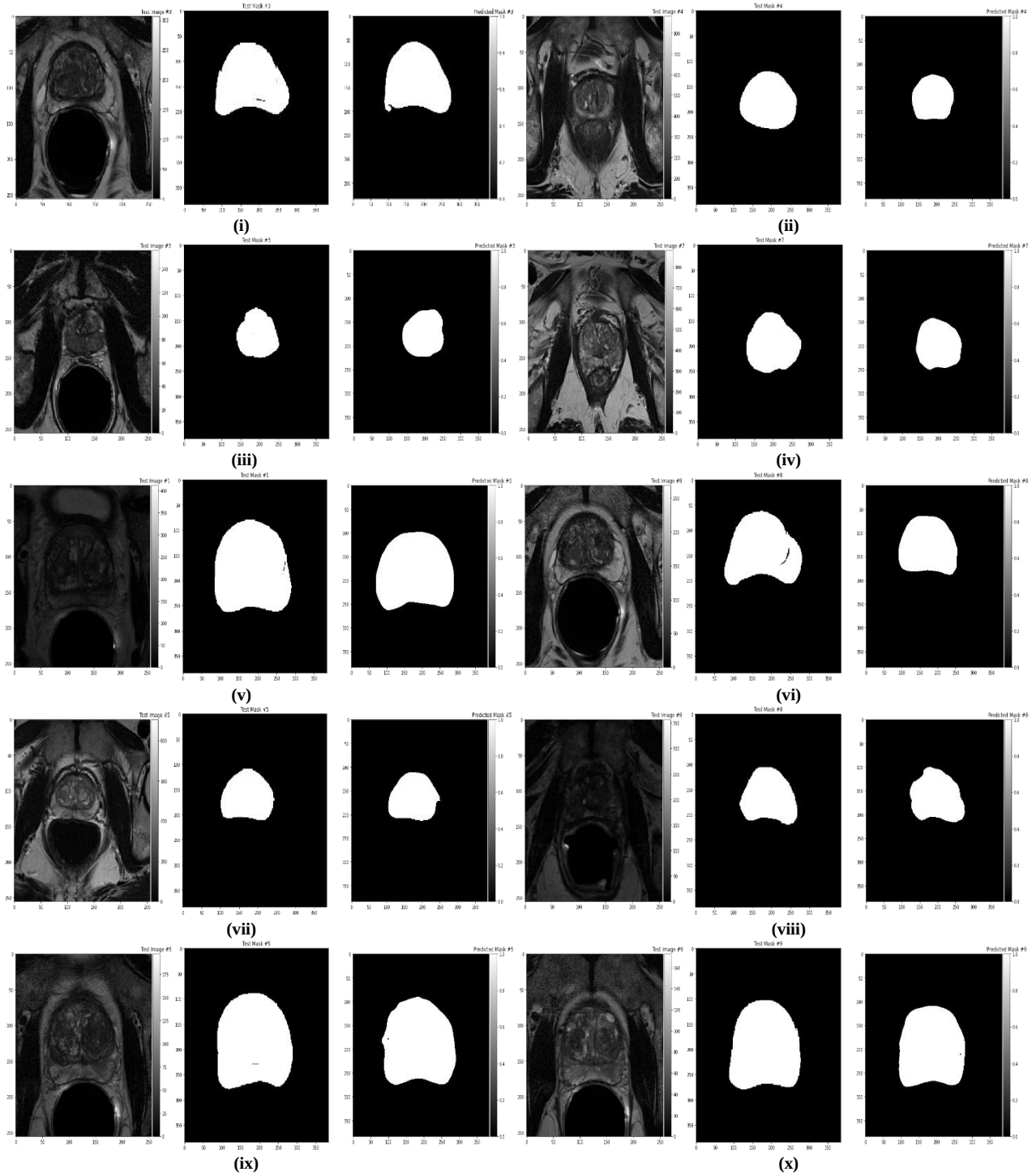


Fig. 11 Prostate MRI Image along with the ground truth or provided mask and the predicted mask is listed for PSP-Net model considering (i) ResNet34 as backbone and Optimizer being Adam, (ii) ResNet34 as backbone and Optimizer being Adamax, (iii) ResNet34 as backbone and Optimizer being SGD, (iv) ResNet34 as backbone and Optimizer being RMSProp, (v) ResNet34 as backbone and Optimizer being Nadam, (vi) SE-ResNet34 as backbone and Optimizer being Adam, (vii) SE-ResNet34 as backbone and Optimizer being Adamax, (viii) SE-ResNet34 as backbone and Optimizer being SGD, (ix) SE-ResNet34 as backbone and Optimizer being RMSProp, (x) SE-ResNet34 as backbone and Optimizer being Nadam

Table 3 shows results obtained while testing the PSP-Net model with ResNet34 and SE-ResNet-34 as the backbone for untrained MRI images. Based on the trials, it can be inferred that the best performance - IoU Score were obtained with SE-ResNet34 as the backbone and Nadam Optimizer 0.709108849, followed by SE-ResNet34 as the

backbone with Adam Optimizer giving 0.699512671, followed by ResNet34 as the backbone with Nadam Optimizer giving 0.694847168, followed by SE-ResNet34 as the backbone with Adamax Optimizer with 0.685369737 and further followed by Resnet34 as the backbone with Adam Optimizer giving 0.68529029. The bar chart in Fig. 12 shows the performance - IoU Score of various combinations of PSP-Net with ResNet34 and SE-ResNet34 backbones and Optimizers.

Table 3. The table indicates IoU Score obtained with PSP-Net Model with ResNet34 and SE-ResNet34 as Backbone with various combinations of optimizers

Architecture and Backbones	Adam	Adamax	SGD	RMSProp	Nadam
PSP-Net – ResNet34	0.68529029	0.680192549	0.605144408	0.669634853	0.694847168
PSP-Net – SEResNet34	0.699512671	0.685369737	0.631567752	0.660367808	0.709108849

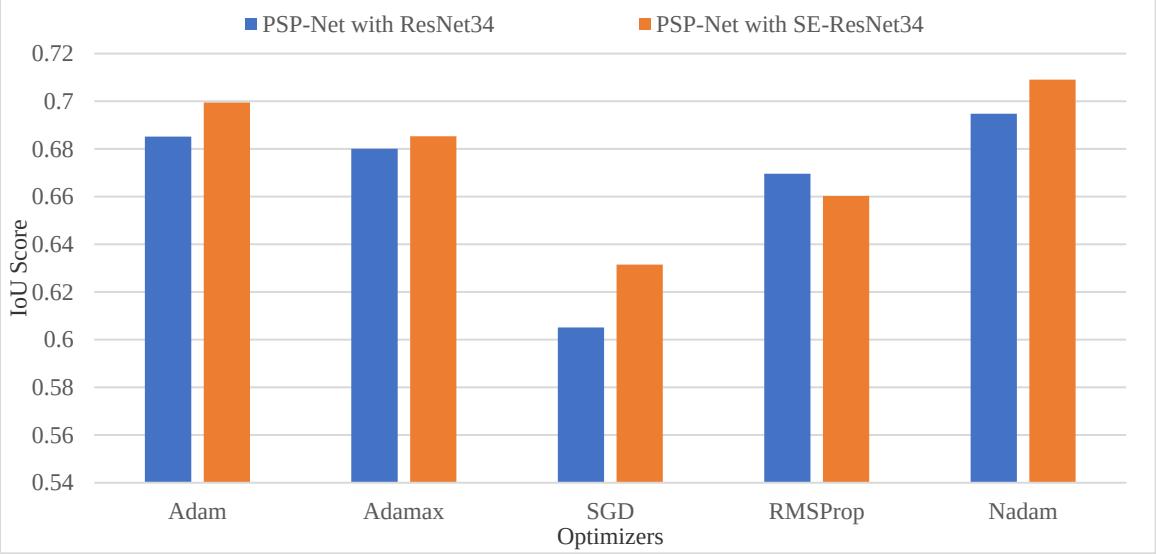
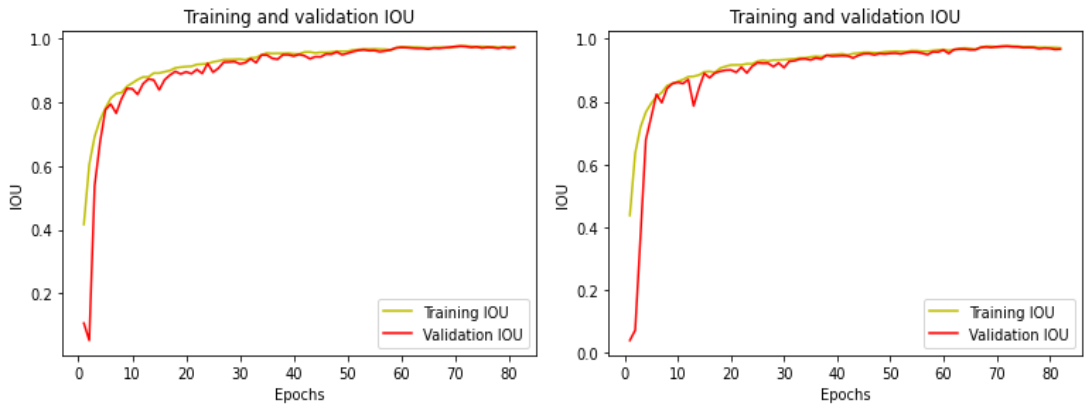


Fig. 12 Bar chart indicates the IoU score obtained plotted against various optimizers used with the PSP-Net model in the combination of ResNet-34 and SE-ResNet-34 as Backbone



(a) **(b)**

Fig. 13 Training and Validation Performance - IOU Vs epochs for PSP-Net with ResNet34 and SE-ResNet34 as Backbones with Adamax Optimizer

Each model was set to 200 epochs for training, 16 as batch size, cyclic learning rates scheduler, and the first learning rates of 1e-3, along with various Optimizer and backbones. Performance of the PSP-Net Model - IoU Score obtained for Training and Validation as against the epochs for the ResNet34 and SE-ResNet34 as Backbones with Adamax Optimizer as in Fig. 13. It was observed that the IoU score for Semantic Segmentation with PSP-Net with SE-ResNet34 as backbone using Nadam as Optimizer during training and validation went up to 0.97639 but while testing resulted in 0.709108849. With SE-ResNet34 as the backbone, using Adam as Optimizer during training and validation went up to 0.98062, while testing resulted in 0.699512671. With ResNet34 as the backbone, using Nadam optimizer during

training and validation went up to 0.97697, but while testing resulted in 0.694847168. With SE-ResNet34 as the backbone, Adamax as Optimizer during training and validation went up to 0.97610 while testing resulted in 0.685369737. The highest IoU score was obtained with RMSProp optimizer, with SE-ResNet34 as backbone got 0.98393 but while testing resulted in 0.660367808 and with ResNet34 as backbone got 0.98335 but while testing resulted in 0.669634853.

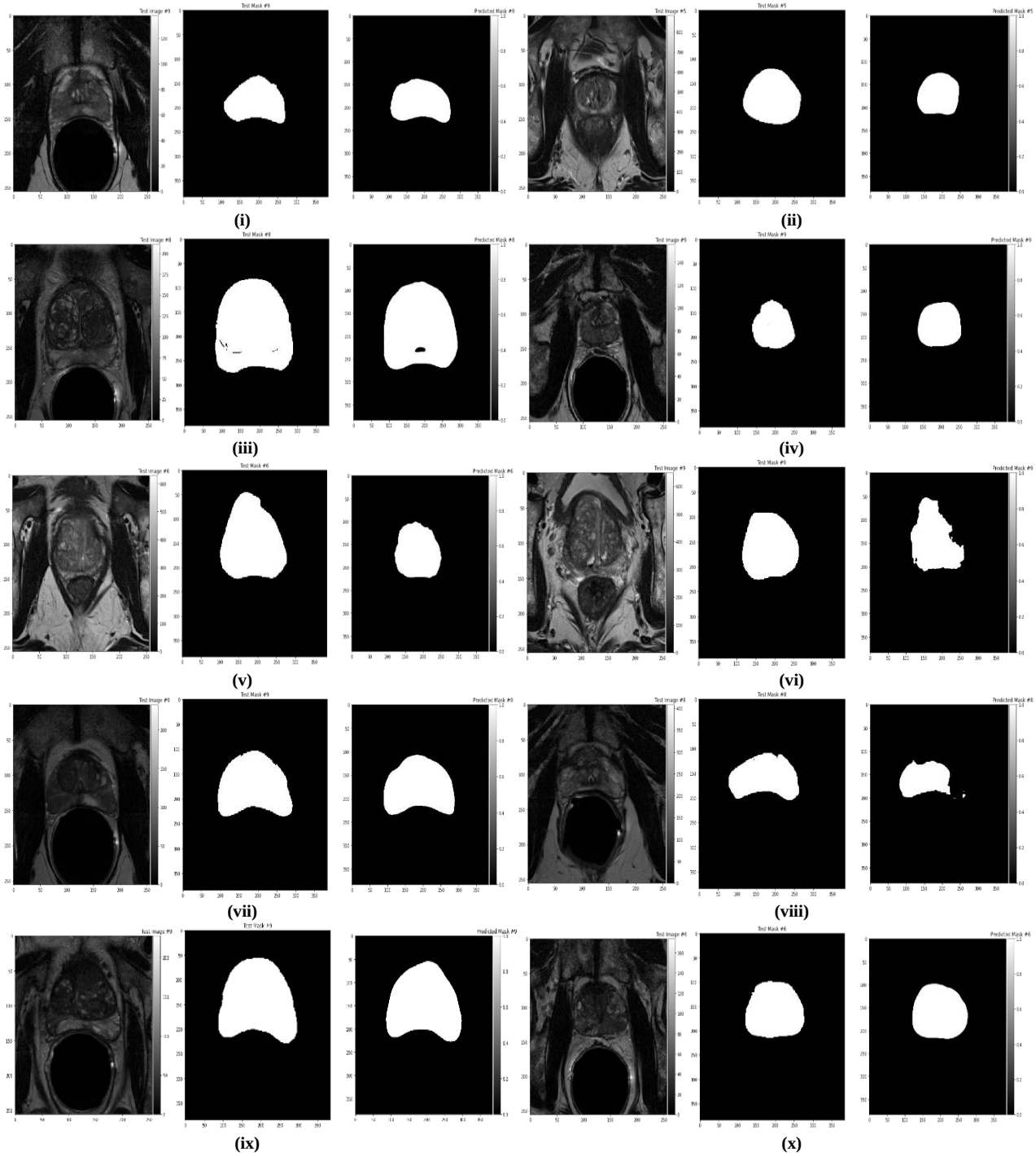


Fig. 14 Prostate MRI Image along with the ground truth or provided mask and the predicted mask is listed for the FPN model considering (i) ResNet34 as backbone and Optimizer being Adam, (ii) ResNet34 as backbone and Optimizer being Adamax, (iii) ResNet34 as backbone and Optimizer being SGD, (iv) ResNet34 as backbone and Optimizer being RMSProp, (v) ResNet34 as backbone and Optimizer being Nadam, (vi) SE-ResNet34 as backbone and Optimizer being Adam, (vii) SE-ResNet34 as backbone and Optimizer being Adamax, (viii) SE-ResNet34 as backbone and Optimizer being SGD, (ix) SE-ResNet34 as backbone and Optimizer being RMSProp, (x) SE-ResNet34 as backbone and Optimizer being Nadam

Table 4. Table indicates IoU Score obtained with FPN Model with ResNet34 and SE-ResNet34 as Backbone with various combinations of Optimizers

Architecture and Backbones	Adam	Adamax	SGD	RMSProp	Nadam
FPN – ResNet34	0.705056016	0.720181423	0.715967565	0.676009491	0.703369762
FPN – SEResNet34	0.72547259	0.735521363	0.698487643	0.684241517	0.67387959

4.4. Results Obtained with FPN

This section describes various approaches for Prostate Semantic Segmentation using FPN architecture with ResNet34 and SE-ResNet34 backbone, along with the performance results obtained. Fig. 14 shows the Original Prostate Gland Image along with the mask that was provided as against the predicted mask by for ResNet34 as backbone and Optimizer being Adamax as in Fig. 14 (i), ResNet34 as backbone and Optimizer being Adamax as in Fig. 14 (ii), ResNet34 as backbone and Optimizer being SGD as in Fig. 14 (iii), ResNet34 as backbone and Optimizer being RMSProp as in Fig. 14 (iv), ResNet34 as backbone and Optimizer being Nadam as in Fig. 14 (v), SE-ResNet34 as backbone and Optimizer being Adam as in Fig. 14 (vi), SE-ResNet34 as backbone and Optimizer being Adamax as in Fig. 14 (vii), SE-ResNet34 as backbone and Optimizer being SGD as in Fig. 14 (viii), SE-ResNet34 as backbone and Optimizer being RMSProp as in Fig. 14 (ix) and SE-ResNet34 as backbone and Optimizer being Nadam as in Fig. 14 (x).

Table 4 shows results obtained while testing the FPN model with ResNet34 and SE-ResNet34 as the backbone for untrained MRI images. Based on the trials, it can be inferred that the best performance - IoU Score were obtained with SE-ResNet34 as the backbone and Adamax Optimizer 0.735521363, followed by SE-ResNet34 as the backbone with Adam Optimizer giving 0.72547259, followed by ResNet34 as the backbone with Adamax Optimizer giving 0.720181423, followed by ResNet34 as the backbone with SGD Optimizer giving 0.715967565, followed by ResNet34 as the backbone with Adam Optimizer with 0.705056016 and further followed by ResNet34 as the backbone with Nadam Optimizer giving 0.703369762. The bar chart in Fig. 15 shows the performance - IoU Score of various combinations of FPN with ResNet34 and SE-ResNet34 backbones and Optimizers. Each model was set to 200 epochs for training, 16 as batch size, a cyclic learning rate scheduler, and the initial learning rate of 1e-3, along with various Optimizer and backbones. Performance of the FPN Model - IoU Score obtained for Training and Validation as against the epochs for the ResNet34 and SE-ResNet34 as Backbones with Adamax Optimizer are shown in Fig. 16.

It was observed that the IoU score for Semantic Segmentation with FPN with SE-ResNet34 as backbone while using RMSProp optimizer during training and validation obtained went up to 0.99177, but while testing resulted in 0.684241517. With SE-ResNet34 as the backbone while using Adamax optimizer during training and validation obtained went up to 0.97560, but while testing resulted in 0.735521363. With SE-ResNet34 as the backbone while using Adam optimizer during training and validation obtained went up to 0.95558, but while testing resulted in 0.72547259. ResNet34 backbone using Adamax Optimizer during training and validation went up to 0.97580, but while testing resulted in 0.720181423. With ResNet34 backbone, using SGD optimizer during training and validation went up to 0.97066 while testing resulted in 0.715967565. Each of these models was pre-set to be trained for 200 iterations with the size of batch 16 and with the starting learning rates of 1e-3 and with cyclic learning rates scheduler to make the network models extricate dense feature maps by managing the region view for accurate localizations; the early sopping was defined by monitoring the loss and for every 10 epochs while restoring the best weights. This has significantly led to the quick and effective encoder and decoder networks that build the deep representation among multi-feature images and covers good spatial data at different ranges against various encoder and decoder networks producing the effective segmentation boundary.

Table 5. Performance Evaluation of the Various Setting of Segmentation Techniques

Model	Backbone	Adam	Adamax	SGD	RMSProp	Nadam
U-Net	**ResNet-34**	0.716165932	0.732374634	0.704223167	0.663843288	0.708662823
U-Net	**SE-ResNet-34**	0.737513949	0.742655936	0.703767031	0.694015578	0.697213093
Link-Net	**ResNet-34**	0.698139869	0.738271933	0.701902515	0.661088051	0.712770765
Link-Net	**SE-ResNet-34**	0.687389449	0.7454453	0.697475474	0.686124772	0.693827738
PSP-Net	**ResNet-34**	0.68529029	0.680192549	0.605144408	0.669634853	0.694847168
PSP-Net	**SE-ResNet-34**	0.699512671	0.685369737	0.631567752	0.660367808	0.709108849
FPN	**ResNet-34**	0.705056016	0.720181423	0.715967565	0.676009491	0.703369762
FPN	**SE-ResNet-34**	0.72547259	0.735521363	0.698487643	0.684241517	0.67387959

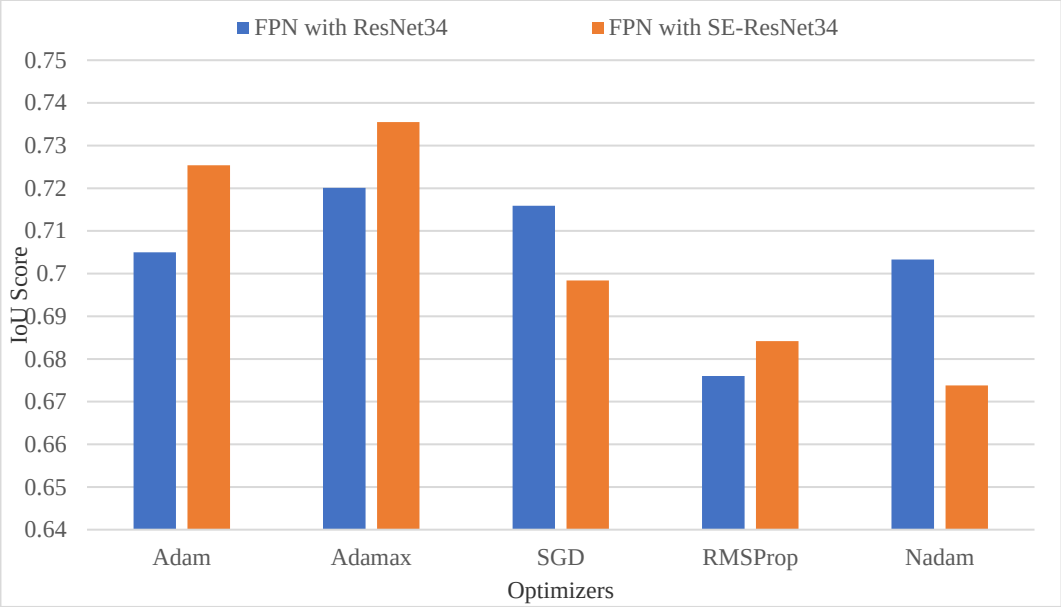
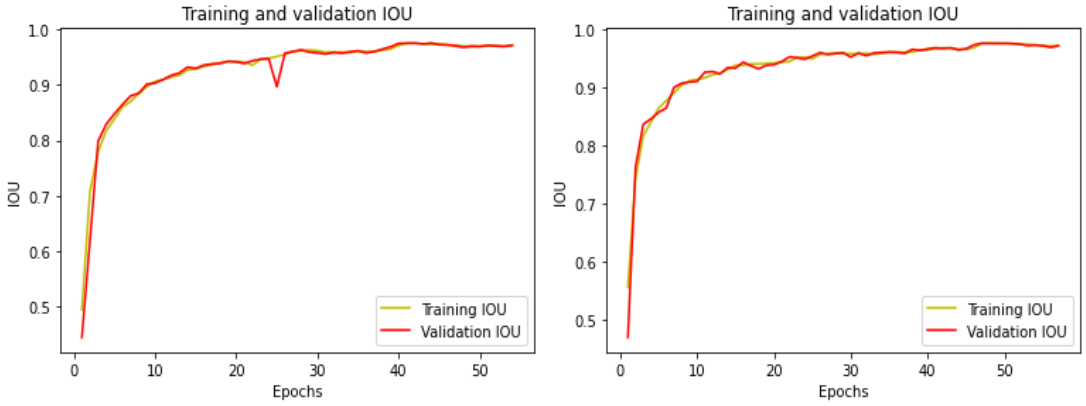


Fig. 15 Bar chart indicates the IoU score obtained plotted against various optimizers used with the FPN Model in the combination of ResNet-34 and SE-ResNet-34 as Backbone



(a) **(b)**

Fig. 16 Training and Validation Performance - IOU Vs epochs for FPN with ResNet34 and SE-ResNet34 as Backbones with Adamax Optimizer

Table 5 represents the performance comparison of the Semantic Segmentation approaches of the model on the CNN models, ResNet34 and SE-Resnet34 on U-Net, LinkNet, PSP-Net and FPN segmentations with various optimizers. Based on the above table, it can be inferred that the best performance measured through IoU score as an index in segmenting the prostate gland was with the LinkNet architecture with SE-ResNet34 as the backbone and with Adamax as Optimizer yielded the better result of 0.7454453, followed by U-Net architecture and SE-ResNet34 as the backbone and with Adamax as Optimizer yielded the result of 0.742655936, followed by LinkNet architecture and ResNet34 as the backbone and with Adamax as Optimizer yielded result of 0.738271933, followed by U-Net architecture and SE-ResNet34 as the backbone and with Adam as Optimizer yielded a result of 0.737513949, followed by FPN architecture and SE-ResNet34 as the backbone and with Adamax as Optimizer yielded the result of 0.735521363.

4.5. Performance Analysis Comparison

The performance of the proposed models was compared with the existing technique called Dynamic Multi-Atlas (DMA) segmentation technique, which is based on deep learning. This DMA technique was proposed to segment the prostate lesions in MRI images using the NCI-ISBI-2013 dataset [19] and compared with 3D and 2D-UNet [35-40].

Table 6 represents the comparison of research models with the existing DMA technique for validation. The IoU values of the proposed models, such as U-Net with SE-ResNet-34 optimized with Adamax, LinkNet with SE-ResNet-34 optimized with Adamax, PSP-Net with SE-ResNet-34 optimized with Nadam, and FPN with SE-ResNet-34 optimized with Adamax models are compared. Compared

to the DMA, 3D-UNet and 2D-UNet techniques, the proposed LinkNet with SE-ResNet-34 optimized with Adamax optimizer achieved 0.7454453 as the best result in this analysis compared to the other proposed models. Following LinkNet, the U-Net with SE-ResNet-34 and FPN with SE-ResNet-34 optimized with Adamax has obtained better results compared to the DMA technique with 0.742655936 and 0.735521363. The proposed models outperformed the compared models mainly because of the optimizers employed with the proposed models. These optimizers help to achieve great results with better performance.

Table 6. Performance Comparison of Segmentation Techniques

Models	IoU
U-Net with SE-ResNet-34	0.7426
LinkNet with SE-ResNet-34	0.7454
PSP-Net with SE-ResNet-34	0.7091
FPN-SE-ResNet-34	0.7355
DMA [19]	0.67±0.07
3D-UNet without batch normalization [35]	0.723
3D-UNet with batch normalization [35]	0.704
2D-UNet with batch normalization [35]	0.547

5. Conclusion

Precise segmentation of the prostate region in MR Imaging improves the treatment opportunities and chances of the patient's prognosis with cancer. Prostate gland segmentation using the DL semantic segmentation models such as U-Net, LinkNet, PSP-Net and FPN with backbones such as ResNet-34 and SE-ResNet-34 has been designed, implemented, and discussed along with various optimizers and other hyperparameters setting. In this work, neural network architecture optimizer Adam, Adamax, RMSProp, Nadam and SGD has been employed along with multiple epoch values along with learning rates to improve learning and reduce loss. Performance of the proposed model has been carried out using NCI-ISBI 2013 dataset to the IoU metric on the various sets of the model. It has proved that LinkNet with SE-ResNet34 architecture yielded better results in predicted segmentation, followed by U-Net with SE-ResNet34 architecture and then followed by LinkNet with ResNet34. The results obtained in this paper are based on some of the hyperparameters tuned to get better results. The results of the proposed models are compared with the DMA technique. However, future work can consider better understanding, developing layers, and tuning actuators of the semantic segmentation models adopted.

Acknowledgments

The authors would like to thank the Department of Electronics and Communication Engineering, Faculty of Engineering & Technology, Jain (Deemed-to-be University), Bangalore, for their constant support and encouragement throughout the completion of this work.

Authors' Contribution

Authors Rajesh M. N. and Chandrasekar B. S. have contributed equally to this work.

References

[1] Narinder Singh Punn, and Sonali Agarwal, "Modality Specifics U Net Variant for Bio Medical Image Segmentations: A Survey," *Artificial Intelligence Review*, vol. 55, pp. 5845–5889, 2022. *Crossref*, https://doi.org/10.1007/s10462-022-10152-1

[2] David Gillespie et al., "Deep Learning in Magnetic Resonances Prostate Segmentations: A Review and New Perspectives," *Image and Video Processing*, *Arxiv Preprint arXiv:2011.07795*. *Crossref*, https://doi.org/10.48550/arXiv.2011.07795

[3] Haozhe Jia et al., "Atlas Registration and Ensembled Deep Convolution Neural Networks-Based Prostate Segmentations Using Magnetic Resonances Imaging," *Neurocomputing*, vol. 274, pp. 1359–1368, 2018. *Crossref*, https://doi.org/10.1016/j.neucom.2017.09.084

[4] Geert Litjens et al., "Evaluations of Prostate Segmentations Algorithm for MRIS: the PROMISE 12 Challenges," *Medical Image Analysis*, vol. 18, no. 2, pp. 359–373, 2014. *Crossref*, https://doi.org/10.1016/j.media.2013.12.002

[5] Qikui. Zhu et al., "Deeply Supervised CNNS for Prostate Segmentations," *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 178–184, 2017. *Crossref*, http://dx.doi.org/10.1109/IJCNN.2017.7965852

[6] Nazia Hameed et al., "An Intelligent Computers-Aided Scheme for Classifying Multiple Skin Lesion," *Computers*, vol. 8, no. 3, pp. 62, 2019. *Crossref*, https://doi.org/10.3390/computers8030062

[7] A.Murugan, S.A.H.Nair, and K.P.S.Kumar, "Detections of Skin Cancers Using SVM, Random Forests and KNN Classifier," *Journal of Medical Systems*, vol. 43, no. 8, pp. 269, 2019. *Crossref*, https://doi.org/10.1007/s10916-019-1400-8

[8] R.D.Seeja, and A.Suresh, "Deep Learning-Based Skin Lesions Segmentations and Classifications of Melanoma Using Support Vector Machines (SVM)," *Asia Pacific Journal of Cancer Prevention*, vol. 21, no. 5, pp. 1554–61, 2019. *Crossref*, https://doi.org/10.31557%2FAPJCP.2019.20.5.1555

[9] S.Liao et al., "Representations Learning: A Unified Deep Learning Framework for Automatic Prostate MR Segmentations," *Medical Image Computing and Computer-Assisted Intervention: MICCAI-International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 16, pp. 254–261, 2013. *Crossref*, https://doi.org/10.1007/978-3-642-40763-5_32

[10] A.Krizhevsky, I.Sutskever, and G.E.Hinton, "Imagenet Classifications with Deep Convolution Neural Networks," *Communication of the ACM*, vol. 61, no. 6, pp. 84-901, 2017.

[11] Achim Hekler et al., "Superior Skin Cancers Classifications by the Combinations of Human and Artificial Intelligences," *European Journal of cancer*, vol. 120, pp. 114–121, 2019. *Crossref,* https://doi.org/10.1016/j.ejca.2019.07.019

[12] Titus Josef Brinker et al., "Skin Cancers Classifications Using Convolution Neural Network: Systematic Review," *Journal of Medical Internet Research*, vol. 20, no. 10, pp. E11936, 2018. *Crossref,* https://doi.org/10.2196/11936

[13] S. R. Guha, and S. R. Haque, "Performance Comparison of Machine Learning-Based Classification of Skin Diseases from Skin Lesion Images," *Proceedings of International Conferences on Communication, Computing and Electronics System*, vol. 637, pp. 15–25, 2020. *Crossref,* https://doi.org/10.1007/978-981-15-2612-1_2

[14] Lei Bi, Dagan Feng, and Jinman Kim, "Dual-Paths Adversarial Learnings for Fully Convolution Network (FCN)-Based Medical Images Segmentations," *The Visual Computer*, vol. 34, pp. 1042–52, 2018. *Crossref,* https://doi.org/10.1007/s00371-018-1519-5

[15] Lei Bi et al., "Dermoscopic Image Segmentation Via Multistage Fully Convolutional Networks," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2065–2074, 2017. doi: 10.1109/TBME.2017.2712771.

[16] Behnaz Abdollahi, Naofumi Tomita, and Saeed Hassanpour, "Data Augmentations in Training Deep Learning Model for Medical Images Analyses," *Intelligent Systems Reference Library*, vol. 186, pp. 167–180, 2020. *Crossref,* http://dx.doi.org/10.1007/978-3-030-42750-4_6

[17] Keerti Maithil, and Tasneem Bano Rehman, "Urban Remote Sensing Image Segmentation Using Dense U-Net+," *SSRG International Journal of Computer Science and Engineering*, vol. 9, no. 3, pp. 21-28, 2022. *Crossref,* https://doi.org/10.14445/23488387/IJCSE-V9I3P103

[18] Z.Khan et al., "Evaluations of Deep Neural Network for Semantics Segmentations of Prostate in T2W-MRI," *Sensors*, vol. 20, No. 11, pp. 1-17, 2020. *Crossref,* https://doi.org/10.3390/s20113183

[19] Hamid Moradi, and Amir Hossein Foruzan, "Integrations of Dynamic Multi-Atlas and Deep Learning Technique to Improve Segmentations of the Prostate in MR-Images," *International Journal of Image and Graphics*, vol. 22, no.4, pp. 1-17, 2022. *Crossref,* https://doi.org/10.1142/S0219467822500310

[20] Justin Lo et al., "Cross Attentions Squeeze Excitations Networks (CASE-Net) for Whole-Body Fetal MRI Segmentations," *Sensor*, vol. 21, no. 13, pp. 4490, 2021. *Crossref,* https://doi.org/10.3390/s21134490

[21] Quande Liu et al., "MS-Net: Multi-Sites Networks for Improving Prostate Segmentations with Heterogeneous MRI Data," *IEEE Transactions on Medical Imaging*, vol. 39, no. 9, pp. 2712-2725, 2020. *Crossref,* https://doi.org/10.1109/TMI.2020.2974574

[22] PrafulHambarde et al., "Prostate Lesions Segmentations in MR Image Using Radiomic Based Deeply Supervised U-Net," *Biocyberneticsand Biomedical Engineering*, vol. 40, no. 4, pp. 1421–1435, 2020. *Crossref,* https://doi.org/10.1016/j.bbe.2020.07.011

[23] Zhiqiang Tian et al., "Psnet: Prostate Segmentations on MRI Based on A Convolution Neural Networks," *Journal of Medical Imaging*, vol. 5, no. 2, pp. 1-6, 2018. *Crossref,* https://doi.org/10.1117%2F1.JMI.5.2.021208

[24] LeonardoRundo et al., "USE-Net: Incorporating Squeeze-and-Excitations Block Into U-net for Prostates Zonal Segmentations of Multi-Institutional MRI Data Sets," *Neurocomputing*, vol. 365, pp. 31-43, 2019. *Crossref,* https://doi.org/10.1016/j.neucom.2019.07.006

[25] Fred Prior et al., "The Public Cancers Radiology Imaging Collection of the Cancers Imaging Archives," *Scientific Data*, vol. 4, no. 170124, pp. 1-7, 2017. *Crossref,* https://doi.org/10.1038/sdata.2017.124

[26] A.Voulodimos et al., "Deep Learning for Computers Vision: A Brief Review," *Computational Intelligence and Neuroscience*, vol. 2018, pp. 1–13, 2018. *Crossref,* https://doi.org/10.1155/2018/7068349

[27] BiLei et al., "Stepwise Integrations of Deep Class-Specifics Learning for Dermoscopy Images Segmentations," *Pattern Recognition*, vol. 85, pp. 78-89, 2019. *Crossref,* https://doi.org/10.1016/j.patcog.2018.08.001

[28] Abhishek Chaurasia, and Eugenio Culurciello, "Linknet: Exploiting Encoders Representation for Efficient Semantics Segmentations," *IEEE Visual Communication and Images Processing*, pp. 1-4, 2017. *Crossref,* https://doi.org/10.1109/VCIP.2017.8305148

[29] Hengshuang Zhao et al., "Pyramid Scenes Parsing Networks," *IEEE Conferences on Computer Visions and Patterns Recognitions*, pp. 2881-2890, 2017. *Crossref,* https://doi.org/10.1109/CVPR.2017.660

[30] Wenchao Zhang et al., "Global Contexts Aware RCNN for Objects Detections," *Neural Computer Application*, vol. 33, pp. 11627–11639, 2021. *Crossref,* https://doi.org/10.48550/arXiv.2012.02637

[31] M.A. Aghdam, A.Sharifi, and M.M.Pedram, "Combinations of rs-fMRI and sMRI Data to Discriminate Autisms Spectrums Disorder in Young Children Using Deep Belief Networks," *Journal of Digital Imaging*, vol. 31, No. 6, pp. 895–903, 2018. *Crossref,* https://doi.org/10.1007/s10278-018-0093-8

[32] Yulian Zhu et al., "MRI-Based Prostate Cancers Detections with High Level Representations and Hierarchical Classifications," *Medical physics*, vol. 44, no. 3, pp. 1028–1039, 2017. *Crossref,* https://doi.org/10.1002/mp.12116

[33] Jie Hu et al., "Squeeze-and-Excitations Network," *Proceedings of IEEE Conferences on Computer Visions and Patterns Recognitions*, pp. 7132-7141, 2018. *Crossref,* https://doi.org/10.1109/CVPR.2018.00745

[34] Olaf Ronneberger et al., "U-Net: Convolution Network for Biomedical Images Segmentations," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, pp. 234-241, 2015. *Crossref,* https://doi.org/10.1007/978-3-319-24574-4_28

[35] Ö.Çiçek et al., "3D U-Net: Learning Dense Volumetric Segmentations From Sparse Annotations," *International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016*, vol. 9901, pp. 424–432, 2016. *Crossref,* https://doi.org/10.1007/978-3-319-46723-8_49

[36] Sitanaboina S L Parvathi, and Harikiran Jonnadula, "A Hybrid Semantic Model for MRI Kidney Objects Segmentations With Stochastic Feature and Edge Detections Technique," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 411-420, 2022. *Crossref,* https://doi.org/10.14445/22315381/IJETT-V70I9P242

[37] P. Poornima, and S. Saranya, "Review on Image Segmentations Techniques to Detect Outlier in Blood Sample," *International Journal of Engineering Trends and Technology*, vol. 53, no. 2, pp. 64-73, 2017. *Crossref,* https://doi.org/10.14445/22315381/IJETT-V53P212

[38] C. Narmatha, and P.M. Surendra, "A Review on Prostate Cancer Detection Using Deep Learning Techniques," *A Review on Prostate Cancer Detection Using Deep Learning Techniques*, vol. 1, no. 2, pp. 26-33, 2020.

[39] M. B. Sudhan et al., "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model," *Journal of Healthcare Engineering*, vol. 2022, no. 1601354, pp. 1–10, 2022. *Crossref,* https://doi.org/10.1155/2022/1601354

[40] S. Sridhar et al., "A Torn ACL Mapping in Knee MRI Images Using Deep Convolution Neural Network with Inception," *Journal of Healthcare Engineering*, vol. 2022, no. 7872500, pp. 1-9, 2022. *Crossref,* https://doi.org/10.1155/2022/7872500

[41] A. K. Kumaraswamy, and Chandrasekar.M.Patil, "Automatic Prostate Segmentations of Magnetic Resonances Imaging Using Resnet," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 35, pp. 621–630, 2022. *Crossref,* https://doi.org/10.1007/s10334-021-00979-0