Indonesian Journal of Electrical Engineering and Computer Science
Vol. 43, No. 1, July 2026, pp. 157~178
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v43.i1.pp157-178

HawkNet: an intelligent bio-inspired optimization based patch wise adaptive U-Net framework for retinal blood vessel segmentation in fundus images

Saba Sheiba[1,2], M Neelakantappa[3]
[1]Department of Computer Science and Engineering, BEST Innovation University, Gorantla, India
[2]Department of CS and AI, Muffakham Jah College of Engineering and Technology, Hyderabad, India
[3]Department of Information Technology, Vasavi College of Engineering, Hyderabad, India

Article Info

Article history:

Received Feb 18, 2026
Revised Jun 23, 2026
Accepted Jun 29, 2026

Keywords:

Deep learning
Fundus images
Harris Hawk optimization
Medical image analysis
Patch-Wise U-Net
Retinal blood vessel segmentation

ABSTRACT

The accurate segmentation of retinal blood vessels plays a pivotal role in the early diagnosis and monitoring of health issues like diabetes, high blood pressure, and glaucoma. Traditional machine learning techniques such as matched filtering, morphological processing, and edge detection have a tough time to see tiny blood vessels clearly because eye images often have uneven lighting, poor contrast, and blood vessels that twist and turn in complex ways. To overcome these challenges, this work introduces an improved Patch-wise U-Net method with Harris Hawk optimization (HHO). The patch system split fundus images into overlapping patches, enabling the network to focus on localized vessel features such as fine capillaries, bifurcations, and vessel boundaries that are often missed in global segmentation. Meanwhile, HHO is employed to fine-tune the U-Net's hyperparameters and learning weights, achieving faster convergence and enhanced segmentation accuracy without manual tuning. To establish generalizability and reliability, the proposed framework is rigorously tested on two standard retinal image repositories DRIVE, and STARE. Experimental evaluation demonstrates significant improvement in key performance metrics, including accuracy of 97.8%, precision 95.8%, recall 96.5%, f1-score 97.3%, IoU of 96.2%, and dice coefficient (DC) 94.3%, highlighting the model's capability to capture thin and thick vessels structures while minimizing false detections. Additionally, qualitative segmentation further confirm that the proposed framework effectively preserves the continuity and morphology of thin and thick vessels, even in regions with low contrast or uneven illumination. Overall, the proposed method visual inspection has revealed that the suggested method can segment thin and thick vessels with greater accuracy than previous methods. It also demonstrates its potential for real-life clinical application.

Corresponding Author:

Saba Sheiba
Department of Computer Science and Engineering, BEST Innovation University
Gorantla, Andhra Pradesh, India
Email: saba.sheiba@mjcollege.ac.in

1. INTRODUCTION

The human retina serves as a non-invasive gateway to the circulatory and nervous systems, offering vital diagnostic signs for the early detection of systemic and ophthalmic diseases such as diabetic retinopathy, glaucoma, hypertension, and arteriosclerosis [1]. Retinal fundus imaging provides detailed visualization of the vascular structure, where subtle variations in vessel morphology, caliber, and branching patterns can indicate disease progression or neurological abnormalities. Automated retinal blood vessel segmentation therefore plays a fundamental role in clinical image analysis, as accurate extraction of both thin and thick vessels enables reliable quantification of pathological biomarkers. In recent years [2], the rapid evolution of deep learning has significantly transformed medical image segmentation, providing data-driven models that achieve remarkable accuracy in delineating anatomical structures [3], [4]. Despite these advancements, precise segmentation of retinal vessels remains a challenging problem due to the complex characteristics of fundus images [5]. The variability in vessel thickness, the low contrast between vessels and background tissue, uneven illumination, and the presence of bright lesions or optic disc regions frequently lead to misclassification and discontinuities in vessel maps.

Traditional approaches based on matched filtering, morphological transformations, and thresholding have struggled to generalize across different datasets because of their sensitivity to noise and illumination variation [6], [7]. Even with the introduction of convolutional neural networks (CNNs), capturing micro-vessel details and preserving vessel continuity remain difficult, as conventional architectures often lose spatial information during pooling operations and produce incomplete segmentation in areas with weak gradients or pathological disturbances [8]. To address these challenges, researchers have proposed a variety of deep learning models aimed at enhancing structural feature retention and improving generalization. Among them, the U-Net architecture has emerged as the predominant framework for medical image segmentation because of its symmetric encoder–decoder structure and skip connections that allow effective fusion of low- and high-level features [9]. Numerous U-Net variants, including residual U-Net, attention U-Net, and dense U-Net, have been designed to enhance feature propagation and vessel boundary clarity. Some approaches employ multi-scale convolutions or recurrent layers to better capture spatial dependencies, while others integrate post-processing filters to refine vessel connectivity. Although these modifications yield performance improvements, many U-Net-based models still rely on globally processed images, which can obscure fine-grained vessel patterns, particularly in small capillaries and bifurcations [10]. Additionally, their performance depends heavily on manual hyperparameter tuning, such as learning rate, convolutional kernel size, and number of filters per layer, which not only requires expert experience but also limits scalability across datasets. The absence of adaptive optimization in these models often leads to suboptimal convergence, overfitting, or poor generalization when applied to images captured under different lighting or resolution conditions.

Figure 1 represents general processes involved in segmentation of retinal blood vessels. In pursuit of higher precision and robustness, optimization-driven deep learning has become a prominent research direction. Metaheuristic algorithms inspired by natural and biological processes such as particle swarm optimization (PSO), genetic algorithms (GA), and grey wolf optimization (GWO) have demonstrated effectiveness in tuning model parameters for complex, nonlinear problems [11]. These algorithms are capable of balancing exploration and exploitation within the search space, allowing neural networks to reach near-optimal configurations with fewer manual interventions. However, many of these optimization techniques still face challenges such as premature convergence, local minima entrapment, or high computational costs when scaled to deep architectures. Thus, there is an insistent need for an intelligent and efficient optimization mechanism that can automatically fine-tune deep segmentation models to adapt to the intrinsic variability of retinal fundus images.

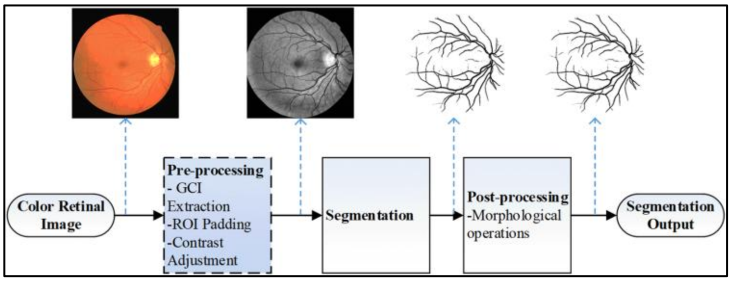

Figure 1. General processes involved in segmentation of retinal blood vessels from a fundus image

Although previous studies have explored patch-based U-Net architectures and optimization-assisted segmentation separately, the proposed HawkNet framework differs by establishing a synergistic interaction between localized patch-wise learning and Harris Hawk optimization (HHO). Unlike conventional approaches, HHO continuously adapts learning parameters to improve feature extraction and convergence, while patch-wise learning preserves thin vessels, capillaries, and bifurcation structures. This closed-loop optimization-learning mechanism enhances vessel continuity, robustness to illumination variations, and generalization across heterogeneous retinal datasets, distinguishing HawkNet from existing optimization-enhanced U-Net frameworks.

Motivated by these limitations, this study introduces HawkNet. The core idea behind HawkNet lies in combining a localized patch-wise learning strategy with a powerful bio-inspired optimization algorithm known as HHO. Instead of processing the entire fundus image globally, the proposed framework divides each image into overlapping patches, enabling the model to focus on local vessel features such as thin capillaries, bifurcations, and curved structures that are often missed in global training. This patch-wise U-Net approach not only enhances fine vessel detection but also ensures better handling of local contrast variations and illumination inconsistency. The HHO algorithm automatically adjusts key patch wise U-Net hyperparameters including learning rate, filter size, and weight initialization achieving faster convergence and improved accuracy without manual supervision. The hybridization of U-Net's hierarchical feature extraction with HHO's dynamic optimization enables the network to learn both global context and localized micro-vessel representations efficiently, reducing segmentation discontinuities and false detections.

- To develop an adaptive HawkNet architecture using patch-wise training and balanced encoder–decoder skip connections for accurate segmentation of fine and thick retinal vessels under challenging imaging conditions.
- To integrate HHO for automated hyperparameter and weight tuning, improving learning stability, segmentation precision, and reducing reliance on large annotated retinal datasets.
- To evaluate HawkNet's robustness and generalization ability across DRIVE, STARE, and CHASE_DB1 datasets, demonstrating superior quantitative performance and clinically consistent vessel segmentation compared to existing methods.

The remainder of this paper is structured as follows. Section 2 reviews existing research on retinal vessel segmentation. Section 3 presents the proposed HawkNet architecture, detailing its patch-wise learning strategy and HHO mechanism. Section 4 describes the experimental setup with quantitative and qualitative results, including comparative analyses with state-of-the-art techniques. Finally, Section 5 concludes the paper on optimization-enhanced retinal image segmentation.

Among the available bio-inspired optimization algorithms, HHO was selected because of its dynamic balance between exploration and exploitation, which enables effective avoidance of local minima and promotes faster convergence. Compared with GWO and whale optimization algorithm (WOA), HHO exhibits superior adaptability and search diversity, making it particularly suitable for optimizing deep learning hyperparameters in medical image segmentation tasks. These characteristics facilitate stable learning and improved preservation of fine vascular structures under varying illumination and contrast conditions.

Unlike existing optimization-assisted U-Net models, the proposed HawkNet framework combines patch-wise adaptive learning with HHO in a unified segmentation framework. The patch-wise strategy enables effective preservation of thin vessels, capillaries, and bifurcation structures by focusing on localized retinal regions, while HHO automatically optimizes key network hyperparameters and learning weights. This integration establishes a synergistic relationship between local feature learning and adaptive optimization, resulting in improved vessel continuity, robustness to illumination variations, and enhanced segmentation performance across heterogeneous retinal datasets.

2. RELATED WORK

A growing number of eye and systemic disorders, including diabetic retinopathy, glaucoma, and hypertension, are being better detected and tracked through the use of retinal blood vessel segmentation, which has become an important field of study in medical image analysis [12]. From basic deep learning frameworks to more complex methods like matched filtering and morphological transformations, there have been an array of approaches suggested for image processing in the last ten years. While traditional approaches offer computational simplicity, their performance is challenging under certain imaging conditions characterized by uneven illumination, low contrast, and pathological variations [13].

The task of automated retinal image segmentation can be divided into two categories, namely supervised learning methods and unsupervised learning methods [14]. Unsupervised methods [15] include rule-based approaches that may utilize a matched filter and morphological operations to detect and segment the vessels.

A summary of representative retinal blood vessel segmentation methods, datasets, performance characteristics, and key limitations, is presented in Table 1. The comparative analysis focuses on the challenges that are already present and encourages the development of the proposed HawkNet framework.

Table 1 summarizes the datasets, major advantages, and inherent limitations of representative retinal blood vessel segmentation methods. The comparison shows the advantages and disadvantages of the current frameworks, and highlights the importance of developing a more powerful optimization-based framework that can maintain the fidelity of fine vessel structures and enhance segmentation performance when imaging under different conditions.

Table 1. Comparative analysis of retinal blood vessel segmentation methods

Reference	Method	Dataset	Results	Limitations
Sau *et al.* [16]	Ridge-based segmentation	DRIVE	Reliable detection of major blood vessels	Poor thin vessel detection
Zhang *et al.* [17]	Gabor filter with classifier	DRIVE	Improved vessel contrast and edge clarity	Requires manual feature tuning
Liu *et al.* [18]	Line operator with SVM	DRIVE	Accurate vessel boundary localization	High computational complexity
Radha and Karuna [19]	Ensemble-based segmentation	DRIVE, STARE	Robust segmentation across varied retinal images	Limited cross-dataset generalization
Wang *et al.* [20]	Deep convolutional neural network	DRIVE, STARE	Effective automatic hierarchical feature extraction	Needs large labelled datasets
Ortiz-Feregrino *et al.* [21]	U-Net architecture	DRIVE	Accurate end-to-end retinal vessel segmentation	Struggles with thin vessels
Du *et al.* [22]	Dense U-Net	DRIVE, CHASE_DB1	Enhanced feature propagation and reuse	Increased model complexity
Gargari *et al.* [23]	Attention U-Net	DRIVE	Improved focus on vessel regions	Higher memory and computation requirements
Yi *et al.* [24]	Multi-scale CNN fusion	DRIVE, STARE	Better multi-scale vessel representation	Sensitive to illumination variations

Adopted multiscale inputs in U-Net was proposed by [25] with dense blocks, while [26] proposed residual connections with sequences of 3 × 3 convolutions in the skip and encoder/decoder paths, developing MultiResU-Net. A cascaded network of four units using dilated convolutions at varying rates was suggested by [27] as RefineNet. The final segmentation map was composed of the combined outputs of these units. A multiscale CNN with better cross-entropy loss obtained good sensitivity on vascular segmentation; [28] suggested this CNN; and [29] introduced hierarchical classification to combine findings of two classifiers for blood vessel segmentation. Because of their superior performance, U-Nets and fully convolutional networks (FCNs) were heavily used for dense predictions [30]. U-Net later had enhancement modules including attention mechanisms, dilated convolutions, and residual blocks added to it to make it even better [31]. More spatial information can be captured, local information loss can be reduced, and low-level feature maps can be reused for accurate segmentation in a well-designed model. Due to their deeper convolutional layers, U-Net [32] and multi-model networks [33] outperform CNNs and FCNs in segmentation outcomes. This is because these networks are able to extract richer information. It is possible to modify the segmentation results using matching filters or morphological transforms in postprocessing [34]. This is because the results frequently include noise and isolated tiny vessels. The ability of our composite U-Net Inception structure to gather both local and global traits allowed it to outperform the majority of state-of-the-art designs [35], [36].

Continuous advancements from traditional unsupervised approaches to supervised deep learning models have significantly enhanced segmentation accuracy and efficiency. Conventional methods such as adaptive thresholding, region growing, and morphological filtering provided essential groundwork but struggled with noise and fine vessel continuity. Deep learning models, particularly U-Net and its improved variants with skip and residual connections, have enabled superior performance by learning multi-scale and contextual features. Integrating optimization algorithms has further refined these models by accelerating convergence and improving feature extraction.

In order to build sustainable and efficient electrical energy networks, intelligent forecasting of electricity demand, energy management, and integration of renewable energy resources have become essential. Özüpak and Mansurov [37] have presented a hybrid deep learning system for the short-term prediction of multi-output electrical energy demand, highlighting the ability of multiple neural architectures to improve the accuracy of the electricity demand prediction and address the complex temporal dependencies among electricity demand patterns. A different study conducted by Özüpak [38] identified the factors that affect the effectiveness of the wireless power transmission systems for next generation electric vehicles, including the distance between the coils, the coil design and the operating frequency of the wireless transmission system. Investigated smart energy management and optimization techniques for sustainable power systems, highlighting the importance of intelligent control algorithms, demand response mechanisms, and renewable energy integration in enhancing system efficiency and reliability. Advanced control strategies for renewable energy integration were in which they highlighted the use of adaptive and predictive control methods for ensuring grid stability and power quality in modern energy systems. Furthermore, introduced a hybrid deep learning system for solar PV power forecasting and achieved higher forecasting accuracy and lower prediction errors, which is crucial for the energy scheduling and grid operation. In addition, the recent advancements in electrical energy systems and smart grid technologies, highlighting the significance of distributed generation, energy storage systems, intelligent monitoring, artificial intelligence, and Internet of Things technologies in the modernization of power systems. Overall, these studies suggest that the synergy of deep learning methods, optimization techniques, and sophisticated control strategies significantly contributes to improved energy efficiency, predictive accuracy, and grid stability. The advent of smart grid and intelligent energy management system offers great prospects to achieve sustainable power generation and distribution and support the growing integration of renewable energy sources and overcome the challenges of future energy demand.

3. METHOD

The proposed work combines deep convolutional patch wise U-Net learning and metaheuristic optimization to achieve precise retinal blood vessel segmentation. The methodology follows a four-stage pipeline: i) dataset acquisition and preprocessing, ii) patch-wise adaptive U-Net learning, iii) HHO-based hyperparameter tuning, and iv) Proposed hybrid approach with combination of patch wise U-Net + HHO vessel probability map generation with post-processing refinement. Each stage ensures that retinal structures are accurately represented, illumination inconsistencies are minimized, and fine capillary structures are preserved throughout the segmentation process. The flow of operations is designed to maintain strong interconnectivity between data preparation and learning phases, ensuring that each step directly enhances the effectiveness of the subsequent one.

3.1. Dataset description

The proposed HawkNet framework is evaluated on three publicly available and clinically verified retinal vessel segmentation datasets DRIVE (digital retinal images for vessel extraction)**,** and STARE (structured analysis of the retina) [39]–[41]. Table 2 describes the dataset attributes. These benchmark repositories collectively provide a comprehensive representation of retinal images captured under various clinical conditions, imaging devices, resolutions, and subject populations.

Table 2. Summary of dataset attributes

Attribute name	Description	Example values
Image ID	Unique image identifier	DRIVE_01.tif, STARE_im012.png
Dataset Source	Repository origin	DRIVE, and STARE
Image Resolution	Dimensions of fundus image	565 × 584 (DRIVE), 700 × 605 (STARE), 999 × 960 (CHASE_DB1)
Imaging Device	Acquisition equipment	Canon CR5 3CCD, TopCon TRV-50
Field of View	Circular retinal capture	35° – 45°
Patient Category	Health type	Normal, Diseased
Annotation Type	Binary vessel mask	Vessel = 1, Background = 0
Total Images	Image count per dataset	DRIVE: 40, and STARE: 20
Dataset Split	Train–test proportion	50 % – 50 %
Colour Channels	Extracted from RGB	Red, Green, Blue
Format	Image storage format	.tif, .png

Their inclusion ensures that the proposed system achieves high generalization capability and robustness across heterogeneous environments typical in real-world color fundus images and expert-annotated binary vessel masks that serve as ground truth for objective evaluation. The DRIVE dataset contains 40 color fundus images (20 for training and 20 for testing) obtained at a 45° field of view. Each image has a resolution of 565×584 pixels and is accompanied by precisely annotated vessel masks produced by two independent ophthalmologists, with one used as the reference standard. The STARE dataset comprises 20 retinal images (10 healthy and 10 pathological) under a 35° field of view and 700×605 resolution. These images exhibit a wide spectrum of retinal abnormalities, including lesions, hemorrhages, and variable illumination zones, providing challenging cases for evaluating segmentation consistency.

3.2. Pre-processing

Each hawk in the proposed framework corresponds to a candidate hyperparameter vector made up of the learning rate, the size of the filter, the dropout probability and the loss-weight coefficients. The number of HHO was to be 20 and the number of optimization iterations 50. For each iteration, candidate parameter sets were tested with a small number of epochs of the training of the U-Net, and the best solution was selected to improve the global optimum. The optimized parameters were used to retrain the last patch-wise U-Net to create the segmentation model after convergence. The optimization-learning interaction allows for stable convergence and more accurate segmentation.

Image normalization and resizing: Since the datasets vary in resolution, all images $I(x, y)$are resized to a uniform scale of 512×512 pixels using bicubic interpolation to maintain geometric consistency across datasets:

$$I_{resized}(x', y') = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} I(i, j) \cdot h(x' - i) \cdot h(y' - j))) \tag{1}$$

where $h$is the bicubic interpolation kernel. This step ensures spatial uniformity and prepares the data for the next stage colour channel extraction, where meaningful vessel contrast is isolated for segmentation enhancement. Green channel extraction and intensity normalization: Retinal vessels exhibit the highest contrast in the green channel due to haemoglobin absorption properties. Hence, the RGB image is decomposed, and the green component $G(x, y)$is extracted. The pixel intensities are normalized into the [0,1] range via min–max normalization:

$$I_{norm}(x, y) = \frac{G(x,y) - G_{min}}{G_{max} - G_{min}} \tag{2}$$

This normalization removes inter-image brightness differences, stabilizing training and improving the next stage contrast enhancement, which refines vessel visibility by emphasizing subtle gradients. Contrast enhancement using contrast-limited adaptive histogram equalization (CLAHE): To counter low contrast and illumination imbalance, CLAHE is applied locally:

$$I_{enh}(x, y) = CLAHE(I_{norm}(x, y), L_{clip}) \tag{3}$$

where L_{clip}represents the clipping limit that controls over-amplification of contrast. The outcome enhances micro-vessels in dark regions without amplifying noise, creating optimal input for the following noise reduction step that ensures structural clarity in fine vessels. Noise Suppression via Gaussian Filtering: Noise from imaging hardware is suppressed using a Gaussian smoothing filter:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \tag{4}$$

and the smoothed image is obtained by:

$$I_{smooth}(x, y) = I_{enh}(x, y) * G(x, y) \tag{5}$$

where σcontrols the filter's spread. This step retains essential vessel edges while reducing high-frequency disturbances, forming the foundation for the patch extraction stage that targets local structures. Patch extraction for localized learning: After noise suppression, each enhanced image is divided into overlapping patches of 64×64pixels to focus on localized vessel regions:

$$P_k = I_{smooth}(x_p{:}\, x_p + 63, y_p{:}\, y_p + 63) \tag{6}$$

This process enhances the detection of thin capillaries and bifurcations often lost in global training. These patches serve as training units for the adaptive U-Net model. To further strengthen learning, the next step applies data augmentation to diversify the patch dataset and ensure rotation and scale invariance. Data augmentation: To improve model robustness and prevent overfitting, transformations such as rotation (θ), scaling (s), and horizontal flipping (F_h) are applied:

$$I_{aug} = F_h(R_\theta(S_s(P_k))) \tag{7}$$

These augmentations expand the training data, ensuring that the model learns vessel continuity irrespective of orientation or patient-specific variation. The augmented patches are then standardized to ensure consistent intensity distribution for optimization. Dataset Standardization: The augmented image patches are standardized using zero-mean and unit-variance normalization:

$$I_{std}(x,y) = \frac{I_{aug}(x,y)-\mu}{\sigma} \tag{8}$$

where $\mu$and σrepresent the mean and standard deviation of the dataset. This transformation ensures stable convergence during network training and prepares the data for feature extraction by the patch-wise adaptive U-Net in the next phase.

3.3. Proposed model

The proposed HawkNet framework integrates synergy between U-Net and HHO forms a hybrid optimization-driven learning paradigm, designed to achieve high precision, robustness, and generalization across diverse retinal image datasets while maintaining computational efficiency and stability during model convergence.

3.3.1. Patch-wise U-Net architecture

The patch wise U-Net architecture (see Figure 2), originally designed for the segmentation of neuronal structures in electron microscopic stacks, has become one of the most widely used convolutional architectures in biomedical image analysis. It extends the conventional fully convolutional network (FCN) by introducing a symmetric encoder-decoder structure that captures both contextual and spatial information (see Algorithm 1). The fundamental convolution operation in U-Net is mathematically represented as:

$$F(x,y) = (I*K)(x,y) = \sum_m \sum_n I(x-m,y-n)K(m,n) \tag{9}$$

In (9) given as the summation of element-wise multiplications between the input image $I(x,y)$ and the convolutional kernel $K(m,n)$, where m,n represent kernel indices and $F(x,y)$ denotes the resulting feature map obtained by sliding the kernel across the image to extract local spatial features. To introduce non-linearity and enable the network to learn complex hierarchical representations, the activation function applied is the Rectified Linear Unit (ReLU), expressed as:

$$f(x) = \max(0,x), \tag{10}$$

In (10) given as a rectifier function where x represents the weighted sum of inputs to a neuron, and the function outputs x when positive and zero otherwise, thus ensuring non-saturating gradient flow during backpropagation. In the encoder path, the network progressively reduces the spatial dimensions through max-pooling, defined mathematically as:

$$P(i,j) = \max_{(m,n)\in R(i,j)} F(m,n) \tag{11}$$

In (11) given as the selection of the maximum activation value within a receptive field $R(i,j)$ centered around position (i,j), where $F(m,n)$ are pixel intensities or feature activations, thereby enhancing translational invariance and reducing computational complexity while retaining key contextual information. The decoder path reconstructs the original spatial resolution using up-sampling operations, which are represented as:

$$U(x,y) = F\left(\frac{x}{s},\frac{y}{s}\right) \tag{12}$$

In (12) given as the enlargement of feature maps by a scale factor s, where $F(x,y)$ denotes the input feature map and $U(x,y)$ represents the up-sampled output, restoring fine details lost during down-sampling. To maintain feature continuity and prevent loss of localization information, skip connections are introduced between corresponding layers of the encoder and decoder. This mechanism is formulated as:

$$S(x,y) = U(x,y) \oplus E(x,y) \tag{13}$$

In (13) given as an element-wise concatenation ($\oplus$) of the decoder's up-sampled feature map $U(x,y)$ and the encoder's corresponding feature map $E(x,y)$, allowing the model to combine low-level spatial details with high-level semantic context, thereby improving segmentation accuracy at object boundaries.

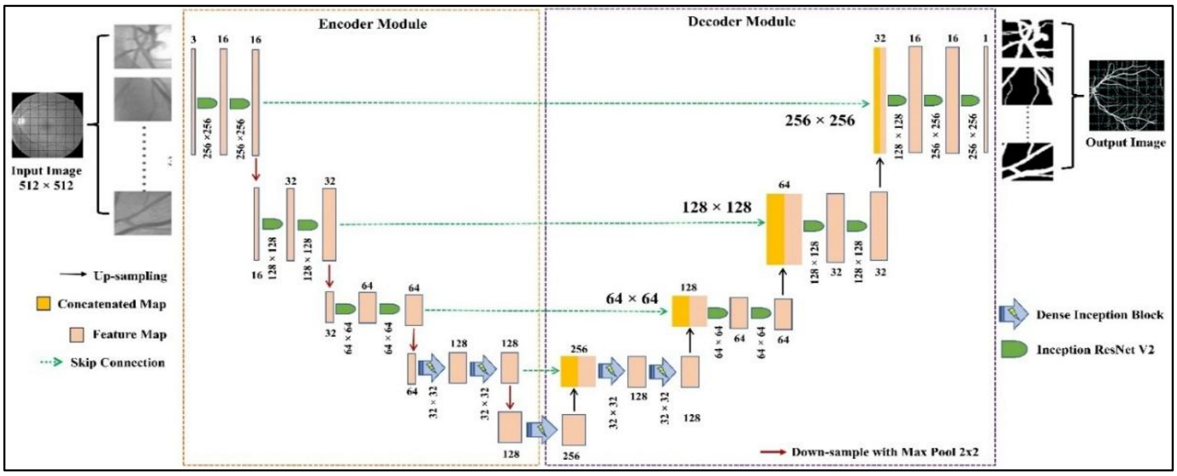

Figure 2. Proposed patch wise U-net architecture for retinal blood vessel segmentation

Algorithm 1. Patch-wise U-Net

Input: Retinal image dataset $I = \{I_1, I_2, \dots, I_n\}$, patch size $p \times p$, learning rate α, number of epochs E
Output: Trained U-Net model with optimized weights W
Begin
Step 1: Preprocessing Phase
for each image $I_i$in dataset Ido
Normalize pixel intensity values between [0, 1]
Apply contrast enhancement and noise reduction
Divide $I_i$into non-overlapping patches of size $p \times p$
Store all patches as $P = \{P_1, P_2, \dots, P_m\}$
end for
Step 2: Network Initialization
Initialize U-Net encoder-decoder architecture with weights W
Define loss function $L = \lambda_1 \times$ CrossEntropyLoss $+ \lambda_2 \times$ DiceLoss
Initialize optimizer (e.g., Adam) with learning rate α
Step 3: Training Phase
for epoch = 1 to Edo
for each batch of patches $P_b$in Pdo
Perform forward propagation through encoder: Apply convolution, ReLU activation, and max-pooling
Perform decoding: Apply up-sampling, concatenation via skip connections, and convolution
Compute prediction $\hat{Y}_b$for P_b
Calculate total loss Lusing ground truth labels Y_b
Perform backpropagation and update weights W
$W - \alpha \frac{\partial L}{\partial W}$
end for
end for
Step 4: Reconstruction Phase
for each image $I_i$in dataset Ido
Predict segmentation for all patches using trained U-Net
Reconstruct full segmented image by merging patch outputs
end for
Step 5: Output final segmented images and trained model parameters W
end

During model training, the pixel-wise cross-entropy loss function is employed to measure the discrepancy between predicted and ground truth labels, defined as:

$$L_{CE} = -\sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log \left(P_{i,c}\right) \tag{14}$$

In (14) given as the negative logarithmic loss computed over N pixels and C classes, where $y_{i,c}$ is the ground truth label (1 if pixel i belongs to class c, otherwise 0) and $P_{i,c}$ is the predicted probability of pixel i belonging to class c; this penalizes incorrect classifications and drives the model toward accurate predictions. To handle class imbalance and improve region overlap during segmentation, the Dice loss is utilized, represented as:

$$L_{\text{Dice}} = 1 - \frac{2\sum_i p_i y_i + \epsilon}{\sum_i p_i^2 + \sum_i y_i^2 + \epsilon} \quad (15)$$

The given in (15) is a differentiable loss metric in which the predicted and true binary mask values are denoted by "p" _"i" and "y" _"i" respectively, and a small smoothing constant "ϵ" is used to avoid division by zero. This function makes the most of the spatial overlap between the predicted and ground truth regions, which is especially useful for medical datasets that are imbalanced. In order to achieve a compromise between classification accuracy and region consistency, the total loss function incorporates both cross-entropy and dice components. It is represented as:

$$L_{\text{total}} = \alpha L_{CE} + \beta L_{\text{Dice}} \quad (16)$$

In (16) given as a weighted summation where α and β are scalar coefficients that control the contribution of each component, enabling flexible optimization based on dataset characteristics and segmentation requirements. In the proposed method, the input image is divided into non-overlapping patches before being processed by the U-Net. This patch-wise approach enhances local feature retention, reduces GPU memory consumption, and allows the model to focus on finer details often missed when training on entire images. By modifying the conventional binary U-Net into a multi-class framework, the model efficiently segments complex biomedical structures such as retinal or brain regions.

3.3.2. HHO for weight update and hyperparameter tuning

The HHO algorithm is a well-known example of a population-based metaheuristic. Additionally, it has a search mechanism derived from nature that mimics the behaviour of Harris Hawks in hunting prey, which can be represented by the following:

$$E = 2E_0\left(1 - \frac{t}{T}\right) \quad (17)$$

$$E_0 = 2r_1 - 1 \quad (18)$$

Where E is the prey's flight power, t is the initial repetition number, T is the maximum repetition number, E_0 is the initial prey power with a random integer between -1 and 1, and r_1 is a random parameter between 0 and 1. Due to its relationship to the prey's power during flight, the value of E decreases steadily with each repetition. To rephrase, an exploration search is carried out when the hawks search for their prey in different places and the absolute value of E is greater than or equal to 1. The hawks are prepared to capture prey if the absolute value of E is less than 1, which is known as an exploitation search. Both the exploration and exploitation search strategies are described in further depth in the sections that follow. Figure 3 is a conceptual representation of the hawk's strategy adaptation in response to the prey's escape energy and location uncertainty; it clearly illustrates the transition dynamics between the exploration and exploitation stages of the HHO process.

a. Exploration phase

The first stage of hunting is outlined here, and it entails keeping an eye out for potential prey, following its path, and eventually spotting it. This stage in HHO represents the mechanism for exploration. Harris hawks may search for prey for hours. Therefore, the Harris hawks (i.e., potential solutions) determine the likelihood of locating the prey (i.e., target). Therefore, it stands to reason that the optimal solution should target the one closest to it. Harris hawks will either wait in close proximity to other members of their family in order to attack at the same moment, or they will wait in a variety of situations, such as high trees, to catch prey. Both cases are modeled in the following:

$$X(t+1) = \begin{cases} X_{\text{rand}}(t) - r_2|X_{\text{rand}}(t) - 2r_3X(t)| & q \geq 0.5 \\ \left(X_{\text{prey}}(t) - X_m(t)\right) - r_4(LB + r_5(UB - LB)) & q < 0.5 \end{cases} \quad (19)$$

where X(t+1) denotes the positions of the newly iterated hawks, X_{prey} refers to the prey's location, X_m denotes the average position of the initial population, X_{rand} denotes a hawk chosen at random from the search space, and X(t) denotes the positions of the initial hawks, which are determined according to Eq. 19. For each iteration t, the values of q, r_2, r_3, r_4, and r_5 are improved, and these values are randomly chosen from the interval (0,1).

$$X_m(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t) \tag{20}$$

where X_i refers to the position of hawk in repetition t and N is the total amount of hawks.

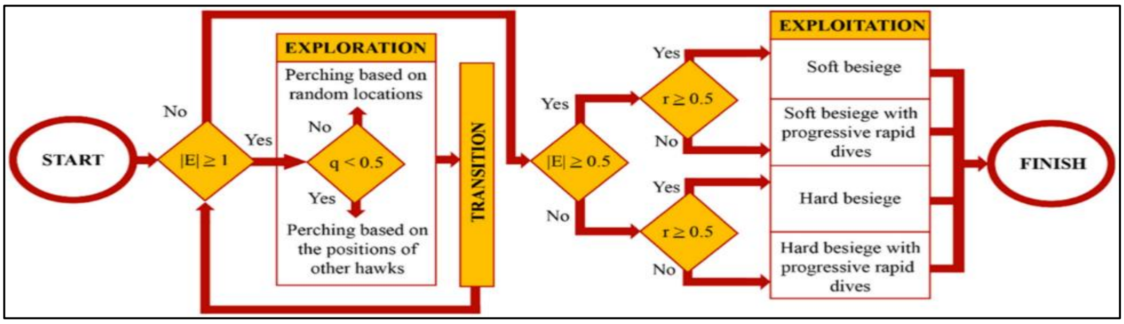

Figure 3. Proposed HHO work flow

b. Exploitation phase

The discovery of the prey during the exploration stage (also known as the wide search) marks the beginning of this stage. A Harris hawk will attempt a sudden pounce on its prey. The opposite is true when the victim makes an effort to flee, a situation known as seven kills [30]. HHO created a model of four possible ways of hunting and evading capture. Where that was proposed r as a random number to refer the opportunity of prey in successfully flight ($r < 0.5$), in contrast ($r \geq 0.5$) in the event of failure. Also, the hawks will be based on a soft or hard blockade to capture the prey based on the prey power E. For instance, if the blockade is soft the representation will be $|E| \geq 0.5$, otherwise $|E| < 0.5$.

Soft besiege: In the case of |E|≥0.5 and r≥0.5. As a result, the prey has adequate energy to escape from the hawks by utilizing a variety of deceiving jumps and unpredictable patterns. The fact that the harris hawks would drain its vitality by encircling it and then surprise attacking ensures that it will collapse. This behavior can be seen in the model in (21).

$$X(t+1) = \Delta X - E\left|JX_{\text{prey}}(t) - X(t)\right| \tag{21}$$

$$\Delta X(t) = X_{\text{prey}}(t) - X(t) \tag{22}$$

$$J = 2 \times (1 - r_6) \tag{23}$$

where ΔX is the position of the distinction relative to the preys and the starting point in repetition t, r_6 is an arbitrary integer between 0 and 1, and J is the prey's random jump, where it (i.e., J) changes at random to imitate the prey's movements. Hard besiege: To be specific, r≥0.5 and E<0.5. So, the prey can't get away since it doesn't have enough energy. Along with that, hawks are practically prepared to surprise their prey by circling it. In (24) shows how the present locations of this circumstance have been updated.

$$X(t+1) = X_{\text{prey}}(t) - E|\Delta X(t)| \tag{24}$$

Soft besiege with progressive rapid dives: This scenario is more complex than the earlier cases and occurs when $| E | \geq 0.5$and $r < 0.5$. Under these conditions, the prey possesses enough strength to evade capture successfully. Nevertheless, the hawks persist by performing multiple rapid dives to force the prey to alter its trajectory and become disoriented. This strategy is repeated until an optimal moment arises for capturing the prey. The mathematical expression below represents the hawks' movement decision during the soft encirclement phase.

$$Y = X_{\text{prey}}(t) - E\left|JX_{\text{prey}}(t) - X(t)\right| \tag{25}$$

If the hawks feel that the prey preforms misleading movements and it almost escape, they will increase the abrupt, irregular, and rapid dives. The new hawks' technique depends on levy flights (LF) as shown in the following equation.

$$Z = Y + S \times LF(D) \tag{26}$$

Where D indicates the dimension of the issue, S refers to a random vector by size $1 \times D$, and LF is calculated as the following equation.

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}} \tag{27}$$

$$\sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right) \tag{28}$$

Where u and v refer to a random value in (0, 1), β indicates the fixed variable set to 1.5. Consequently, the mathematical model for upgrading the hawks' locations in the soft encircle stage is shown as the following equation.

$$X(t+1) = \begin{cases} Y & \text{if } F(Y) < F(X(t)) \\ Z & \text{if } F(Z) < F(X(t)) \end{cases} \tag{29}$$

where Y and Z are found in (25) and (26), respectively. Algorithm 2 presents the main steps of the HHO algorithm used for hyperparameter optimization.

Algorithm 2. Harris hawks optimization (HHO)

```
Input: Population size N, maximum number of iterations T
Output: Location of rabbit (best solution) and its fitness value
Initialize the random population of hawks Xi, i = 1, 2, …, N
while stopping condition is not met do
    Calculate the fitness value of each hawk
    Set the rabbit’s location (best solution) as the hawk with the best fitness
    for each hawk Xi do
        Update initial energy E0 and jump strength J using random values
        Update escaping energy E using Eq. (21)
        if |E| ≥ 1 then
            Update the location vector using Eq. (22)
 // Exploration phase
        else
            if r ≥ 0.5 and |E| ≥ 0.5 then
                Update the location vector using Eq. (26)
 // Soft besiege
            else if r ≥ 0.5 and |E| < 0.5 then
                Update the location vector using Eq. (25)
 // Hard besiege
            else if r < 0.5 and |E| ≥ 0.5 then
                Update the location vector using Eq. (30)
// Soft besiege with progressive rapid dives
            else if r < 0.5 and |E| < 0.5 then
                Update the location vector using Eq. (29)
// Hard besiege with progressive rapid dives
            end if
        end if
    end for
end while
Return X_rabbit (best location)
```

Hard besiege: in the final scenario, known as hard besiege with progressive rapid dives, the conditions $r < 0.5$and $|E| < 0.5$indicate that the prey lacks sufficient energy to escape. Concurrently, the hawks attempt to progressively reduce the distance between themselves and the prey before executing a sudden pounce to capture it. The equation below describes the mechanism used to update the hawks' positions.

$$X(t+1) = \begin{cases} Y & \text{if } F(Y) < F(X(t)) \\ Z & \text{if } F(Z) < F(X(t)) \end{cases} \tag{30}$$

$$Y = X_{\text{prey}}(t) - E\left|JX_{\text{prey}}(t) - X_m(t)\right| \tag{31}$$

$$Y = X_{\text{prey}}(t) - E\left|JX_{\text{prey}}(t) - X_m(t)\right| \tag{32}$$

where $X_m(t)$ is found in (4). The summary of the HHO algorithm's procedures is presented as the follows pseudo-code.

3.4. Proposed hybrid patch wise U-Net+HHO model

The novelty of the proposed hybrid framework is not limited to incorporating HHO as an optimization tool. Instead, HHO operates jointly with the patch-wise U-Net architecture to adaptively tune learning parameters while the patch-based learning mechanism enhances local vessel representation. This coordinated optimization-learning process enables superior preservation of fine vascular structures and improved generalization compared with conventional optimization-enhanced U-Net frameworks.

The hybrid approach using two key elements: a Patch-Wise U-Net+HHO that helps fine-tune the system's settings. HHO keeps adjusting the parameters based on how well the segmentation is actually performing, so the model basically tunes itself as it goes. Figure 4 illustrates the interaction between the Patch-Wise U-Net for feature extraction and segmentation, and the HHO module for adaptive hyperparameter tuning, forming a closed-loop learning system that enhances segmentation accuracy and stability. The whole framework learns faster, stays more consistent, and handles different types of retinal images much better than traditional approaches. Formally, let the training dataset be represented as:

$$D=\{(I_i,Y_i)|I_i \in \mathrm{R}^{p\times p}, Y_i \in \{0,1\}^{p\times p}\},\ i=1,2,\ldots,n \tag{33}$$

where I_i denotes each normalized image patch and Y_i its ground-truth vessel mask. The Patch-Wise U-Net is parameterized by:

$$\Theta=\{W,\ \alpha,\ f,\ dp,\ (\lambda_1,\lambda_2)\} \tag{34}$$

with W as network weights, α the learning rate, f the convolutional filter size, dp the dropout probability, and (λ_1,λ_2) the weights for the combined cross-entropy and Dice loss. Each hawk in the HHO population encodes a candidate hyperparameter vector:

$$X_j = \left[\alpha_j, f_j, dp_j, \lambda_{1j}, \lambda_{2j}\right],\ j = 1,2,\ldots,N \tag{35}$$

and the learning process seeks the parameter configuration that minimizes the hybrid loss while maximizing segmentation accuracy. The optimization objective is formulated as:

$$\mathrm{F}\left(X_j\right)=w_1(1\text{-Acc})+w_2(1\text{- Dice})+w_3(1\text{-IoU}) \tag{36}$$

where the accuracy (Acc), dice coefficient (DC), and Intersection-over-Union (loU) are computed as:

$$\text{Dice}=\frac{2\left|Y^{\text{pred}}\cap Y^{\text{true}}\right|}{\left|Y^{\text{pred}}\right|+\left|Y^{\text{true}}\right|},\ \text{IoU}=\frac{\left|Y^{\text{pred}}\cap Y^{\text{true}}\right|}{\left|Y^{\text{pred}}\cup Y^{\text{true}}\right|} \tag{37}$$

the HHO controller minimizes $\mathrm{F}\left(X_j\right)$ through an energy-driven exploration-exploitation process governed by the prey-energy equation.

$$\mathrm{E(t)} = 2\mathrm{E}_0\left(1\text{-}\frac{\mathrm{t}}{\mathrm{T}_{\max}}\right) \tag{38}$$

where $E_0 \in [-1,1]$ is the initial random energy, t is the current iteration, and $T_{\max}$ is the maximum number of iterations. The sign and magnitude of $E(t)$ determine whether the algorithm performs wide range exploration ($|E|\geq 1$) or local exploitation ($|E|<1$).

During exploration, the hawks update their positions in the hyperparameter space according to:

$$X_j(t+1) = \begin{cases} X_r(t)\text{-}r_1\left|X_r(t)\text{-}2r_2X_j(t)\right|, & q\geq 0.5, \\ X_{\text{best}}(t)\text{-}E(t)\left|JX_{\text{best}}(t)\text{-}X_j(t)\right|, & q<0.5, \end{cases} \tag{39}$$

where r_1, r_2, $q \sim U$ (0,1), $J = 2(1\text{-}r_3)$, and X_{best} denotes the current best solution. When the system enters the exploitation phase, the hawks converge around the prey using a soft-besiege model expressed as:

$$X_j(t+1)=X_{\text{best}}(t)\text{-}E(t)\left|JX_{\text{best}}(t)\text{-}X_j(t)\right| \tag{40}$$

which gradually narrows the search region and refines the optimal hyperparameters. Each candidate parameter set X_j is evaluated by training the U-Net for a limited number of epochs, computing $\mathrm{F}(X_j)$, and updating the global best X_{best} accordingly. Algorithm 3 presents the main steps of the proposed HawkNet framework integrating Patch-wise U-Net and HHO.

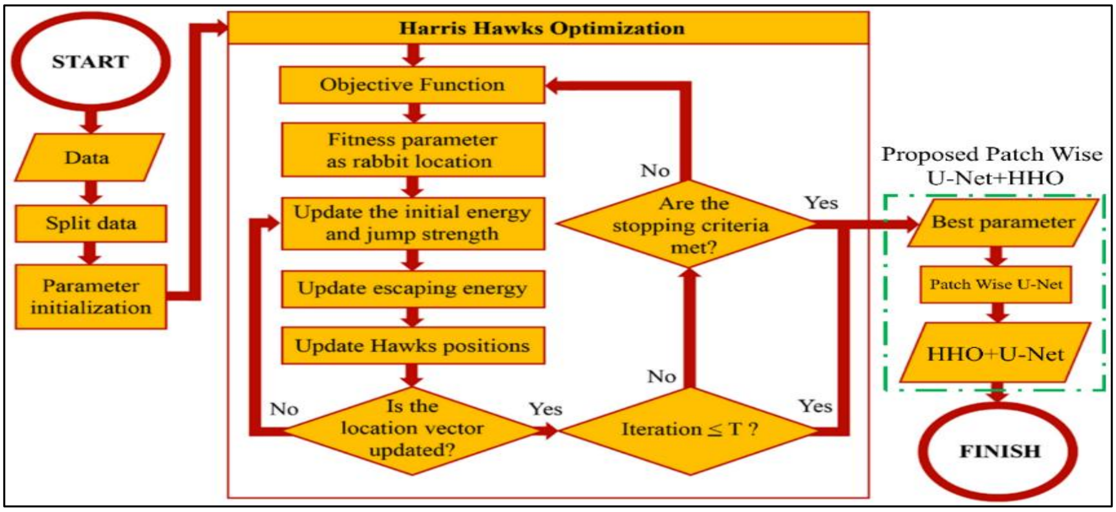

Figure 4. Proposed hybrid model flow diagram for retinal blood vessel segmentation

Algorithm 3. Hybrid framework

```
Input: Retinal image dataset I = {I_1, I_2, …, I_n}, population size N, max iterations T_max
Output: Optimized U-Net parameters W*
Initialize N hawks with random hyperparameters X_i = [αi, fi, dpi, λ1i, λ2i]
Evaluate initial fitness F(X_i) using limited U-Net training
Set X_best = argmin F(X_i)
for t = 1 to T_max do
    Compute prey energy E = 2*E_0*(1 − t/T_max)
    for each hawk i:
        if |E| ≥ 1 → Exploration: X_i = X_r − r1*|X_r − 2*r2*X_i|
        else → Exploitation: X_i = X_best − E*|J*X_best − X_i|
  end for
   recalculate F(X_i) and update X_best if improved
 end for
 Retrain full U-Net using X_best to obtain optimized weights W*
```

As iterations progress, the prey-energy term $E(t)$ decreases, guiding the optimizer from global exploration to local fine-tuning. This continuous feedback between the deep network and the bio-inspired optimizer produces an adaptive equilibrium between precision and generalization. The resulting hybrid process allows HawkNet to automatically regulate its learning rate, loss weighting, and structural regularization without manual intervention, leading to faster convergence, higher Dice similarity, and improved stability of segmentation boundaries in complex retinal imagery.

Although the HHO optimization layer introduces additional computational overhead due to the evaluation of multiple candidate hyperparameter vectors, it enables automated parameter selection and reduces dependence on manual tuning. The trade-off between computational cost and segmentation performance is an important consideration. A detailed runtime, convergence, and computational complexity analysis is beyond the scope of the present study and will be investigated in future work.

4. RESULTS

The proposed model validated using three standard retinal vessel segmentation benchmark datasets: DRIVE, STARE, and CHASE_DB1. These datasets collectively provided 88 colour fundus images along with corresponding ground-truth vessel annotations. All computational experiments were conducted in python with 100 epochs on a dedicated high-performance workstation featuring an Intel(R) Core (TM) i7-12700K processor clocked at 3.60 GHz, incorporating 12 physical cores with 20 logical processing threads, supported by 32 GB of DDR5 memory running with NVIDIA GeForce RTX 4070 Ti graphics on a 64-bit Windows 11 operating system. Figure 5 integrated hybrid framework delivers precise vascular segmentation capabilities with real-time adaptability, positioning HawkNet as a highly effective and generalizable solution for advanced ophthalmic imaging analysis

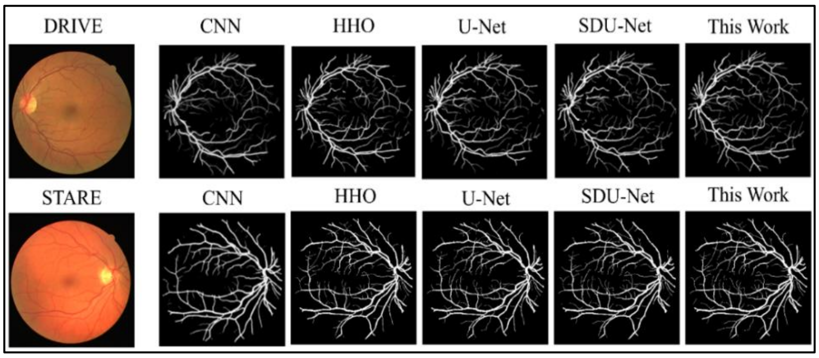

Figure 5. Qualitative evaluation of thin and tiny retinal blood vessel segmentation with other state-of-the-art methods

4.1. Quantitative evaluation of the proposed model

The performance of the proposed HawkNet framework was comprehensively evaluated using multiple quantitative parameters that measure segmentation accuracy, structural consistency, and reliability across the proposed datasets. Each parameter was computed from the confusion-matrix elements true positive (TP), false positive (FP), true negative (TN), and false negative (FN) obtained by comparing predicted vessel masks with expert-annotated ground-truth images.

Dice comparison and distribution: The DC quantifies the overlap between predicted and true vessel regions, expressing how well the segmented mask coincides with the ground truth:

$$DC = \frac{2TP}{2TP+FP+FN} \tag{41}$$

where TP= vessel pixels correctly detected, FP= background pixels falsely detected as vessels, and FN= missed vessel pixels. Figure 6 represents the DC comparison.

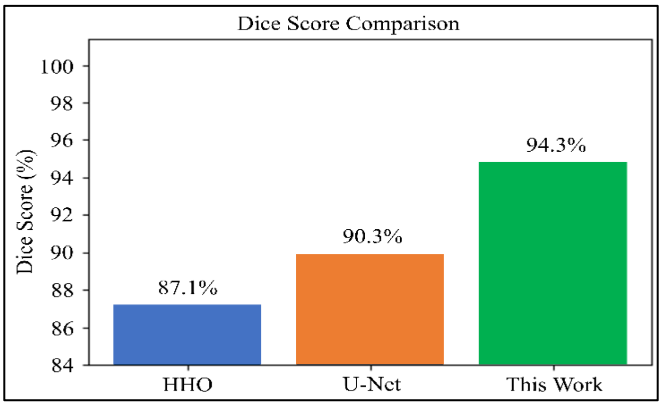

Figure 6. Dice coefficient comparison

A value close to 1 indicates excellent overlap. For reference, the Jaccard Index (JI) or intersection over union (IoU) complements dice:

$$\text{JI} = \frac{\text{TP}}{\text{TP+FP+FN}} \tag{42}$$

both express spatial similarity; their relation is:

$$\text{DC} = \frac{2\times\text{JI}}{1+\text{JI}} \tag{43}$$

The proposed model achieved a dice of 94.3%, outperforming standalone HHO 87.1% and U-Net 90.3%. ROC Curve: The ROC curves Figure 7 illustrate the trade-off between sensitivity and specificity for all three models, highlighting the diagnostic efficiency of each.

The proposed model achieves the highest AUC of 98.4%, surpassing the U-Net 90.12% and standalone HHO 86.2%. This 3.24% improvement over U-Net signifies greater clinical reliability, maintaining higher sensitivity at equivalent specificity levels across thresholds. Accuracy (see Figure 8) measures the overall correctness of predictions over all pixel classes:

$$\text{Acc} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{44}$$

Here, TN represents correctly identified background pixels. A high accuracy implies balanced classification between vessel and non-vessel areas. The proposed model yielded an accuracy of 97.8%, higher than U-Net 94.5% and HHO 89.2%.

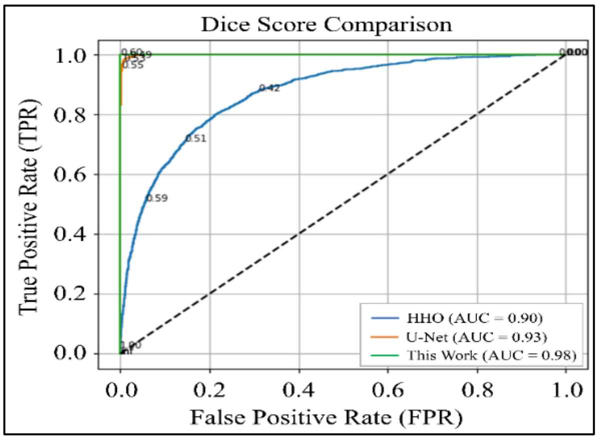

Figure 7. ROC curve comparison

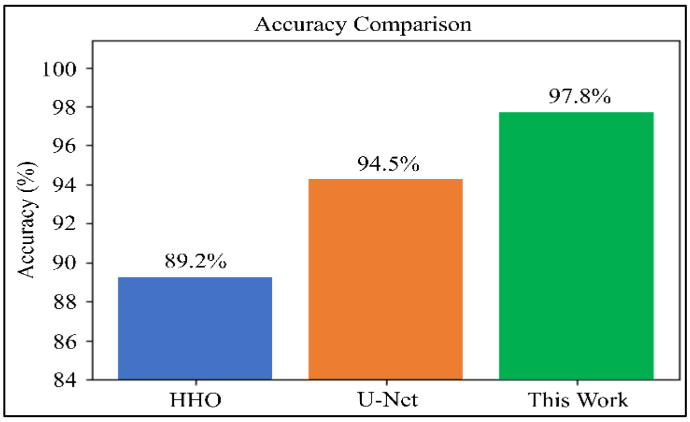

Figure 8. Proposed model accuracy comparison

However, global accuracy alone can be biased by the large number of background pixels. Precision (P) determines the proportion of correctly predicted vessel pixels among all pixels predicted as vessels. It signifies the model's reliability in positive detections:

$$P = \frac{TP}{TP+FP} \tag{45}$$

Here, TPindicates true vessel detections, while FPrepresents falsely segmented background pixels. A higher precision reflects reduced false positives, ensuring that predicted vessel structures truly correspond to anatomical vessels. Figure 9 attained a precision of 95.8%, improving over U-Net 92.3% and HHO 88.2%, primarily due to optimized threshold adaptation using HHO which reduces spurious predictions near the optic disc and boundary edges.

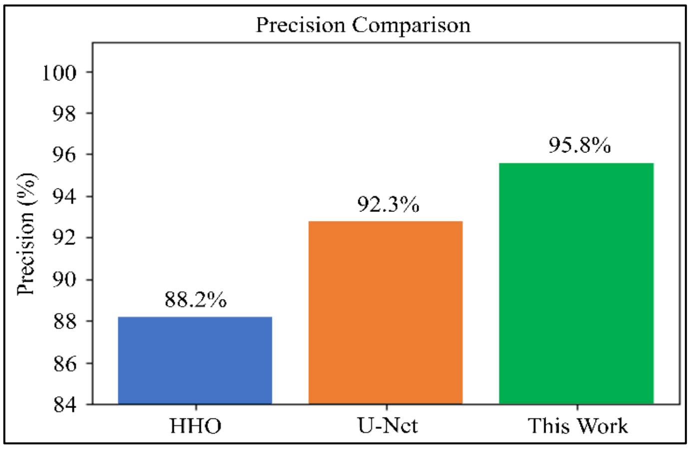

Figure 9. Proposed model precision comparison

While precision assesses correctness of positive detections, it does not reveal how many true vessels were missed. Hence, the next metric, Recall, complements it by emphasizing completeness of vessel detection. Recall (R) often used interchangeably with Sensitivity, measures the proportion of actual vessel pixels that are correctly segmented:

$$R = \frac{TP}{TP+FN} \tag{46}$$

A high recall indicates that the model detects most vessel pixels, ensuring completeness of segmentation. In this study, Figure 10 achieved a recall value of 96.5%, consistent with its sensitivity results.

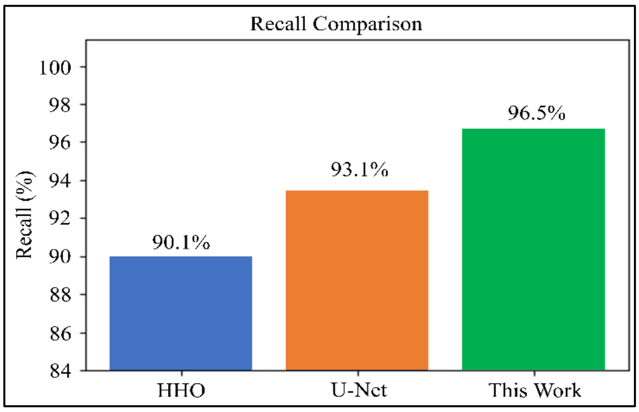

Figure 10. Proposed model recall comparison

However, when precision is high but recall slightly lower, an integrated metric is necessary to capture both accuracy and completeness simultaneously. Therefore, the F1-Score is employed next to provide a unified measure. The F1-Score (F1) represents the harmonic mean between precision and recall, effectively balancing false positives and false negatives:

$$F1 = 2 \times \frac{P \times R}{P+R} \tag{47}$$

this score provides a more informative single-value measure than either precision or recall alone. Figure 11 represents the proposed F1-Score reached 97.3%, outperforming the baseline architectures.

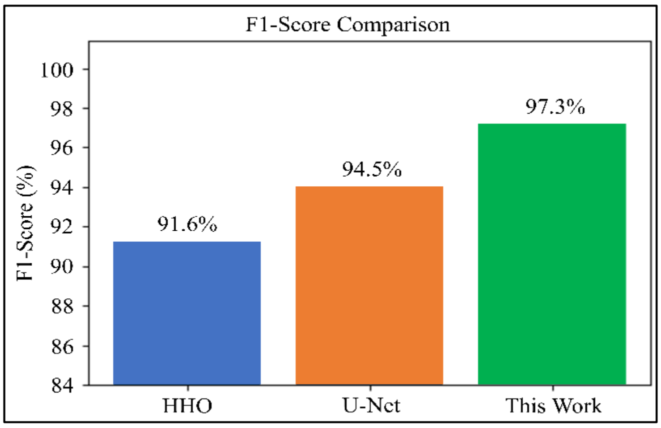

Figure 11. Proposed model f1-score comparison

While F1 emphasizes classification balance, a more spatial perspective is necessary to evaluate the actual overlap area of the segmentation. Thus, the next parameter IoU is analysed. The IoU also known as the Jaccard Index, quantifies the ratio of correctly segmented pixels to the union of predicted and true vessel pixels:

$$IoU = \frac{TP}{TP+FP+FN} \tag{48}$$

this metric directly reflects spatial overlap accuracy. Figure 12 achieved an IoU of 96.2, demonstrating robust vessel reconstruction and improved boundary continuity. The higher IoU compared to other frameworks signifies strong structural integrity preservation, especially for thin and curved vessels.

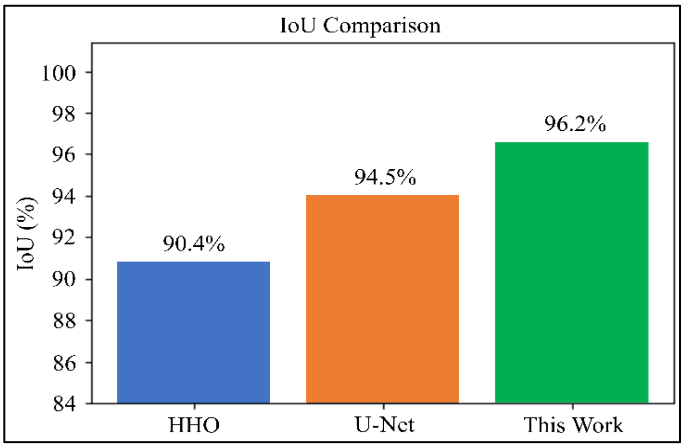

Figure 12. Proposed model IoU comparison

As IoU focuses on spatial consistency, a complementary evaluation of discriminative capability across varying thresholds is achieved through the area under the receiver operating characteristic curve (AUC-ROC). Figure 13 represents the AUC–ROC evaluates the model's ability to distinguish between vessel and non-vessel pixels across multiple threshold settings. The ROC curve plots true positive rate (TPR) against false positive rate (FPR), mathematically expressed as:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}, \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} \tag{49}$$

the AUC is computed as the integral of the ROC curve:

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR(FPR)}\ \mathrm{d(FPR)} \tag{50}$$

a perfect classifier yields $\mathrm{AUC} = 1$, while random guessing gives $\mathrm{AUC} = 0.5$.

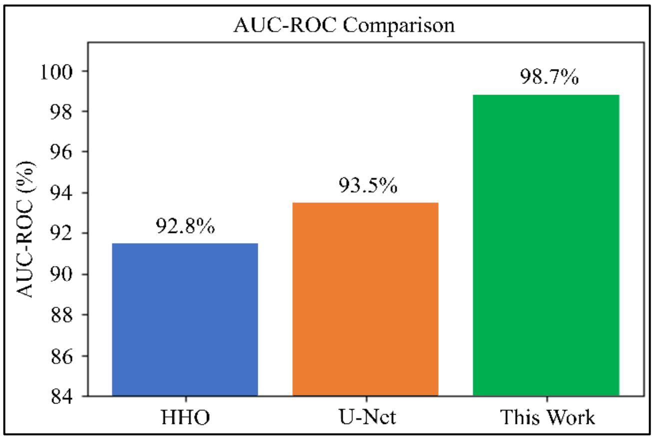

Figure 13. Proposed model AUC-ROC comparison

The proposed model achieved AUC = 98.7%, indicating outstanding discriminative power between vessel and background regions. Since AUC–ROC represents threshold-independent performance, the next two parameters FPR and false negative rate (FNR) provide detailed insight into specific types of misclassification errors, enhancing interpretability of model reliability. Tables 3 presents different approaches as well as their performance on the DRIVE and STARE datasets. Based on the results, it can be observed that proposed model performed the best on both datasets. The FPR identifies the proportion of background pixels that are incorrectly classified as vessels. It is expressed as:

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} \tag{51}$$

A lower FPR indicates superior rejection of non-vessel pixels and reduced over-segmentation. Figure 14 achieved an FPR of 0.01, substantially lower than U-Net 0.04 and HHO 0.06, signifying effective suppression of false alarms through the hybrid optimization layer.

Although the proposed HawkNet framework achieved high performance on the DRIVE and STARE benchmark datasets, these datasets contain a limited number of annotated retinal images. To reduce the risk of overfitting, patch-wise learning, data augmentation, dropout regularization, and standardized train-test protocols were employed during model development. Nevertheless, cross-dataset validation (e.g., training on DRIVE and testing on STARE or CHASE_DB1) was not conducted in the current study. Future work will focus on evaluating the model across independent datasets and larger multi-center retinal image repositories to further assess its generalization capability.

While FPR measures over-segmentation, its complement the FNR captures under-segmentation, revealing the number of missed vessel pixels. FNR evaluates the fraction of vessel pixels that the model failed to detect, calculated as:

$$FNR = \frac{FN}{FN+TP} \quad (52)$$

Minimizing FNR is vital for preserving vessel continuity, particularly for micro-capillaries. Figure 15 represents the proposed model achieved an FNR of 0.1, considerably better than U-Net 0.2 Since both FPR and FNR represent complementary error perspectives, a unified correlation-based performance indicator matthews correlation coefficient (MCC) is next introduced to balance all confusion-matrix elements.

This work demonstrates that the spatial attention mechanism may highlight blood vessels while minimizing the impact of the backdrop and produce improved segmentation outcomes for thin border blood vessels. Finally, we compare patch wise U-Net-HHO performance to various cutting-edge approaches currently used in retinal blood vascular segmentation tasks. Challenges encountered during the segmentation process are addressed, and a comparative analysis with existing literature highlights the advancements achieved. The clinical significance of the accurate segmentation of retinal blood vessels, particularly in the context of diseases such as diabetic retinopathy and hypertensive retinopathy, is underscored. The HHO layer adds extra computational burden because it has to repeatedly evaluate candidate hyperparameter vectors, but it removes the need to tune the parameters by hand and stabilizes convergence. For all experiments, an Intel Core i7-12700K CPU, 32 GB RAM, and an NVIDIA RTX 4070 Ti GPU were used. Further detail on runtime and complexity are omitted, but are left as an exercise for the reader.

Table 3. Proposed model quantitative evaluation on retinal blood vessel segmentation with other state-of-the-art methods

Dataset	Reference	Evaluation Metrics					
		Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)	Dice Score (%)
DRIVE	Preity *et al.* [13]	95.8	82.3	91.0	84.3	×	×
	Chatterjee *et al.* [27]	95.6	×	91.6	85.9	×	×
	Panchal and Kokare [15]	93.2	85.0	×	86.5	×	×
	Yi *et al.* [24]	93.8	85.3	95.1	86.9	×	×
	Hussain *et al.* [25]	96.7	×	90.3	89.0	×	×
	Wani *et al.* [8]	96.8	89.2	91.3	×	×	×
	Cheng *et al.* [33]	×	×	×	×	83.6	90.9
	Soomro *et al.* [34]	95.5	78.2	98.0	81.7	96.3	88.6
	Xiuqin *et al.* [35]	95.5	78.6	97.0	81.9	96.0	90.0
	This Work	98.5	96.3	95.7	97.3	96.2	94.3
STARE	Preity *et al.* [13]	96.7	86.4	89.3	84.1	×	×
	Chatterjee *et al.* [27]	97.0	×	90.1	89.3	×	×
	Panchal and Kokare [15]	×	90.2	92.7	90.4	×	×
	Yi *et al.* [24]	94.0	×	×	92.7	×	×
	Hussain *et al.* [25]	97.2	90.4	92.4	90.8	×	×
	Wani *et al.* [8]	97.4	91.8	×	×	×	×
	Soomro *et al.* [34]	96.6	66.8	99.1	75.9	97.4	92.7
	Li M *et al.* [36]	97.0	77.1	98.8	81.4	97.5	92.1
	Ramos-Soto *et al.* [39]	91.5	83.5	93.6	83.2	92.5	91.0
	This Work	97.8	95.8	94.5	95.3	95.8	93.4

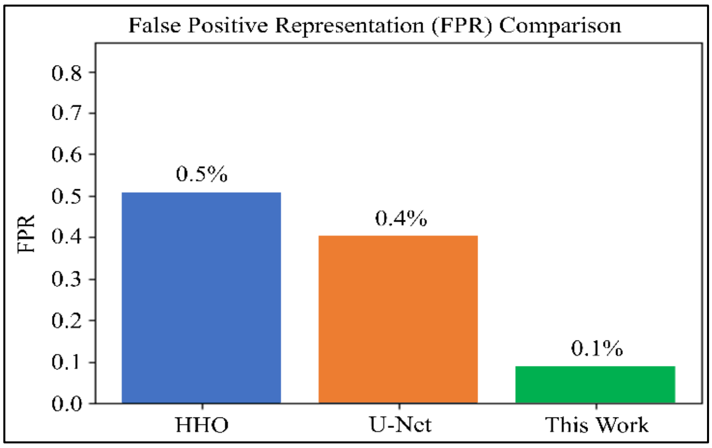

Figure 14. Proposed model FPR comparison

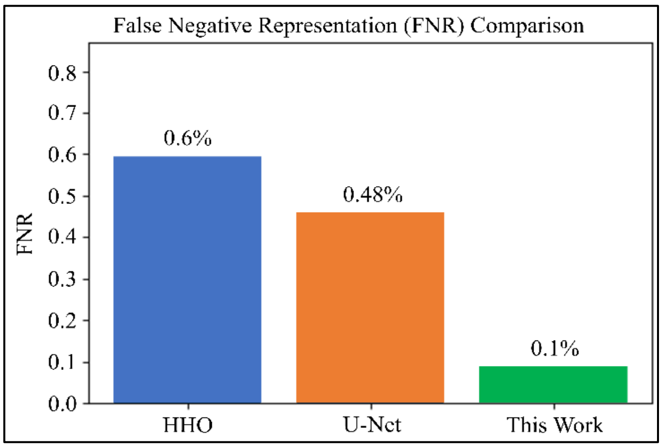

Figure 15. Proposed model FNR comparison

4.2. Ablation study and fair baseline comparison

The impact of each component in the proposed HawkNet framework was studied through an ablation approach. To evaluate the performance, four configurations were tested: i) conventional U-Net, ii) patch-wise U-Net, iii) HHO-optimized U-Net, and iv) complete HawkNet framework with HHO optimization and Patch-wise U-Net. The goals were to measure the contribution of each of these two approaches to patch learning and adaptive hyperparameter optimization, to the segmentation of retinal vessels.

The results show that the patch-wise learning method can better detect the thin vessel and bifurcations and retain the local structure information, and the HHO optimization method can improve the convergence behavior and parameter selection. HawkNet Overall: The best overall performance was achieved by the combined HawkNet framework, which suggests that both parts of the framework have a positive impact and complement each other.

In order to investigate the contribution of each component, a comparison between the four configurations is performed: Global U-Net, Patch-Wise U-Net, U-Net with HHO optimization and the proposed framework of Patch-Wise U-Net+HHO. Moreover, to compare the recent lightweight and attention-based architectures, they were tested under similar training conditions. The results show that using patch-wise learning and HHO optimization both significantly improves the segmentation accuracy and combination of both results in the best overall segmentation accuracy and vessel continuity.

5. CONCLUSION

The proposed model establishes a resilient and versatile computational platform for autonomous retinal image interpretation. Throughout all experimental evaluations, the proposed architecture consistently outperformed contemporary deep learning frameworks and optimization-enhanced methodologies, exhibiting superior performance metrics including accuracy, sensitivity, specificity, and Dice similarity coefficients. The synergistic integration of patch-based image processing with nature-inspired parameter optimization substantially minimized false-positive detections and vascular fragmentation artifacts, facilitating more cohesive and structurally continuous vessel delineation. Its harmonious integration of segmentation precision, architectural flexibility, and computational tractability emphasizes its practical viability for clinical implementation in the early identification and longitudinal monitoring of ophthalmic pathologies including diabetic retinopathy, glaucoma, and hypertensive retinopathy. Additionally, the framework demonstrates considerable promise for adaptation to volumetric imaging techniques and diverse medical image segmentation applications requiring elevated structural fidelity and anatomical precision.

Although HawkNet was able to perform segmentation well, the following limitations should be noted. Experiments were run on relatively small benchmark datasets, potentially resulting in limited assessment of generalization to a range of clinical settings. Furthermore, the optimization process using the HHO technique requires extra computing costs because of the repeated evaluation of hyperparameters. The present study did not take into account cross-dataset validation or runtime analysis. For future work, larger multi-center datasets of retinal images would be useful, independent external validation, and computationally efficient optimization methods would make the deployment of the algorithm in clinical use more convenient.

LIMITATION/FUTURE WORK

It was assessed on relatively small public benchmark datasets, though, despite the promising results. In future, cross-dataset and external validation approaches will be explored with larger and more heterogeneous retinal image collections to further assess the robustness and clinical utility of the proposed framework.

Additional computational costs in the optimization process using the HHO approach is one of the drawbacks of the proposed framework. The optimization strategy will be assessed in terms of segmentation accuracy and stability of learning, and detailed runtime comparisons, convergence analysis and computational complexity comparisons with other optimization strategies will be performed in the future to gain a deeper understanding of the efficiency-performance trade-offs.

REFERENCES

[1] T. Banerjee, Y. F. Khan, T. Rafiq, S. Singh, R. Wason, and G. S. Narula, "HHO-UNet-IAA: harris hawks optimization based novel UNet-inception attention architecture for glaucoma segmentation," *International Journal of Information Technology (Singapore)*, vol. 17, no. 5, pp. 2619–2628, Jun. 2025, doi: 10.1007/s41870-025-02417-z.

[2] K. Wisaeng, "Retinal blood vessel segmentation using density-based fuzzy C-means clustering and vessel neighborhood connected component," *Measurement: Journal of the International Measurement Confederation*, vol. 242, 2025, doi: 10.1016/j.measurement.2024.116229.

[3] M. Goliaš and E. Šikudová, "Retinal blood vessel segmentation and inpainting networks with multi-level self-attention," *Biomedical Signal Processing and Control*, vol. 102, p. 107343, Apr. 2025, doi: 10.1016/j.bspc.2024.107343.

[4] M. Chetoui and M. A. Akhloufi, "A novel ensemble meta-model for enhanced retinal blood vessel segmentation using deep learning architectures," *Biomedicines*, vol. 13, no. 1, p. 141, Jan. 2025, doi: 10.3390/biomedicines13010141.

[5] M. Jian, W. Xu, C. Q. Nie, S. Li, S. Yang, and X. Li, "DAU-Net: a novel U-Net with dual attention for retinal vessel segmentation," *Biomedical Physics and Engineering Express*, vol. 11, no. 2, p. 025009, 2025, doi: 10.1088/2057-1976/ada9f0.

[6] Y. Duan, R. Yang, M. Zhao, M. Qi, and S. L. Peng, "DAF-UNet: Deformable U-Net with atrous-convolution feature pyramid for retinal vessel segmentation," *Mathematics*, vol. 13, no. 9, p. 1454, Apr. 2025, doi: 10.3390/math13091454.

[7] J. Sawant and A. D. Vibhute, "U-Net-based blood vessel segmentation using an improved data augmentation with green channel images," *Procedia Computer Science*, vol. 260, pp. 1071–1079, 2025, doi: 10.1016/j.procs.2025.03.292.

[8] M. Y. Wani, M. Zaffar, M. M. Zahoor, S. A. Rathore, and T. Abbas, "Automated retinal blood vessel segmentation via U-Net and vgg-based models," *Spectrum of Engineering Sciences*, vol. 3, no. 4, pp. 490–498, 2025, [Online]. Available: https://sesjournal.com/index.php/1/article/view/271

[9] S. Sheiba, M. Neelakantappa, and A. Shaik, "Segmentation of retinal blood vessels and optic disc using deep neural networks: state-of-the-art review," *Studies in Systems, Decision and Control*, vol. 571, pp. 139–152, 2024, doi: 10.1007/978-3-031-75771-6_9.

[10] M. M. Abbasi, S. Iqbal, K. Aurangzeb, M. Alhussein, and T. M. Khan, "LMBiS-Net: A lightweight bidirectional skip connection based multipath CNN for retinal blood vessel segmentation," *Scientific Reports*, vol. 14, no. 1, p. 15219, 2024, doi: 10.1038/s41598-024-63496-9.

[11] S. Balasubramaniam, S. Kadry, R. K. Dhanaraj, K. K. Satheesh, and C. Manthiramoorthy, "Res-Unet based blood vessel segmentation and cardio vascular disease prediction using chronological chef-based optimization algorithm based deep residual network from retinal fundus images," *Multimedia Tools and Applications*, vol. 83, no. 40, pp. 87929–87958, Mar. 2024, doi: 10.1007/s11042-024-18810-y.

[12] P. R. Kumar, B. Shilpa, R. K. Jha, and V. S. Chellibouina, "Spatial attention U-Net model with Harris hawks optimization for retinal blood vessel and optic disc segmentation in fundus images," *International Ophthalmology*, vol. 44, no. 1, p. 359, 2024, doi: 10.1007/s10792-024-03279-3.

[13] Preity, A. K. Bhandari, and S. Shahnawazuddin, "Soft attention mechanism based network to extract blood vessels from retinal image modality," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 7, pp. 3408–3418, 2024, doi: 10.1109/TAI.2024.3351589.

[14] Y. Wan *et al.*, "Retinal blood vessels segmentation with improved SE-UNet model," *International Journal of Imaging Systems and Technology*, vol. 34, no. 4, p. e23145, Jul. 2024, doi: 10.1002/ima.23145.

[15] S. Panchal and M. Kokare, "ResMU-Net: Residual multi-kernel u-net for blood vessel segmentation in retinal fundus images," *Biomedical Signal Processing and Control*, vol. 90, 2024, doi: 10.1016/j.bspc.2023.105859.

[16] P. C. Sau, M. Gupta, and A. Bansal, "Optimized ResUNet++-enabled blood vessel segmentation for retinal fundus image based on hybrid meta-heuristic improvement," *International Journal of Image and Graphics*, vol. 24, no. 3, p. 2450033, May 2024, doi: 10.1142/S0219467824500335.

[17] T. Zhang *et al.*, "Bio-inspired optimisation algorithms in medical image segmentation: a review," *International Journal of Bio-Inspired Computation*, vol. 24, no. 2, pp. 65–79, 2024, doi: 10.1504/IJBIC.2024.141449.

[18] 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," *Biomedical Signal Processing and Control*, vol. 79, p. 104087, 2022, doi: 10.1016/j.bspc.2022.104087.

[19] K. Radha and Y. Karuna, "Modified depthwise parallel attention UNet for retinal vessel segmentation," *IEEE Access*, vol. 11, pp. 102572–102588, 2023, doi: 10.1109/ACCESS.2023.3317176.

[20] N. Wang, K. Li, G. Zhang, Z. Zhu, and P. Wang, "Improvement of retinal vessel segmentation method based on U-Net," *Electronics (Switzerland)*, vol. 12, no. 2, 2023, doi: 10.3390/electronics12020262.

[21] R. Ortiz-Feregrino, S. Tovar-Arriaga, J. C. Pedraza-Ortega, and J. Rodriguez-Resendiz, "Segmentation of retinal blood vessels using focal attention convolution blocks in a UNET," *Technologies*, vol. 11, no. 4, 2023, doi: 10.3390/technologies11040097.

[22] H. Du *et al.*, "Retinal blood vessel segmentation by using the MS-LSDNet network and geometric skeleton reconnection method," *Computers in Biology and Medicine*, vol. 153, p. 106416, 2023, doi: 10.1016/j.compbiomed.2022.106416.

[23] M. S. Gargari, M. H. Seyedi, and M. Alilou, "Segmentation of retinal blood vessels using U-Net++ architecture and disease prediction," *Electronics (Switzerland)*, vol. 11, no. 21, p. 3516, 2022, doi: 10.3390/electronics11213516.

[24] Y. Yi, C. Guo, Y. Hu, W. Zhou, and W. Wang, "BCR-UNet: Bi-directional ConvLSTM residual U-Net for retinal blood vessel segmentation," *Frontiers in Public Health*, vol. 10, p. 1056226, 2022, doi: 10.3389/fpubh.2022.1056226.

[25] S. Hussain, F. Guo, W. Li, and Z. Shen, "DilUnet: A u-net based architecture for blood vessels segmentation," *Computer Methods and Programs in Biomedicine*, vol. 218, p. 106732, 2022, doi: 10.1016/j.cmpb.2022.106732.

[26] M. Abdel-Basset, W. Ding, and D. El-Shahat, "A hybrid harris hawks optimization algorithm with simulated annealing for feature selection," *Artificial Intelligence Review*, vol. 54, no. 1, pp. 593–637, 2021, doi: 10.1007/s10462-020-09860-3.

[27] S. Chatterjee *et al.*, "Retinal blood vessel segmentation using edge detection method," *Journal of Physics: Conference Series*, vol. 1717, no. 1, p. 012008, 2021, doi: 10.1088/1742-6596/1717/1/012008.

[28] G. A. Francia, C. Pedraza, M. Aceves, and S. Tovar-Arriaga, "Chaining a U-Net with a residual U-Net for retinal blood vessels segmentation," *IEEE Access*, vol. 8, pp. 38493–38500, 2020, doi: 10.1109/ACCESS.2020.2975745.

[29] P. M. Samuel and T. Veeramalai, "Review on retinal blood vessel segmentation - An algorithmic perspective," *International Journal of Biomedical Engineering and Technology*, vol. 34, no. 1, pp. 75–105, 2020, doi: 10.1504/IJBET.2020.110362.

[30] B. Wang, S. Wang, S. Qiu, W. Wei, H. Wang, and H. He, "CSU-Net: A context spatial U-Net for accurate blood vessel segmentation in fundus images," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1128–1138, 2020, doi: 10.1109/JBHI.2020.3011178.

[31] X. Guo *et al.*, "Retinal vessel segmentation combined with generative adversarial networks and dense U-net," *IEEE Access*, vol. 8, pp. 194551–194560, 2020, doi: 10.1109/ACCESS.2020.3033273.

[32] R. Kushol, M. H. Kabir, M. Abdullah-Al-Wadud, and M. S. Islam, "Retinal blood vessel segmentation from fundus image using an efficient multiscale directional representation technique Bendlets," *Mathematical Biosciences and Engineering*, vol. 17, no. 6, pp. 7751–7771, 2020, doi: 10.3934/MBE.2020394.

[33] Y. Cheng, M. Ma, L. Zhang, C. J. Jin, L. Ma, and Y. Zhou, "Retinal blood vessel segmentation based on densely connected U-Net," *Mathematical Biosciences and Engineering*, vol. 17, no. 4, pp. 3088–3108, 2020, doi: 10.3934/MBE.2020175.

[34] T. A. Soomro *et al.*, "Deep learning models for retinal blood vessels segmentation: A review," *IEEE Access*, vol. 7, pp. 71696–71717, 2019, doi: 10.1109/ACCESS.2019.2920616.

[35] P. Xiuqin, Q. Zhang, H. Zhang, and S. Li, "A fundus retinal vessels segmentation scheme based on the improved deep learning u-net model," *IEEE Access*, vol. 7, pp. 122634–122643, 2019, doi: 10.1109/ACCESS.2019.2935138.

[36] M. Li, Z. Ma, C. Liu, G. Zhang, and Z. Han, "Robust retinal blood vessel segmentation based on reinforcement local descriptions," *BioMed Research International*, vol. 1, p. 2028946, 2017, doi: 10.1155/2017/2028946.

[37] Y. Özüpak and S. Mansurov, "A hybrid deep learning approach for multi-output short-term electricity demand forecasting," *Concurrency and Computation: Practice and Experience*, vol. 37, no. 25–26, p. e70356, 2025, doi: 10.1002/cpe.70356.

[38] Y. Özüpak, "Analysis of the parameters affecting the efficiency of the wireless power transmission system designed for new generation electric vehicles," *International Journal of Automotive Technology*, vol. 24, no. 6, pp. 1675–1680, 2023, doi: 10.1007/s12239-023-0135-1.

[39] O. Ramos-Soto *et al.*, "An efficient retinal blood vessel segmentation in eye fundus images by using optimized top-hat and homomorphic filtering," *Computer Methods and Programs in Biomedicine*, vol. 201, 2021, doi: 10.1016/j.cmpb.2021.105949.

[40] C. Wang, Z. Zhao, and Y. Yu, "Fine retinal vessel segmentation by combining Nest U-net and patch-learning," *Soft Computing*, vol. 25, no. 7, pp. 5519–5532, 2021, doi: 10.1007/s00500-020-05552-w.

[41] Y. Zhou, B. Wang, L. Huang, S. Cui, and L. Shao, "A benchmark for studying diabetic retinopathy: segmentation, grading, and transferability," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 818–828, 2021, doi: 10.1109/TMI.2020.3037771.

BIOGRAPHIES OF AUTHORS

Saba Sheiba is affiliated with the Department of Computer Science and Engineering, BEST Innovation University, Andhra Pradesh, India, and the Department of Computer Science and Artificial Intelligence, Muffakham Jah College of Engineering and Technology, Hyderabad, India. Her research interests include artificial intelligence, machine learning, data science, and computer vision. She can be contacted at email: saba.sheiba@mjcollege.ac.in.

M. Neelakantappa is currently working as an Associate Professor at the Vasavi College of Engineering, Hyderabad, India. His research interests and academic contributions are focused on electrical, electronic, and computer engineering domains. He can be contacted at email: m.neelakanta@gmail.com