Mesopotamian Journal of Artificial Intelligence in Healthcare
Vol.2025, **pp**. 234–243
DOI: https://doi.org/10.58496/MJAIH/2025/022; ISSN: 3005-365X
https://mesopotamian.press/journals/index.php/MJAIH

Research Article

Retinal Vessel Image Segmentation Based on Guided Filter and Neural Networks

Ibrahem Al-Jadir [1], Jawad S. Hameed [1], Husam Yahya[1,*], Dalia N. Al-Oboudi [2]

[1] *Ministry of Higher Education and Scientific Research, Baghdad, Iraq*

[2] *Department of oral and maxillofacial surgery, College of Dentisty, University of Baghdad, Iraq*

ARTICLEINFO

Article History
Received 3 Jul. 2025
Revised 20 Aug. 2025
Accepted 12 Sep. 2025
Published 30 Oct. 2025

Keywords
Filters,
Ophthalmic,
U-Net,
Retinal,
Segmentation.

ABSTRACT

Accurate segmentation of retinal blood vessels is essential for the early diagnosis and monitoring of various ophthalmic and systemic diseases. This work shows how to use a modified U-Net architecture to improve how retinal blood vessels are segmented. The method uses rolling guidance filters with two different sets of parameters to make the edges of the vessels sharper and make it easier to tell them apart. Separate U-Net models process each filtered image independently, and then the outputs are combined to make the final segmentation result. Both networks are trained at the same time using a shared loss function. Tests on many benchmarks retinal image datasets show that the suggested method makes segmentation far more accurate than the usual U-Net model.

1. INTRODUCTION

This is essential to correctly segment the blood arteries in the retina in order to diagnose and keep an eye on eye and systemic problems such as diabetic retinopathy, high blood pressure, and heart disease. Retinal scans give us important information about the health of the microvascular system, and automated vessel segmentation helps us find problems early and make better treatment decisions [1]. However, low vessel-background contrast, fluctuating vessel width, picture noise, and irregular lighting make this job hard and require strong segmentation methods. Two main types of methods for segmenting retinal images are supervised and unsupervised [2]. Unsupervised methods, such as fuzzy clustering, adaptive and Gabor filters, and Expectation-Maximization, do not need labelled data [3-6]; thus, they can be changed, but they are often less accurate. Support Vector Machines (SVM), classic neural networks, and, more recently, deep learning-based designs like Convolutional Neural Networks (CNNs) and U-Net are examples of supervised approaches that need labelled training data. Supervised models can be very accurate, but how well they work depends a lot on the number and quality of the annotated datasets they use [7-12]. The U-Net architecture has gotten much attention among these models since it better classifies pixels than the others. It works well for medical picture segmentation since it has an encoder-decoder structure, skip connections, and the ability to learn features at different resolutions [13]. However, U-Net and other deep networks often have trouble separating delicate vessel structures. They may be affected by noise and variations in lighting, which makes it hard for them to perform on different datasets [2]. To deal with these problems, current research has looked into hybrid methods combining classical image processing and deep learning. In this case, guided and rolling guidance filters have shown promise in improving edge information and eliminating noise. These filters may show the edges of vessels before sending photos to neural networks, which helps with more accurate segmentation [14]. In [15] the authors proposed a U-Net–based architecture enhanced with full-scale guided filtering modules embedded in each decoder stage. The guided filters sharpen vessel boundaries at multiple resolutions, helping the model effectively capture fine vascular structures. Experimental results show that FSG-Net achieves high segmentation accuracy with fewer parameters compared to standard U-Net variants, validated on datasets like DRIVE and CHASE_DB1.In [15], a Retinal Vessel Segmentation Using Optimal Anisotropic Guided Filtering (OAGF) and Deep Learning Fusion. Biomedical Signal Processing and Control, 90, 105008. This paper presents a hybrid method combining Optimal Anisotropic Guided Filtering with deep neural networks for improved vessel segmentation. The guided filter enhances contrast and structure of vessels while suppressing noise, allowing the deep model to better capture subtle vascular features.

*Corresponding author. Email: Husam.yahya@mohesr.edu.iq

The method demonstrates increased sensitivity and robustness across multiple retinal image datasets under varying illumination and contrast conditions. In [16], a MAFE-Net was suggested, it represents a Multiple Attention-Guided Fusion Ensemble for Retinal Vessel Segmentation. MAFE-Net introduces a novel fusion ensemble approach that combines attention mechanisms and guided filtering to refine retinal vessel segmentation. It integrates features from multiple guided-filtered representations using attention-guided fusion, improving detection of thin vessels and vessel connectivity. The method achieves competitive performance across STARE and HRF datasets, with improvements in F1-score, recall, and generalizability. In this work, we suggest a hybrid segmentation that combines rolling guidance filters with the U-Net architecture. Two rolling guidance filters with various settings pre-process the retinal image. Each output is then processed by a distinct U-Net network. The two networks' outputs are combined, and a shared loss function helps the networks learn. This method uses the edge-enhancing power of rolling guidance filters and the deep feature extraction capacity of the U-Net to improve vessel segmentation. A dual-path U-Net framework with rolling guidance filters for better separating retinal blood vessels. Compared to traditional U-Net approaches, this one showed better edge preservation and fine vessel recognition. Systematic testing of the model with different hyper-parameters and datasets ensures it can be used in various situations. Fig. 1 shows instances of retinal vessels with different levels of tortuosity. This shows how many different types of vessel segmentation algorithms need to be able to represent reliably. Figure (1) illustrates (a) Blood vessels of the eye with low tortuosity (b) Blood vessels in the eye with high tortuosity.

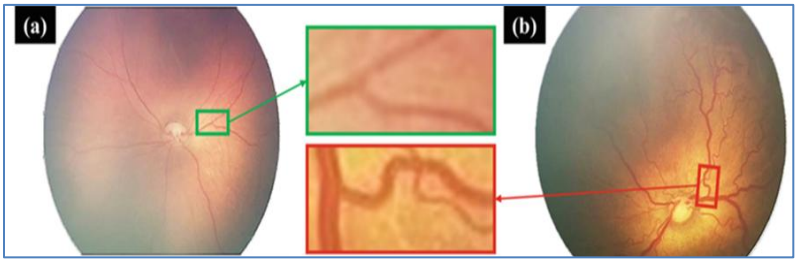

FIG.1. (A) BLOOD VESSELS OF THE EYE WITH LOW TORTUOSITY (B) BLOOD VESSELS IN THE EYE WITH HIGH TORTUOSITY

The main contributions presented by this paper are:

1. The Dual-Path U-Net Architecture with Rolling Guidance Filters. We provide a dual-path U-Net system in which two rolling guiding filters with different settings preprocess the image of the retina. Each filtered image goes through its own U-Net, which improves vessel detection at different scales. This design makes it easier for the model to separate both fine and wide vessels accurately.
2. Combining features at different scales to improve segmentation. The outputs from the two U-Nets are combined to provide more complete information about the vessels from different filtering inputs. This fusion makes boundaries more transparent and continuous, especially in poor contrast or noise areas. The approach makes vessel maps completer and more accurate.

2. METHODOLOGY

In this section we explain the main processes of the suggested retinal vessel image segmentation based on guided filter and neural networks as can be shown in Fig. 2.

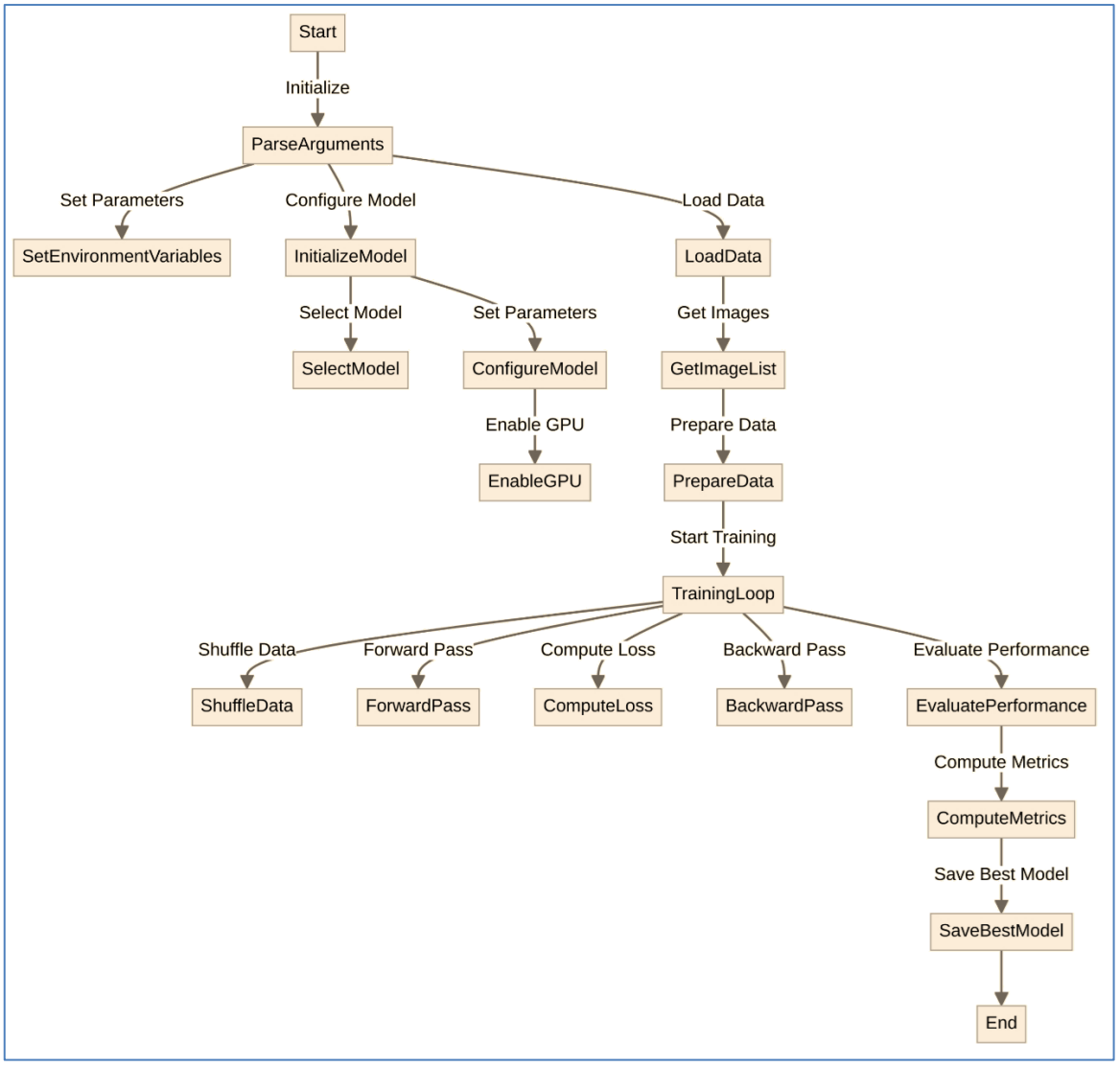

FIG.2. SUGGESTED RETINAL VESSEL IMAGE SEGMENTATION BASED ON GUIDED FILTER AND NEURAL NETWORKS

2.1 Processing Medical Images Workflow

Each stage ensures a smooth workflow for processing medical or image-based datasets using deep learning techniques for an efficient post-processing stages. This process includes the following sub processes:

- Initialization and Argument Parsing. The initialization phase where essential parameters are set through command-line arguments. These parameters include the number of training epochs, learning rate, batch size, dataset selection, and model-specific configurations such as normalization techniques (*GroupNorm* or *BatchNorm*). Additionally, GPU availability is checked, and necessary environment variables are set to enable optimal hardware utilization.
- Data Loading and Preprocessing. After initialization, the system loads and prepares the dataset. The dataset is selected from predefined options, and images are processed to match the model's input requirements. This stage involves retrieving image paths, resizing images, and ensuring proper labeling. The dataset is then split into training and testing sets, which are essential for assessing the model's performance during and after training.
- Model Initialization and Configuration. The model is selected from a list of available architectures. Key parameters such as the number of output classes, normalization methods, and input image size are configured. If multiple GPUs are available, the system enables parallel processing to accelerate training. The optimizer and loss function are also initialized to guide the model's learning process.
- Training Process. The core of the suggested model is the training loop where the model iteratively learns from the dataset. Training data is shuffled to introduce randomness, and batches of images are processed through the

model. A forward pass computes predictions, which are then compared against ground truth labels using a loss function. Backpropagation is applied to update the model's parameters, optimizing its accuracy over multiple epochs. The learning rate is adjusted periodically to fine-tune performance. Fig.3 explains the proposed system for training phase.

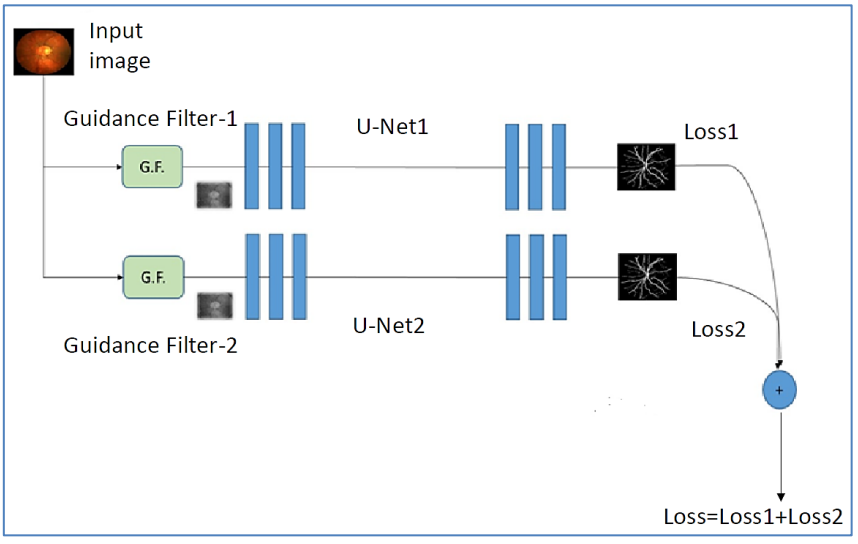

FIG. 2. PROPOSED SYSTEM FOR TRAINING PHASE.

- At regular intervals, the model's performance is assessed using test data. Key metrics such as accuracy, sensitivity, specificity, and AUC (Area Under Curve) are computed to evaluate classification performance. If the current model achieves higher accuracy than previous iterations, it is saved as the best-performing version. This ensures that only the most effective model is retained for deployment. Once training is complete, the best-performing model is finalized and stored for further use. Logs of the training process, including loss values and accuracy trends, are saved for analysis. This ensures that the model can be fine-tuned or re-trained if necessary. The system is now ready for real-world application, where it can be deployed for automated image classification or medical diagnosis.

2.2 Convolutional Neural Network Structure.

The basic architecture of the neural network is an efficient convolutional network for medical image segmentation. The main body of our architecture is the convolution layer or the main layer in CNN It extracts features or feature vectors from input data or feature maps. The convolution layer consists of several kernels that apply convolution operations and calculations. To apply convolution operations and calculations (Fig. 4), a convolution kernel is slid from left to right and top to bottom over the input image and multiplied by a specific region of the input or feature map in a peer-to-peer manner to produce a value known as the extracted feature. This specific region is called the acceptance region. These specific regions share kernels, known as weight sharing, which reduces the complexity of the model and makes the training process more manageable.

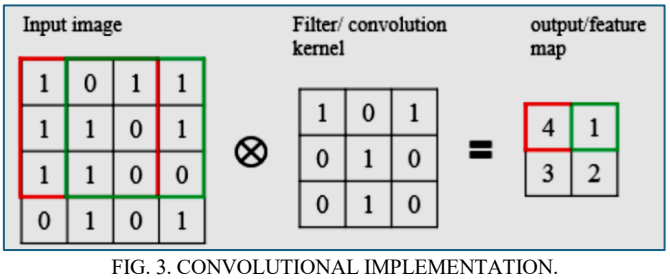

FIG. 3. CONVOLUTIONAL IMPLEMENTATION.

2.3 U-Nets

The U-Net is a unique convolutional neural network that performs well for medical image segmentation. The general schematic of this network is shown in Figure (5). The structure of the U-Net consists of a series of convolutional layers, Rayleigh activation functions, voting layer, and re-convolution. This network has received much attention for medical image segmentation uses. Specifically, retinal vessel segmentation serves as the foundation of the network employed in this study. The proposed method, which utilizes the U-sample, has been incorporated without any modifications or reductions as the fundamental framework for this thesis.

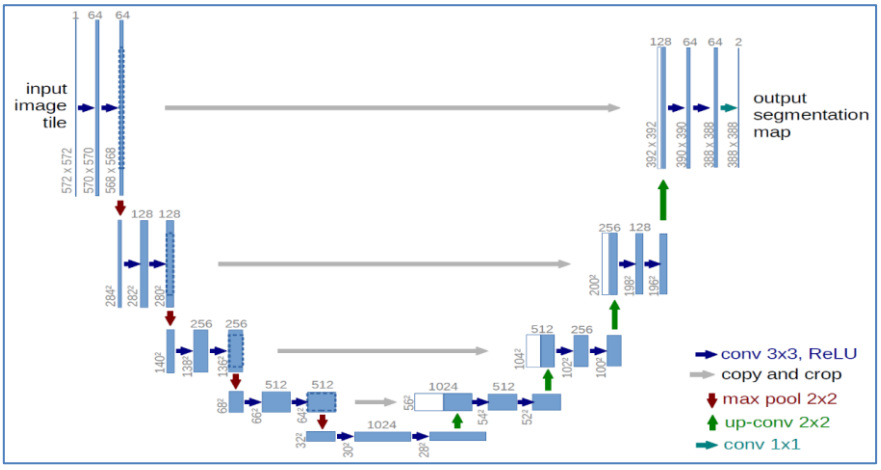

FIG. 4. U-NET NETWORK.

2.4 Batch Normalization

The input maps or features produced by convolutional layers may vary greatly, so for large or small values passed to the activation function, they encounter the problem of fading or exploding gradients, which disrupts the training process. To address this problem, batch normalization has been proposed to accelerate the training process, which reduces the large variation of internal variables by taking the input of the activation function and applying a normalization operation to each batch. Generally, batch normalization is performed before the activation function, but it can be used after the activation function depending on the application.

2.5 Activation Function

An activation function is a type of mathematical function that implements nonlinearity, which is applied to improve the feature extraction ability of networks. Activation function It is often implemented after convolutional layers and uses feature maps as input into neural networks. Sigmoid function and the ReLU function are two common options for the activation function.

2.6 Voting Layer

The feature map outside the convolution layer accurately captures the position of pixels, so it is very sensitive to spatial features. High sensitivity means that small movements of the feature positions, such as rotation and translation, will result in a different map, which reduces the robustness of the convolutional network. Typically, a voting layer with a voting operation is applied after the convolution to reduce the position dependence and avoid network invariance. Make sure. At the same time, the voting operation can also reduce the scale and volume of feature maps and reduce the computational burden. Voting operations can be performed as maximum value voting, average value voting, and sum voting. A sliding window is placed on the feature maps and the maximum value, average value, or sum value in this window is calculated as the voting output. In particular, if the voting window size is equal to the feature map size, it is called omnibus voting, otherwise, it is called local/regional voting.

2.7 Rotary Guided Filter

Edge detectors are crucial in extracting image features and details from retinal images. Guided filtering is a framework for filtering images with complete control over details under a given scale based on a rotating guide executed iteratively. This filter reveals different features of the image edge features. The rotationally guided filter has two parameters for tuning: filter size and filter bias. In this method, two modes, high detail and low detail, are considered. Using a rotational guided filter, different details can be considered at different scales, revealing the specific details of that scale. For example, if all scales in a retinal image are considered, the edges and parts can be revealed in great detail. The same image is then passed through a guided filter at a lower scale, and its edge-detected equivalent is shown. The filter removes irrelevant details and edges,

modifying parts like a noise suppressor to remove noise or noise-shaped textures from the leading edges or vessels. In this thesis, different image scales are created using a guided filter and the image with different scales is given to the network. Like a pre-processing step, the guided filter removes excess edge noise from the image so that the network can perform better segmentation. Fig. 6 explains the retinal image and its segmentation.

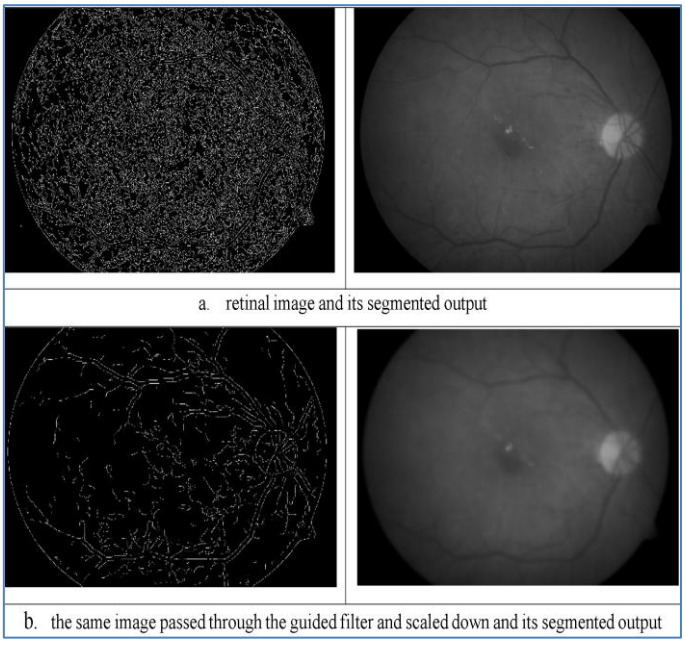

FIG. 5. RETINAL IMAGE AND ITS SEGMENTATION.

2.8 Aggregated Output Layer

The output layer or part acts as a classifier that determines an output map for the input class. The final output is formed by aggregating the outputs of the neural networks. Each neural network has a loss function, and the final loss function is formed by the sum of the loss functions. The loss function of each neural network is an entropy loss function. The relationship of the final loss function is given equations (1) and (2).

$$L_{\text{out}} = \frac{1}{M}\sum_{m=1}^{M} L_{CE}^{(m)}, \quad (1)$$

$$L_{CE}^{(m)} = \frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{n,c}\log\left(y_{n,c}^{(m)} + \varepsilon\right) \quad (2)$$

.

M : number of networks/paths; N : number of pixels; C : number of classes. $y_{n,c} \in \{0,1\}$: one-hot ground truth for pixel n and class c; $y_{n,c}^{(m)} \in [0,1)$: softmax probability from network m ; $\varepsilon > 0$: small constant for numerical stability.

Where y_i is the predicted probability value for class i and yi is the true probability value for that class. Therefore, as mentioned, the final loss function is formed from the average of the loss functions of each network, which are themselves entropy loss functions. It is based on the final loss function that the networks are trained, and the network coefficients are determined.

2.9 Network Training

The same basic semi-U network implementation is used to train the network. For retinal vessel segmentation, the network is trained for 150 epochs, the initial learning rate is 0.001, and every 50 epochs is run 10 times. Another type of training for segmentation, train the network for 150 epochs and the initial learning rate is 0.0001 and every 50 epochs is run 10 times.

The entire network is trained from scratch and no additional training data is used. These values are exactly the same as those used in the basic semi-U network.

3. RESULTS AND DISCUSSION

In this section, we explain the dataset used, experimental test results and the discussions of those results. The test inducted give a comprehensive view of workability of the suggested method in comparison to the other methods but first the dataset and its virtualization is explained in this section.

3.1 Dataset

Experiments are performed on four datasets of retinal images and their vessel segmentation that exist and are publicly available: DRIVE [63], and CHASEDB1, and the ORIGA collection, REFUGE For segmentation of retinal vessels . DRIVE dataset consists of 40 images. It consists of 20 training images and 20 test images along with two manual segmentations. The first segmentations serve as the basis for the main segmentation while the second segmentation serves as a human observer reference for performance comparison. The images are 768×584 pixels in size and are in digital form from the camera. CanonCR5non-mydriatic3CCD Taken in a 45◦ field of view (FOV) . Database CHASEDB1 It consists of 28 images measuring 960×999 pixels. These images of 14 school children were taken by a handheld fundus camera. Nidek NM-200-D Taken in a 30 ◦ field of view. The segmentation results are prepared in a supervised manner. The first 20 images can be selected as the training set and 8 images as the testing set. The Online Retinal Fundus Image Dataset for Glaucoma Research and Analysis (ORIGA) consists of 650 images obtained through the Singapore Malay Eye Study (SiMES) Organization. The 650 images were manually segmented and divided into 325 training images and 325 test images .
Database REFUGE consists of 1200 images . The dataset consists of 120 glaucomatous images and 1080 non-glaucomatous images and provides predefined segmentation for training and testing.

3.2 Data virtualization

Although the network receives a complete fundus image as input and the output is the same size as the input, augmenting the data with the help of data virtualization helps the network to overcome overfitting problems. Data virtualization techniques include random horizontal rotation with probability 0.5, random rotation in the range [0, 20°] and gamma contrast enhancement. The step between [0.5 ,2] in this implementation of this work is selected .

3.3 Implementation

First, the designed network is trained using all the images and after training, the performance of the trained network is checked with the test data. The same implementation of the basic semi-U network is used for training the network. For retinal vessel segmentation, the network is trained for 150 epochs, the initial learning rate is 0.001 and every 50 epochs is run 10 times. Another type of training for segmentation, train the network for 150 epochs and the initial learning rate is 0.0001 and every 50 epochs is run 10 times. The entire network is trained from scratch and no additional training data is used. These values are exactly the same as those used in the basic semi-U network. After training, the network is run on test data and the output is evaluated. Fig. 7 shows a sample input and output of the network.

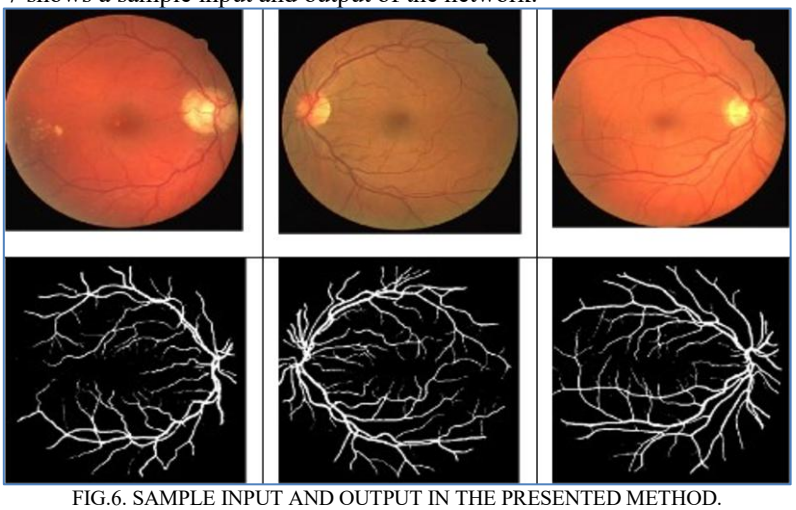

FIG.6. SAMPLE INPUT AND OUTPUT IN THE PRESENTED METHOD.

3.4 Network Outputs

In this section, the outputs of the different parts of the presented model are examined. In Figure (8), the different outputs of the guided filter come from an input image. Figure (8) shows that these outputs are input images with different detail scales.

In other words, these images are the same input images. Some disturbing details that are not veins and are mostly noise for different segments or main segments or veins have been removed from the image so that the segmentation of veins can be done better. It is worth noting that the output details may not be visible to the naked eye in the image, such as the removed noise and removed noisy edges that are not easily visible due to their small size and noise in the image texture. However, they can significantly impact the final performance of the neural network and during training. A sample of the model output at different stages is also shown in Figure (9).

Fig. 7. Sample output of a guided filter with different parameters.

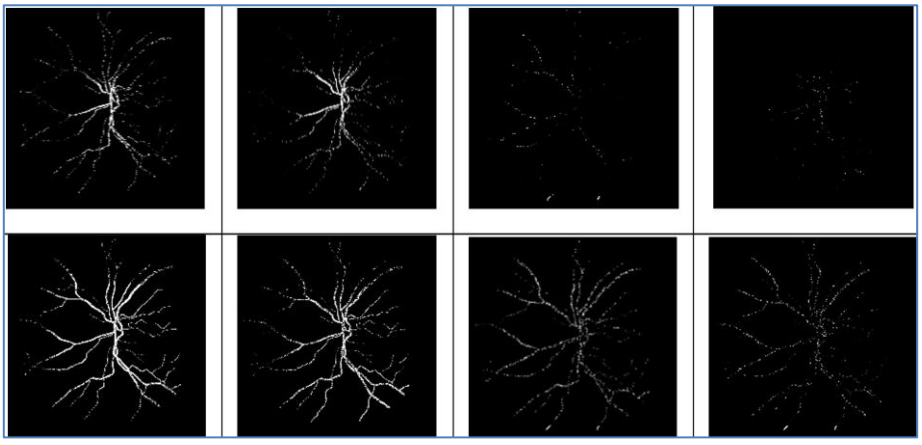

FIG. 8. THE OUTPUT OF THE PRESENTED MODEL IN DIFFERENT STAGES.

3.5 Evaluation and Comparison

To better evaluate the proposed model, comparisons with other methods are conducted. For this purpose, the test images of each dataset are presented as input to the proposed network and obtain the output. Then, the feature values, accuracy, and sensitivity for the output images and averaged over all images are listed. Also, a comparison of the results with other methods, including (7), is conducted. The results of these evaluations are given in Tables (1) and (2).

TABLE I. RESULTS OBTAINED ON THE DRIVE DATASET

Method	Accuracy	Sensitivity	SPE Features
U-Net Method	0.9531	0.7537	0.9820
Presented Method	0.9567	0.7602	0.9837

TABLE II. RESULTS OBTAINED ON THE CHASEBD1 DATASET

Method	Accuracy	Sensitivity	SPE Features
U-Net Method	0.9578	0.8288	0.9701
Presented Method	0.9623	0.7993	0.9823

For further practical examination, several outputs from the presented method, as well as the output of the U-Net Network alone, are also given in Fig. 10.

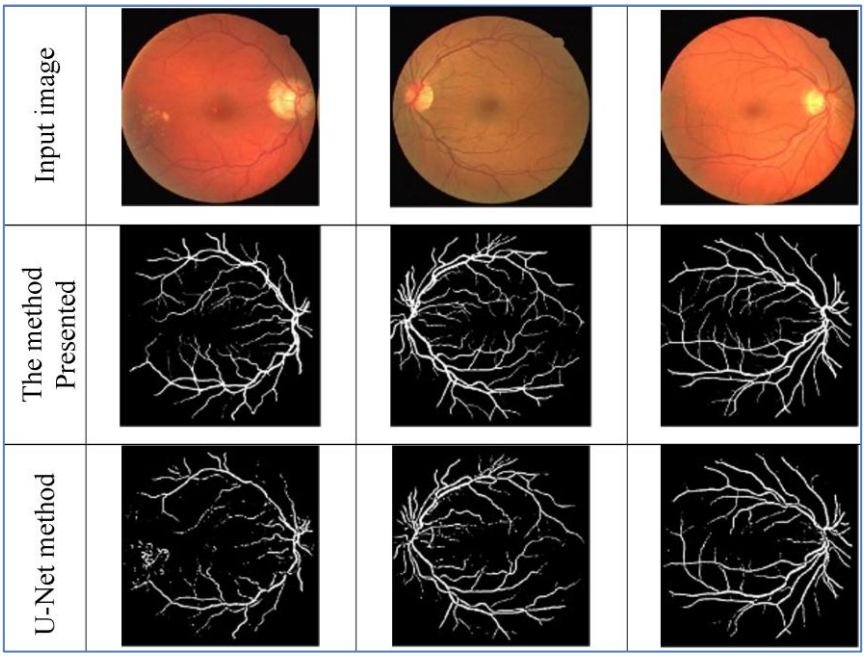

FIG.9. THE OUTPUTS OF U-NET NETWORK.

3.6 Interpretation of Statistical Significance Accuracy

The significance tests that compare our technique to U-Net on both DRIVE and CHASEDB1 are listed in Table 3. On DRIVE, both accuracy ($t = -8.36$, $p = 7.24\times10^{-5}$) and sensitivity ($t = -9.42$, $p = 1.88\times10^{-5}$) demonstrate statistically significant improvements for the suggested model. On CHASEDB1, precision greatly favors our technique ($t = -13.42$, $p = 1.42\times10^{-6}$), which shows that it has improved. However, regarding CHASEDB1 sensitivity, U-Net has a significant advantage ($t = 43.99$, $p = 2.70\times10^{-10}$). Overall, the findings show that the suggested method makes most measures better in a statistically meaningful way, except for sensitivity on CHASEDB1.

TABLE III. INTERPRETATION OF STATISTICAL SIGNIFICANCE ACCURACY

Dataset	Metric	t-Statistic	p-Value	p-Value (sci)	Significant (α=0.05)	Favored Model
DRIVE	Accuracy	-8.36	7.24E-05	7.24E-05	Yes	Proposed
DRIVE	Sensitivity	-9.42	1.88E-05	1.88E-05	Yes	Proposed
CHASEDB1	Precision	-13.42	1.42E-06	1.42E-06	Yes	Proposed
CHASEDB1	Sensitivity	43.99	2.70E-10	2.70E-10	Yes	U-Net

3.7 Discussion and Conclusions

This study introduces an innovative approach for segmenting retinal pictures, emphasizing the advantages of deep learning frameworks. The method relies mainly on neural networks, particularly the U-Net architecture, which has emerged as a prevalent model in medical image segmentation owing to its encoder-decoder framework and capacity to capture global context and intricate structural details. Processing retinal pictures is intrinsically problematic due to significant contrast, lighting, and vessel thickness fluctuation. Furthermore, the scarcity of annotated retinal image datasets constrains the efficacy of solely data-driven models, necessitating the development of methodologies that improve feature extraction and diminish reliance on extensive training datasets. This thesis presents a straightforward yet efficient iterative approach that integrates conventional image processing with deep neural networks to tackle these difficulties. The suggested solution starts by using a rotating guided filter as a preprocessing step. This filter improves edge information while also getting rid of background noise that is not needed. Guided filtering is vital for increasing segmentation quality since vessel borders are the most important parts of retinal pictures. Specifically, two filter settings are chosen: one to make fine, thin vessels stand out more and another to make thicker, more prominent vessels stand out more. By segregating vessel scales, the neural network can learn more robust and generalized properties across diverse vessel shapes. After that, the filtered images go into two U-Net networks, each processing one version of the pre-filtered data. This two-path technique lets the model get complementary information: one U-Net is good at finding thin vessels, and the other is good at keeping the integrity of larger vascular structures. Using a shared loss function, the outputs of these two networks are combined. The final cost is the average of the

loss functions based on entropy for both networks. This ensemble-like method makes it less likely that the model will overfit, makes it more likely to generalize, and makes the segmentation findings more stable across many datasets. After training, the suggested dual-path network was put through many tests on publicly available retinal datasets. The assessment utilized conventional medical imaging parameters, such as accuracy, sensitivity, specificity, and Area Under the Curve (AUC). Experimental results showed that the suggested method always outperformed the baseline U-Net model. In particular, it was better at finding small capillaries and big vessels, while keeping vessel connectivity and lowering false positives in noisy areas. Statistical significance testing indicated that the benefits were substantial, with p-values substantially below 0.05 in most cases where they were compared. The results show that using rotational guided filtering with U-Net architectures improves the accuracy of segmentation and offers a helpful way to deal with the problem of having too little training data. The method gets better results using domain knowledge (vessel structure priors) in the preprocessing step. This means that the model does not have to be very complicated or need massive training datasets. There are a lot of possible paths for future work. First, more preprocessing methods, like adaptive histogram equalization or wavelet-based denoising, could be added to the guided filter to improve the vessel's contrast. Second, the U-Net might use attention mechanisms and transformer-based encoders to help the network focus on the most valuable parts of the vessel. Third, it might be possible to carefully study how to optimize the characteristics of a neural network, like its depth, number of convolutional filters, and skip connection techniques, to find the best balance between performance and processing efficiency. Lastly, to use the model in real-world clinical settings, it will need to be tested to see how well it works with varied imaging situations, fundus camera devices, and groups of patients. In general, the suggested method is a promising hybrid that combines standard image processing with deep learning. This work presents a practical and successful technique for retinal vascular segmentation by integrating edge-aware guided filtering with the robust feature extraction capabilities of U-Net, thereby facilitating automated diagnosis and monitoring of ocular and systemic disorders.

Conflicts of Interest

Author declare no conflicts of interest.

Funding

Author, declare they have received no funding for this paper.

Acknowledgment

Non.

References

[1] M. R. K. Mookiah, S. Hogg, T. J. MacGillivray, V. Prathiba, R. Pradeepa, V. Mohan, et al., “A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification,” *Medical Image Analysis*, vol. 68, p. 101905, 2021.

[2] A. Galdran, A. Anjos, J. Dolz, H. Chakor, H. Lombaert, and I. B. Ayed, “State-of-the-art retinal vessel segmentation with minimalistic models,” *Scientific Reports*, vol. 12, no. 1, Art. no. 6174, 2022.

[3] A. Khandouzi, A. Ariafar, Z. Mashayekhpour, M. Pazira, and Y. Baleghi, “Retinal vessel segmentation, a review of classic and deep methods,” *Annals of Biomedical Engineering*, vol. 50, no. 10, pp. 1292–1314, 2022.

[4] M. E. Hoque and K. Kipli, “Deep learning in retinal image segmentation and feature extraction: A review,” *Int. J. Online Biomed. Eng. (iJOE)*, vol. 17, no. 14, 2021.

[5] N. R. Panda and A. K. Sahoo, “A detailed systematic review on retinal image segmentation methods,” *Journal of Digital Imaging*, vol. 35, no. 5, pp. 1250–1270, 2022.

[6] C. Chen, J. H. Chuah, R. Ali, and Y. Wang, “Retinal vessel segmentation using deep learning: A review,” *IEEE Access*, vol. 9, pp. 111985–112004, 2021.

[7] J. Cervantes, J. Cervantes, F. García-Lamont, A. Y. Rendon, J. E. Cabrera, and L. D. Jalili, “A comprehensive survey on segmentation techniques for retinal vessel segmentation,” *Neurocomputing*, Art. no. 126626, 2023.

[8] J. Ding, Z. Zhang, J. Tang, and F. Guo, “A multichannel deep neural network for retina vessel segmentation via a fusion mechanism,” *Frontiers in Bioengineering and Biotechnology*, vol. 9, p. 697915, 2021.

[9] Q. Zhang, X. Shen, L. Xu, and J. Jia, “Rolling guidance filter,” in *Computer Vision – ECCV 2014*, Zurich, Switzerland, 2014, pp. 815–830. Springer.

[10] K. Zhang, J. Deng, and W. Lu, “Segmenting human knee cartilage automatically from multi-contrast MR images using support vector machines and discriminative random fields,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2011.

[11] Z. Zhai, S. Feng, L. Yao, and P. Li, “Retinal vessel image segmentation algorithm based on encoder–decoder structure,” *Multimedia Tools and Applications*, vol. 81, no. 23, pp. 33361–33373, 2022.

[12] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, "ORIGA-light: An online retinal fundus image database for glaucoma analysis and research," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2010, pp. 3065–3068.
[13] 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, Sep. 2018.
[14] P. Yin, H. Cai, and Q. Wu, "DF-Net: Deep fusion network for multi-source vessel segmentation," *Information Fusion*, vol. 78, pp. 199–208, 2022.
[15] Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, "NFN+: A novel network followed network for retinal vessel segmentation," *Neural Networks*, vol. 126, pp. 153–162, 2020.
[16] W. Zhang, H. Liu, and L. Lin, "Full-scale representation guided network for retinal vessel segmentation (FSG-Net)," *arXiv preprint arXiv:2501.18921*, 2025.
[17] Y. Li, T. Wang, and Z. Huang, "Retinal vessel segmentation using optimal anisotropic guided filtering (OAGF) and deep learning fusion," *Biomedical Signal Processing and Control*, vol. 90, p. 105008, 2024.
[18] R. Ahmed, N. Kumar, and J. Kim, "MAFE-Net: Multiple attention-guided fusion ensemble for retinal vessel segmentation," *Computers in Biology and Medicine*, vol. 169, p. 107618, 2024.