Received: September 26, 2023. Revised: November 2, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Channel and Spatial Attention Aware UNet Architecture for Segmentation of Blood Vessels, Exudates and Microaneurysms in Diabetic Retinopathy

Anand M[1]* **Meenakshi Sundaram A[1]**

[1]*School of Computer Science and Engineering, REVA University, Bengaluru, India*
* Corresponding author's Email: r21pcs01@reva.edu.in

Abstract: Diabetic retinopathy stands out as one of the highly prevalent causes of vision loss in working people worldwide. In computer vision, deep learning based strategies are seen as a viable solution for efficient diabetic retinopathy detection. We present a UNet-based deep learning architecture for diabetic retinopathy segmentation of blood vessels, exudates, and microaneurysms. Traditional methods often consider the features only from the last convolution unit and discard the remaining features, resulting in low-quality feature maps. However, boundary information plays important role in medical image segmentation. To overcome this, we introduce a skip connection mechanism to concatenate all attributes from each layer. Additionally, we utilize an upsampling layer to aggregate the features at the final sigmoid layer. Finally, we apply channel and spatial attention mechanisms to generate the semantic feature map. Therefore, the proposed approach overcomes the issues of existing methods by incorporating dense skip connection along with channel and spatial attention mechanism which helps to retain the substantial information of image. We tested proposed approach on several publicly available datasets such as IDRiD, DIARETDB1, STARE, ChaseDB1, DRIVE, and HRF datasets. The comparative analysis shows that the proposed approach achieves superior performance, with an average accuracy of 98.10%, average sensitivity of 97.60%, and average specificity of 98.2% for segmentation.

Keywords: Diabetic retinopathy, Segmentation, UNet, Blood vessels, Exudates, Microaneurysms.

1. Introduction

This article provides a comprehensive study of DR, a dangerous, progressive vascular and neurodegenerative condition that damages retinal cells without impairing vision visibly and is first difficult to diagnose. Unregulated blood sugar levels can cause diabetic retinopathy (DR), which impairs vision and causes gradual impairment. Therefore, it is essential to identify DR early to stop retinal damage that might eventually result in blindness. According to a report presented by Saeedi et al. [1], approximately 463 million individuals worldwide had DR as of 2019, and 31 million of them lived in the United States. Researchers estimate that by 2030, this number will increase to 578 million (34.4 million in the US), and by 2045, it will reach 700 million (367 million in the US). Furthermore, the occurrence of DR has been estimated at 4.1 million in the United States, according to the centers for disease control and prevention (CDC) [2]. Additionally, the U.S. spends around $500 million annually on diabetes-related blindness [2]. There are 899,000 blind people in the United States. Therefore, it is essential to identify DR as soon as possible.

The human body consists of several sensory organs, and the eye is one of the vital organs of the body. The pathology of the eye is becoming a serious concern in the medical field. Blindness is one of the main challenges in eye pathology, and it can occur due to different types of eye diseases, such as diabetic retinopathy, glaucoma, macular degeneration, etc. However, these conditions often develop silently, making continuous eye check-ups essential for early diagnosis. Diabetes is considered one of the most chronic diseases, and it has a high prevalence globally. A recent study presented in [1] reported a total of 463 million adults aged between

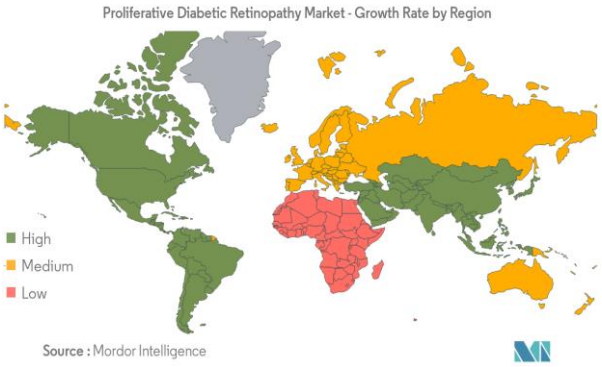

Figure. 1 Worldwide prevalence of diabetes

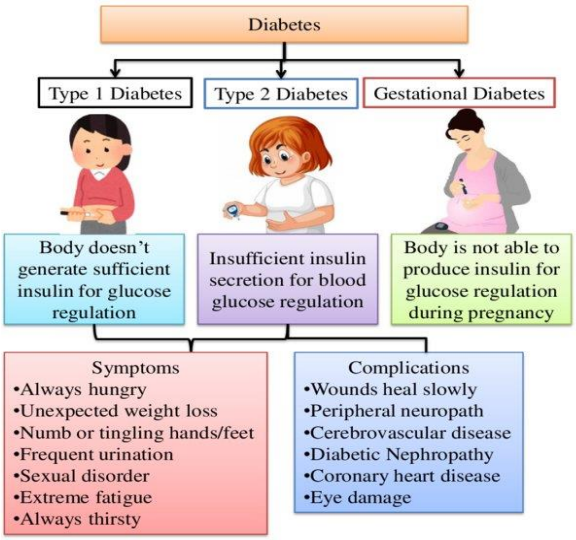

Figure. 2 Diabetes categories

20 to 79 years were suffering from diabetes globally. Similarly, authors in [2, 3] reported that by 2040, diabetic patients are expected to increase by 600 million. Fig. 1 depicts the current status of diabetes in the worldwide scenario.

Tripathy et al. [4] reported that China has the highest number of diabetic cases, and India is in the second position in the world. The international diabetes federation [5] surveyed and reported that India had an estimated 77,005,600 diabetic-affected people in 2019. Diabetes mellitus (commonly known as diabetes) is characterized by a metabolic disease related to hyperglycaemia, resulting in high blood sugar levels. In the human body, the insulin hormone is responsible for the movement of sugar from the blood to cells where it may be stored and used for energy. However, diabetic abnormality affects this process, wherein the body either fails to produce enough insulin or fails to utilize the insulin effectively. This impaired insulin secretion and struggle with the peripheral actions of insulin lead to beta-cell failure. The figure given below, Fig. 2, depicts the classification of different types of diabetes, including type-1, type-2, and gestational

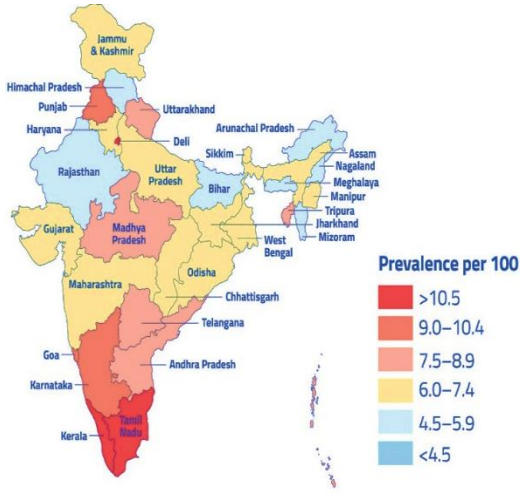

Figure. 3 Diabetic retinopathy in India

diabetes.

Diabetic patients are highly vulnerable to various short-term and long-term complications due to metabolic anomalies that can damage the organ systems and lead to aggravating life-threatening health complications, such as retinopathy, neuropathy, nephropathy, macrovascular complications, etc. In this context, diabetic retinopathy is considered one of the most commonly occurring microvascular complications, which is caused due to high levels of sugar affecting the blood vessels inside the retina [6]. The worsen conditions of diabetic retinopathy may lead to blindness. According to the recent studies, the DR was estimated 126.6 million in 2010, expected to grow to 191 million by 2030. It is a leading cause of vision loss in adults worldwide and has serious impact on life of the individual. Generally, DR is classified as non-proliferative DR (NPDR) which is a mild phase and considered as early stage. The early phase of DR is characterized by microaneurysms. On the other hand, another type is known as proliferative DR (PDR) which is a moderate stage and can lead to severe vision loss. However, these DRs are caused mainly due to Type-1 diabetes and Type-2 diabetes thus it is important to identify and diagnose the diabetes in its early stage. Regular DR screening can be helpful to prevent the vision loss.

In the Indian context, the diabetes rate varies between 5% and 16%, and due to the high population, scrutiny of the diabetes population becomes a burden on the country. Moreover, undiagnosed diabetes also remains a challenging problem in India. Currently, it is observed that at least one person in every ten adults is affected by diabetes in several states such as Kerala, Tamil Nadu, Chandigarh, Punjab, etc. [7]. Fig. 3 depicts

the prevalence of diabetic retinopathy in India.

Due to the serious impacts and constant increase in cases, identifying DR in its early stages becomes an important task. Current technological advancements have helped develop digital image processing-based automated diabetic retinopathy detection techniques. These techniques include blood vessel detection, optic disc segmentation, vessel extraction, lesion segmentation, etc. However, the accuracy of DR detection is reliant on the segmentation of various parts of the retinal images, such as microaneurysms (MAs), haemorrhages (HEs), soft exudates (SEs), and hard exudates (EXs). These parts provide detailed information that can be useful for diagnosis.Currently, deep learning-based techniques are widely adopted, enabling computers to learn diverse and complex patterns in a way that exceeds human capabilities in many areas [8]. Several deep learning schemes have been introduced for retina segmentation, DR detection, and other tasks related to retina images [9, 10]. The next section discusses some of the recent techniques.

1.1 Problem statement

Generally, the medical image segmentation is a crucial task in the field of medical image analysis, where the goal is to automatically identify and delineate regions of interest in medical images. The U-Net architecture is a popular choice for medical image segmentation due to its ability to capture both local and global features effectively. As discussed before, the DR images consist of several components. Therefore, the segmentation becomes challenging issue. The U-Net architecture, which consists of an encoder and decoder network, has been proven to be effective in this task as it can learn to localize and segment structures of interest.

However, in some cases, traditional U-Net models may struggle to handle complex cases where multiple structures overlap or when the target structures have significant variations in shape and appearance. This is where the attention mechanism comes into play. The attention mechanism can be incorporated into the U-Net architecture to enhance its performance by selectively focusing on more relevant image regions during the encoding and decoding process. By doing so, the model can give higher importance to salient features and suppress irrelevant or noisy information.

The problem definition, therefore, is to develop a U-Net based model with an attention mechanism to perform accurate and robust medical image segmentation. The model should be able to handle complex cases, such as overlapping structures, ambiguous boundaries, and variations in shape and size, while producing high-quality segmentations for clinical use. The effectiveness of the model can be measured using metrics such as dice coefficient, intersection over union (IoU), or hausdorff distance, which evaluate the similarity between the predicted segmentation and the ground truth masks provided by medical experts. The goal is to achieve state-of-the-art performance and provide reliable segmentation results to assist medical professionals in diagnosis, treatment planning, and disease monitoring.

1.2 Contributions

In this work, we focus on the segmentation task and present a new deep learning based approach for segmentation in diabetic retinopathy images. The main contribution of this work are as follows:

We have presented Unet-based deep learning architecture for segmentation. This architecture contains a skip connection module between the encoder and decoder modules of UNet. This unit of skip connection aggregates the low-level attributes of the encoder model with the high-level attributes of the decoder side. Moreover, it helps to concatenate the coarse-grain attributes with fine-grained features.

The proposed UNet architecture considers the dilated convolution which helps to capture the information without increasing the kernel parameters.

We have applied upsampling model where the remaining feature maps of the decoder block are combined with the final block of the decoder to generate the rich feature map. Finally, channel and spatial attention mechanisms are also applied to generate the final semantic map.

1.3 Novelty of the work

The proposed approach is a combination of three different methods which are dense UNet, attention mechanism where channel and spatial attention modules are employed, and finally, future fusion scheme is added which captures the global and contextual information from the images. Each module contributes in segmentation as follows:

- **Attention mechanism:** The introduction of attention mechanism allows the model to focus on relevant parts of the retinal image, enhancing the accuracy of segmentation by prioritizing important features.
- **Densely connected UNet:** The densely

connected UNet architecture ensures the seamless flow of information, enabling the network to capture intricate patterns and minute details in retinal images.

- **Hybrid fusion:** A carefully designed fusion strategy integrates the attention-enhanced features with the dense blocks of the UNet, creating a synergy that captures both global context and fine-grained details.

1.4 Organization

Rest of the article consists of following sections: section II presents the brief literature review about existing techniques, section III presents the proposed deep learning based segmentation model, section IV presents the outcome of proposed approach and its comparative analysis, finally, section V presents the concluding remarks and future scope of the research.

2. Literature review

This section presents a brief discussion of existing deep learning-based techniques for diabetic retinopathy. Deep learning is a promising technique belonging to the broad family of machine learning used to learn data patterns based on their low-level attributes.

Deep learning techniques utilize several hidden layers to identify these low-level attributes. Currently, UNet-based deep learning segmentation methods are being developed in the field of biomedical image segmentation. Ye et al. [11] developed a U-shape-based algorithm that uses an attention mechanism, a method known as the group attention network, to segment the fundus images with diabetic retinopathy. This approach uses channel group and spatial group attention modules. The channel-based model allocates resources according to the importance of feature channels. The attention mechanism helps to handle the diverse information obtained from images. The traditional U-Net merges the high and low-level features in the decoder stage to generate semantic segmentation and obtain an efficient outcome. The authors incorporated an attention mechanism to focus on the lesions. Thus, the attention mechanism presented here is used to improve the segmentation accuracy.

Hasan et al. [12] discussed the issues to detect optic disc, localization of OD, and Fovea centers due to small dataset size, inconsistency in spatial, texture and shape related information which also affects the training process of machine learning techniques. In order to overcome these issues, authors introduced encode-decoder based architecture for segmentation and localization of ODs and Fovea centers. Generally, pooling operation in decoder phaseaffects the spatial information which is handled by the residual skip connection to preserve the resolution and feature information.

Xuet al. [13] adopted U-Net based model which is further improved by incorporating a feature fusion module. The traditional UNet architecture uses a pooling layer to reduce the dimensions of the feature map. However, it affects the quality of features. Thus, the authors replaced this pooling layer with a convolution layer to minimize spatial loss. Later, it considers channel attention and skip connection mechanism which focuses on combining the multiscale features and fusing them into the encoder stage to generate the enriched feature maps. However, these models suffer from the data imbalance problem which is handled by introducing a balanced focal loss function.

Huang et al. [14] reported that existing techniques don't pay much attention to the lesions and indicated that certain lesions have a significant impact on eye pathology. Therefore, the authors presented a new deep learning-based approach that adopts the attention mechanism. In these attention mechanisms, the self-attention schemes focus on global dependencies, whereas cross-attention schemes interact between vessels and lesions. This cross-attention model uses vascular information for interaction and mitigates the ambiguity in lesion detection. Moreover, the global transformer block (GTB) is also used to identify the small lesion patterns, resulting in preserving the detailed information of the network.

Guo et al. [15] reported several issues related to lesions, such as structural complexity, variation in size, and interclass similarity between lesions. In order to address these issues, the authors introduced a dual-input attentive RefineNet to segment the lesions. This network architecture uses two encoder modules for image encoding: global and local image encoding, and the decoder module uses an attention refinement decoder. Furthermore, it utilizes ResNet50 for the whole image, and the ResNet101 model is used for the patch image. To perform multiscale feature fusion, it employs a high-level attention refinement decoder based on the dual attention mechanism.

Similarly, in [17], the authors also focused on multi-lesion segmentation and presented the cascade attentive RefineNet, a new deep learning architecture for segmentation. This architecture consists of a global image encoder, a local image encoder, and an attention mechanism-based

Table 1. Comparison between existing and proposed method

Article	Existing	Proposed
Ye et al. [11]	This mechanism allocates weights to features based on channel and spatial groups but this allocation requires additional computational resources. Moreover, limited dataset can lead to weight ambiguity resulting in misclassification.	Proposed Model uses dense skip connection model which by default carries significant attributes to the decoder side.
Hasan et al. [12]	This model uses attention mechanism but works well with only OD segmentation and Fovea localization.	The proposed mechanism is able to solve for Exudates, and Microaneurysm, OD and vessels. This makes is more robust to various segmentation in DR.
Xu et al. [13]	This approach suffers from overfitting, incomplete and contaminated labels	Proposed approach uses customized loss function to reduce the segmentation error
Huang et al.[14]	This approach suffers from various issues such as lack of spatial information, high memory requirement, and inefficiency in dense prediction.	Proposed approach has efficient learning model by using spatial and channel attention which aggregates the spatial information.
Li et al. [16]	It presents the improved UNet model in which the max-pooling operation in original U-net model was replaced by the convolution operation to obtain the robust feature information. However, it may fail to extract the fine grained information	Proposed attention helps to obtain the robust feature vector

refinement decoder. It uses the whole image, which is fed into ResNet50, and the patch image, which is fed into ResNet101, constructing a dual input system to extract the lesion features. Furthermore, the decoders use the attention mechanism, such as dual attention and low-level attention mechanism, for feature fusion, resulting in high-resolution feature generation for final predictions.

Li et al. [16] discussed that the insufficient number of samples creates difficulty in generalizing the network's performance. Therefore, a novel segmentation method is required to obtain reliable performance. Thus, the authors proposed a new approach called improved U-Net segmentation. The new architecture replaces the max-pooling layer with a convolution operation to contain more feature information. In the first stage, it extracts 128x128 patches from all slices, and later these samples are divided into different training and testing sets. The obtained training set is processed through the data augmentation phase to generate augmented data.A brief comparison of existing methods is discussed in below given table where existing approach and their drawbacks are discussed. Also, the novelty of the proposed approach is also discussed.

3. Proposed model

This section presents the proposed solution for diabetic retinopathy segmentation by using a deep learning-based strategy. As discussed before, UNet is considered one of the most promising techniques to achieve a reliable outcome of segmentation.

3.1 Basics of UNet

Unet is a deep learning-based network that was developed in 2015 for biomedical image segmentation. The traditional Unet model performs segmentation by processing the input image through two paths: the down-sampling path and the up-sampling path. The down-sampling path consists of four major blocks, as depicted in the figure given below (Fig. 4). Each of these blocks performs 3x3 unpadded convolutions with Leaky ReLu (α value considered as 0.3). Furthermore, it uses a max-pooling operation of size 2x2 with a stride of 2. After processing through the first block, the number of feature channels doubles.

Similarly, the up-sampling path also uses 4 blocks as depicted on the right side in Fig. 4. The main task of these blocks is to up-sample the feature maps and apply a 2x2 convolution on feature channels. These feature maps are concatenated with the feature maps obtained from the down-sampling path, and a 3x3 convolution operation is performed along with Leaky ReLU. The final layer performs a 1x1 convolution with the softmax function, which helps map the feature vector into two classes. Moreover, during the training phase, it uses Adam

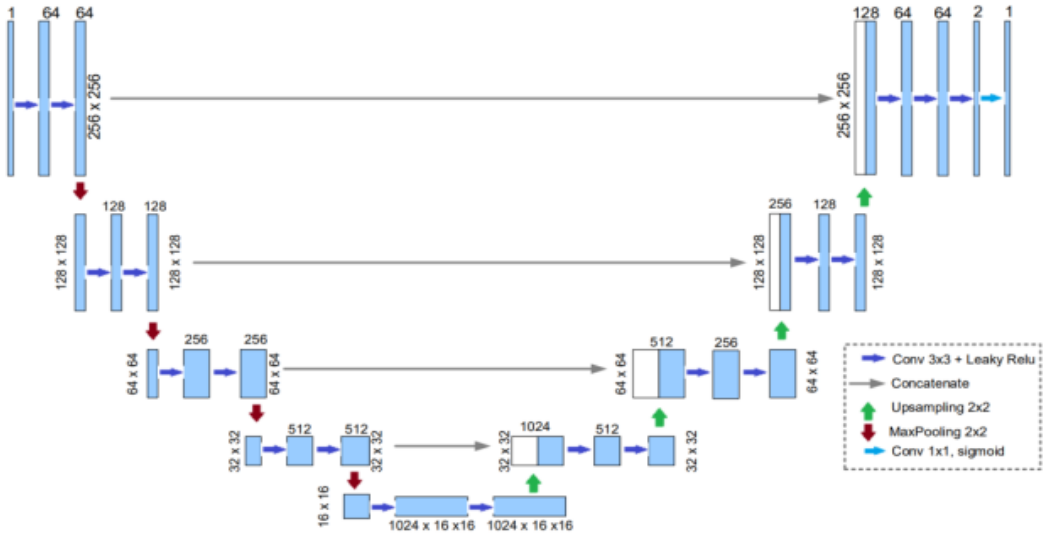

Figure. 4 UNet architecture

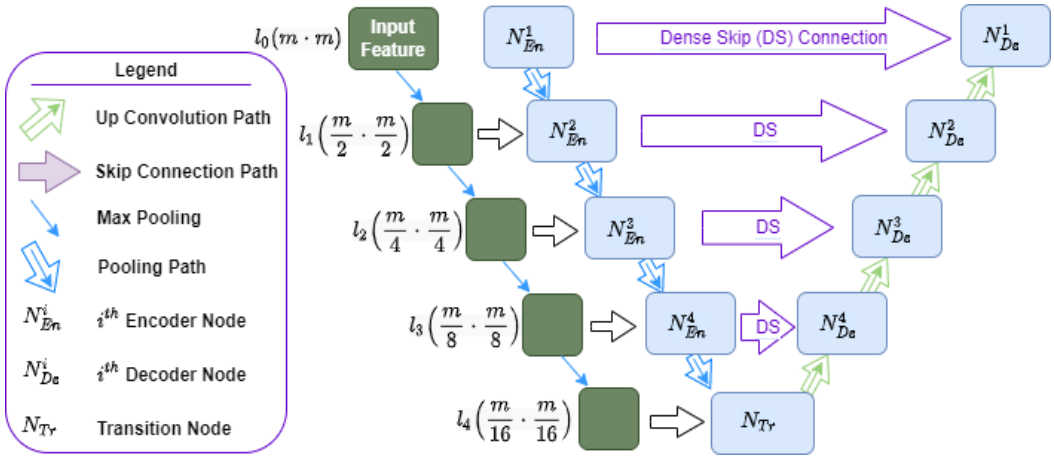

Figure. 5 Proposed UNet architecture

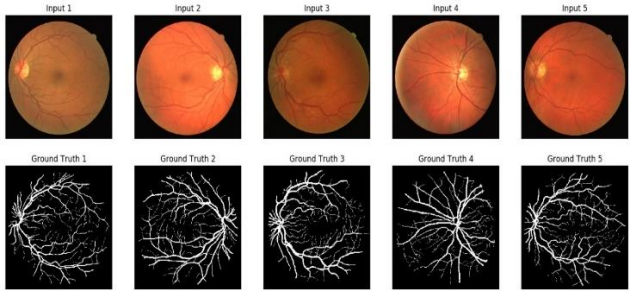

Figure. 6 Blood vessel segmentation (original and groundtruth sample images)

optimization with a loss function to generate the final output

3.2 Proposed architecture for segmentation

UNet has several advantages in the field of biomedical image segmentation. However, it suffers from various challenges, such as its failure to address the heterogeneity in organ segmentation, where the region of interest (ROI) is inconsistent. In our case, we focus on blood vessel segmentation, hard exudates, soft exudate, and Microaneurysms, which pose inconsistencies in ROI. The figure given below depicts the original image and corresponding ground truth for vessel segmentation. This shows that the segmented architecture must handle these inconsistencies.

Moreover, the traditional UNet architectures require a huge number of parameters for training. These architectures utilize residual blocks where the convolution layers store redundant features, resulting in inaccurate feature map generation. Meanwhile, these layers discard the low-level features of previous layers, thus generating low-quality feature maps. The traditional UNet models fail to retrieve the information of multiple receptive scales efficiently. Several researches have been carried out to deal with these issues but fail to handle the diverse modalities and variations [18].

In order to overcome these issues, we present a new deep learning-based architecture for diabetic

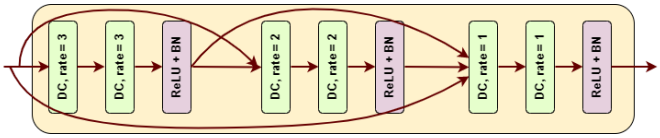

Figure. 7 Single block of the encoder module

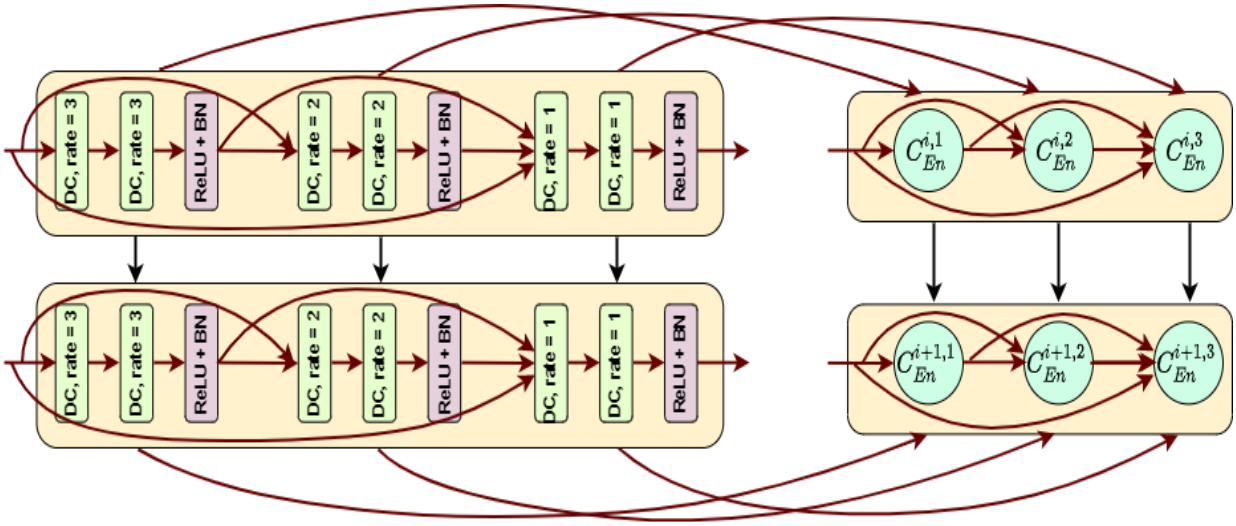

Figure. 8 Detailed model of working of encoder model

retinopathy images. The proposed approach is based on the UNet architecture; however, we have incorporated several modifications, such as modifying the UNet encoder model, considering the output features of convolution, applying dilated convolution, and incorporating the skip connection model to improve the segmentation performance. Moreover, we have incorporated Spatial and channel attention models to improve the system's performance. The figure given below (Fig. 6) depicts the overall architecture of the proposed model. Similar to the traditional UNet, this model also consists of four encoder and decoder units.

Left branch of the Proposed UNet

Let us consider $\mathbb{I}_0$ is the input obtained from the convolution for input tensor given as $\mathbb{I}_0 \in \mathbb{R}^{m\times m\times n}$where n denotes the total number of filters applied on the features and $m \times m$ denotes the size of input image. Thus, the output of the first node is expressed as:

$$\mathcal{X}_1^i(U) = \mathcal{C}^{4-1}\left(\left[\mathcal{X}_1^k\right]_{k=1}^{i-1}, \mathbb{I}_0\right), where i = [1,3] \quad (1)$$

Here, i denotes the convolution layer considered in the encoder blocks, $\mathcal{X}_1^i(U)$ denotes the i^{th} feature output obtained from the encoder node of proposed UNet. Here, $\mathcal{X}_1^0 = \phi.\mathbb{C}^r$ represents the dilated convolution with a rate r and concatenation of features is represented as $[.]$. below given Fig. 7 illustrates the detailed model of a block of encoder node architecture of proposed UNet. According to this architecture, we initially, we have considered two dilated convolution layer with dilation rate 3, in next stage, rate is reduced to 2 and later it is reduced to 1. However, these dilation layers use ReLU and Batch normalization functions.

Based on this, we extent the encoder model and illustrate the process to input the features in next block of encoder. Below given Fig. 8 shows this process with the help of max pooling operation.

This process to generate the features for next encoder node is obtained by applying max-pooling operation. This pooling operation reduces the size of features to half of the actual features. This function can be expressed as:

$$\mathbb{I}_j = \mathbb{M}\left(\mathbb{I}_{j-1}\right) with j = [1,4]$$
$$\mathcal{X}_h^i(U) = \mathbb{C}^{4-i}\left(\left[\left[\mathcal{X}_1^k\right]_{k=1}^{i-1}, \mathbb{M}_0\left(\mathcal{X}_{h-1}^i\right), \mathbb{I}_h\right]\right) \quad (2)$$

Here, j varies from 1 to 4 because four max pooling operations are performed at the encoder side,h denotes the order of encoder with $\mathcal{X}_5$ as the output obtained from the transition node present at the bottom of the network and $\mathbb{M}(.)$ denotes the max pooling operation. The complete process of feature map generation through various encoder model, processing these feature maps through transition phase to provide as input to the decoder is connected process which can be illustrated as:

Thus, overall, we have four encoder model, four decoder model connected through the transition

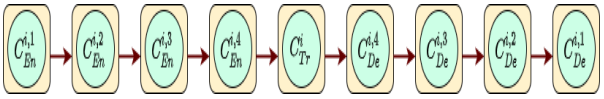

Figure. 9 Encoder to decoder module blocks

model.

- ***Right branch of the Proposed UNet and Skip connections***

Similar to the encoder, model, we perform some specific operations at the decoder phase. This module also uses upsampling and up-convolution operations to restore the features to generate the final output. According to the main architecture, we use deep skip connections to concatenate the features of encoder model to the up-sampled features. This process is done through the skip connection mechanism as depicted in Fig. 6. Generally, the traditional UNet models also use skip connection but it considers output feature of the last convolution unit. This results in generation of less significant features and some of the important features which are generated by other convolution units are discarded from the process. To mitigate this challenge, we incorporate skip connection to consider the features from all convolution units. Below given Fig. 10 depicts the process of skip connection where output of each convolution is processed and fed to the next decoder module.

Let γ_h^i denotes the output obtained from the i^{th} convolution unit in the h^{th}decoder node. This can be expressed as follows:

$$\gamma_h^i(U) = \mathbb{C}^{4-i}\left(\left[[\gamma_h^k]_{k=1}^{i-1}, \tau(\gamma_{h+1}^i), [\mathcal{X}_h^m]_{m=1}^i\right]\right), \quad where\ i \in [1,3] \qquad (3)$$

where $\tau(.)$ Denotes the up convolution operation and γ_h^i is the output feature of i^{th} convolution. Each model (total three model because of three units) at the decoder end generates output where 1x1 convolution and sigmoid functions will be applied and finally. These maps are concatenated to obtain the output. This can be expressed as:

$$\mathbb{F} = \delta\left(\Gamma\left(\left[\delta\left(\Gamma(\gamma_1^i)\right)\right]_{i=1}^3\right)\right) \qquad (4)$$

Where $\mathbb{F}$ denotes the final output of the proposed UNet model, $\delta(.)$represents the sigmoid function, and $\Gamma(.)$ represents 1×1 convolution.

- ***Attention mechanism***

However, the inconsistency of ROI in these images cause a severe challenge which affects the segmentation performance. To overcome this issue, attention mechanism plays an important role. Thus, we incorporate a novel attention mechanism which performs channel and spatial attention mechanism. Below given figure depicts the overall architecture of proposed approach where feature maps are generated through the UNet which also considers an upsampling layer.

The channel attention mechanism uses channel attention matrix to combined the outputs of final feature map, similarly, the spatial attention module uses context modelling block. The output map of these blocks is fused by applying sum fusion method. Thus, the final segmentation map is generated.

The context model scheme uses 1x1 convolution, softmax layer, layer normalization with ReLU and 1x1 convolution in sequence. The channel attention map is denoted as $X \in R^{C\times C}$ obtained from original features $\mathbb{A} \in R^{C\times H\times W}$. these features are reshaped from $\mathbb{A}$ to $R^{C\times N}$ followed by matrix multiplication between $\mathbb{A}$ and $\mathbb{A}'$ i.e. transposed features. Finally, softmax layer is applied which produces channel attention map, expressed as:

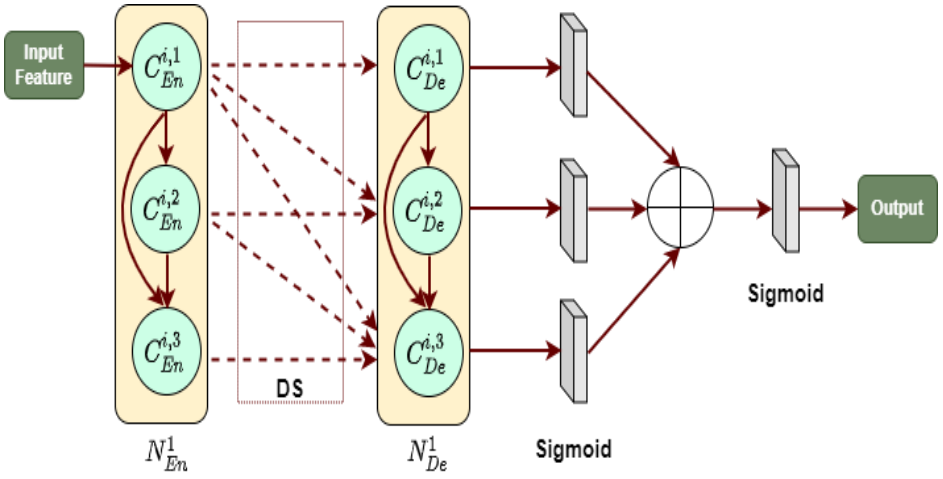

Figure. 10 Skip connection from encoder to decoder module

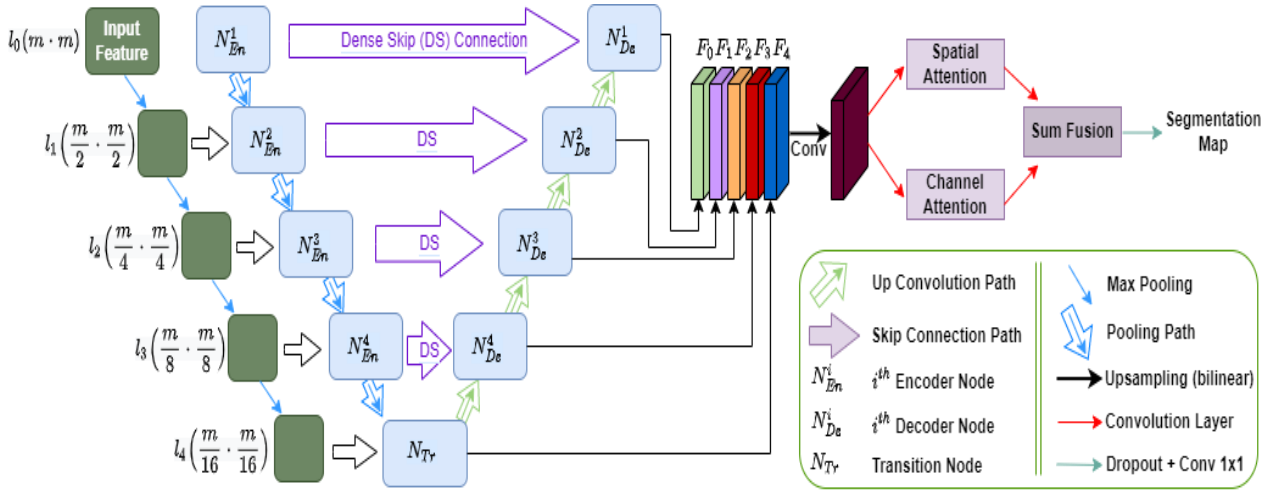

Figure. 11 Final architecture of proposed model

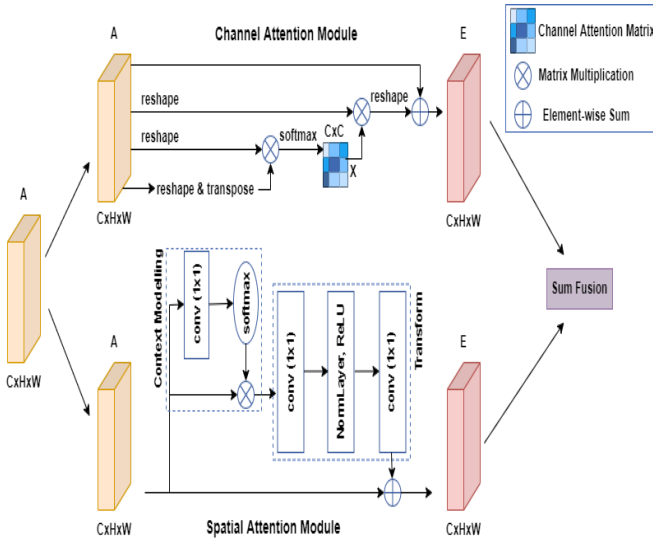

Figure. 12 Channel and spatial attention mechanism architecture

$$x_{ij} = \frac{\exp(\mathbb{A}_i.\mathbb{A}_j)}{\sum_{i=1}^{C}\exp(\mathbb{A}_i.\mathbb{A}_j)} \quad (5)$$

Later, we perform multiplication of x_{ij} with $\mathbb{A}$and scale parameter $\beta$to generate the final output as:

$$\mathbb{E}_j = \beta\sum_{i=1}^{\mathbb{C}}(x_{ji}\mathbb{A}_i) + \mathbb{A}_j \quad (6)$$

Similarly, spatial attention helps to generate the rich-local features, layer normalization improves the object detection and segmentation to simplify the optimization process. The output of attention module can be computed as:

$$\mathbb{Z}_j = \mathbb{A}_i + Wv2\ ReLU\left(LN\left(Wv1\sum_{j=1}^{N_p}\frac{e^{W_k\mathbb{A}_j}}{\sum_{m=1}^{N_p}e^{W_k\mathbb{A}_m}}\right)\right) \quad (7)$$

Where $\frac{e^{W_k\mathbb{A}_j}}{\sum_{m=1}^{N_p}e^{W_k\mathbb{A}_m}}$ represents the weight of global attention, N_p is the number of positions in the feature map, and $\mathbb{A}$ is the input feature map.

Finally, we define a loss function based on the predicted and ground truth dataset. The binary cross entropy loss is defined as:

$$\mathbb{L}_{BCE} = BinCrELoss(GT_i, Pred_i) = -\sum_{i=1}^{W}\sum_{j=1}^{H}\left[\left(GT_{i,j}\times\log Pred_{i,j}\right) + \left(1-GT_{i,j}\right)\times\log\left(1-Pred_{i,j}\right)\right] \quad (8)$$

4. Results and discussion

This section presents the complete experimental analysis of proposed approach where we evaluate the performance of proposed approach and compare its performance with the existing schemes. These experimentsare performed on windows 10 operating system with NVDIA 2060 graphic card with 8 GB memory and 16GB RAM.

4.1 Dataset description

In order to measure the performance of proposed approach, we have considered 8 standard datasets which are HRF [19], ChaseDB1 [20], DIARETDB0 [21], DIARETDB1 [22], STARE [23], DRIVE [24], and IDRiD [25].

DRIVE dataset [24]: this dataset contains total 40 retinal images where 20 images are considered for training and 20 images are considered for testing. In this dataset, seven cases are diagnosed as diabetic retinopathy case whereas other 33 cases are healthy cases.

High-resolution fundus dataset [19]: this dataset contains total 30 retinal images where 15 images belong to healthy category and other images belong

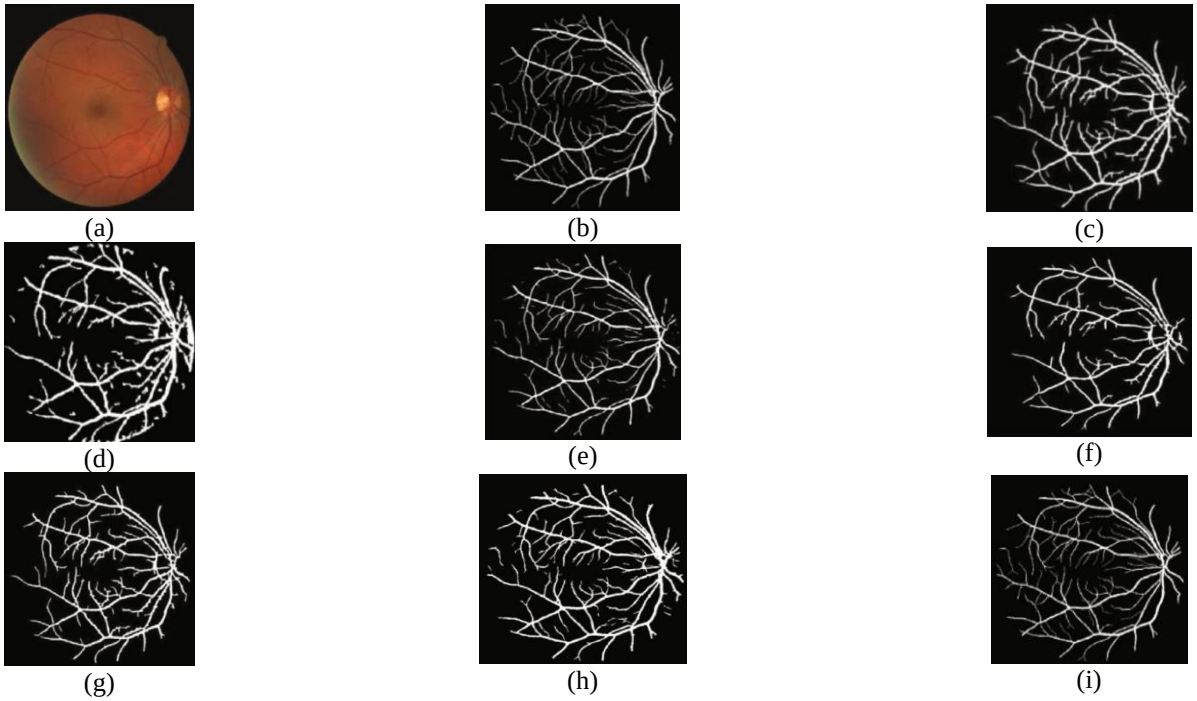

Figure. 13 Blood vessel segmentation results for DRIVE dataset: (a) Input image, (b) Ground truth, (c) Technique in [27], (d) Technique in [28], (e) Technique in [29], (f) Technique in [30], (g) Technique in [31], (h) Technique in [26], and (i) Proposed approach

Table 2. Performance matrices

Performance Metrics	Computation formula
SensitivityS_n	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
DSC	$\frac{2\times TP}{2\times TP+FP+FN}$
Accuracy	$\frac{TP+TN}{TN+TP+FP+FN}$
PPV	$\frac{TP}{TP+FP}$
AUC	$\approx 0.5(Sn+Sp)$

to healthy categories. This dataset also contains blood vessels ground truth for both healthy and diabetic categories.

ChaseDB1 [20]: this dataset contains total 28 images and their blood vessels are manually segmented by the clinical experts. Moreover, this dataset consists of samples of 14 children. The training portion consist of 20 images and testing portion consist of 8 images.

DIARETDB0 [21]: in this dataset total 130 images are present where 20 images belong to normal category and rest of the images belong to the diabetic class.

DIARETDB1 [22]: it contains total 89 images where experts have categorised five images as normal and rest of the images are categorised as mild NPDR. This dataset contains all type of abnormalities such as hard exudates, soft exudates, haemorrhages and microaneurysms.

STARE [23]: Total 400 images are provided in this this dataset. Out of these image, 20 images have groundtruth of blood vessels, thus, 10 images are considered in healthy category and 10 images are considered for diabetic case.

IDRiD [25]: this dataset contains total 81 images with 4 types of lesion groundtruth such as HM, MA, Hard EX, soft EX. These groundtruth are annotated at pixel-level. Out of these 81 images, 54 images are considered for training and 27 images are considered for testing.

4.2 Performance measurement

In order to measure the performance of proposed approach, we have considered several metrics such as sensitivity, specificity, dice score, accuracy, positive predicted value and area under curve. These metrics can be computed as follows:

Here, TP, FP, FN, and TN denote true positive, false positive, false negative and true negative, respectively. Sensitivity (S_n) is the true positive rate, Specificity (S_p) specificity is proportion of true negative and false positive, accuracy represents the rate of actual prediction, false positive represents the incorrect positive predictions, false negative denotes incorrect negative prediction, DSC denotes similarity between actual ground truth and predicted

Table 3. Comparative performance for DRIVE dataset

Method	Se	Sp	Acc	AUC
Khan et al. [32]	0.7380	0.9675	0.950	0.8530
Khan et al. [33]	0.7820	0.9725	0.9530	0.8770
Soomro et al. [28]	0.7455	0.917	0.946	0.8315
Ngo and Han et al. [34]	0.7465	0.9840	0.9540	0.8655
Biswal et al. [35]	0.71	0.97	0.95	0.84
Yan et al. [36]	0.7650	0.9820	0.9540	0.8740
Oliveria et al. [37]	0.8039	0.9804	0.9576	0.8922
Wang et al. [38]	0.7648	0.9817	0.9541	0.8733
Feng et al. [29]	0.7625	0.9809	0.9528	0.8717
Ribeiro et al. [39]	0.7880	0.9819	0.9569	0.8850
Dharmawan et al [40]	0.8314	0.9726	-	0.902
Saroj et al. [27]	0.7307	0.9761	0.9544	0.8534
Dash and Senapati et al. [41]	0.7403	0.9905	0.9661	0.8654
Biswas et al. [42]	0.7823	0.9814	0.9561	0.8819
Budak et al. [43]	0.7439	0.9900	0.9685	0.8670
Ma et al. [26]	0.8745	0.9624	0.9546	0.9185
AbdelMaksoud et al. [45]	62.45	98.79	95.61	-
B-COSFIRE filter [46]	76.5	97.04	94.4	96.1
Gao et al [47]	78	98.7	96.3	97.7
Adapa et al. [48]	69.9	98.1	94.5	93.9
AbdelMaksoud et al. [44]	72.58	98.89	96.58	97.84
Proposed approach	89.96	98.96	98.15	98.50

segmentation mask, PPP denotes the correct and incorrect true positive classes over the incorrect positive prediction. Finally, AUC is computed which is approximately near to half of the sum of sensitivity and specificity.

4.3 Comparative analysis for blood vessel segmentation

This section presents the comparative analysis for blood vessel segmentation. First of all, we measure the qualitative performance of proposed approach and compare its performance with existing schemes. Further, we validate the outcome of proposed approach by presenting quantitative comparative analysis.

Above given Fig. 13 illustrates the comparative analysis for blood vessel segmentation. For one input image, we have compared the output of several techniques such as mentioned in [26-31]. The outcome is presented in the extracted vessels. The quantitative measurement performance is presented in below given Table 2 where outcome of proposed approach is compared with the several existing schemes in terms of sensitivity, specificity, accuracy and AUC.

Table 4. Comparative analysis for STARE dataset

Method	Se	Sp	Acc	AUC
Khan et al. [32]	0.7359	0.9708	0.9502	0.8534
Khan et al. [33]	0.7728	0.9649	0.9518	0.8689
Soomro et al. [28]	0.748	0.922	0.948	0.835
Yan et al. [36]	0.7581	0.9846	0.9612	0.8714
Oliveria et al. [37]	0.8315	0.9858	0.9694	0.9087
Wang et al. [38]	0.7523	0.9885	0.9664	0.8704
Feng et al. [29]	0.7709	0.9848	0.9633	0.8779
Dharmawan et al [40]	0.7924	0.9828	-	0.8876
Saroj et al. [27]	0.7278	0.9724	0.9509	0.8501
AbdelMaksoud et al. [44]	0.8903	0.9744	0.9699	0.9323
Proposed model	0.9256	0.9840	0.9854	0.9640

This comparative study shows that the proposed approach achieves the overall performance as 89.96%, 98.96%, 98.15%, and 98.50% in terms of sensitivity, specificity, accuracy, and AUC, respectively. The existing approach [32] only considers contrast enhancement mechanism but it leads to diminish the fine grain information of vessels which impacts on overall accuracy. In [34] spatial dropout is used to reduce the overfitting and generalize the training process but it fails to achieve the improved sensitivity performance due to loss to spatial and contextual information. Similarly, the technique mentioned in [35] uses line detector with

Table 5. Comparative analysis for CHASE_DB1 dataset

Method	Se	Sp	ACC	AUC
Biswal et al. [35]	0.76	0.97	-	0.865
Yan et al. [36]	0.7633	0.9809	0.9610	0.8721
Oliveria et al. [37]	0.7779	0.9864	0.9653	0.8822
Wang et al. [38]	0.7730	0.9792	0.9603	0.8761
Soomro et al. [28]	0.8020	0.968	0.891	0.885
AbdelMaksoud et al. [44]	0.8916	0.9596	0.9561	0.9256
Proposed approach	0.9560	0.9683	0.9750	0.9710

Table 6. Average performance for blood vessel segmentation for different dataset

Method	DB	Acc	AUC	Sn	Sp	PPV	DSC
AbdelMaksoud et al. [45]	DRIVE	95.6	-	62.4	98.7	-	71
Soares et al.		94.6	95.9	-	-	-	-
BCOSFIRE filter [46]		94.4	96.1	76.5	97.0	-	-
Gao et al [47]		96.3	97.7	78	98.7	89	-
Adapa et al. [48]		94.5	93.9	69.9	98.1	-	-
AbdelMaksoud et al. [44]		96.5	97.8	72.5	98.8	86.2	78.8
Proposed Approach		98.5	98.9	81.2	99.2	95.2	86.2
AbdelMaksoud et al. [45]	STARE	96.1	-	69.8	97.6	-	68.2
Soares et al.		97.7	96.5	-	-	-	-
BCOSFIRE filter [46]		94.9	95.6	77.1	97.0	-	-
Adapa et al.[48]		94.8	95	62.9	98.3	-	-
AbdelMaksoud et al. [44]		95.5	94.9	66.1	97.9	72.2	69.0
Proposed Approach		96.2	97.5	86.2	98.5	86.9	82.5
B-COSFIRE filter [46]	CHASEDB1	93.8	94.8	75.8	95.8	-	-
AbdelMaksoud et al. [44]		96.1	95.0	56.7	98.9	79.0	69.0
Proposed approach		98.2	96.0	85.3	99.1	89.8	85.2
B-COSFIRE filter [46]	HRF	96.4	95	75	97.4	-	-
AbdelMaksoud et al. [44]		95.6	95.3	70.1	98.2	85.1	76.2
Proposed approach		98.5	96.3	84.2	99.5	91.2	82.5

Table 7. Comparative analysis for Ex and MA segmentation for different dataset

Technique	Dataset Used	Lesion type	Accuracy	AUC	Sn	Sp	PPV	DSC
Kou et al [49]	IDRiD	Ex	98	98.6	95	93.9	-	-
		MA	98	98.01	92.9	93.5	-	-
Abdelmaksoud et al [44]	IDRiD	Ex	100	-	100	100	-	100
		MA	97.8	-	100	80	-	98.7
Khojasteh et al [50]	DIARETDB1	Ex	98	-	95	98	94	-
		MA	94	-	85	96	83	-
Yan et al [36]	DIARETDB1	Ex	-	96.4	89.1	98	-	-
		MA	-	96.4	97.8	95.5	-	-
Abdelmaksoud et [44]	IDRiD	Ex	100	100	100	100	100	100
		MA	99.9	98	99.8	99.5	98.5	99.9
	DIARETDB1	Ex	98.72	97.8	97	96.2	95	96
		MA	98	97	96	98	94	95
Proposed Approach	IDRiD	Ex	99.86	99.9	99.9	99.9	99.9	99.9
		MA	99.9	99.1	99.8	99.6	99.2	99.9
	DIARETDB1	Ex	99.80	98.5	98.2	97.5	96.3	97.8
		MA	98.56	98.40	98.1	98.8	97.3	98.5

multiple masks but fails to detect bright pathological regions which affects the sensitivity performance.

Similarly, we measure the performance for STARE dataset and compared the obtained performance with existing schemes. Below given Table 3 shows the performance for STARE dataset.

According to this experiment, the proposed approach achieves the performance as 0.9256, 0.9840, 0.9854, and 0.9640 in terms of sensitivity, specificity, accuracy and AUC, respectively.

Further, we present the experimental analysis CHASE_DB1 dataset and compared the obtained performance with existing schemes. Below given Table 4 shows the comparative performance.

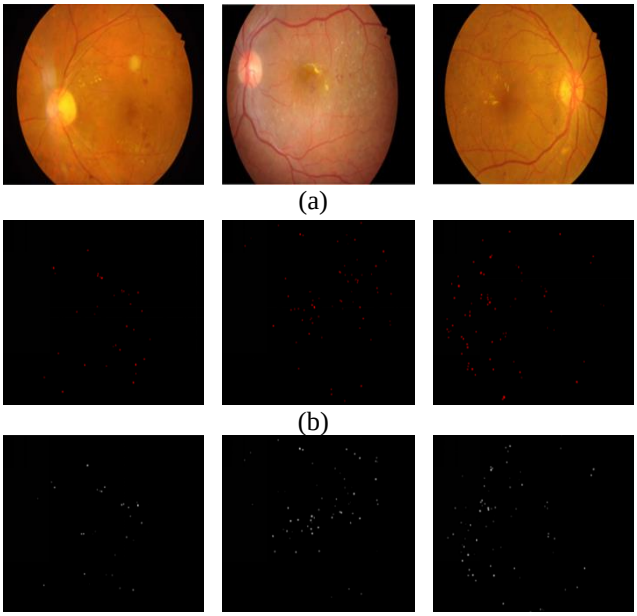

Figure. 14 MA segmentation by using proposed approach: (a) Original images, (b) Ground truth images, and (c) Output using proposed approach

The average performance for blood vessel segmentation is presented in below given table where we compare the performance of proposed approach with state-of-art segmentation techniques.

4.4 Comparative analysis for exudates, and microaneurysm (MA) segmentation

In this subsection, we present a comparative analysis for exudate and Microaneurysm segmentation. In order to achieve this objective, we have considered retraining the network for exudate and MA data. The obtained performance is presented in below given figure.

First row of the Fig. 14 shows the original images, second row shows the groundtruth and 3rd row shows the final output of segmentation.

Further, we extend this experiment to measure the performance for exudate detection and performed experiment for exudate detection. The comparative performance for different dataset is presented in blow given table.

This experiment shows that the proposed approach achieves exudate detection accuracy as 99.86 for IDRiD dataset and 99.9% accuracy as MA detection whereas EX and MA detection accuracies are obtained as 99.80, 98.56 for DIARETDB1 dataset.

5. Conclusion

The growing impact of diabetes globally and increasing number of cases of diabetic retinopathy at an alarming rate has drawn attention of research community. Several techniques have been introduced for DR detection, however, applying machine learning schemes directly on the raw images doesn't provide a reliable outcome. Thus, segmentation plays an important role to facilitate better diagnosis system for clinicians and medical experts. In this work, we have focused on deep learning based technique for segmentation. The proposed approach is based on the well-known UNet segmentation model; however, we have modified the UNet architecture by incorporating, skip connection, channel attention, and spatial attention. The skip connection helps to generate the rich feature maps. Further, upsampling layers are also connected to final map, and attention mechanism is also used to generate the final semantic feature map. The proposed approach is trained for blood vessel detection, exudates detection and Microaneurysmsegmentation. This approach is tested for publicly available datasets such as IDRiD, DIARETDB1, STARE, ChaseDB1, DRIVE and HRF dataset. The comparative analysis shows that proposed approach achieves improved segmentation accuracy for diabetic retinopathy cases.

Conflicts of interest

The authors declare there is no conflict of interest.

Author contributions

Following are the contribution of authors: "Conceptualization, Anand M and Dr. Meenakshi Sundaram A; methodology, Anand M; software, Anand M; validation, Anand M, Dr. Meenakshi Sundaram A; formal analysis, Anand M; investigation, Anand M; resources, Anand M; data curation, Anand M writing—original draft preparation, Anand M; writing—review and editing, Anand M; visualization, Anand M; supervision, Dr. Meenakshi Sundaram A.

Acknowledgement

This endeavour would not have been possible without the generous support from REVA University, Bengaluru. Words cannot express my gratitude to my Research Supervisor Dr. Meenakshi Sundaram A, Associate Professor, School of Computer Science & Engineering, REVA University, Bengaluru and chair of my committee for her invaluable patience and feedback. I also could not have undertaken this journey without my defense committee, who generously

provided knowledge and expertise.

References

[1] P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin,S. Colagiuri, L. Guariguata, A. A. Motala, K. Ogurtsova et al., "Globaland regional diabetes prevalence estimates for 2019 and projectionsfor 2030 and 2045: Results from the international diabetes federationdiabetes atlas", *Diabetes Research and Clinical Practice*, Vol. 157, p. 107843, 2019.

[2] J. P. O. Li, H. Liu, D. S. Ting, S. Jeon, R. P. Chan, J. E. Kim, D. A. Sim, P. B. Thomas, H. Lin, Y. Chen et al., "Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective", *Progress in Retinal and Eye Research*, Vol. 82, p. 100900, 2021.

[3] K. Oh, H. M. Kang, D. Leem, H. Lee, K. Y. Seo, and S. Yoon, "Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images", *Scientific Reports*, Vol. 11, No. 1, p. 1897, 2021.

[4] J. P. Tripathy, J. Thakur, G. Jeet, S. Chawla, S. Jain, A. Pal, R. Prasad, and R. Saran, "Prevalence and risk factors of diabetes in a large community-based study in north India: results from a steps survey inpunjab, India", *Diabetology & Metabolic Syndrome*, Vol. 9, No. 1, pp. 1–8, 2017.

[5] I. D. Atlas et al., "Idf diabetes atlas", *International Diabetes Federation (9th edition)*, Retrieved from http://www. idf. org/about-diabetes/facts-figures, 2019.

[6] A. Gupta and R. Chhikara, "Diabetic retinopathy: Present and past", *Procedia Computer Science*, Vol. 132, pp. 1432–1440, 2018.

[7] R. Raman, K. Ramasamy, R. Rajalakshmi, S. Sivaprasad, and S. Natarajan, "Diabetic retinopathy screening guidelines in india: All India ophthalmological society diabetic retinopathy task force and vitreoretinal society of India consensus statement", *Indian Journal of Ophthalmology*, Vol. 69, No. 3, p. 678, 2021.

[8] L. Dai, L. Wu, H. Li, C. Cai, Q. Wu, H. Kong, R. Liu, X. Wang, X. Hou, Y. Liu, X. Long, Y. Wen, L. Lu, Y. Shen, Y. Chen, D. Shen, X. Yang, H. Zou, B. Sheng, and W. Jia, "A deep learning system for detecting diabetic retinopathyacross the disease spectrum", *Nature Communications*, Vol. 12, No. 1, p. 3242, 2021.

[9] A. Bilal, G. Sun, and S. Mazhar, "Survey on recent developmentsin automatic detection of diabetic retinopathy", *Journal Franc¸, Ophtalmologie*, Vol. 44, No. 3, pp. 420–440, 2021.

[10] M. Z. Atwany, A. H. Sahyoun, and M. Yaqub, "Deep learning techniques for diabetic retinopathy classification: A survey", *IEEE Access*, Vol. 10, pp. 28-642 p. 655, 2022.

[11] L. Ye, W. Zhu, S. Feng, and X. Chen, "Ganet: group attention network for diabetic retinopathy image segmentation", *in Medical Imaging 2020: Image Processing*, Vol. 11313, pp. 14–19, SPIE, 2020.

[12] M. K. Hasan, M. A. Alam, M. T. E. Elahi, S. Roy, and R. Mart´ı, "Drnet: Segmentation and localization of optic disc and fovea from diabetic retinopathy image", *Artificial Intelligence in Medicine*, Vol. 111, p. 102001, 2021.

[13] Y. Xu, Z. Zhou, X. Li, N. Zhang, M. Zhang, P. Wei et al., "Ffu-net: Feature fusion u-net for lesion segmentation of diabetic retinopathy", *BioMed Research International*, Vol. 2021, 2021.

[14] S. Huang, J. Li, Y. Xiao, N. Shen, and T. Xu, "Rtnet: relation transformer network for diabetic retinopathy multi-lesion segmentation", *IEEE Transactions on Medical Imaging*, Vol. 41, No. 6, pp. 1596–1607,2022.

[15] Y. Guo and Y. Peng, "Multiple lesion segmentation in diabetic retinopathy with dual-input attentive refinenet", *Applied Intelligence*, Vol. 52, No. 12, pp. 14-440 p. 464, 2022.

[16] Q. Li, S. Fan, and C. Chen, "An intelligent segmentation and diagnosis method for diabetic retinopathy based on improved u-net network", *Journal of Medical Systems*, Vol. 43, pp. 1–9, 2019.

[17] Y. Guo and Y. Peng, "Carnet: Cascade attentive refinenet for multi-lesion segmentation of diabetic retinopathy images", *Complex & Intelligent Systems*, Vol. 8, No. 2, pp. 1681–1701, 2022.

[18] P. Ahmad, H. Jin, R. Alroobaea, S. Qamar, R. Zheng, F. Alnajjar, and F. Aboudi, "Mhunet: A multi-scale hierarchical based architecture for medical image segmentation", *IEEE Access*, Vol. 9, pp. 148-384 p. 408, 2021.

[19] "High-resolution fundus (hrf) image database", https://www5.cs.fau.de/research/data/fundus-images/, (Accessed on07/31/2023).

[20] "Chasedb1kaggle", https://www.kaggle.com/datasets/khoongweihao/chasedb1, (Accessed on 07/31/2023).

[21] "www.it.lut.fi", https://www.it.lut.fi/project/imageret/diaretdb0/, (Accessedon 07/31/2023).
[22] "www.it.lut.fi", https://www.it.lut.fi/project/imageret/diaretdb1/index.html, (Accessed on 07/31/2023).
[23] "The stare project", https://cecas.clemson.edu/ahoover/stare/, (Accessedon 07/31/2023).
[24] "Drive dataset -machine learning datasets", https://datasets.activeloop.ai/docs/ml/datasets/drive-dataset, (Accessed on 07/31/2023).
[25] "Indian diabetic retinopathy image dataset — kaggle", https://www.kaggle.com/datasets/aaryapatel98/indian-diabetic-retinopathy-image-dataset, (Accessed on 07/31/2023).
[26] Y. Ma, Z. Zhu, Z. Dong, T. Shen, M. Sun, and W. Kong, "Multichannel retinal blood vessel segmentation based on the combination of matched filter and u-net network", *BioMed Research International*, Vol. 2021, pp. 1–18, 2021.
[27] S. K. Saroj, R. Kumar, and N. P. Singh, "Frechet pdf based matched filter approach for retinal blood vessels segmentation", *Computer Methods And programs in Biomedicine*, Vol. 194, p. 105490, 2020.
[28] T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, M. A. Khan, M. Paul, and L. Zheng, "Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network", in *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, *IEEE*, 2017, pp. 1–8.
[29] S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "Ccnet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features", *Neurocomputing*, Vol. 392, pp. 268–276, 2020.
[30] N. P. Singh and R. Srivastava, "Retinal blood vessels segmentation by using gumbel probability distribution function based matched filter", *Computer Methods and Programs in Biomedicine*, Vol. 129, pp. 40–50, 2016.
[31] "Weibull probability distribution function-based matched filter approach for retinal blood vessels segmentation", in *Advances in Computational Intelligence: Proceedings of International Conference on Computational Intelligence 2015*, Springer, pp. 427–437, 2017.
[32] M. A. Khan, T. A. Soomro, T. M. Khan, D. G. Bailey, J. Gao, and N. Mir, "Automatic retinal vessel extraction algorithm based on contrast-sensitive schemes", in *2016 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. *IEEE*, pp. 1–5, 2016.
[33] K. B. Khan, A. A. Khaliq, M. Shahid, and S. Khan, "An efficient technique for retinal vessel segmentation and denoising using modified iso data and clahe", *IIUM Engineering Journal*, Vol. 17, No. 2, pp. 31–46, 2016.
[34] L. Ngo and J. H. Han, "Multi-level deep neural network for efficient segmentation of blood vessels in fundus images", *Electronics Letters*, Vol. 53, No. 16, pp. 1096–1098, 2017.
[35] B. Biswal, T. Pooja, and N. B. Subrahmanyam, "Robust retinal blood vessel segmentation using line detectors with multiple masks", *IET Image Processing*, Vol. 12, No. 3, pp. 389–399, 2018.
[36] Z. Yan, X. Yang, and K. T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation", *IEEE Transactions on Biomedical Engineering*, Vol. 65, No. 9, pp. 1912–1923, 2018.
[37] A. Oliveira, S. Pereira, and C. A. Silva, "Retinal vessel segmentation based on fully convolutional neural networks", *Expert Systems with Applications*, Vol. 112, pp. 229–242, 2018.
[38] X. Wang, X. Jiang, and J. Ren, "Blood vessel segmentation from fundus image by a cascade classification framework", *Pattern Recognition*, Vol. 88, pp. 331–341, 2019.
[39] A. Ribeiro, A. P. Lopes, and C. A. Silva, "Ensemble learning approaches for retinal vessel segmentation", in *2019 IEEE 6th Portuguese Meeting on Bioengineering (ENBENG)*. *IEEE*, pp. 1–4, 2019.
[40] D. A. Dharmawan, D. Li, B. P. Ng, and S. Rahardja, "A new hybrid algorithm for retinal vessels segmentation on fundus images", *IEEE Access*, Vol. 7, pp. 41-885 p. 896, 2019.
[41] S. Dash and M. R. Senapati, "Enhancing detection of retinal blood vessels by combined approach of dwt, tylercoye and gamma correction", *Biomedical Signal Processing and Control*, Vol. 57, p. 101740, 2020.
[42] R. Biswas, A. Vasan, and S. S. Roy, "Dilated deep neural network for segmentation of retinal blood vessels in fundus images", *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, Vol. 44, pp. 505–518, 2020.
[43] U. Budak, Z. Comert, M. Cıbuk, and A. Seng¨ur, "Dccmed-net: Densely connected and concatenated multi encoder-decoder CNNs for retinal vessel extraction from fundus images", *Medical Hypotheses*, Vol. 134, p. 109426, 2020.
[44] E. Abdelmaksoud, S. E. Sappagh, S. Barakat, T. Abuhmed, and M. Elmogy, "Automatic

diabetic retinopathy grading system based on detecting multiple retinal lesions", *IEEE Access*, Vol. 9, No. 960, pp. 15-939, 2021.

[45] E. Abdelmaksoud, S. Barakat, and M. Elmogy, "A comprehensive diagnosis system for early signs and different diabetic retinopathy grades using fundus retinal images based on pathological changes detection", *Computers in Biology and Medicine*, Vol. 126, No. 11, p. 104039, 2020.

[46] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable cosfire filters for vessel delineation with application to retinal images", *Medical Image Analysis*, Vol. 19, No. 1, pp. 46–57, 2015, [Online]. Available: https://api.semanticscholar.org/CorpusID:10684069.

[47] X. Gao, Y. Cai, C. Qiu, and Y. Cui, "Retinal blood vessel segmentation based on the gaussian matched filter and u-net", in *2017 10thInternational Congress on Image and Signal Processing, Bio Medical Engineering and Informatics (CISP-BMEI)*, pp. 1–5, 2017.

[48] D. Adapa, A. N. J. Raj, S. N. Alisetti, Z. Zhuang, G. K, and G. R. Naik, "A supervised blood vessel segmentation technique for digital fundus images using zernike moment based features", *PLoS ONE*, Vol. 15, 2020, [Online]. Available: https://api.semanticscholar.org/CorpusID:212621276.

[49] C. Kou, W. Li, W. Liang, Z. Yu, and J. Hao, "Microaneurysms segmentation with a U-Net based on recurrent residual convolutional neural network", *J Med Imaging (Bellingham)*, Vol. 6, No. 2, p. 025008, 2019, Apr, doi: 10.1117/1.JMI.6.2.025008. Epub 2019 Jun 19. PMID: 31259200; PMCID: PMC6582229.

[50] H. Khojasteh, H. Faghihi, M. Mirghorbani, and F. Hajizadeh, "Diabetic Retinopathy and Retinal Vascular Diseases", *Diagnostics in Ocular Imaging: Cornea, Retina, Glaucoma and Orbit*, pp. 499-515, 2021.