AVE Trends in Intelligent Health Letters

U-Net Driven Deep Learning Segmentation for Retinal Vasculature for Early Disease Detection

D. Sumathi[1,*], Kuruva Sree Sowmya[2], Dhivya Kalaiselvan[3], R. Kesavamoorthy[4], Madhan Raj Gopi Akila[5], Jaime Alfonso Flores Navas[6]

[1,2,3,4]Department of Computer Science and Engineering, CMR Institute of Technology, Bengaluru, Karnataka, India.
[5]Department of Computer Science, University of Stuttgart, Stuttgart, Baden-Württemberg, Germany.
[6]Faculty of Sciences, National Autonomous University of Mexico, Coyoacán, Mexico.
d.sumathiphd@gmail.com[1], appskuruva@gmail.com[2], dhivyaakalaiselvan@gmail.com[3], kesavamoorthy.r@cmrit.ac.in[4], st199709@stud.uni-stuttgart.de[5], jaguarnegro175@gmail.com[6]

*Corresponding author

Abstract: To develop a robust and effective design, the proposed research uses a model that analyses blood vessels in retinal images by leveraging Deep Learning and Image Processing techniques. Deep learning, one of the AI techniques, has experienced rapid growth and is widely used in the medical field to analyse medical images, driving advancements in technology and medicine. The U-Net semantic segmentation neural network is the primary deep learning model used in our study. Training data consists of a set of retinal images, and the model's performance is assessed on the training set. Once testing is complete, the results help evaluate the accuracy and effectiveness of the model and the techniques utilised in this research. This study indicates that deep learning segmentation techniques can offer valuable insights beyond standard pixel-based detection techniques. The proposed method gives a better understanding of retinal blood vessel analysis. Researchers used the DRIVE dataset for this research. It has 20,512x512 training and testing images for the model's training and testing. The neural network's blood vessel shape recognition helps diagnose eye diseases. Diseases include hypertension, diabetic retinopathy, etc. The neural network's effectiveness makes the system suited for real-time blood vessel detection. This paper accurately mapped major and minor blood arteries. Thus, 96.53% of this research succeeded. The mean intersection over union is 8184. This exceeds the previous E-NET model of 7744.

Keywords: Blood Vessel Segmentation; Deep Learning (DL); Diagnosis of Retinal Diseases; Digital Image Recognition; Retinal Vasculature Analysis; Retinal Vessel Segmentation.

Cite as: D. Sumathi, K. S. Sowmya, D. Kalaiselvan, R. Kesavamoorthy, M. R. G. Akila, and J. A. F. Navas, "U-Net Driven Deep Learning Segmentation for Retinal Vasculature for Early Disease Detection," *AVE Trends in Intelligent Health Letters*, vol. 3, no. 1, pp. 17–26, 2026.

Journal Homepage: https://avepubs.com/user/journals/details/ATIHL

Received on: 19/01/2025, **Revised on:** 11/05/2025, **Accepted on:** 20/08/2025, **Published on:** 05/01/2026

DOI: https://doi.org/10.64091/ATIHL.2026.000262

1. Introduction

The Retinal blood vessel segmentation is an important task in medical image analysis, as the structure and shape of retinal blood vessels provide key diagnostic signals for a range of eye and systemic diseases, including diabetic retinopathy, hypertension, glaucoma, and various forms of age-related macular degeneration. The retina is considered an important part of

the eye, a sensitive tissue with a network of blood vessels that reflects local eye and broader cardiovascular issues. For early disease detection, analysing retinal blood vessels is important for tracking disease progression and planning treatments accordingly. Usually, experienced ophthalmologists manually study the retinal vessels in fundus images. Over time, this method is time-consuming, requires significant effort, and may vary across observers. Usually, pixel-level thresholds and standard image processing methods are difficult for identifying thin vessels and complex branching patterns, particularly in noisy, unevenly lit images. As there is a growing demand for automated, definite, non-invasive retinal vessel segmentation methods to help clinicians detect disease early, these objections lower their reliability and their use in large-scale screening. To address these challenges, a modified U-Net is designed to train semantic segmentation models for retinal blood vessels [1].

Over the past few years, CNNs have been powerful tools for correlating medical images. In the digital era, medical data volumes are growing rapidly, so disease identification and prediction are becoming more efficient with deep learning techniques [2]. Semantic Segmentation, a method within CNNs, is more focused and assigns every pixel in an image to a specific task. For accurate clinical interpretation, this detailed classification is important for analysing retinal blood vessels, including both major vessels and thin capillaries. U-Net is a popular CNN model for image segmentation and uses an encoder-decoder architecture. The encoder extracts high-level contextual features through a series of convolutional down-sampling operations. The decoder improves spatial resolution using up-sampling layers. This design makes the U-Net mainly suitable for medical imaging tasks. The main objective of this research is to build and evaluate a U-Net-based deep learning framework for precise segmentation of retinal blood vessels from fundus images. The proposed method aims to address challenges in thin vessel detection, class imbalance between vessel and background pixels, and loss of spatial information during feature extraction. This study enhances segmentation accuracy and calculates how well the proposed model captures detailed vascular structures, with the use of standard retinal datasets and evaluation metrics The main aim of this research is to know for example, that diabetes may be the cause of retinal hemorrhages might aid in the early diagnosis and research of these conditions, providing high-quality care and attention much earlier and more successfully [4].

2. Related Works

Currently, Artificial Intelligence is dominating medical image processing. Recent research by Khan et al. [11] explains how deep learning algorithms can solve medical problems. The customised CNN model is used for medical diagnosis and the effective prediction of required information in health care applications [17]. The GAN technique has a multi-scale generative adversarial network to improve vessel continuity. These networks reduce noise and improve accuracy, especially in complex training and computationally intensive tasks. Retinal imaging is a tool for diagnosing various diseases [16]. The fundamental steps towards computerised retinal image search are vessel segmentation and artery/vein classification, which provide distinct information on susceptibilities [3]. For retinal image analysis, performance improves when a U-Net-based model is combined with a GAN. The medical datasets are very limited and challenging to improve models, especially for retinal vessel segmentation. Uysal et al. [6] demonstrated that aggressive augmentation strategies improved model performance. The adem optimiser techniques, adapted for fast convergence in Guo et al. [10], yield improved analysis results and scalability.

The feature aggregation and the full-resolution processing are implemented in Liu et al. [7], which provides better preservation of vessel boundaries [15]. G-CASCADE, Rahman and Marculescu [8] was introduced as a graph convolutional decoding framework for modelling long-range dependencies, suitable for 2D medical images, achieving reduced computational cost and high performance. The ET-Net (Edge-aware architecture) by Zhang et al. [14] enhanced boundary delineation, helping detect narrow retinal vessels. A full-resolution network proposed in Liu et al. [7] explores the full-resolution processing and feature aggregation. It improves image segmentation accuracy. The U-NET architecture, proposed in Ronneberger et al. [9], was adapted for biomedical segmentation of medical imaging. To preserve special information, the encoder-decoder and skip structures are used. The SA-UNET in Guo et al. [5] provides an enhanced feature representation for vessels with low contrast and small size. The MA-UNET proposed in Cai and Wang [12] explains multi-scale feature extraction, and attention methods demonstrate enhanced segmentation results. This method uses a diverse medical imaging dataset and provides better results by reducing false negatives in complex vascular regions of images.

3. Proposed Methodology

The U-Net CNN explains that the supplied code describes the architecture for semantic blood vessel segmentation in medical analysis. The functions and processes of each code block are described as follows:

3.1. Dataset and Images

The DRIVE dataset has 20 training and 20 testing images, each of size 512x512. A mask of the same number of images is used to convert it into 80 training images for data preprocessing, performing 3 different actions: horizontal flip, vertical flip, and 45-

degree rotation. Additionally, this dataset provides more information on retinal blood vessels, which supports the study's results (Figure 1).

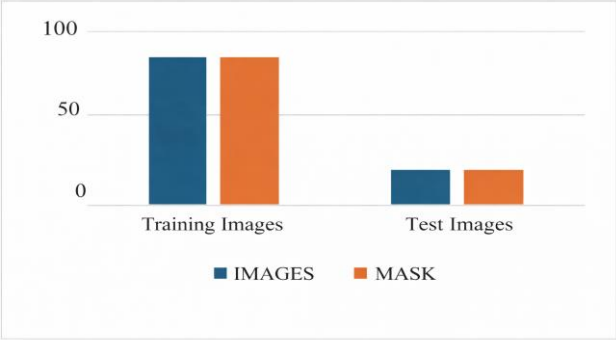

Figure 1: Image dataset

3.2. Division Based on Semantics

The goal of semantic segmentation is to assign each pixel in an image to a specific category. In this instance, researchers wish to categorise each pixel as either a background or a blood vessel (foreground). U-Net is a powerful deep learning technique that effectively extracts both high-level and low-level information from images [8].

3.3. Architecture of U-Net

The UNET architecture employs both encoding and decoding. Convolutional layers are applied in the encoding phase to extract features and downsample the image. In contrast, deconvolutional layers are used in the decoding phase to upsample the features and create the segmentation map (Figure 2). To improve contextual information, the model includes an additional feature aggregation module that merges multiscale feature maps [9].

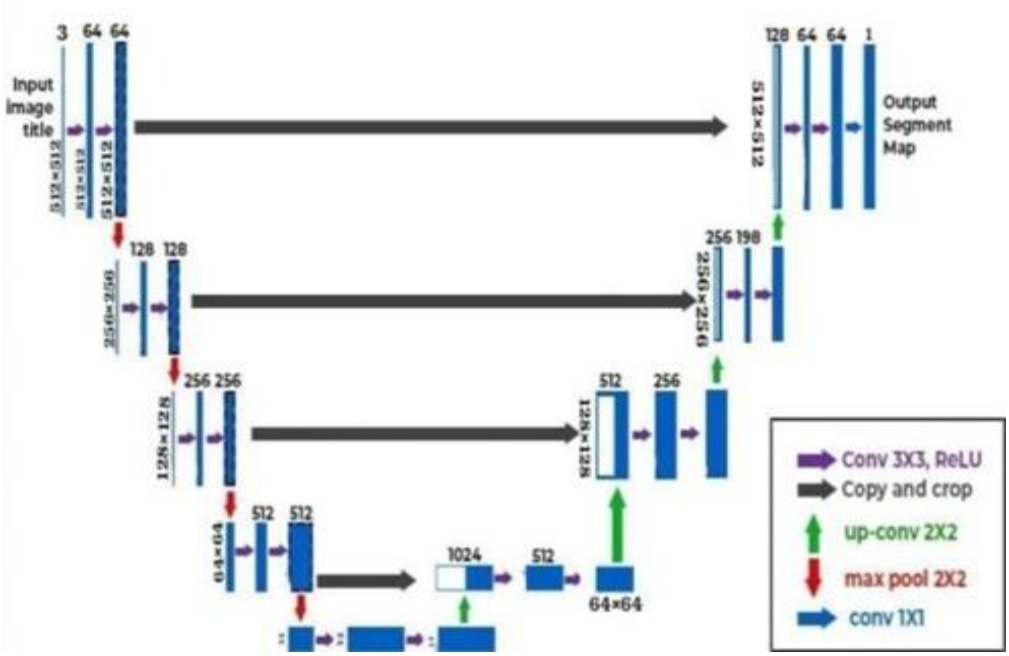

Figure 2: U-Net architecture

This shows that the model can leverage both local and global information for accurate segmentation. The code lists the components of a U-Net (Table 1).

Table 1: Segmentation model comparison

Segmentation Model	Our Model	Unet Model
Convolutional Layer	25	23
Training Parameters	35,661,633	31,402,501

3.3.1. Convolution Block

This block executes a ReLU activation function, batch normalisation, and three successive convolutional processes. $X \in R^{H\times W\times C}$ be the input feature map, * denote the convolution operation, $W$and b denote learnable weights and bias, $\sigma(\cdot)$denote the ReLU activation, and BN $(\cdot)$ denote Batch Normalisation. Given an input feature map X, it is given in equation (1):

$$F_1 = \sigma\big(BN(W_1 * X + b_1)\big)\, F_2 = \sigma\big(BN(W_2 * F_1 + b_2)\big)\, F_3 = \sigma\big(BN(W_3 * F_2 + b_3)\big)\, F_{conv} = F_3 \quad (1)$$

Where:

- $W_i \in R^{k\times k\times C_i\times C_{i+1}}$
- Kernel size $k = 3$
- Padding preserves spatial dimensions.

This facilitates local feature extraction from the image.

3.3.2. Encoder Block

This block combines a conv block with a max pooling operation. The convolution block harvests features, and the encoding begins with three input channels and a 64-bit input size. Max pooling is applied at 512, reducing the image size while preserving the dominant features. The encoder repeats this process to extract features at different resolutions. Let X_eThe input to the encoder block is given in Equation (2):

$$F_e = ConvBlock(X_e) P_e = MaxPool(F_e) E = (P_e) \quad (2)$$

Where:

- $F_e$is forwarded to the decoder via skip connections
- $P_e \in R^{\frac{H}{2}\times\frac{W}{2}\times 2C}$

The encoder gradually identifies high-level semantic features while simultaneously reducing spatial resolution.

3.3.3. The Layer of the Bottle Neck

Conv-block (512,1024) does this, after which the features are passed to the decoder block. The deepest level of the U-Net is given in equation (3):

$$F_b = ConvBlock(P_{e4}) F_b \in R^{\frac{H}{16}\times\frac{W}{16}\times 1024} \quad (3)$$

Conserves the global contextual information before decoding.

3.3.4. Decoder Block

This block concatenates the image with the relevant feature map from the encoder after up-sampling the photograph to improve its resolution. This guarantees that spatial information is preserved while the decoder recovers pooling. Another conv block refines the features. The decoder iterates through this procedure to gradually produce a high-resolution segmentation map. This hands over the extracted features to the classifier. Each decoder block performs three operations. The first is upsampling with transposed convolution, the second is concatenation with encoder feature maps, and the last is feature refinement with a convolutional block. Let:

- $F_d$be the decoder input

- F_e be the corresponding encoder feature map

Then the up-sampling can be like equation (4):

$$U_d = W_u \uparrow F_d + b_u \quad (4)$$

Skip connection can be like equation (5):

$$C_d = U_d \oplus F_e \quad (5)$$

Then the feature Refinement would be as per equation (6):

$$F'_d = ConvBlock(C_d) F'_d \in R^{H \times W \times C} \quad (6)$$

3.4. Execution of Code

The build UNET class defines the general U-Net architecture. The feature map depth is gradually increased by building an encoder with four encoder blocks. Incorporating a convolution block into a bottleneck block preserves high-level information. The decoder incorporates four blocks that successively reduce depth and increase resolution via skip connections [10]. The final forecast is earned using a 1×1 convolution proceeded by a sigmoid activation given in the equation below (7):

$$\hat{Y} = \sigma(W_o * F_{d4} + b_o) \hat{Y} \in [0,1]^{H \times W} \quad (7)$$

Where:

- $\hat{Y}$ represents vessel probability per pixel

3.5. Training the Data

For formatting the data for the built model, the Adam Optimiser is used with the torch module in Python, and seeding is used to randomise the input. The model was trained with a batch size of 2 for 50 epochs to achieve optimal results (Table 2). The trained model at a learning rate, lr=lr-4. The learning rate is reduced to 25% of its prior value after each epoch, and the model employs both Binary Cross Entropy and Dice Binary Cross Entropy loss to enhance performance [11].

Table 2: Layers of model

Layer Name	Layer Type	Output Shape	Number of Param	Connected To
e1	Encoder Block	[1, 64, 256, 256]	12,352	inputs
e1.conv	Convolution Block	[1, 64, 256, 256]	18,496	e1.inputs
e1.pool	MaxPool2d	[1, 64, 128, 128]	0	e1.conv
e2	Encoder Block	[1, 128, 128, 128]	249,088	e1.pool
e2.conv	Convolution Block	[1, 128, 128, 128]	271,488	e2.inputs
e2.pool	MaxPool2d	[1, 128, 64, 64]	0	e2.conv
e3	Encoder Block	[1, 256, 64, 64]	993,984	e2.pool
e3.conv	Convolution Block	[1, 256, 64, 64]	1,101,056	e3.inputs
e3.pool	MaxPool2d	[1, 256, 32, 32]	0	e3.conv
e4	Encoder Block	[1, 512, 32, 32]	3,937,280	e3.pool
e4.conv	Convolution Block	[1, 512, 32, 32]	4,196,352	e4.inputs
B	Constriction of the Bottle Neck	[1, 1024, 32, 32]	8,393,856	b.inputs
d1	Decoder Block	[1, 512, 64, 64]	4,196,352	b.outputs
d1.up	ConvTranspose2d	[1, 512, 64, 64]	2,097,152	d1.inputs
d1.conv	Convolution Block	[1, 512, 64, 64]	5,120,448	d1.up, d1.skip
d2	Decoder Block	[1, 256, 128, 128]	1,282,432	d1.outputs
d2.up	ConvTranspose2d	[1, 256, 128, 128]	524,288	d2.inputs
d2.conv	Convolution Block	[1, 256, 128, 128]	1,761,728	d2.up, d2.skip
d3	Decoder Block	[1, 128, 256, 256]	344,064	d2.outputs
d3.up	ConvTranspose2d	[1, 128, 256, 256]	131,072	d3.inputs
d3.conv	Convolution Block	[1, 128, 256, 256]	589,824	d3.up, d3.skip

d4	Decoder Block	[1, 64, 512, 512]	92,160	d3.outputs
d4.up	ConvTranspose2d	[1, 64, 512, 512]	32,768	d4.inputs
d4.conv	Convolution Block	[1, 64, 512,512]	249,856	d4.up, d4.skip
Output	Conv2d	[1, 1, 512, 512]	65,537	d4.outputs

3.6. Foundation Data Set for Retinal Vasculature

The DRIVE database, obtained from the Dutch Diabetic Retinopathy Screening Initiative, is a key benchmark for studies of retinal vasculature segmentation. It comprises 40 colour fundus images (768 × 584 pixels) acquired using a Canon CR5 non-mydriatic 3CCD camera with a 45° field of view. The dataset includes 33 normal cases and 7 showing mild diabetic retinopathy, divided equally into training and test sets. A comparable field-of-view mask is applied to each image, and the vascular structures are directly segmented—once for training and twice for testing (one as the gold standard). For retinal vessel detection, this dataset supports applications such as disease screening, diagnosis of diabetic and hypertensive retinopathy, vascular morphology analysis, image registration, and biometric identification, thereby promoting the development and evaluation of automated algorithms. For detailed segmentation of retinal vessels, the FIVES data set is used, comprising 800 high-resolution colour fundus images with great pixel-level detail for manual annotation. This data set consists of age data from 4 to 83 years, spanning a wide demographic range, and includes various ocular pathologies, including diabetic retinopathy, glaucoma, and age-related macular degeneration. FIVES is a significant openly accessible dataset for segmenting retinal vessels, providing a crucial resource for developing and evaluating AI models for automated retinal vasculature analysis.

4. Results

The proposed U–Net–based retinal vessel segmentation model delivers high–quality results. Figure 3 shows the model's proficiency at distinguishing large and small retinal blood vessels (eye nerves) from fundus images. From the segmented output, it extracts continuous vessels, along with branching patterns and thin capillaries, which are repeatedly difficult to detect.

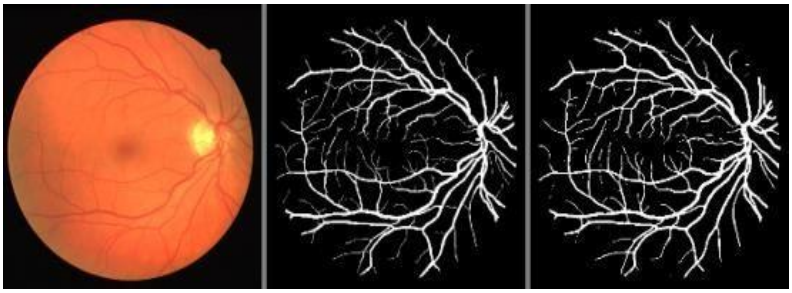

Figure 3: Retinal eye nerves image

The proposed architecture completely determines the vessel structures from the background by leveraging hierarchical side extraction and connections, thereby minimising loss of spatial information during downsampling. Researchers observe that the segmentation output coordinates closely match the ground truth annotations, affirming the model's robustness in capturing complex vascular morphology. The proposed system can be quantified with standard metrics and performance in a measurable format. The metrics are assessed using standard pixel-level segmentation, comparing the predicted segmentation mask to the actual values. Using a confusion matrix, researchers can obtain binary segmentation, where each pixel belongs to one of two classes: either the vessel (foreground) or the background.

- TP – True Positives (Vessel Pixels Correctly Detected)
- TN – True Negatives (Background Pixels Correctly Detected)
- FP – False Positives (Background Pixels Misclassified as Vessel)
- FN – False Negatives (Vessel Pixels Missed)

The spatial overlap is calculated between ground-truth masks and predicted values with the given equation (8) of Dice Similarity Coefficient (DSC):

$$Dice = \frac{2TP}{2TP+FP+FN} \quad (8)$$

Overlap ratio between ground truth and prediction is evaluated by equation (9), Intersection over Union (IoU):

$$I = A \cap B, U = A \cup BIoU = \frac{I}{U+\epsilon} \quad (9)$$

Mean Intersection over Union (mIoU) given in equation (10), for C classes, the mean IoU is computed, which is given in equation (10):

$$mIoU = \frac{1}{C}\sum_{c=1}^{C} \quad IoU(c) \quad (10)$$

Where $\epsilon$is a small constant to prevent division by zero. The system can detect most of the retinal nerve with a precision of 0.7909, while our model attained an accuracy of 0.9653. The sensitivity is 0.8255, the F1 score is 0.8052, and the dice coefficient is 0.6742. This uses a graphic illustration to illustrate the eye's nerves. It provides an accurate image of the nerve and makes it easy to identify the eye's blood vessels. Our model's F1 score is higher than the DR_2021 model's, which is 0.75. Researchers also outperform them in terms of accuracy. Additionally, our sensitivity for retinal segmentation is higher than that of the earlier UNET model. Researchers have compared the model in Table 3.

Table 3: Quantitative performance summary

Metric	Value
Accuracy	0.9653
Precision	0.7909
Sensitivity	0.8255
F1-Score	0.8052
Dice Coefficient	0.6742
mIoU	0.8184

The function returns the mean IoU as a single float value. The Precision–Recall (PR) illustrates the trade-off between precision and recall (sensitivity) of the proposed U-Net–based retinal vessel segmentation model across varying decision thresholds (Figure 4).

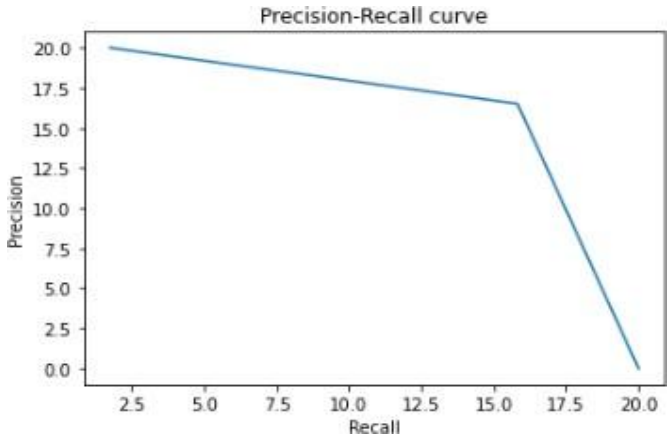

Figure 4: Precision–Recall (pr) curve for retinal vessel segmentation

Binary Cross-Entropy (BCE) and Dice Loss, combined as a loss function, are used for training the model, confirming pixel-wise accuracy and region consistency:

$$L_{total} = L_{BCE} + L_{Dice}$$

Figure 5 confirms the stable junction by the decreasing trend in training and validation loss over 50 epochs, without overfitting or effective generalisation.

For evaluating imbalanced datasets, the precision-recall curve provides a means to balance them. Also, to achieve robustness, particularly in handling class imbalance, vessel pixels occupy a smaller portion of the image than background pixels. In the medical field of image delineation, this behaviour is important, as high recall confirms minimal vessel omission, whereas high precision prevents false findings that may affect clinical understanding.

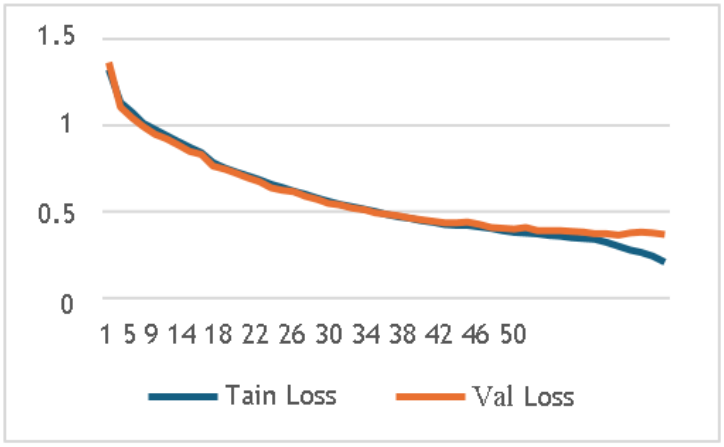

Figure 5: Graph of training and validation loss

The proposed segmentation model, using the DRIVE dataset, is shown in the ROC curve below (Figure 6). The curve remains near the upper-left corner of the ROC across a range of thresholds, indicating a high true positive rate and a low false positive rate. This verifies the robustness by identifying retinal blood vessels from background regions.

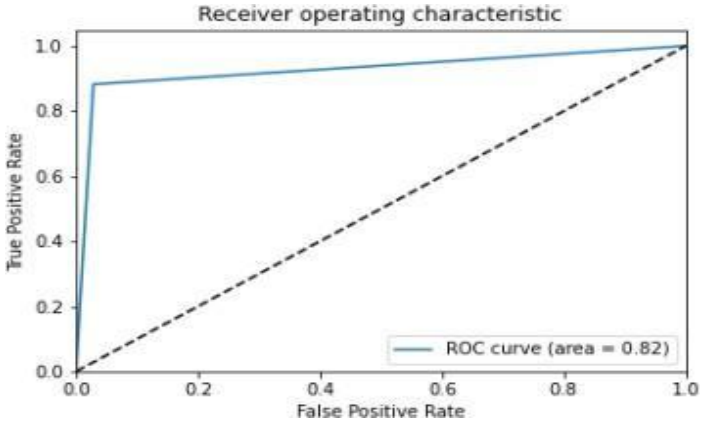

Figure 6: Receiver operating characteristic (ROC) curve

Regarding the suggested model, it achieves a high ROC–AUC score, enabling it to effectively detect retinal blood vessels with minimal false alarms, making it suitable for initial screening of retinal conditions such as hypertension and diabetic retinopathy. The quantitative evaluation of the projected UNet–based retinal vessel segmentation model against several state-of-the-art methods is shown in Table 4, including DR_2021, E-NET, MERIT-GCASCADE, and PVT-GCASCADE. Using the standard evaluation metrics, the comparison is based on Intersection over Accuracy, F1-Score, and Sensitivity.

Table 4: Results comparison

Model	Mi Ou	Accuracy	F1-Score	Sensitivity
OUR MODEL	0.8184	0.9653	0.8052	0.8255
DR_2021 [13]	-	0.9593	0.75	0.7119
E- NET [14]	0.774	0.956	-	-
MERIT- GCASCA DE [8]	0.7081	0.970	0.82	0.828
PVT- GCASCA DE [8]	0.6970	0.9689	0.82	0.83

The presented paradigm achieves an F1 score of 0.8052, surpassing that of DR_2021 (0.75) and indicating a better balance between precision and sensitivity. This model achieves a sensitivity of 0.8255, corroborating its validity in recognising retinal blood vessels and thin low-contrast shapes that are frequently overlooked by conventional methods. This U-Net-based architecture constantly surpasses or remains competitive. Deep learning's knack for retinal segmentation improves reliability and support in preliminary clinical assessments and enables early disease identification, thereby enhancing treatment strategies.

5. Conclusion

Studies in medical image analysis repeatedly show that the U-Net architecture and its many variants are the most effective for accurately segmenting retinal blood vessels. These deep learning models are designed to capture both the big picture and the small details, both of which are necessary for accurate vascular segmentation. Architectures like U-Net use encoder-decoder structures, and when combined with advanced techniques such as skip connections, they facilitate the propagation of spatial information across layers. This makes it easier to maintain features and improves segmentation accuracy. Adding components such as batch normalization also helps stabilize and speed up the training process. Spatial attention modules, on the other hand, let the network focus on the most important parts of retinal images. These improvements not only make segmentation more accurate but also more efficient, making the models more useful in the real world. As a result, the system can now identify specific blood vessels and obtain more accurate biomarker measurements, leading to better diagnostic results. Using U-Net variants to separate vessels can be achieved with remarkable detail and consistency. The suggested U-Net-based methods with additional branching mechanisms perform better than typical algorithms when handling complex vessel structures. This progress is especially important for finding and screening for common retinal illnesses, including diabetic retinopathy, hypertension, and glaucoma, early on. Better segmentation quality ensures that minor changes in pathology are not missed. These improvements demonstrate the importance of U-Net-based models in real-world applications. Ongoing advancements in segmentation methodologies are crucial to the evolution of reliable, automated retinal screening systems, ultimately facilitating early diagnosis and enhanced patient care.

Acknowledgement: N/A

Data Availability Statement: Relevant data supporting this research can be shared by the corresponding author upon request, ensuring openness and reproducibility.

Funding Statement: No external funding or sponsorship was received by the authors for conducting this research or preparing the manuscript.

Conflicts of Interest Statement: The authors affirm that there are no competing interests that could have influenced the outcomes or interpretations of this study.

Ethics and Consent Statement: The study complies with ethical research standards, and all authors have provided their consent for publication and dissemination of the work.

References

1. L. Luo, D. Chen, and D. Xue, "Retinal blood vessels semantic segmentation method based on modified U-Net," in *Proc. 2018 Chinese Control and Decision Conference (CCDC)*, Shenyang, China, 2018.
2. U. Niyaz and A. S. Sambyal, "Advances in deep learning techniques for medical image analysis," in *Proc. 2018 Fifth Int. Conf. on Parallel, Distributed and Grid Computing (PDGC)*, Solan, India, 2018.
3. L. Li, M. Verma, Y. Nakashima, R. Kawasaki, and H. Nagahara, "Joint learning of vessel segmentation and artery/vein classification with post-processing," in *Proc. Medical Imaging with Deep Learning (MIDL), ser. Proc. Machine Learning Research*, vol. 121, no. 7, pp. 440-453, 2020.
4. O. Gorgulu and A. Akilli, "Use of fuzzy logic based decision support systems in medicine," *Ethno Med*, vol. 10, no. 4, 393-403, 2016.
5. C. Guo, M. Szemenyei, Y. Yi, W. Wang, B. Chen, and C. Fan, "SA-UNet: Spatial attention U-Net for retinal vessel segmentation," in *Proc. 2020 25th Int. Conf. on Pattern Recognition (ICPR)*, Milan, Italy, 2021.
6. E. S. Uysal, M. Ş. Bilici, B. S. Zaza, M. Y. Özgenç, and O. Boyar, "Exploring the limits of data augmentation for retinal vessel segmentation," *arXiv preprint*, 2021. [Accessed by 18/11/2024].
7. W. Liu, H. Yang, T. Tian, Z. Cao, X. Pan, W. Xu, Y. Jin, and F. Gao, "Full-resolution network and dual-threshold iteration for retinal vessel and coronary angiograph segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 9, pp. 4623–4634, 2022.
8. M. M. Rahman and R. Marculescu, "G-Cascade: Efficient cascaded graph convolutional decoding for 2D medical image segmentation," in *Proc. IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV)*, Waikoloa, Hawaii, United States of America, 2024.
9. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Cham, Switzerland, 2015.

10. Z. Guo, Y. Xu, W. Yin, R. Jin, and T. Yang, "A novel convergence analysis for algorithms of the Adam family," *arXiv preprint*, 2021. [Accessed by 28/12/2024].
11. A. H. Khan, A. Khan, D. N. F. A. Iskandar, H. Mewada, S. Saeed, F. Algarni, F. Ullah, M. A. Khan, N. Iqbal, and A. A. Hussain, "BrainNet: A custom-designed CNN and transfer learning-based models for diagnosing brain tumors from MRI images," *PeerJ Computer Science*, vol. 11, no. 10, p. e3154, 2025.
12. Y. Cai and Y. Wang, "MA-UNet: An improved version of UNet based on multi-scale and attention mechanism for medical image segmentation," in *Proc. Third Int. Conf. on Electronics and Communication; Network and Computer Technology (ECNCT)*, Harbin, China, 2022.
13. S. Basu, S. Mukherjee, A. Bhattacharya, and A. Sen, "Segmentation of blood vessels, optic disc localization, detection of exudates, and diabetic retinopathy diagnosis from digital fundus images," in *Proc. Research and Applications in Artificial Intelligence (RAAI)*, Singapore, 2021.
14. Z. Zhang, H. Fu, H. Dai, J. Shen, Y. Pang, and L. Shao, "ET-Net: A generic edge-attention guidance network for medical image segmentation," in *Proc. Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Cham, Switzerland, 2019.
15. S. A. Kamran, K. F. Hossain, A. Tavakkoli, S. L. Zuckerbrod, K. M. Sanders, and S. A. Baker, "RV-GAN: Segmenting retinal vascular structure in fundus photographs using a novel multi-scale generative adversarial network," in *Proc. Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Cham, Switzerland, 2021.
16. M. Senbagavalli, R. C. Karpagalakshmi, D. Sumathi, J. Lenin, G. R. K. Prasad, and A. Manikandan, "A novel heart disease monitoring and prediction using machine learning algorithm," in *Proc. Int. Conf. on Internet of Things and Connected Technologies*, Singapore, 2023.
17. K. Ćosić, S. Popović, M. Šarlija, I. Kesedžić, M. Gambiraža, B. Dropuljić, I. Mijić, N. Henigsberg, and T. Jovanovic, "AI-Based Prediction and Prevention of Psychological and Behavioral Changes in Ex-COVID-19 Patients," *Front Psychol.*, vol. 12, no. 12, p. 782866, 2021.