IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 21 October 2023, accepted 20 December 2023, date of publication 25 December 2023,
date of current version 3 January 2024.

Digital Object Identifier 10.1109/ACCESS.2023.3347196

RESEARCH ARTICLE

MSR U-Net: An Improved U-Net Model for Retinal Blood Vessel Segmentation

GIRI BABU KANDE[1], (Senior Member, IEEE), LOGESH RAVI[2,3], NITYA KANDE[4], MADHUSUDANA RAO NALLURI[5], (Member, IEEE), HOSSAM KOTB[6], KAREEM M. ABORAS[6], AMR YOUSEF[7,8], (Member, IEEE), YAZEED YASIN GHADI[9], AND A. SASIKUMAR[10]

[1]Vasireddy Venkatadri Institute of Technology, Nambur, Guntur 522508, India
[2]Centre for Advanced Data Science, Vellore Institute of Technology, Chennai, Tamil Nadu 600127, India
[3]School of Electronics Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu 600127, India
[4]VIT-AP University, Amaravathi 522237, India
[5]School of Computing, Amrita Vishwa Vidyapeetham, Kuragallu, Guntur 522503, India
[6]Department of Electrical Power and Machines, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt
[7]Engineering Mathematics Department, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt
[8]Electrical Engineering Department, University of Business and Technology, Al Rawdah, Jeddah 23435, Saudi Arabia
[9]Department of Computer Science and Software Engineering, Al Ain University, Abu Dhabi, United Arab Emirates
[10]Department of Data Science and Business Systems, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India

Corresponding authors: A. Sasikumar (sasikumaraphd@gmail.com), Amr Yousef (amr.yousef@alexu.edu.eg), and Madhusudana Rao Nalluri (n_madhusudanarao@av.amrita.edu)

ABSTRACT For the proper diagnosis and treatment of a variety of retinal conditions, retinal blood vessel segmentation is crucial. Delineation of vessels with varying thicknesses is critical for detecting disease symptoms. However, this task is challenging due to inadequate contextual information, complex vessel morphology, and lesion confusion. Many recent works employed several variations of CNNs with U-Net as baseline model for segmenting blood vessels from the fundus images. However, the existing methods still lack in generalizing the vessels well enough, indicating scope for improvement in this challenging problem of vessel segmentation. We introduce a novel Multi-Scale Residual (MSR) U-Net model in this study replacing convolution block and skip connections with an improved Multi-Scale Residual (MSR) convolution block and Bottleneck residual paths (B-Res paths) respectively. Specifically, STARE, DRIVE, and CHASE_DB1 datasets of fundus images are used to validate the proposed segmentation method. Our experimental results consistently showcase better/comparable performances when compared with current approaches, achieving higher area under receiver operator characteristic (AUC), accuracy, and F1 score. in segmenting blood vessels of varying thicknesses, even in scenarios with diverse contextual information, the presence of coexisting lesions, and complex vessel morphologies.

INDEX TERMS Blood vessels, fundus, retina, segmentation, U-Net.

I. INTRODUCTION

Retinal image assessment is extensively employed within the medical community to analyze vascular and non-vascular pathologies. One of the fundamental aspects of this assessment involves extracting salient morphological features from the retinal vasculature, including attributes such as tortuosity, width, length, curvature, branching patterns, and branch angles. These features serve as the foundational elements for the screening, diagnosing, and assessment of various chronic diseases, encompassing optical and cardiovascular conditions [1]. For instance, the presence of narrow arterioles with bright reflexes may indicate hypertension, and measurements of tortuosity and bifurcation angles can help identify diseases like cardiovascular ailments and diabetic retinopathy [2], [3]. Additionally, it is frequently believed that the vascular tree segmentation is the most suitable representation for registration applications, as fundus image tends to remain relatively

The associate editor coordinating the review of this manuscript and approving it for publication was Okyay Kaynak.

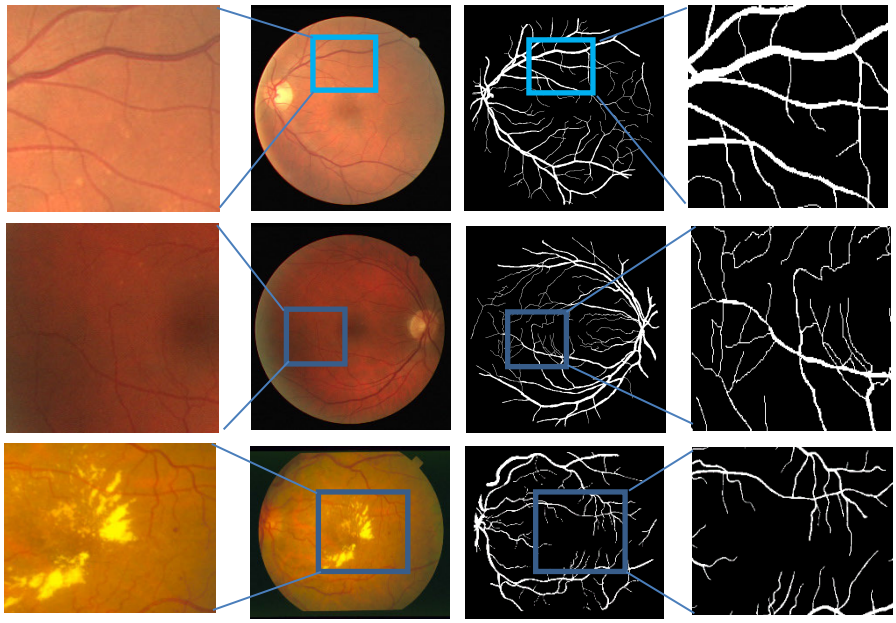


FIGURE 1. Example retinal images containing challenging regions for segmentation of blood vessels. (Row wise) 1. Blood vessels with variable thicknesses and corresponding ground truth, 2. Blood vessels with blurred boundaries and corresponding ground truth, 3. Blood vessels along with pathologies and corresponding ground truth. Challenges such as variations in thickness, indistinct boundaries, and the presence of pathologies contribute to the difficulty in accurately segmenting blood vessels.

stable, except in specific disease cases, and provides adequate details for localization of important anchor points. Additionally, vascular tree can assist as a valuable landmark for laser treatment of choroidal neovascularization. Furthermore, the unique characteristics of an individual's retinal vascular tree make it suitable for biometric identification, where it can be used as a distinctive identifier [4]. Hence, the development of reliable methods for the segmentation of vasculature is of paramount importance.

Technically, retinal vessel segmentation entails precisely extracting pixel-wise vessel locations, regardless of variations in vessel diameters, angles, and branching patterns. Nevertheless, as depicted in Figure 1, retinal vessel segmentation poses significant challenges for several reasons:

1. Varying Vessel Thickness: Blood vessel thickness in retinal images can vary widely from 3 to 20 pixels. This variation depends on factors such as vessel width and image resolution.
2. Inadequate Contextual Information: Inaccuracies can arise due to insufficient contextual visual information stemming from the inherent characteristics of diseases and uneven illuminations. This inadequacy makes learning diverse features critical for accurately identifying thin and edge vessels more challenging.
3. Complex Vessel Morphology: The occurrence of vessel crossings, central light reflexes, bifurcations, and other morphological features further complicates the segmentation of thin vessels.
4. Lesion Confusion: Fundus images often contain several lesion regions, some of which bear a striking resemblance to blood vessels. Consequently, these vessel-like lesions are susceptible to being mistakenly segmented as actual blood vessels.

These challenges highlight the necessity for reliable and accurate retinal vascular segmentation methods in medical image analysis.

Depending on whether predefined annotations are utilized, vessel segmentation methods can be classified into unsupervised and supervised approaches. Unsupervised methods do not rely on manually annotated information and can be categorized as Edge Detection based [5], [6], Match Filter based [7], [8], Morphology based [9], Vessel Tracking based [10], Clustering based [11], Multilevel Threshold based [12] and active contour based [13], [14] methods. Unsupervised methods, owing to the absence of supervision and expert annotations, exhibit significant variations in their performance when applied to retinal images with diverse backgrounds. Notably, in cases involving irregular illumination and diabetic retinopathy, unsupervised approaches are susceptible to identifying wrong edges, causing in a high rate of false detections.

Supervised methods, unlike their unsupervised counterparts, build vascular segmentation models using annotated data. The two steps of feature extraction and classification were the norm for supervised approaches prior to the usage of deep learning methods in vascular segmentation. In the

first step, features were extracted using hand-crafted rules. The second phase then involved using supervised classifiers to divide these features extracted into vessels and non-vessels.

Deep learning techniques can generate highly discriminative features by combining feature extraction and classification processes. Deep learning methods have transformed the area of fundus vessel segmentation and analysis of ophthalmological ailments using fundus images, achieving high accuracy [15], [16], [17], [18]. Notable deep learning models commonly employed for vessel segmentation include SegNet [15], U-Net [16], and their modified variants. U-Net, in particular, has gained immense popularity in the biomedical image segmentation domain. Recent advancements in this area have introduced novel architectures to enhance performance. For instance, a Recurrent convolution neural network (RUNet) [17] is proposed based upon the effectiveness of the U-Net. DUNet [18] presents deformable convolution within the U-Net architecture to effectively identify retinal vasculature with varying scales. Wang et al. proposed a modified UNet with double encoders. One encoder captures spatial data through a large kernel, and the second encoder conserves semantic information with multiple convolutions [19]. Spatial context data capture was enhanced in Dense UNet [20], while S-UNet [21] incorporates a saliency mechanism within a minimal UNet architecture named Blood Vessel Segmentation CNN, inspired by U-Net [16]. A robust Orientation and Context Entangled Network having the ability to get intricate orientation and context data of the retinal vasculature is used in OCE-Net [22]. In contrast, a combination of a multitask segmentation network and a fusion network was used to segment retinal vasculature [23]. SPNet [24] introduced a shared decoder and pyramid-like loss for superior segmentation of vessels. A directed graph search-based multi-attentive neural network approach is used for automatic vascular network segmentation [25]. Lightweight convolution blocks were proposed to reduce computation complexity and get better segmentation results [26].

In the pursuit of refined blood vessel segmentation, a significant challenge lies in detecting vessels of varying thicknesses. Recent medical image segmentation methods have shifted towards employing multi-scale networks to grasp a more extensive range of information [27], [28], [29]. In one approach, Hu et al. [27] presented a multi-scale architecture aimed at acquiring more comprehensive multi-scale vascular features. They also devised an enhanced cross-entropy loss function, which prioritizes challenging examples. In a similar vein, Fu et al. [28] proposed a multi-label deep network that integrates multi-scale inputs and side-output layers, concentrating on the combined boundary estimation of the optic cup and optic disc. However, fully convolutional networks (FCNs) often introduce a drawback during the encoding phase, as convolutions and pooling operations may need to pay more attention to structural information. To overcome this limitation, a filter which is attention guided is integrated into the network by Zhang et al. [29]. To get multiscale features, a fusion block based on the attention mechanism is employed in ResDO-Unet [30]. A dice-loss based on multi-scales was used to combine features from various layers [31]. Filters with multi-scales are used to cope with the diversity of vessel thickness [32]. A segment-level loss was introduced to enhance the segmentation results of thin vessels [33]. Although multi-scale models are segmenting more blood vessels effectively, they often face a classic trade-off: There may be a slight drop in the detection of thick blood vessels while emphasizing thin vessels.

The motivation for carrying out this research is to build a segmentation model to generalize more efficiently to blood vessels in fundus images, overcoming the challenges involved. Despite the efforts made to address the aforementioned challenges, there is still significant room for improvement in the effectiveness of the models. Most existing retinal blood vessel segmentation approaches at multiple scales employ multi-scale blocks, which are trained individually and combined into an ensemble model. However, this strategy lacks mechanisms for sharing information between sub-models and fails to establish interaction mechanisms among sub-models for detecting vessels with varying thicknesses. This drawback can result in sub-models that do not effectively collaborate and may exhibit over-specialization towards specific blood vessel thicknesses, rendering them less robust against noise and image variations.

These observed facts motivated us to integrate novel advancements in network architecture, adopt diverse preprocessing techniques, and customize training strategies to better suit the employed datasets in achieving better generalization in delineating vessels from color retinal images. Following are the contributions of this paper towards achieving that goal.

1. The proposed MSR block of the modified U-net architecture has multi-scale kernels, which allows the analysis of identical local regions of the fundus image through various receptive fields. This feature enhances the model's sensitivity to the different blood vessel sizes.
2. The network is trained on a combination of diverse pre-processed images such as grayscale, CLAHE, and the shade-corrected versions of the input images. This adaption is meant to improve the model's generalization power against inadequate contextual information.
3. Deeper neural networks frequently encounter issues in the propagation of features towards the penultimate layers, resulting in weaker feature propagation. Additionally, these networks often suffer from the vanishing gradient problem towards the initial layers. These issues make it challenging to preserve complex morphological characteristics of the vessels across depths, particularly in the context of the U-net architecture, where propagating features from the encoder to the decoder at various depths can prove to be complicated. Using the B-Res path, we successfully alleviated such issues and yielded better results.

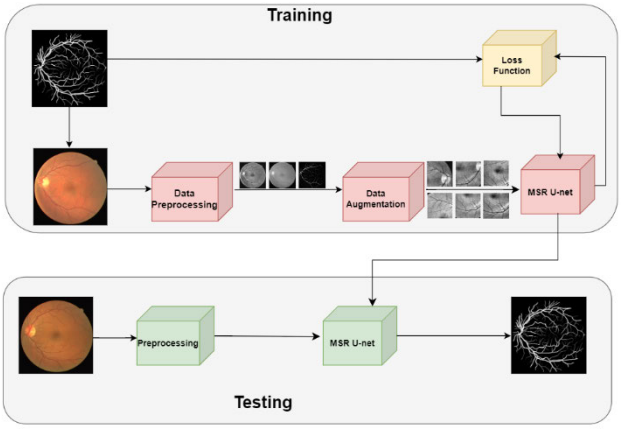


FIGURE 2. The proposed system architecture.

4. In fundus images, particular lesions bear a striking resemblance to blood vessels, which can appear fraudulent before an authentication system. As a result, a customized input pipeline strategy is necessary during training, requiring the model to train on a greater amount of data that is sufficient to learn features capable of discriminating the actual and fraudulent vessels. We generated such data using a random cropping mechanism.

II. THE METHOD

In this paper, we introduce the MSR U-Net, a model designed to tackle the challenges associated with vessel segmentation. This model adopts the U-Net architecture [16] as its foundation, with significant structural modifications aimed at enhancing algorithm performance. The acronym 'MSR' encapsulates the essence of our model, highlighting its key features: multi-scale convolutional blocks with residual learning, expanded utilization of the inception module, and the incorporation of the bottleneck residual blocks for copy and crop path inspired by U-Net. Our enhancements primarily focus on two aspects: the MSR block and the B-Res path. To assess the efficacy of the MSR U-Net, we conducted training and experimental testing. The process began with data preprocessing and augmentation to enhance image diversity. Subsequently, model training is executed to obtain optimal parameter weights. Finally, the model underwent testing and analysis using a set of testing images. The entire workflow is visually depicted in Figure 2.

A. PRE-PROCESSING

Although deep learning architectures frequently perform well when working with unprocessed input data, segmentation performance can be improved with the proper preprocessing methods. To investigate the impact of preprocessing, we conducted tests with and without preprocessing to assess its effects. Our proposed framework sequentially incorporates three image preprocessing strategies: 1. Grayscale Image, 2. Contrast Limited Adaptive Histogram Equalization

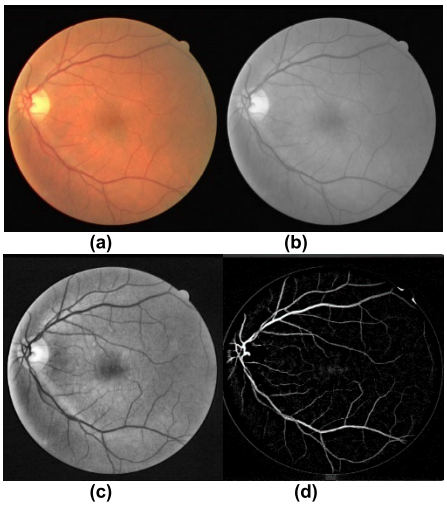


FIGURE 3. (a)The original fundus image. Results of each preprocessing approach: (b) grayscale; (c) CLAHE; (d) Shade corrected.

(CLAHE) Image, and 3. Shade Corrected (SC) Image. These preprocessing techniques are applied to the input images, and their influence on segmentation performance is evaluated. The first preprocessing strategy involves the conversion of RGB images into single-channel grayscale images. This conversion decomposes the RGB color image into a grayscale image, which enhances discrimination between retinal vasculature and background compared to the original RGB image. The second preprocessing strategy involves enhancing the contrast between the foreground and background of the entire dataset using CLAHE [34]. The resulting image after applying CLAHE is shown in Figure 3. (c). The third preprocessing step involves estimating the background image I_{bg} and subtracting it from I_{green}, the green channel image of the corresponding RGB image. To generate I_{bg}, we apply a median filter to I_{green} using a 25×25 pixel kernel. The filter size selection ensures that it is broader than the widest blood vessel in our dataset of images. The resulting image may have negative values for pixels with intensities lower than the background. All pixels with positive values are set to zero to get the shade-corrected image.

B. MSR U-NET

U-Net is a prominent semantic segmentation architecture, subsequently becoming a baseline model for many downstream tasks requiring tailor-made modifications to the baseline model. For the task of vessel segmentation from fundus images, we propose MSR U-Net, which incorporates MSR blocks in the encoder and decoder, and the flow of features between them at all depths is accomplished by B-Res paths, as shown in Figure 4. Such modifications are necessitated by the challenges posed by the diversified visual patterns associated with blood vessels present in fundus images.

1) MSR BLOCK

Utilizing multi-scale features is beneficial for improving vessel segmentation. With a primary goal of feature extraction

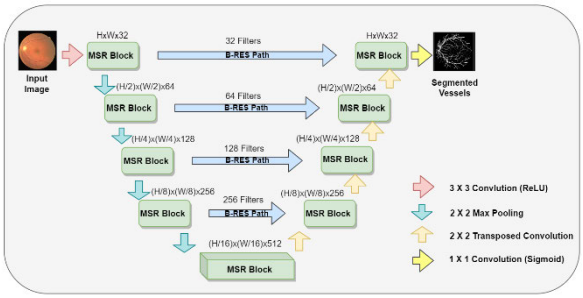


FIGURE 4. Proposed MSR U-Net architecture. The two convolution layer sequence in U-Net is replaced with MSR blocks. Moreover, we employed proposed B-Res paths rather than just using ordinary skip connections.

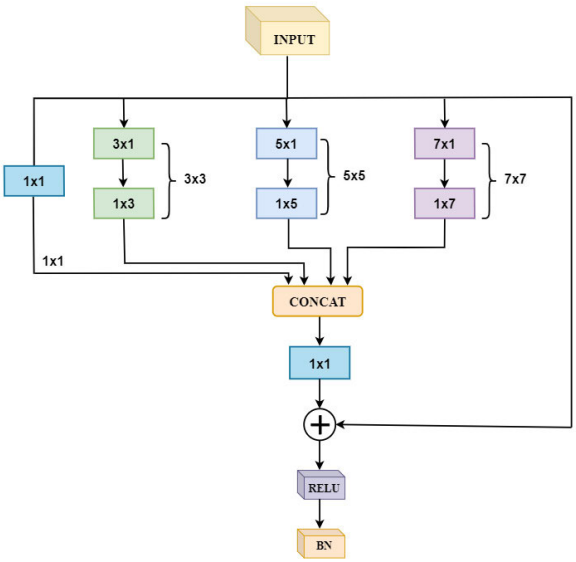


FIGURE 5. MSR convolutional block. To extract the features at different scales, Convolutions with different kernels are performed in parallel. A residual connection is used as it makes the learning easier and keep some unmodified information.

at different scales, the proposed Multi-Scale Residual (MSR) block is illustrated in Figure 5. The MSR block, drawing inspiration from the inception module [35], enhances the convolution module of the U-Net. Serving as the fundamental building block of our network, this MSR block orchestrates the divergence of data into parallel branches, each hosting convolutions with distinct kernel sizes. The input is simultaneously processed by a set of parallel convolutions with kernel sizes of 1×1, 3×3, 5×5, and 7×7. Subsequently, the outputs from these diverse branches are concatenated and merged using a 1×1 convolution layer. Furthermore, drawing inspiration from the results, we incorporate a residual connection alongside the parallel branches. This addition is inspired by the remarkable performance it has demonstrated in medical image segmentation [36]. The residual structure effectively mitigates the degradation problem and enhances the model's capacity to grasp complex spatial information. Lastly, the ReLU (Rectified Linear Unit) activation function is used to activate the output of the block and is subjected to batch normalization.

We employ convolutional factorization to address the challenge of the significant improvement in the number of parameters and memory requirements arising from the parallel branches [37]. This technique decomposes a convolution of size $N \times N$ into two consecutive convolutions: $N \times 1$ and $1 \times N$. This factorization sequence approximates the original convolution while significantly reducing the number of parameters. By Integrating Multi-Scale Residual Convolution blocks in place of conventional convolutional layers within the U-Net architecture enables U-Net to incorporate features learned from images at various scales effectively. This enhancement contributes to improved feature reconciliation within the model.

2) B-RES PATH

U-Net is a typical architecture which involves the gradual compression of features along the encoder path and the gradual reconstruction of the targeted information along the decoder path. However, to avoid the semantic gap, U-Net has skip connections at the corresponding depths of the encoder and decoder, which are meant to fuse the features representing inherently different types of information. But these skip connections, which to some extent can reduce the semantic gap, need to be replaced by learning paths to avoid possible discrepancies that would be introduced in the reconstruction of the masks that correspond to objects as complex as blood vessels [38].

To accomplish this task, we propose an approach that uses bottleneck residual blocks instead of skip connections, which we refer to as 'B-Res Path'. The structure of the B-Res Path is illustrated in Figure 6.

The bottleneck residual block is a combination of the basic Residual block and dimensionality reduction, as described in the Inception module [39]. This block starts by reducing the input tensor's dimension using a 1×1 convolution. Then, a 3×3 convolution is applied to extract deeper features. Finally, the original dimension is restored through another 1×1 convolution. The underlying idea is to keep these blocks as slim as possible to increase depth while minimizing the number of parameters.

Our choice to gradually lower the number of bottleneck residual blocks along the B-Res paths is influenced by the potential reduction in the semantic gap between encoder and decoder features as we travel the inner shortcut paths. We employed 4, 3, 2, and 1 bottleneck residual blocks along the four B-Res paths. Furthermore, we used 32, 64, 128, and 256 filters in the four B-Res paths, respectively, to maintain alignment with the number of feature maps in the encoder-decoder of the proposed architecture.

3) LOSS FUNCTION

Among the number of loss functions used in machine learning pertaining to various tasks such as classification and

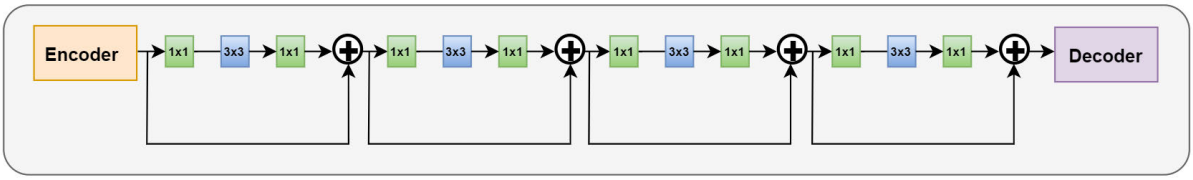


FIGURE 6. Proposed B-Res Path replaced skip connections that merge feature maps of the encoder straight forwardly with that of the decoder, we opt for a more intricate approach. The features from encoder undergo a sequence of Bottleneck Residual layers, introduced to narrow the semantic gap between the features of decoder with the corresponding features of encoder. Moreover, the incorporation of residual connections facilitates easier learning and proves highly beneficial in deep convolutional networks.

regression, we chose to use binary cross-entropy, anticipating the task at hand is a binary classification, which is defined as:

$$L_{CE}\left(\hat{y}, y\right) = -\sum_{k}\left[y_k \log \hat{y}_k + (1 - y_k) \log\left(1 - \hat{y}_k\right)\right] \quad (1)$$

where the equation (1) penalizes the miss classification more, thus contributing more to the total value of the loss with $\hat{y}_k$ being prediction in the scale of [0,1], y_k being the binary representation of the true label, and k indicating the index of the training sample.

III. EXPERIMENTAL DESIGN

The experimental design of the proposed work is discussed in this section. At first, we present insights into the diagnostic predicament and the data involved. Next, we explain the process of the training data preparation that includes preprocessing and random cropping. The third subsection comprehensively expounds on the implementation of MSR U-Net. In the last subsection, the evaluation metrics that have been employed to measure the model performance are discussed.

A. DATASETS

We utilized three benchmark databases: DRIVE [40], STARE [41], and CHASE_DB1 [42]. The three datasets under consideration encompass a diverse range of content, incorporating challenging images. This provides a chance to benchmark the performance of the proposed MSR U-Net method against existing approaches.

- The DRIVE [40] database comprises forty colored fundus images depicting varying levels of different disease progression in each of them across which some of the blood vessels spread. Accurate segmentation of every such vessel is very important, though it is difficult to do so. These images were acquired by a 3CCD camera, with the anatomical structure primarily situated in the central area having a radius of about 540 pixels. It is divided into two sets: a train set and a test set, each containing twenty images sourced from distinct patients. Manual segmentation of blood vessel ground truth serves as the gold standard for each image.
- The STARE [41] database comprises twenty colored fundus images accompanied by with corresponding retinal masks and manual segmentation of blood vessel images. The STARE database was captured using the TRV-50 Topcon fundus camera, resulting in images having a size of 605 × 700 pixels. Within the STARE [41] database, 50% of the images feature various pathologies, making the assessment of retinal blood vessel segmentation methods particularly challenging.
- The CHASE_DB1 [42] database comprises twenty eight colored fundus images collected from a diverse group of schoolchildren of different ethnic backgrounds. These images were acquired by a hand-held NM-200-D Nidek camera, utilizing a 30-degree field of view (FOV). The resolution of the images is 960 × 999 pixels. Within this database, manually segmented ground truth references are also provided.

B. DATA PREPARATION

Considering the three datasets mentioned in previous section, the proposed network model is supposed to be trained on relatively small datasets compared to the size of the datasets used in other downstream tasks as demanded by the training of deep neural network models. This is a constraint that may result in the overfitting of our model to the data. However, as all three datasets provide us with high-resolution images, we decided to train the model on the patches extracted from those images instead of on the entire image. This allows the model to acquire a sufficient amount of generalizability.

The basic idea is to enhance the training dataset by extracting the patches from each image within the training set. To maintain consistency with the original fundus image sizes, we set the dimensions of the extracted patches to be equal to 64 × 64 pixels. Each patch is generated by randomly selecting its center.

For the DRIVE [40] dataset, which comprises only 20 fundus images, a total of 200,000 image patches were randomly selected for training. The same patch extraction methodology is applied to both the STARE [41] and CHASE_DB1 [42] databases. Regarding the STARE [41] dataset, during our experiments, 250,000 patches are randomly chosen for the training. In the case of the CHASE_DB1 [42] dataset, 250,000 patches are randomly selected for the training. Upon extracting these patches, it became apparent that they comprehensively cover nearly all regions of the image. Many of the extracted patches encompass vessel-containing regions and border regions, providing a thorough representation of

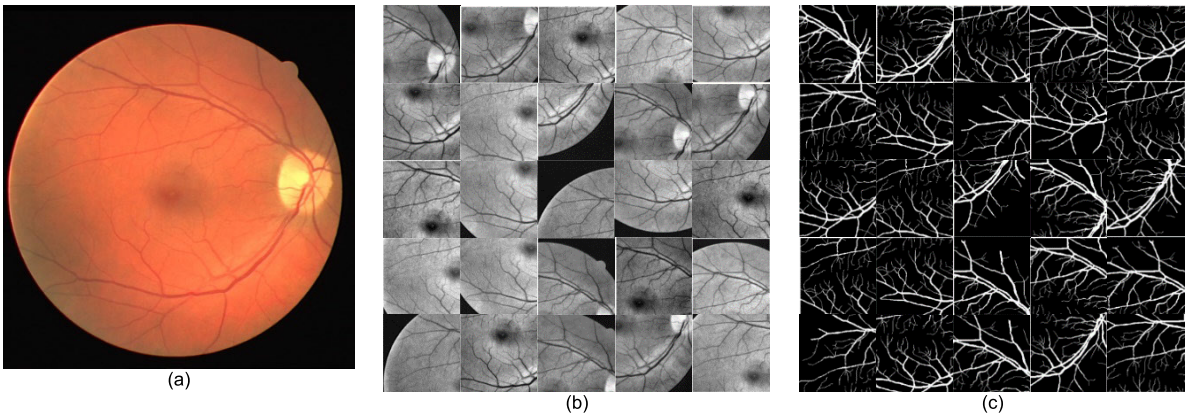


FIGURE 7. A color fundus image and samples of patches extracted from the corresponding preprocessed image. In the fundus image (a), a random center for the patches is chosen. (b) Patches for Training. (c) Corresponding ground truths.

the complete image. The model is trained on all these regions, offering exposure to the full spectrum of image features. Additionally, we extracted corresponding ground truth patches to accompany these patches for training. Visualization of sample patches can be seen in Figure. 7, indicating their varied origins from diverse regions within the complete images.

C. TRAINING

The proposed MSR U-Net underwent training and evaluation using NVidia GeForce RTX 3070TI GPU, Intel Core i9-12900HQ CPU@3.80GHz processor, 32 GB of RAM based on Keras framework with tensorflow as backend. For weight initialization, the Xavier initialization method was employed. The optimization of the network was carried out using the Adam optimizer. The specific optimization settings included a momentum value of 0.9 and a weight decay of 10^{-5}. The learning rate is set to 0.001 and is adjusted as needed during training. If the validation error did not show improvement, the learning rate is decreased. The training procedure iterated till the validation accuracy plateaued, at which point the best-performing model is used for evaluation on the test data. A mini-batch size of 256 is employed.

Train test split of the dataset is a common practice in machine learning which distinguishes it from optimization. As we are considering three different datasets in this work, this splitting is done with little more care in order to ensure better performance of the model. We allocate 20 images for training and remaining 20 images for testing for the DRIVE dataset in accordance with the guidelines given by the data provider. We divide the STARE database's 20 images into training and testing groups by allocating 10 of the 20 images to each. Among the 28 photos in the CHASE_DB1 database, 20 are employed for training and the other eight are held back for testing. TABLE 1 presents a summary of the statistics of the databases and their corresponding split sets.

TABLE 1. Statistics of the databases: The total number of images available in every database is displayed in Column#2. Column#3 and #4 indicates the train-test split. Column#5 indicates the images with bad quality due to several reasons as discussed in introductory section while #6 indicates images that are considered to be having normal or appropriate visible quality without pathologies. Column#5 and #6 together are equally distributed into the column#3 and #4.

Databases	Number of images	Training images	Testing images	Challenging images	Normal images
DRIVE	40	20	20	20	20
STARE	20	10	10	10	10
CHASE_DB1	28	20	08	0	28

D. MEASURING METRICS

Five criteria are used to gauge the proposed method's effectiveness, which are:

- Sensitivity (Se): Sensitivity measures the method's capability to correctly classify vessel pixels as vessels, representing its ability to detect true positive vessel pixels.
- Specificity (Sp): Specificity gauges the algorithm's capacity to correctly recognize non-vessel pixels as non-vessels, reflecting its effectiveness in detecting true negative non-vessel pixels.
- Accuracy (Acc): Accuracy computes the ratio of correctly classified vessel pixels to the total number of identified vessel pixels. It quantifies the overall correctness of the vessel pixel classification.
- Area under the Curve (AUC): A significant quantitative statistic generated from Receiver Operating Characteristic (ROC) curves is AUC. When creating binary segmentations using probability maps, ROC curves are produced by comparing the true positive rate against the false positive rate. An ideal classifier consistently produces an AUC of 1. The ROC curve and AUC are especially valuable in situations with imbalanced class distributions, offering a robust assessment of classification accuracy.
- F1 Score: It is defined as the harmonic mean of precision and recall.

It's worth noting that for the computation of these metrics, for the three datasets, only the pixels inside the field of view (FOV) are taken into account. This focus on the FOV ensures

TABLE 2. Test results with different image preprocessing.

Database	Number (No.)	Gray Scale	CLAHE	Shade Corrected	AUC	Acc	Se	Sp	F1
DRIVE	0	Color Image			0.9771	0.9630	0.7237	0.9866	0.7772
	1	✓			0.9804	0.9657	0.7450	0.9927	0.8048
	2		✓		0.9835	0.9685	0.7756	0.9889	0.8106
	3			✓	0.9839	0.9681	0.7266	0.9910	0.8083
	4	✓	✓		0.9872	0.9709	0.8199	0.9852	0.8304
	5		✓	✓	0.9853	0.9703	0.8022	0.9862	0.8244
	6	✓		✓	0.9871	0.9692	0.8473	0.9807	0.8269
	7	✓	✓	✓	0.9883	0.9722	0.8717	0.9886	0.8344
STARE	8	✓	✓	✓	0.9952	0.9813	0.8585	0.9914	0.8748
CHASE_DB1	9	✓	✓	✓	0.9835	0.9781	0.8551	0.9891	0.8683

TABLE 3. Average performance metrics with standard deviation for DRIVE [40], STARE [41], and CHASE_DB1 [42].

Dataset	Expert/Method	Acc	AUC	F1	Se	Sp
DRIVE[40]	Second human expert	0.9473 ± 0.0049	--	--	0.7760 ± 0.0594	0.9725 ± 0.0083
	Proposed method	0.9722	0.9883	0.8344	0.8717	0.9886
STARE[41]	Second human expert	0.9353 ± 0.0168	--	--	0.8851 ± 0.1085	0.9388 ± 0.0258
	Proposed method	0.9813	0.9952	0.8748	0.8585	0.9914
CHASE_ DB1[42]	Second human expert	0.9560 ± 0.0076	--	--	0.7686 ± 0.0765	0.9779 ± 0.0085
	Proposed method	0.9781	0.9835	0.8683	0.8551	0.9891

that the evaluation is conducted on the relevant regions of interest within the retinal images.

$$Se = \frac{TP}{TP + FN} \tag{2}$$

$$Sp = \frac{TN}{TN + FP} \tag{3}$$

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$\Pr ecision = \frac{TP}{TP + FP} \tag{5}$$

$$\text{Re}call = \frac{TP}{TP + FN} \tag{6}$$

$$F1 = 2 \times \frac{\Pr ecision.\text{Re}call}{\Pr ecision + \text{Re}call} \tag{7}$$

The terms TP (true positive), TN (true negative), FP (false positive), and FN (false negative) are defined in the context of vessel pixel detection as follows: A vessel pixel is considered TP when correctly identified as a vessel, TN when correctly identified as a non-vessel, FP when identified as a vessel but is actually a non-vessel pixel, and FN when identified as a non-vessel but is actually a vessel pixel. The cumulative counts of these four detection conditions are employed as criteria for evaluating the system's performance on an image.

E. EXPERIMENTS

1) EXPERIMENT 1: INFLUENCE OF DIFFERENT COMBINATIONS OF PREPROCESSING

The influence of various preprocessing steps is observed, and TABLE 2 presents a summary of the outcomes for each phase. When comparing No. 0 with No. 1, it becomes evident that the experimental results significantly improved after converting to grayscale images compared to RGB color images, with a notable increase in the F1 score by 1.035%. Furthermore, when examining the effects of CLAHE and shade correction, it is observed that these methods have yielded positive outcomes. However, it's important to note that performance suffers when grayscale and shade correction are applied without CLAHE, as seen in No. 6. Remarkably, No. 5 and No. 7 serve as examples of how CLAHE and shade correction work better together. Especially the F1 score improved by 2.73% over No. 0, reaching 0.8344 when all three preprocessing techniques were used simultaneously (No. 7). In subsequent experiments, these three preprocessing methods were used together.

2) EXPERIMENT 2: OUTCOMES OF THE PROPOSED MSR U-NET METHOD

Performance metrics are derived using the first human expert's annotations as the ground truth after vessel segmentation images have been obtained in each database. TABLE 3 provides a summary of average performance metrics from all test images in the DRIVE [40], STARE [41], and CHASE_DB1 [42] databases. All average performance metrics of the proposed algorithm consistently surpass the second human expert, as given in the TABLE. This shows that the method when compared to the second human expert has a low FP rate and a robust capacity to discern between vascular and non-vessel pixels. Furthermore, as shown in Figure 8, the proposed method produces AUC values that

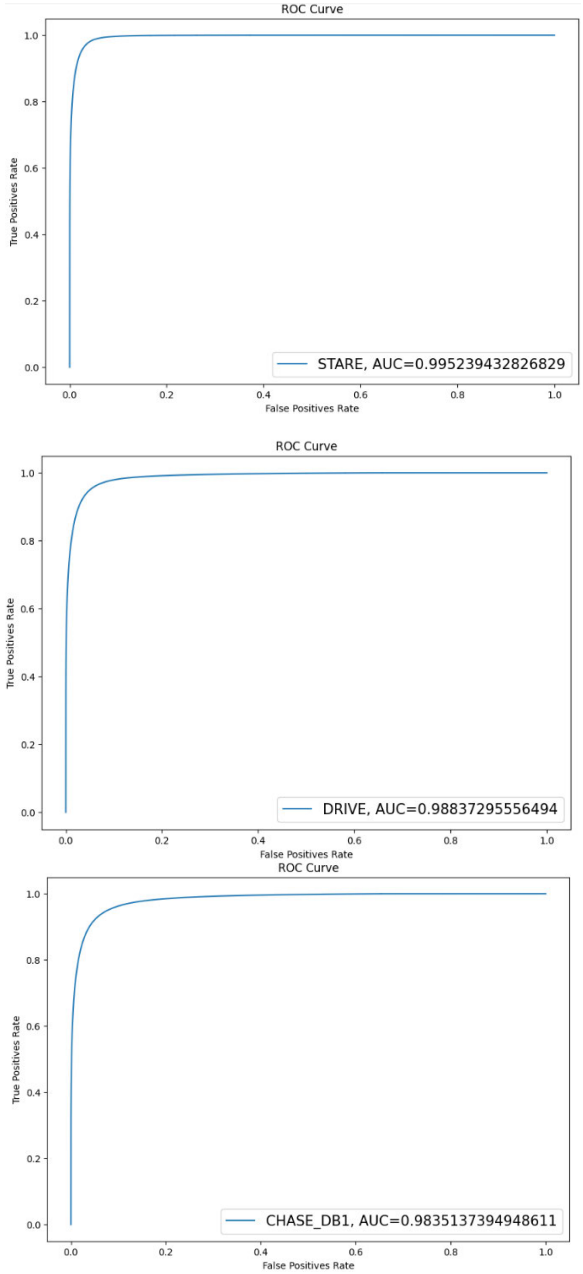


FIGURE 8. ROC curves for the datasets STARE [41], DRIVE [40], and CHASE_DB1 [42]. AUC for the STARE [41] database is 0.9952, 0.9883 for the DRIVE [40] database and 0.98351 for CHASE_DB1 [42] database.

are greater than 0.97 (particularly 0.9883 for the DRIVE, 0.9952 for the STARE, and 0.9835 for the CHASE_DB1 databases). These high AUC values illustrate the proposed method's robust generalization capability, indicating its proficiency in vessel segmentation across different datasets. The segmentation results of the proposed MSR U-Net method are presented in Figure 9.

3) EXPERIMENT 3: COMPARATIVE ANALYSIS WITH EXISTING METHODS

Acc, AUC, Se, Sp, and F1 score are used to evaluate the performance of the proposed MSR U-Net method using DRIVE [40], STARE [41], and CHASE_DB1 [42] databases. The average performances are reported in TABLES 4, 5, and 6, along with a comparison to existing methods. The performance metrics of the comparison methodologies are taken verbatim from the publications where they were published. Bold fonts are used to highlight the best findings, while dashes are used to denote results that were not accessible (i.e., a method did not use a specific dataset in its tests). The proposed MSR U-Net achieved AUC rates of 0.9883, 0.9952, and 0.9835 for the DRIVE [40], STARE [41], and CHASE_DB1 [42] databases, respectively. These results demonstrate that the proposed method works better on the DRIVE [40] and STARE [41] databases than any other method. The AUC of the proposed method for the CHASE_DB1 [42] database is slightly below the AUCs of MRC NET [47], WA-Net [57], and Bridge-Net [58]. However, it is noteworthy that the F1 score, Sensitivity, and Specificity obtained by our method on the CHASE_DB1 database surpass those of MRC NET [47].

Looking at TABLES 4, 5, and 6, it is evident that MSR U-Net has achieved the highest F1 score when compared to other methods. This observation suggests that our approach is more adept at accurately segmenting both the vasculature and the background. Our approach shows a superior balance between recall and precision, and so is the F1 score, a comprehensive statistic that balances both.

MSR U-Net has attained the highest sensitivity, indicating its ability to accurately label more vessel pixels. In most cases, vascular pixels make up only around 10% of all pixels in the image. This imbalance in categories makes it challenging to segment vessels using a classifier. Hence, our method's high sensitivity is crucial, especially in the framework of a computer-aided diagnosis system, where the detection of blood vessels needs to be accurate without introducing false positives. Furthermore, the results of the proposed MSR U-Net method demonstrate a higher level of competitiveness in terms of specificity and accuracy. This means that MSR U-Net effectively differentiates the background and blood vessels. Significantly, it achieves this without an increase in the number of FPs and FNs. Hence, the proposed method has a lower false positive ratio.

4) EXPERIMENT 4: CROSS-TRAINING RESULTS

In clinical practice, retraining the model for every new patient's fundus image analysis is impractical. Additionally, medical equipment from different hospitals is often manufactured by different companies. Therefore, a reliable method must be capable of effectively analyzing images acquired from various equipment sources. Evaluating the method's suitability for application to any retinal image in real-world scenarios becomes crucial. Thus, the model's practical applicability hinges on the crucial evaluation criteria of robustness and generalization.

When training on the STARE [41] dataset, the performance metrics for the DRIVE [40] and CHASE_DB1 [42] datasets are reported in TABLE 7. Similarly, when using the DRIVE [40] dataset for training, the performance results for the

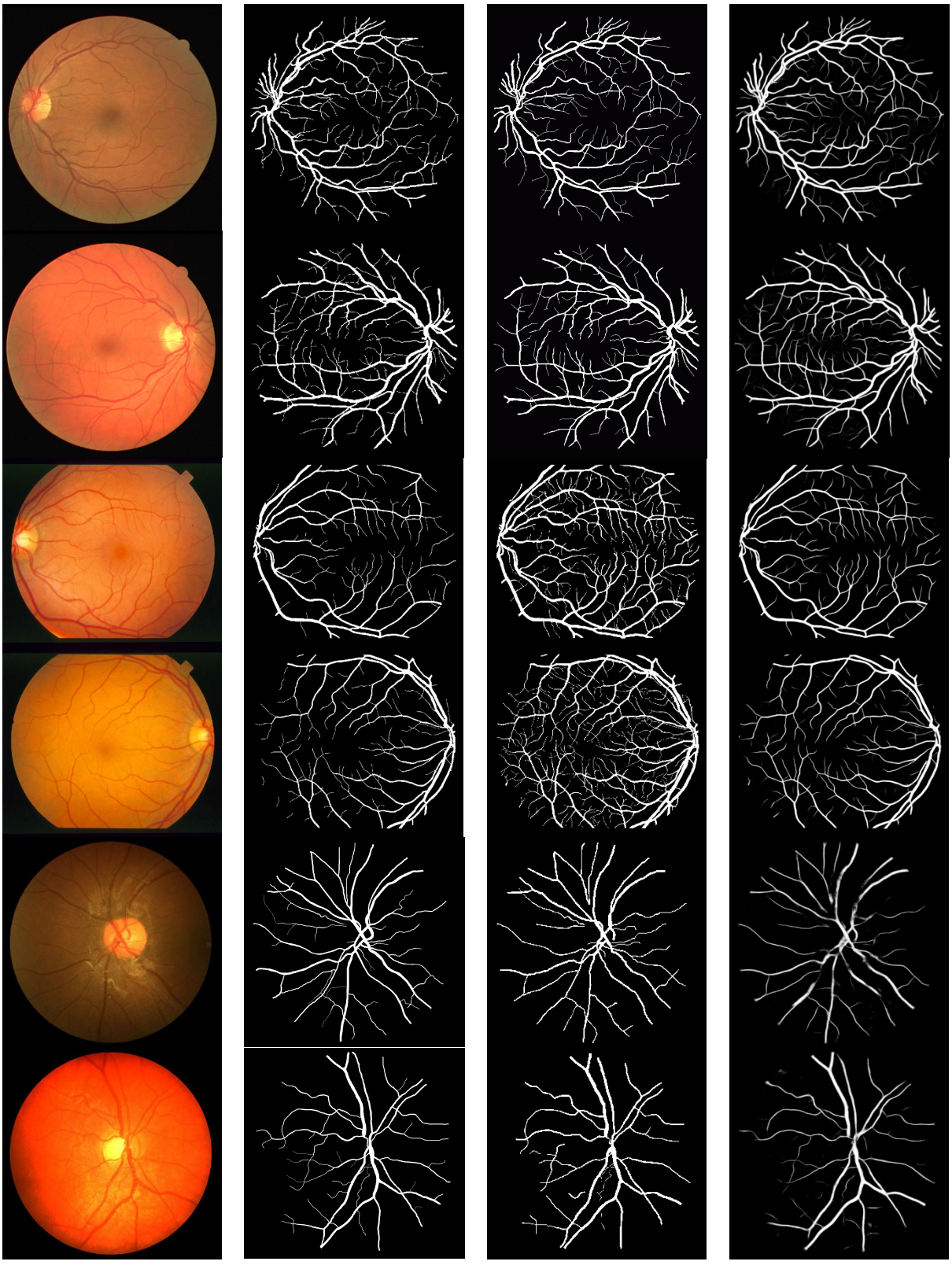


FIGURE 9. Randomly selected images along with their manual ground truth images corresponding segmented images. (Column wise) 1. Original color retinal fundus images, 2. 1st manual ground truth images. 3. 2nd manual ground truth images. 4. Segmentation results of the proposed MSR U-Net method.

TABLE 4. Quantitative performance analysis of the proposed MSR U-Net method on the DRIVE [40] dataset.

S.No	Methods	Year	AUC	Acc	Se	Sp	F1
1	2nd human expert		0.8781	0.9637	0.7743	0.9819	0.7889
2	MSR-Net [59]	2022	0.9650	0.9742	0.8936	0.9876	
3	Zhang [47]	2017	0.9703	0.9466	0.7861	0.9712	0.7953
4	Dasgupta [42]	2017	0.9744	0.9533	0.7691	0.9801	0.8074
5	Ronneberger [16]	2015	0.9755	0.9531	0.7537	0.9820	0.8142
6	Alom [17]	2018	0.9784	0.9556	0.7792	0.9813	0.8171
7	Zhuang [48]	2018	0.9793	0.9561	0.7856	0.9810	0.8202
8	Jin [18]	2019	0.9802	0.9566	0.7963	0.9800	0.8237
9	Khan [51]	2022	0.9812	0.9698	-	-	0.8270
10	Oliveira [49]	2018	0.9821	0.9576	0.8039	0.9804	
11	Zhang [58]	2022	0.9834	0.9565	0.7853	0.9818	0.8203
12	Dora [57]	2022	0.9849	0.9708	0.7962	0.9875	0.8269
13	Du [52]	2023	0.9460	0.9597	0.6745	0.9872	-
14	Maninis [40]	2016	0.9861	0.9541	0.8261	0.9115	0.8210
15	Jiang [50]	2019	0.9874	0.9703	0.8197	0.9848	0.8286
16	Liu [53]	2023	0.9877	0.9704	0.8150	0.9856	0.8277
17	Jiang [50]	2019	0.9880	0.9706	0.8325	0.9838	0.8321
18	Proposed MSR U-Net Method	2023	**0.9883**	**0.9722**	**0.8717**	**0.9886**	**0.8344**

TABLE 5. Quantitative performance analysis of the proposed MSR U-Net method on the STARE [41] dataset.

S.No	Methods	Year	AUC	Acc	Se	Sp	F1
1	2nd human expert		0.9291	0.9522	0.9017	0.9564	0.7417
2	MSR Net [59]	2022	0.9532	0.9486	0.9640	0.8441	-
3	Khan [51]	2022	0.9706	0.9747	0.8190	0.9874	0.8286
4	Du [52]	2023	0.9726	0.9686	0.7935	0.9830	-
5	Zhang [47]	2017	0.9740	0.9547	0.7882	0.9729	0.7815
6	Lu [45]	-	0.9801	0.9628	0.8090	0.9770	-
7	Jin [18]	2019	0.9832	0.9641	0.7595	0.9878	0.8143
8	Li [48]	-	0.9834	0.9673	0.8465	-	0.8435
9	Maninis [40]	2016	0.9861	0.9541	0.8261	0.9115	0.8210
10	Liu [53]	2023	0.9883	0.9765	0.8269	0.9885	0.8396
11	Ronneberger [16]	2015	0.9898	0.9690	0.8270	0.9842	0.8373
12	Zhang [58]	2022	0.9901	0.9668	0.8002	0.9864	0.8289
13	Oliveira [49]	2018	0.9905	0.9694	0.8315	0.9858	-
14	Dora [57]	2022	0.9906	0.9752	0.7834	0.9908	0.8216
15	Jiang [50]	2019	0.9917	0.9773	0.8369	0.9888	0.8485
16	Jiang [50]	2019	0.9923	0.9777	0.8522	0.9880	0.8531
17	Proposed MSR U-Net Method	2023	**0.9952**	**0.9813**	**0.8585**	**0.9914**	**0.8748**

STARE [41] and CHASE_DB1 [42] datasets are presented in TABLE 8.

Finally, in TABLE 9, evaluation results on DRIVE [40] and STARE [41] datasets are presented when CHASE_DB1 [42] dataset is used for training. Tables 4 and 7, respectively, illustrate that the generalization power of the proposed MSR U-Net model, trained on the STARE [41] dataset, declines slightly when tested on DRIVE [40] dataset compared to when it is tested on the same dataset. The same is true with other training and testing combinations of the three datasets as well, which is depicted in Tables 5 and 8 and Tables 6 and 9.

From TABLES 7, 8, and 9, it is evident that the proposed method outperforms other methods in all cross-training results. In the instance of training with DRIVE [40] dataset and testing on STARE [41] dataset, the proposed MSR U-Net has the highest AUC and accuracy, but its specificity is a little lower compared to the approach by Mo and Zhang [55]. The experimental results obtained through cross-training highlight the superior generalization and robustness of the

TABLE 6. Quantitative performance analysis of the proposed MSR U-Net method on the CHASE_DB1 [42] dataset.

S.No	Methods	Year	AUC	Acc	Se	Sp	F1
1	2nd human expert	-	0.9071	0.9733	0.8313	0.9829	0.7969
2	Du [52]	2023	0.9385	0.9708	0.8550	0.9785	-
3	Zhang [47]	2017	0.9706	0.9502	0.7644	0.9716	0.7581
4	Ronneberger [16]	2015	0.9772	0.9578	0.8288	0.9701	0.7783
5	Jin [18]	2019	0.9804	0.9610	0.8155	0.9752	0.7883
6	Liu [53]	2023	0.9812	0.9770	0.8318	0.9867	0.8188
7	Alom [17]	2018	0.9815	0.9634	0.7756	0.9820	0.7928
8	Zhuang [48]	2018	0.9830	0.9656	0.7978	0.9818	0.8031
9	Khan [51]	2022	0.9857	0.9779	0.8485	0.9887	0.8548
10	Dora [57]	2022	0.9890	0.9758	0.8041	0.9873	0.8069
11	Zhang [58]	2022	**0.9893**	0.9667	0.8132	0.9840	0.8293
12	Proposed MSR U-Net method	2023	0.9835	**0.9781**	**0.8551**	**0.9891**	**0.8683**

TABLE 7. Comparison of experimental results: STARE dataset is used for training, then testing on DRIVE and CHASE datasets.

S.No	Train Set	Test Set	Methods	AUC	Acc	Se	Sp
1	STARE[41]	DRIVE[40]	Fraz [42]	0.9697	0.9456	0.7242	0.9792
2			Li [54]	0.9677	0.9486	0.7273	0.9810
3			Mo [55]	0.9653	0.9492	0.7412	0.9799
4			Proposed MSR U-Net Method	0.9827	0.9684	0.7904	0.9822
1		CHASE_DB1 [42]	Li [54]	0.9565	0.9415	0.7103	0.9665
2			Li [54]	0.9553	0.9417	0.7240	0.9768
3			Mo [55]	0.9690	0.9515	0.7032	0.9794
4			Proposed MSR U-Net Method	0.9758	0.9657	0.7668	0.9805

TABLE 8. Comparison of experimental results: DRIVE dataset is training, then testing on the STARE and CHASE datasets.

S.No	Train Set	Test Set	Methods	AUC	Acc	Se	Sp
1	DRIVE[40]	STARE[41]	Fraz [42]	0.9660	0.9493	0.7010	0.9770
2			Li [54]	0.9671	0.9545	0.7027	0.9820
3			Mo [55]	0.9751	0.9570	0.7009	0.9843
4			Proposed MSR U-Net Method	0.9807	0.9686	0.7883	0.9822
1		CHASE_DB1[42]	Li [54]	0.9628	0.9429	0.7118	0.9791
2			Mo [55]	0.9671	0.9478	0.7003	0.9750
3			Proposed MSR U-Net Method	0.9709	0.9633	0.7492	0.9786

TABLE 9. Comparison of experimental results: CHASE dataset is used for training, then testing on the STARE and DRIVE datasets.

S.No	Train Set	Methods	Year	AUC	Acc	Se	Sp
1	CHASE_DB1[42]	DRIVE[40]	Li [54]	0.9628	0.9429	0.7118	0.9791
2			Mo [55]	0.9671	0.9478	0.7003	0.9750
3			Proposed MSR U-Net Method	0.9701	0.9598	0.7316	0.9812
1		STARE[41]	Fraz[42]	0.9565	0.9415	0.7103	0.9665
2			Li [54]	0.9553	0.9417	0.7240	0.9768
3			Mo [55]	0.9690	0.9515	0.7032	0.9794
4			Proposed MSR U-Net Method	0.9799	0.9654	0.7583	0.9857

proposed MSR U-Net framework when confronted with new data.

In addition to the experimental results presented in comparison to existing research, we contend that the proposed

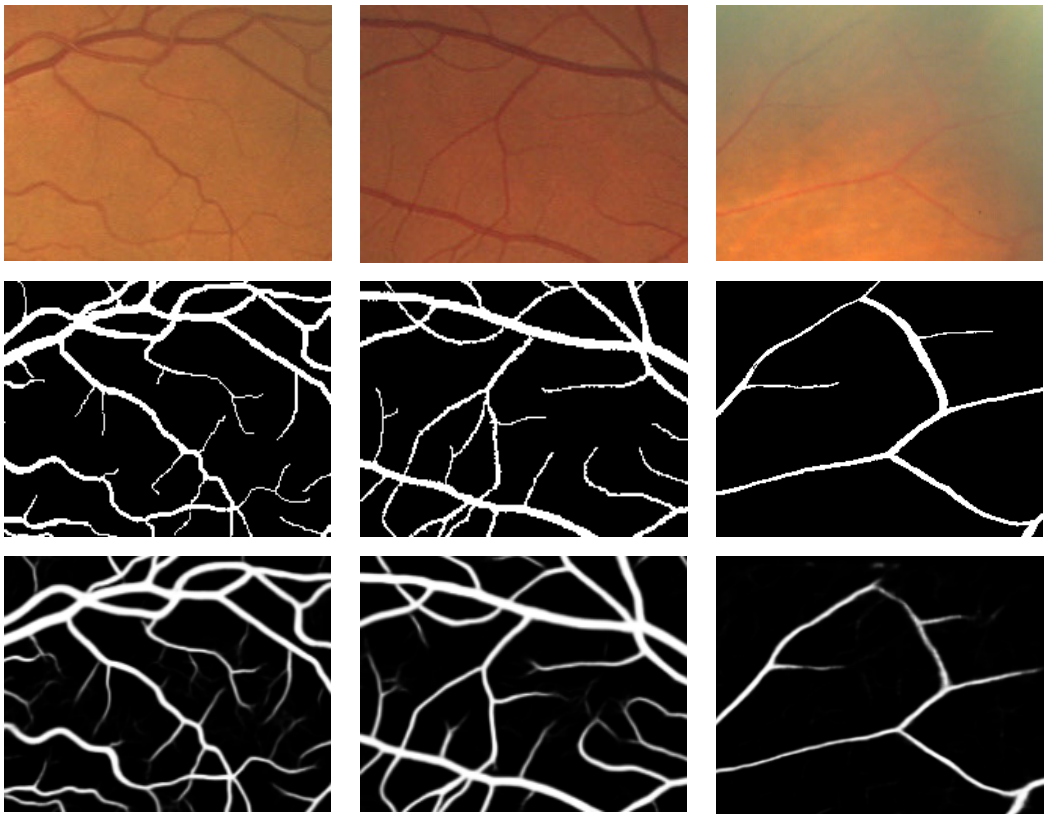


FIGURE 10. Segmentation of vasculature in challenging images. (Row wise) 1. Fundus image patches with low-varying contrast and central light reflex, 2. Manual segmentation of the respective patches, and 3. Segmented vessels of our MSR U-Net method.

framework exhibits superior generalization and robustness for the following reasons: 1. Multi-Scale Feature Extraction: The MSR Block effectively extracts features at different scales, allowing the model to capture information at various levels of granularity. This multi-scale approach contributes to the model's ability to generalize across a wide range of inputs. 2. B-Res Paths as Skip Connections: The incorporation of B-Res paths as skip connections is instrumental in reducing training complexity. These B-Res paths maintain gradient flow, enabling the training of deeper neural networks. This design choice not only simplifies the training process but also enhances the model's ability to generalize effectively. By employing a proper training strategy, we can significantly enhance the model's generalization. In our proposed framework, we have incorporated an effective data preprocessing method and employed random patch extraction for data augmentation, thereby enriching the training samples. Collectively, these solutions contribute to improved model generalization and can be applied to enhance the performance of other related works.

5) EXPERIMENT 5: PERFORMANCE ON CHALLENGING IMAGES

Numerous techniques have been put forth in the area of vascular segmentation. However, there are still important issues that need to be addressed, such as vascular segmentation when a central light reflex is present, segmentation when pathologies are present, and segmentation in low contrast situations. An exhaustive study of each of these scenarios is conducted and found that retinal vessel segmentation becomes notably more challenging due to these factors. As depicted in Figure 10 and the results presented in TABLE 10, our proposed method successfully addresses these challenges, allowing precise segmentation of vessels.

TABLE 10. Evaluation of proposed method on challenging images.

Database	Acc	AUC	Se	Sp
DRIVE [40]	0.9686	0.9794	0.84122	0.9810
STARE [41]	0.9786	0.9852	0.8503	0.9646
CHASE_DB1 [42]	0.9640	0.9726	0.8240	0.9767

6) EXPERIMENT 6: ANALYSIS ON PATHOLOGICAL IMAGES

Nguyen method [6] and Soomro method [56] are used to compare with our method, and the results of each method are illustrated in Figure 11. The proposed method yields better vessel segmentation results than Nguyen's and Soomro's methods. Following the comparison of segmented images

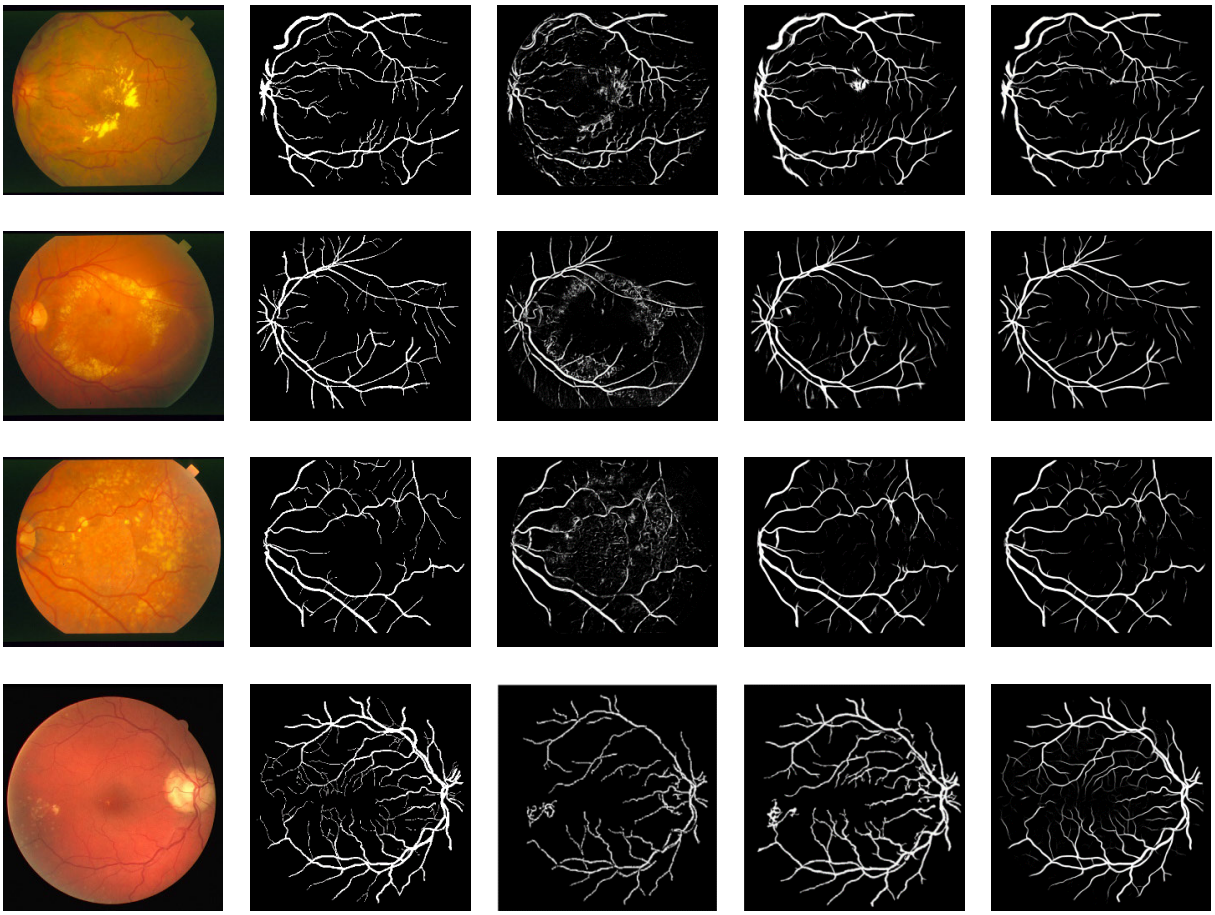


FIGURE 11. Comparison of Vessel Observation in Images with Pathologies. (Column wise) 1. Original fundus images, 2. Corresponding ground truth images, 3. Results from Nuygen's method [6], 4. Results from Soomro's method [56], and 5. Output images from our MSR U-Net method.

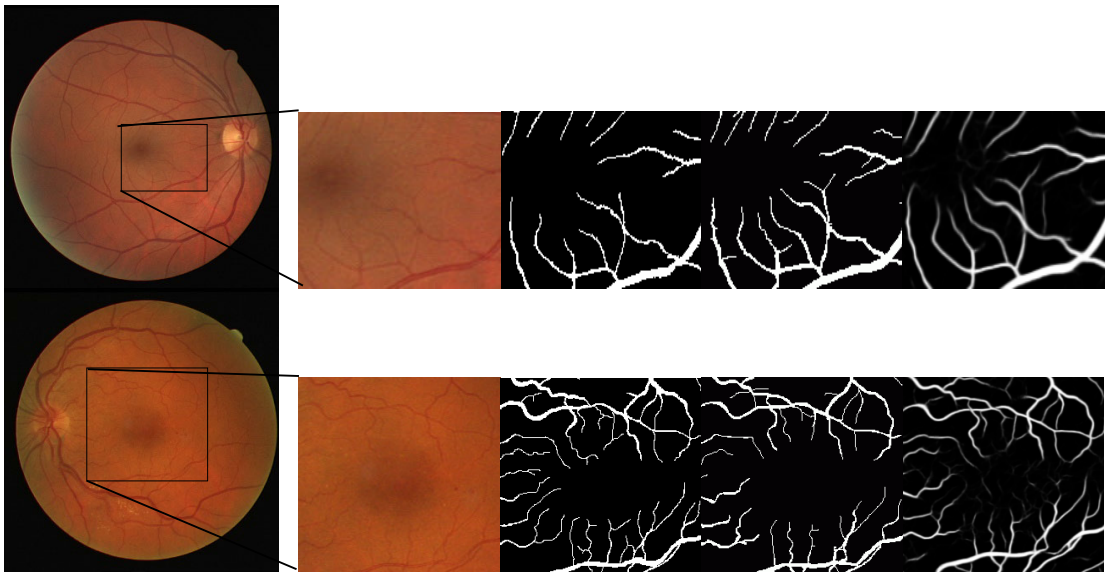


FIGURE 12. Tiny Vessel Observation. The color fundus images and cropped regions with obscure blood vessel appearance from them are presented in the first and second columns, followed by the equivalent ground truth from the first and second observers, respectively, in the third and fourth columns. The segmented output of our method is shown in fifth column. This visualization makes it possible to see tiny vessels clearly inside the cropped area, allowing for a detailed comparison with the ground truths.

produced by our method and these approaches, we analyzed their performances and compared them to our method. The performance comparison indicates a significant improvement in vessel segmentation results with our method. These

findings prove the robustness of our method in handling pathological fundus images.

7) EXPERIMENT 7: SEGMENTATION OF TINY VESSELS

Despite the fact that numerous segmentation techniques have been put forth, two visual limitations have been consistently observed across all approaches: the first is the inability to detect tiny vessels; and the second is reduced sensitivity. These drawbacks can only be addressed by effectively segmenting tiny vessels that play a significant role in achieving good sensitivity. It is evident that our proposed method excels in detecting even the smallest vessels. The output of our method can be directly compared to the gold standard image. Experimental results confirm that the proposed approach efficaciously identifies vessels of both diminutive and substantial sizes, yielding superior performance. It gives ophthalmologists a perfect chance to monitor the progress of eye ailments. Figure 12 compares the ground truth to tiny vessels produced by the proposed method. The third column shows the result of our method and clearly displays tiny vessels, and the accurate detection of these small vessels significantly enhances the effectiveness of our approach. Contrarily, the segmentation performance as a whole declines when thick vessels are recognized at the expense of tiny vessels.

IV. DISCUSSION AND CONCLUSION

The precise delineation of retinal vasculature in fundus images is crucial for identifying and analyzing diverse ocular pathologies. However, fundus images often present various challenging factors, including: Variability in vessel thickness as they extent to the periphery of the retina, insufficient contextual information, complex vessel morphology, and confusion with lesions, etc. will influence the accurate segmentation of retinal vasculature.

Combined with their strong ability to represent contextual features, CNNs have demonstrated promising results in segmenting medical images, notably fundus images, with the U-Net framework leading the way. Nonetheless, there are still several drawbacks in achieving precise segmentation of vasculature in retinal images. These include the limited effectiveness in extracting multiscale features, the loss of information due to multiple pooling operations, and the inadequate handling of local context features through skip connections. Most research in this area typically relies on three well-known benchmark datasets: DRIVE, STARE, and CHASE_DB1. However, these datasets have a limited number of images, which poses a further challenge in learning from a relatively small training dataset.

In order to tackle these concerns, we have devised an innovative framework for the automatic segmentation of retinal vessels. The framework is founded on a comprehensive deep learning approach, encompassing enhancements in data preprocessing, data augmentation, and network architecture.

1. We explored the performance change brought by combining diverse pre-processed images such as shade-corrected, CLAHE, and grayscale versions of the input images. This adaptation primarily aims to improve the network's ability to generalize in the face of insufficient contextual information, thus leading to an improvement in network performance.
2. A customized input pipeline strategy is necessary during training, requiring the model to train on a greater amount of data that is sufficient enough for the model to learn features that are capable of discriminating the actual and fraudulent vessels. We generated such data using a random cropping mechanism from the preprocessed images. Subsequently, we train our MSR U-Net model using these sampled patches.
3. The proposed MSR block of the modified U-net architecture has multi-scale kernels, which allows the analysis of identical local regions of the fundus image through various receptive fields. The purpose of this characteristic is to enhance the model's capacity to detect blood vessel sizes that vary.
4. Preserving complex morphological characteristics of the vessels across depths, particularly in the context of the U-net architecture, where propagating features from the encoder to the decoder at various depths can prove to be difficult. Using the B-Res path, we successfully alleviated such issues and yielded better results.

AUC, Acc, Se, Sp, and F1 score are the metrics used to evaluate the effectiveness of the proposed MSR U-Net.

Blood vessels occupy a lesser proportion of the total area (in terms of number of pixels) of the fundus images. Thus, the effect of vessel extraction is highly profound on performance metrics, indicating a large improvement in the identification of vessel pixels with a small increase in the metric values. In experiments 5, 6, and 7, we identified challenging areas, such as thin and edge vessels, across the databases and investigated the efficacy of the proposed approach. Evidently, the effectiveness of our MSR U-Net method is superior over existing U-Net variants not only in the identified challenging thin and edge vessel areas but also in the other areas of the fundus images as well. The performance gain achieved with the use of MSR U-Net reflects the various benefits of the proposed approach, including the following. (1) Complexity of the model: We are able to design a novel, innovative model architecture with less increase in the number of parameters. That is, the proposed model has the same level of complexity comparable to basic U-Net, though the convolution layers are replaced with the introduction of MSR block. The main objective of the MSR block is to extract features that are at multiple scale levels and more diverse. (2) Variations in the vessel thickness and their positions: The proposed model has successfully dealt with vessels from thin to thick and is present across various positions in the fundus images i.e., edge vessels. (3) Cross-validation: Since three publicly available datasets are utilized for the evaluation, it's fair to infer that the proposed approach is more generalizable than other approaches currently in use based on how well

it performs when subjected to training on one dataset and subsequently evaluated on the remaining two. Harnessing the innovations at every stage of the proposed method, such as in the input pipeline, model architecture, and training strategies, we believe that the aforementioned advantages are achieved in our work.

The proposed MSR U-Net is presently employed in the identification of blood vessels within the retinal region. Nevertheless, we are of the opinion that this approach possesses the capacity to precisely and effectively segment vasculature in additional organs that are similarly afflicted by the complications of indistinct boundaries and fluctuations in scale. Our upcoming efforts will be directed toward expanding the applicability of our methodology to a larger variety of OCT images and clinical angiography data.

REFERENCES

[1] J. Flamme, K. Konieczka, R. M. Bruno, A. Virdis, A. J. Flammer, and S. Taddei, "The eye and the heart," *Eur. Heart J.*, vol. 34, no. 17, pp. 1270–1278, May 2013.

[2] M. Li, Y. Yang, H. Jiang, G. Gregori, L. Roisman, F. Zheng, B. Ke, D. Qu, and J. Wang, "Retinal microvascular network and microcirculation assessments in high myopia," *Amer. J. Ophthalmol.*, vol. 174, pp. 56–67, Feb. 2017.

[3] O. Gishti, V. W. Jaddoe, J. F. Felix, C. C. Klaver, A. Hofman, T. Y. Wong, M. K. Ikram, and R. Gaillard, "Retinal microvasculature and cardiovascular health in childhood," *Pediatrics*, vol. 135, pp. 678–685, Apr. 2015.

[4] Z. Waheed, M. U. Akram, A. Waheed, M. A. Khan, A. Shaukat, and M. Ishaq, "Person identification using vascular and non-vascular retinal features," *Comput. Electr. Eng.*, vol. 53, pp. 359–371, Jul. 2016.

[5] S. Iqbal, K. Naveed, S. S. Naqvi, A. Naveed, and T. M. Khan, "Robust retinal blood vessel segmentation using a patch-based statistical adaptive multi-scale line detector," *Digit. Signal Process.*, vol. 139, Jul. 2023, Art. no. 104075.

[6] U. T. V. Nguyen, A. Bhuiyan, L. A. F. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern Recognit.*, vol. 46, no. 3, pp. 703–715, Mar. 2013.

[7] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–269, Sep. 1989.

[8] G. B. Kande, P. V. Subbaiah, and T. S. Savithri, "Unsupervised fuzzy based vessel segmentation in pathological digital fundus images," *J. Med. Syst.*, vol. 34, no. 5, pp. 849–858, Oct. 2010.

[9] W. Wang, W. Wang, and Z. Hu, "Retinal vessel segmentation approach based on corrected morphological transformation and fractal dimension," *IET Image Process.*, vol. 13, no. 13, pp. 2538–2547, Nov. 2019.

[10] Y. Yin, M. Adel, and S. Bourennane, "Retinal vessel segmentation using a probabilistic tracking method," *Pattern Recognit.*, vol. 45, no. 4, pp. 1235–1244, Apr. 2012, doi: 10.1016/j.patcog.2011.09.019.

[11] N. M. Salem, S. A. Salem, and A. K. Nandi, "Segmentation of retinal blood vessels based on analysis of the Hessian matrix and clustering algorithm," in *Proc. 15th Eur. Signal Process. Conf.*, Sep. 2007, pp. 428–432.

[12] M. D. Saleh, C. Eswaran, and A. Mueen, "An automated blood vessel segmentation algorithm using histogram equalization and automatic threshold selection," *J. Digit. Imag.*, vol. 24, no. 4, pp. 564–572, Aug. 2011.

[13] B. Al-Diri, A. Hunter, and D. Steel, "An active contour model for segmenting and measuring retinal vessels," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1488–1497, Sep. 2009.

[14] Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images," *IEEE Trans. Med. Imag.*, vol. 34, no. 9, pp. 1797–1807, Sep. 2015.

[15] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[17] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation," 2018, *arXiv:1802.06955*.

[18] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "DUNet: A deformable network for retinal vessel segmentation," *Knowl.-Based Syst.*, vol. 178, pp. 149–162, Aug. 2019.

[19] B. Wang, S. Qiu, and H. He, "Dual encoding U-Net for retinal vessel segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2019, pp. 84–92.

[20] K. Yue, B. Zou, Z. Chen, and Q. Liu, "Retinal vessel segmentation using dense U-net with multiscale inputs," *J. Med. Imag.*, vol. 6, no. 3, Sep. 2019, Art. no. 034004.

[21] J. Hu, H. Wang, S. Gao, M. Bao, T. Liu, Y. Wang, and J. Zhang, "S-UNet: A bridge-style U-Net framework with a saliency mechanism for retinal vessel segmentation," *IEEE Access*, vol. 7, pp. 174167–174177, 2019.

[22] X. Wei, K. Yang, D. Bzdok, and Y. Li, "Orientation and context entangled network for retinal vessel segmentation," *Expert Syst. Appl.*, vol. 217, May 2023, Art. no. 119443.

[23] L. Yang, H. Wang, Q. Zeng, Y. Liu, and G. Bian, "A hybrid deep segmentation network for fundus vessels via deep-learning framework," *Neurocomputing*, vol. 448, pp. 168–178, Aug. 2021.

[24] G.-X. Xu and C.-X. Ren, "SPNet: A novel deep neural network for retinal vessel segmentation based on shared decoder and pyramid-like loss," *Neurocomputing*, vol. 523, pp. 199–212, Feb. 2023.

[25] G. Wang, Y. Huang, K. Ma, Z. Duan, Z. Luo, P. Xiao, and J. Yuan, "Automatic vessel crossing and bifurcation detection based on multi-attention network vessel segmentation and directed graph search," *Comput. Biol. Med.*, vol. 155, Mar. 2023, Art. no. 106647.

[26] H. Boudegga, Y. Elloumi, M. Akil, M. H. Bedoui, R. Kachouri, and A. B. Abdallah, "Fast and efficient retinal blood vessel segmentation method based on deep learning network," *Computerized Med. Imag. Graph.*, vol. 90, Jun. 2021, Art. no. 101902.

[27] K. Hu, Z. Zhang, X. Niu, Y. Zhang, C. Cao, F. Xiao, and X. Gao, "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, Oct. 2018.

[28] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1597–1605, Jul. 2018.

[29] S. Zhang, H. Fu, Y. Yan, Y. Zhang, Q. Wu, M. Yang, M. Tan, and Y. Xu, "Attention guided network for retinal image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2019, pp. 797–805.

[30] Y. Liu, J. Shen, L. Yang, G. Bian, and H. Yu, "ResDO-UNet: A deep residual network for accurate retinal vessel segmentation from fundus images," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104087.

[31] L. Mou, L. Chen, J. Cheng, Z. Gu, Y. Zhao, and J. Liu, "Dense dilated network with probability regularized walk for vessel detection," *IEEE Trans. Med. Imag.*, vol. 39, no. 5, pp. 1392–1403, May 2020.

[32] V. Cherukuri, B. G. V. Kumar, R. Bala, and V. Monga, "Deep retinal image segmentation with regularization under geometric priors," *IEEE Trans. Image Process.*, vol. 29, pp. 2552–2567, 2020.

[33] Z. Yan, X. Yang, and K.-T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 1912–1923, Sep. 2018.

[34] S. M. Pizer, "Adaptive histogram equalization and its variations," *Comput. Vis. Graph. Image Process*, vol. 39, pp. 355–368, Sep. 1987.

[35] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," 2016, *arXiv:1602.07261*.

[36] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The importance of skip connections in biomedical image segmentation," in *Deep Learning and Data Labeling for Medical Applications*. Cham, Switzerland: Springer, 2016, pp. 179–187.

[37] H. Yang, C. Yuan, B. Li, Y. Du, J. Xing, W. Hu, and S. J. Maybank, "Asymmetric 3D convolutional neural networks for action recognition," *Pattern Recognit.*, vol. 85, pp. 1–12, Jan. 2019.
[38] N. Ibtehaz and M. S. Rahman, "MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Netw.*, vol. 121, pp. 74–87, Jan. 2020.
[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.
[40] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
[41] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.
[42] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, Sep. 2012.
[43] R. Li, M. Li, J. Li, and Y. Zhou, "Connection sensitive attention U-NET for accurate retinal vessel segmentation," 2019, *arXiv:1903.05558*.
[44] J. Lu, Y. Xu, M. Chen, and Y. Luo, "A coarse-to-fine fully convolutional neural network for fundus vessel segmentation," *Symmetry*, vol. 10, no. 11, p. 607, Nov. 2018.
[45] K. K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, Athens, Greece, Oct. 2016, pp. 140–148.
[46] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Melbourne, VIC, Australia, Apr. 2017, pp. 248–251.
[47] B. Zhang, S. Huang, and S. Hu, "Multi-scale neural networks for retinal blood vessels segmentation," 2018, *arXiv:1804.04206*.
[48] J. Zhuang, "LadderNet: Multi-path networks based on U-Net for medical image segmentation," 2018, *arXiv:1810.07810*.
[49] A. Oliveira, S. Pereira, and C. A. Silva, "Retinal vessel segmentation based on fully convolutional neural networks," *Expert Syst. Appl.*, vol. 112, pp. 229–242, Dec. 2018.
[50] Y. Jiang, H. Zhang, N. Tan, and L. Chen, "Automatic retinal blood vessel segmentation based on fully convolutional neural networks," *Symmetry*, vol. 11, no. 9, p. 1112, Sep. 2019.
[51] T. M. Khan, S. S. Naqvi, A. Robles-Kelly, and I. Razzak, "Retinal vessel segmentation via a multi-resolution contextual network and adversarial learning," *Neural Netw.*, vol. 165, pp. 310–320, Aug. 2023.
[52] H. Du, X. Zhang, G. Song, F. Bao, Y. Zhang, W. Wu, and P. Liu, "Retinal blood vessel segmentation by using the MS-LSDNet network and geometric skeleton reconnection method," *Comput. Biol. Med.*, vol. 153, Feb. 2023, Art. no. 106416.
[53] R. Liu, T. Wang, X. Zhang, and X. Zhou, "DA-Res2UNet: Explainable blood vessel segmentation from fundus images," *Alexandria Eng. J.*, vol. 68, pp. 539–549, Apr. 2023.
[54] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 109–118, Jan. 2016.
[55] J. Mo and L. Zhang, "Multi-level deep supervised networks for retinal vessel segmentation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 12, pp. 2181–2193, Dec. 2017.
[56] T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, L. Zheng, and M. Paul, "Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation," *Expert Syst. Appl.*, vol. 134, pp. 36–52, Nov. 2019.
[57] D. E. Alvarado-Carrillo and O. S. Dalmau-Cedeño, "Width attention based convolutional neural network for retinal vessel segmentation," *Expert Syst. Appl.*, vol. 209, Dec. 2022, Art. no. 118313.
[58] Y. Zhang, M. He, Z. Chen, K. Hu, X. Li, and X. Gao, "Bridge-Net: Context-involved U-Net with patch-based loss weight mapping for retinal blood vessel segmentation," *Expert Syst. Appl.*, vol. 195, Jun. 2022, Art. no. 116526.
[59] M. K. Kar, D. R. Neog, and M. K. Nath, "Retinal vessel segmentation using multi-scale residual convolutional neural network (MSR-Net) combined with generative adversarial networks," *Circuits, Syst., Signal Process.*, vol. 42, no. 2, pp. 1206–1235, Feb. 2023.

GIRI BABU KANDE (Senior Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Nagarjuna University, Guntur, in 1996, the M.E. degree in electronics and communication engineering from Andhra University, Visakhapatnam, in 2000, and the Ph.D. degree in electronics and communication engineering from Jawaharlal Nehru Technological University, Hyderabad, in 2010. From 2008 to 2019, he was a Professor and the Head of the Department of Electronics and Communication Engineering Department. Since 2019, he has been a Professor and the Dean of Academics with the Vasireddy Venkatadri Institute of Technology, Nambur. He is the author of one book and more than 75 publications. His research interests include computer vision, statistical modeling and learning, and medical image analysis. He was a recipient of the Best Researcher by Jawaharlal Nehru Technological University, Kakinada, in 2018, and the ISTE AP Best Teacher Award, in 2021.

LOGESH RAVI is associated with the Centre for Advanced Data Science, Vellore Institute of Technology, Chennai, India. He has published more than 100 papers in reputable international journals and conferences. His research interests include artificial intelligence, recommender systems, big data, information retrieval, fintech, and social computing. He is listed and ranked in the prestigious top 2% scientists worldwide by Stanford University and Elsevier B.V.

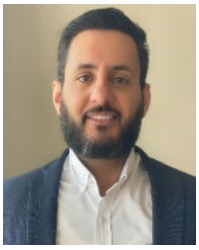
NITYA KANDE received the B.S. degree in data science from IIT Chennai. She is currently pursuing the integrated M.Tech. degree in computer science and the Engineering degree with VIT-AP University, Amaravathi. Her research interests include deep learning, computer vision, and medical image analysis.

MADHUSUDANA RAO NALLURI (Member, IEEE) received the M.Sc. degree in pure mathematics from Andhra University, India, the M.Tech. degree in scientific computing from the Birla Institute of Technology, India, and the Ph.D. degree from Sastra University, in February 2019. He was a Senior Lecturer with the Computer Science and Engineering Department, National Institute of Science and Technology, from 2009 to 2010. Then, he was an Assistant Professor-III with the School of Computing, Sastra University, for nine years. He is currently an Associate Professor with the Department of Computer Science and Engineering, School of Computing, Amaravati Campus. He has obtained 84th rank in GATE-2007. He has authored around 17 technical papers in reputed conferences and journals indexed in Scopus and SCI. His research interests include evolutionary algorithms, heuristics for NP-HARD/complete problems, machine learning, multi-objective optimization, and medical image analysis.

HOSSAM KOTB received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Faculty of Engineering, Alexandria University, Alexandria, Egypt, in 2009, 2013, and 2020, respectively. The Ph.D. research work is focused on the performance enhancement of renewable energy conversion systems. Currently, he is an Assistant Professor with the Electrical Power and Machines Department, Faculty of Engineering, Alexandria University. His research interests include power system analysis, electrical drives, modern control techniques, smart grids, optimization, electric vehicles, and renewable energy systems. He is an Associate Editor of *Alexandria Engineering Journal* (AEJ).

KAREEM M. ABORAS received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Faculty of Engineering, Alexandria University, Alexandria, Egypt, in 2010, 2015, and 2019, respectively. The Ph.D. research work is focused on the performance enhancement of renewable energy conversion systems. Currently, he is an Assistant Professor with the Electrical Power and Machines Department, Faculty of Engineering, Alexandria University. His research interests include power electronics, control, drives, power systems, smart grids, microgrids, power quality, optimizations, electric vehicles, machine learning, modeling, fuel cells, HVDC, and renewable energy systems. He is a Reviewer of IEEE TRANSACTIONS ON ENERGY CONVERSION, *Electric Power Systems Research*, *Smart Science*, *Alexandria Engineering Journal* (IET), *Energy Reports*, IEEE ACCESS, *Cybernetics and Systems*, *Protection and Control of Modern Power Systems* (MDPI), *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *Cogent Engineering*, and Hindawi journals.

AMR YOUSEF (Member, IEEE) received the B.Sc. degree from the Engineering Mathematics Department, Alexandria University, in 2001, the M.Sc. degree from the Electrical Engineering Department, Alexandria University, in 2006, and the Ph.D. degree in electrical and computer engineering from Old Dominion University (ODU), in May 2012. He was a Postdoctoral Research Associate with the Old Dominion Vision Laboratory, USA. He is currently an Assistant Professor with the Electrical Engineering Department, University of Business and Technology, Saudi Arabia. His research interests include optimization techniques, image processing/computer vision, and machine learning algorithms. He is a member of SPIE and OSA.



YAZEED YASIN GHADI received the Ph.D. degree in electrical and computer engineering from Queensland University. He was a Postdoctoral Researcher with Queensland University. He is currently an Assistant Professor of software engineering with Al Ain University. He has published more than 25 peer-reviewed journals and conference papers. He holds three pending patents. His current research interests include developing novel electro-acoustic-optic neural interfaces for large-scale high-resolution electrophysiology and distributed optogenetic stimulation. He was a recipient of several awards. His dissertation on developing novel hybrid plasmonic photonic on-chip biochemical sensors received the Sigma Xi Best Ph.D. Thesis Award.

A. SASIKUMAR received the B.E. and M.E. degrees from Anna University, in 2011 and 2013, respectively, and the Ph.D. degree from SASTRA, in 2020. He is currently an Assistant Professor with the Department of Data Science and Business Systems, School of Computing, SRM Institute of Science and Technology, Chennai, India. He has published more than 25 journal articles. His research interests include blockchain, analog VLSI, digital VLSI, and swarm intelligence.