International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Comparison of PSO-based U-Net and SegNet for Automatic Fundus Image Segmentation

Audrey Huong

Submitted: 28/01/2024 **Revised**: 06/03/2024 **Accepted**: 14/03/2024

Abstract: The retinal vascularization morphology can be an important cue for various eye pathologies. Past studies have focused on exploring complex image processing and enhancement methods to improve vessel segmentation and detection for screening eye diseases. This research explores the potential of the Particle Swarm Optimization (PSO) method for optimizing segmentation of vessel images without requiring expensive data or computing resources. This optimization framework searches for important hyperparameters for the efficient training of deep-convolutional U-Net and SegNet on a small fundus dataset. The comparison results showed that U-Net achieves better segmentation of fundus photographs with mean overlap measures of 0.74-0.83 than its competing model. A comparison with the state-of-the-art methods showed considerably high classification accuracy and sensitivity scores ranging from 0.93-0.98 were achieved by the proposed networks. This study identified the insufficiency of the employed data augmentation strategies as the main factor responsible for the poor segmentation sensitivity of 0.52-0.63. Future works include optimizing network parameters and adopting effective image preprocessing processes to improve the detection results.

Keywords: *Blood vessel, Fundus, Particle-swarm optimization, Segmentation*

1. Introduction

Recent progress in understanding vasculature-related diseases, known as the risk factors for many noncommunicable diseases (NCDs), such as cancer, diabetes, hypertension, congestive heart failure, and stroke, is focused primarily on investigating microcirculation disturbance [1, 2]. Diabetic retinopathy (DR), hypertensive retinopathy, retinal vein and artery occlusions, and wet age-related macular degeneration (AMD) are among the NCD-related vascular diseases affecting the eyes [2, 3]. These vascular-related eye pathologies can result in total blindness if not properly managed. These diseases change vascular network morphology and physical features, which are used as key traits and markers for diagnosing and understanding disease progression [4]. These characteristics are also used in designing treatment and prevention interventions.

The standard imaging modality uses fundus photographs to examine the optic disc, retinal vessels, and macula in the vitreous cavity to identify anatomic signs of abnormal vasculature within the inner retina when screening for vision-threatening diseases [5]. This physical examination is performed based on real-time video streaming or recorded photographs; the process is manual, time-consuming, highly subjective, and depends on the skills and experience of the medical experts [3, 5]. General physical network characteristics of the vasculature, such as vessel shape, tortuosity, and size, are commonly associated with pathologic development in the study of disease processes. These investigations emphasized the need for in-situ microenvironment visualization to systematically reveal the underlying pathophysiological mechanism and uncover predictive biomarkers for disease progression [6]. Clinical suspect is confirmed using contrast dye injection to visualize functional microvascular networks [7]. Therefore, the procedure lacks patient acceptance.

Many researchers in this domain have explored artificial intelligence (AI) techniques for instance classification or disease prediction based on the input image. Recent efforts have included identifying fundus vessel regions for understanding retinal disease progression using deep learning-based segmentation approaches [4, 8-11]. The studies in [4, 8, 10-12] employed convolutional-based deep learning networks, the state-of-the-art methods, for fundus vessel segmentation. Unlike the one-stage detection model, i.e., YOLO and its variants and single-shot detector (SSD), convolutional-based segmentation networks can effectively detect sparse and small objects [13]. This research is important for assisting medical professionals in their diagnosis based on vascular morphology and changes, hence increasing their confidence in making decisions. To improve the detection accuracy of deep learning models, [8, 10-11] proposed using a complex image preprocessing process to enhance visualization contrast before extracting relevant image features for learning tasks. Image enhancement and filtering techniques can improve visual quality, but they also accentuate noise and introduce artifacts and distortions in the image [14], producing over/underenhanced results. In addition, these preprocessing operations can lead to high memory storage and computing resource demands. Therefore, it is challenging to obtain a large amount of high-quality data to guarantee the learning results of efficient deep-learning models. Similarly, Singh *et al.* [4] suggested an elaborate preprocessing process to extract small patches randomly from the original images to enlarge the training dataset while preserving the image quality. Shi *et al.* [12] introduced image quality assessment and strict selection criteria to overcome varying image quality from combining different public datasets for generalization. Considering the importance of this practical issue, an innovative

Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia
ORCID ID : 0000-0002-4505-5860
Corresponding Author Email: audrey@uthm.edu.my

solution is to use optimization algorithms, such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Bayesian Optimization (BO), to increase a deep learning model's generalization performance when trained on a limited and inferior-quality data [15-16]. The ability of optimization to enhance fundus segmentation performance has yet to be extensively explored, especially given the practical, real-world issues of limited labelled imaging data such as fundus photographs. Such an important goal may be achieved by exploiting the effective convergence of PSO, whose superiority was proven in [15], in training different segmentation models from scratch to detect vessel pixels in a fundus image. The outcomes can be useful in understanding and predicting the clinical progression of vascular-related diseases without the expense of expensive medical data and computing resources.

2. Materials and Methods

The following subsections describe the dataset used for the study and preprocessing of the data, the segmentation models, and the strategy adopted to optimize the network learning efficiency. All the image processing and analysis procedures were implemented with program codes written in MATLAB R2023a and executed on a single GPU (Tesla K80, 256 GB RAM).

2.1. Retinal Vessel Dataset and Image Preparation

Unlike signal or image classification or prediction tasks, segmentation is a more complex and computationally expensive process that requires meticulously annotated ground truths to classify each image pixel. The annotated ground truth data, created through a human-annotated process, can be difficult to obtain. The fundus image quality can be compromised by different factors, including poor illumination, motion artifacts, and glare, so at least two trained professionals typically carry out the manual process to increase confidence in their decision. Changes in retinal blood vessels are among the most common characteristics for detecting vascular malignancies. Despite different resources of retinal photographs being available, these datasets usually comprise a small number of samples of different image qualities, which are insufficient for training deep learning models. For simplicity and reproducibility, this work used only the Retinal Images vessel Tree Extraction (RITE) dataset [17] downloaded from Kaggle repository to validate the proposed strategy. This public dataset contains one hundred original fundus images and their ground-truth (GT) masks labeled with the location of blood vessels; some examples are shown in Fig. 1.

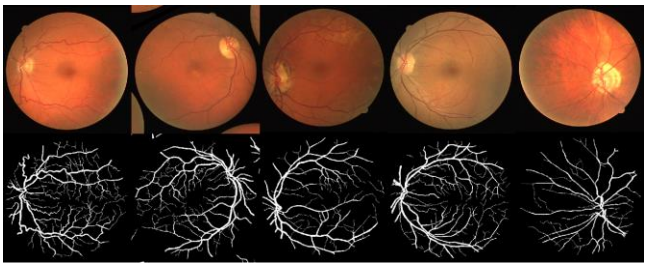


Fig. 1. Example of RITE training images (top) and the corresponding ground truth mask (bottom).

Fig. 2 shows the overall data flow diagram. The original RITE dataset contained 80 training images, while the validation and testing sets contained 10 images each. The original data are color (RGB) PNG images with dimensions of 512 × 512 pixels, and their ground truth mask consists of 2D image pixels with grayscale intensities. Some of these images and their masks contain segments of recursive and repetitive pictures that appear at the corner of the photograph, as illustrated in the second image of Fig. 1. Thus, image cleansing was carried out, wherein each image and the corresponding mask were manually screened for their presence and blocked using manual manipulation (covering the regions with black pixels) in Fig. 3(a). The processed images were saved in PNG format. Since this is a two-class problem, binary masks were generated by setting the object pixels in the image with grayscale values smaller than a threshold of 0.01 as white with a value of "0" and the rest as black with a pixel value of "1".

For stricter testing, twenty images from the original training set were randomly chosen and moved to the test set, giving a training/validation/testing dataset split of 60/10/30 %. Since this dataset is small, image augmentation techniques have been adopted to improve the model generalization ability by enriching the training and validation datasets. The employed augmentation methods included horizontal (Hor.) and vertical flips and four angle rotations, ± 10° and ± 20°, respectively, which produced six additional images from the original fundus image, as shown in Fig. 3(b). All images (original and augmentation) were resized to 128 × 128 × 3 using an automatic program written in MATLAB. The same geometrical transformation is applied to their corresponding

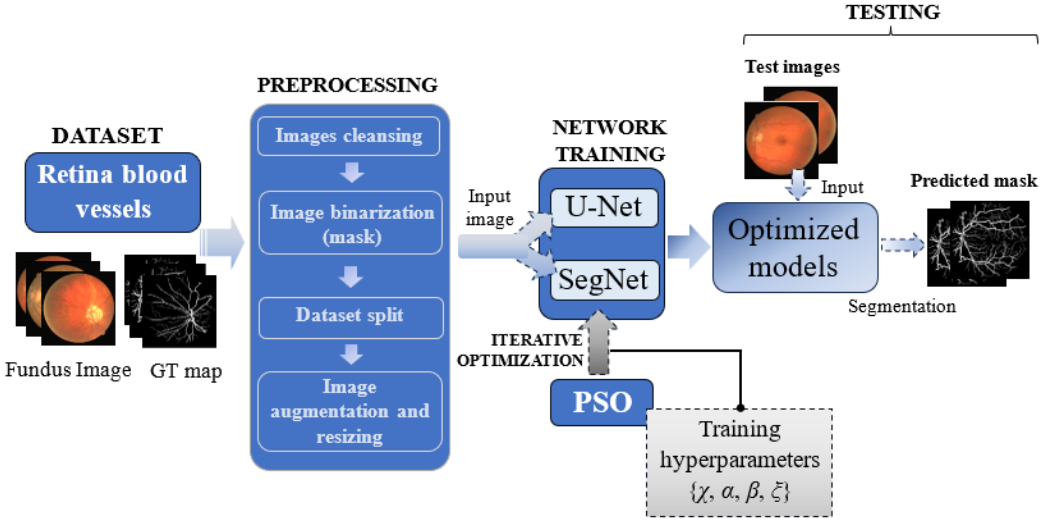
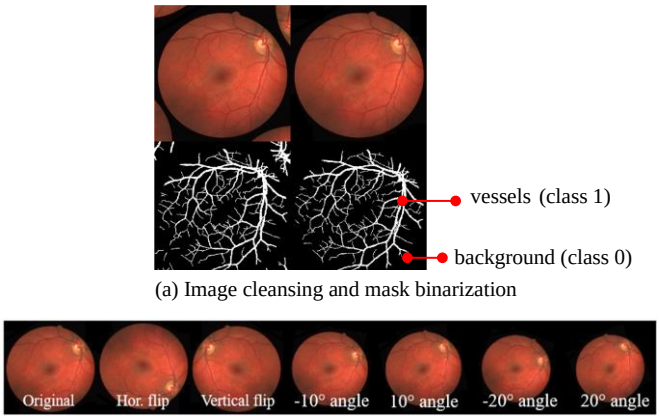


Fig. 2. The overall data workflow.

mask. This size was chosen because it can fit into the GPU used for training. The preprocessed data were fed into the chosen networks' input layer for learning the segmentation features.



(a) Image cleansing and mask binarization

(b) Augmented images

Fig. 3. (a) RITE images cleansing and (b) the produced augmented images.

2.2. Segmentation Models

This study considered fully convolutional U-Net and SegNet models for vessel segmentation tasks. These networks use the encoder-decoder architecture shown in Figs. 4 and 5 to learn the pixel-level feature representations from scratch based on the input image. The encoder depth of the proposed network was arbitrarily chosen to be five, producing 70-layer and 73-layer networks for U-net and SegNet, respectively.

In the encoder stage, the input image of dimensions 128 × 128 × 3 is reduced by maxpooling as it passes through five dense blocks (shown on the left of the figures) to extract various features; then, localization information is restored in the decoding stage. There is a difference in the image restoration technology used in these techniques. U-Net adopts skip connections to aggregate feature maps in the encoding stage with the decoding section using depth concatenation (DepCon) layers to recover spatial information, whereas SegNet uses maxpooling indices (or convolution processes) to upsample low-resolution feature maps during the decoding stage. Their encoder and decoder networks are connected through bridge blocks, as shown in Fig. 4, and pooling operations (pooling-unpooling), as shown in Fig. 5. The last convolutional layer in the decoder stage is fed into the Softmax layer for two-class labeling. The output is a binary-class mask. Also indicated in the figures is the last ReLU layer with 64 channels output, whose activations are chosen for comparison and discussion in section 3.

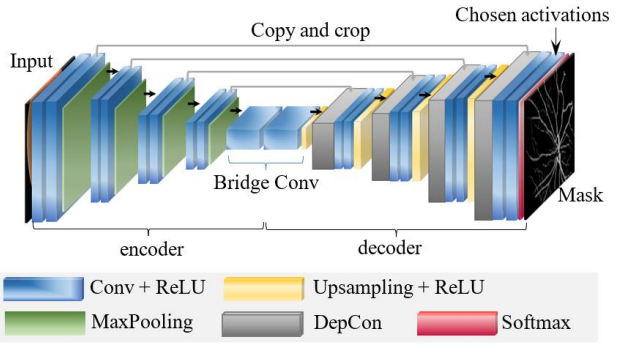


Fig. 4. U-Net architecture and concatenation process for feature maps reconstruction.

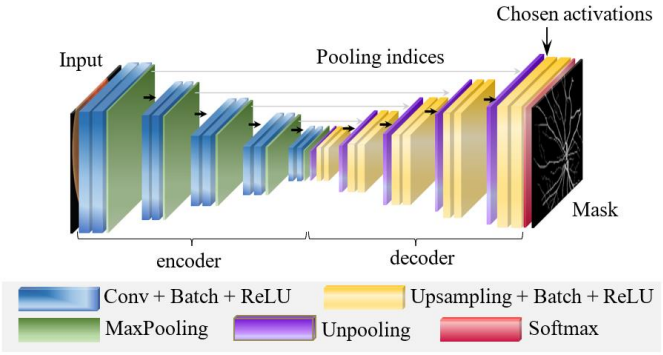


Fig. 5. SegNet architecture and transferred pool indices process for upsampling.

2.3. PSO-based Network Training

The learning efficiency of convolutional networks depends on the appropriateness and combination of the chosen training hyperparameters. Several training hyperparameters can be adjusted, but the most significant parameters are the training optimizer, number of epochs, mini-batch number, and initial learning rate. The traditional means of determining these values are using a manual grid or random search method, which starts with a random initial point in the search space before each parameter value is manually adjusted based on the resulting model training and validation results. This method is tedious and time-consuming, and it requires knowledge of the search objective and its changes as the search moves toward the space where optimal hyperparameters may be located.

An efficient and more probable method is to incorporate optimization techniques in searching for these variables. PSO was shown in [15] to perform remarkably well with comparatively faster convergence for the same solution quality compared to other commonly used methods. This method improves deep learning networks' learning efficiency with considerably lower computing time. This work focuses on the search for the optimizer type (χ), epoch number (α), mini-batch size (β), and initial learning rate (ξ). The process began with twenty particles (i.e., solution points) randomly launched into the four-dimensional search space, whose limits are defined as follows: the χ varies from 1 to 3 (value '1' represents *Adam*, '2' denotes *Sgdm*, and '3' is *RMSProp*), the α value range is between 50 and 200, the lower and upper limits of β and ξ are 32 and $1e^{-4}$, and 256 and $1e^{-1}$, respectively. Each particle is allowed to iterate five times around its position based on the results of the objective function defined in (1). This function is made up of the sum of errors in training ($100\text{-}A_T$) and validation accuracies ($100\text{-}A_v$) and training time (t). Different penalty weights are assigned to each factor, and the validation error with a weight factor of 10^3 dominates the penalty cost. The solution that yields the best-fitted value (i.e., lowest function value) among its previous performance, P_{best}, and best global performance among its neighbor, G_{best}, is chosen as the reference point for the subsequent iteration, where the position and velocity of the particles are stochastically updated toward the corresponding point. The termination criteria are either when there is no improvement in the evaluated validation accuracies for over 20 iterations or when the maximum number of iterations is reached. This iterative optimization process is summarized in Fig. 6.

$$f(A_T, A_V, t) = (100 - A_T)10^2 + (100 - A_V)10^3 + t/100 \quad (1)$$

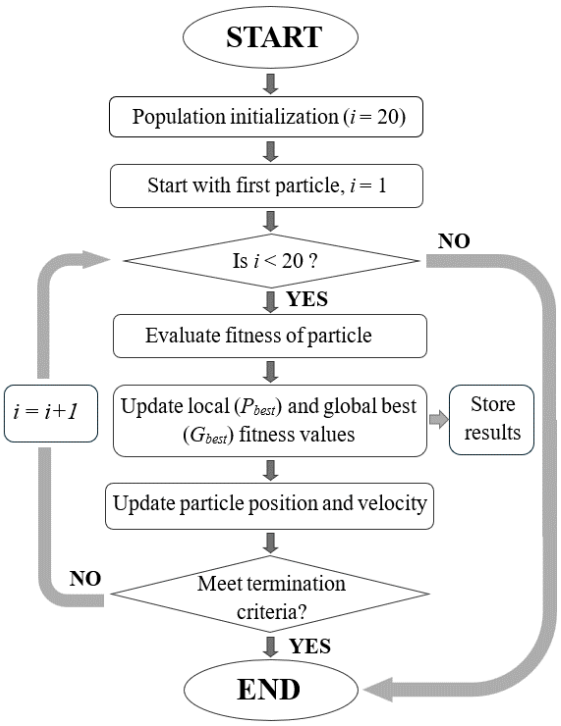


Fig. 6. PSO optimization flow diagram.

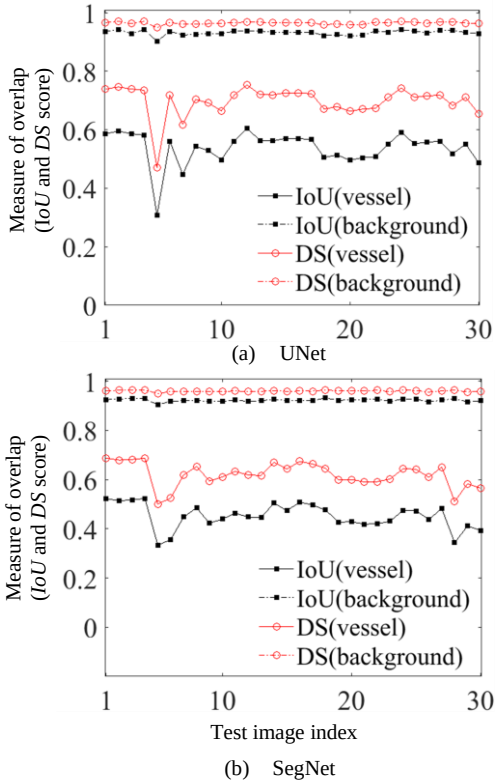


Fig.7. *IoU* and *DS* scores for (a) UNet and (b) SegNet evaluated on the test set.

3. Results and Analysis

The best hyperparameter solutions from the iterative process in Fig. 6 for U-Net and SegNet are χ = *RMSProp*, α = 166, β = 81, and $\xi = 1e^{-4}$, and χ = *Sgdm*, α = 84, β = 52, and ξ = 0.1, respectively, using the RITE dataset. The networks in Figs. 4 and 5 optimized based on these settings were evaluated using an independent testing set that has no role in the training and validation process. The effectiveness of their detection is examined by measuring the area of overlap between the ground truth mask and their predicted mask. For this purpose, the intersection over union (*IoU*), also known as the Jaccard index, Dice similarity coefficient (*DS*), and pixel classification accuracy, precision, sensitivity, and specificity metrics shown in (2) - (7) are used to assess the performance of binary semantic segmentation networks. The correctness of the segmented background and object (i.e., vessel) pixels evaluated on all test images are shown in Fig. 7.

$$IoU\ (GT, Pred) = \frac{|GT \cap Pred|}{|GT \cup Pred|} \quad (2)$$

$$DSC\ (GT, Pred) = \frac{2(GT \cap Pred)}{GT + Pred} \quad (3)$$

$$ACC = \frac{(TP + TN)}{(FN + FP + TP + TN)} \quad (4)$$

$$SENS = \frac{TP}{(TP + FN)} \quad (5)$$

$$SPEC = \frac{TN}{(TN + FP)} \quad (6)$$

$$PREC = \frac{TP}{(TP + FP)} \quad (7)$$

GT ∩ *Pred* and *GT* ∪ *Pred* represent the intersection and union of the ground truth (*GT*) and predicted (*Pred*) masks, respectively. *TP* and *TN* represent the correct classification of vessel and background pixels, respectively. *FN* is an incorrect classification of vessel pixels as background, while *FP* is the opposite of *FN*.

The mean and standard deviation (SD) of the *IoU*, *DS*, and other metrics in (4)-(7) across all test images in Fig. 7 are calculated and tabulated in Table 1. The average inference time recorded during the testing stage for U-Net and SegNet is given by 0.01 and 0.013 seconds, respectively. The best and worst-performing images identified based on the average highest and lowest *IoU* and *DS* performance (combining the background and object detection results) are plotted in Figs. 8 and 9, respectively. The boundaries of the GT mask and the mask predicted by the optimized models are also shown in the figures. The GT and predicted masks of the worst-performing image are compared for the strongest activations chosen from the output of the ReLU layer of the networks (indicated by an arrow in Figs. 4 and 5) in Fig. 10. The total search and training process involving 220 search iterations took 73 and 22 hours to complete in optimizing the weights of U-Net and SegNet, which contained 124.3 million and 0.67 million parameters, respectively. Meanwhile, Table 2 compares the performance of our optimized networks against several other researchers using

Table 1. Detection performance of the optimized U-Net and SegNet

Model	Performance metrics (mean ± SD)					
	IoU	*DS*	ACC	SPEC	SENS	PREC
U-Net	0.74±0.03	0.83±0.02	0.94±0.01	0.98±0.01	0.63±0.09	0.81±0.08
SegNet	0.7±0.02	0.79±0.03	0.93±0.01	0.98±0.01	0.52±0.08	0.80±0.07

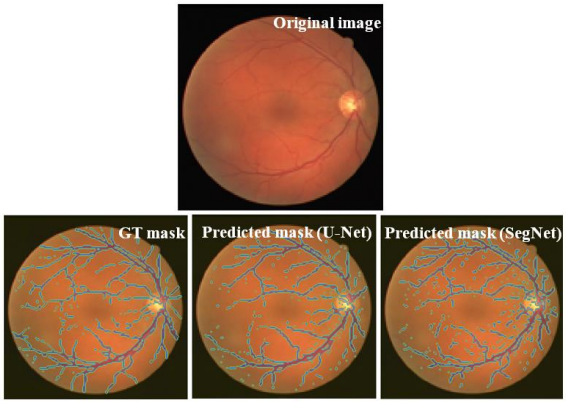


Fig. 8. The best performing image (index 4), its GT, and mask predicted by U-Net and SegNet.

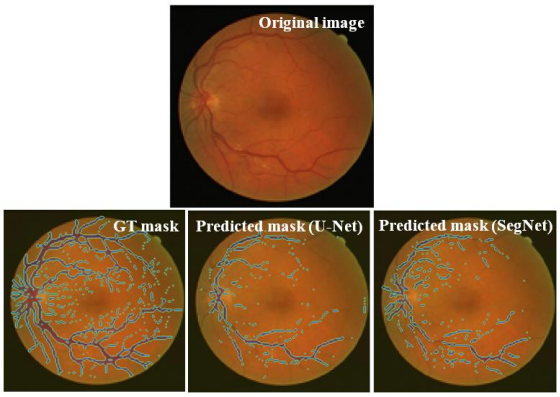


Fig. 9. The worst performing image (index 5), its GT, and predicted masks by U-Net and SegNet.

different strategies.

4. Discussion

This study is motivated to overcome one of the common challenges in machine learning of scarce labeled data by enhancing the learning performance of deep learning networks using the PSO method. The method is included in training the employed U-Net and SegNet to improve the models' generalization ability in segmenting blood vessels from fundus images by optimizing the best hyperparameter settings. This strategy is less complex and offers a faster solution to the problem than state-of-the-art techniques, such as combining different image enhancing and preprocessing algorithms, and training set enlargement processes. The current experiment adopted a stricter data split ratio compared

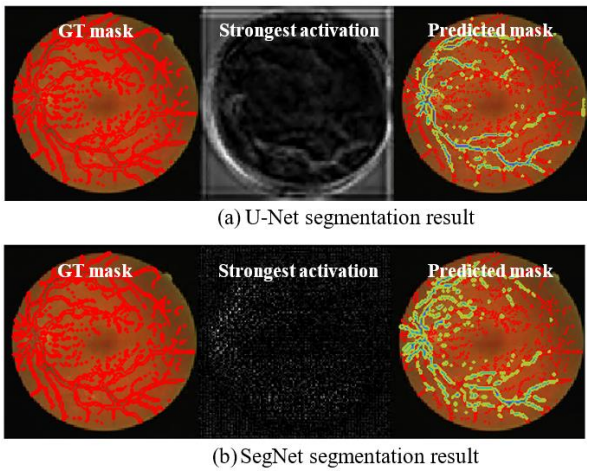


Fig. 10. (Left) Worst performing image (index 5) and the GT mask, (center) highest activation channel, (right) GT and predicted masks overlaid on the original image.

to the previous works in Table 2 to test the robustness of the proposed optimization framework and evaluate sufficiency of the augmentation methods. Table 1 shows consistent segmentation accuracy, precision, and specificity performance, ranging between 0.8 and 0.98 for the networks used. The recognition of fundus background pixels is much greater than that of blood vessels, as shown in Fig. 7, giving an acceptable overlap area of *IoU* and *DS* >0.7 achieved by U-Net and SegNet. A comparison with the existing studies in Table 2 also shows that the PSO-incorporated training process produced good model generalization ability comparable to that of recent studies adopting different techniques, with segmentation accuracy and specificity scores reaching 0.98. However, sensitivity scores that evaluate correctness in detecting vessel location (*TP*) in these tables, ranging between 0.52 and 0.63, suggest room for improvement in the models' vessel detection performance. This study shows that the U-Net in Fig. 4, whose size is approximately 185 times larger than that of the employed SegNet, adopted depth concatenation technology that achieved noticeably better segmentation scores, especially with higher sensitivity, as shown in Table 1, than did the pooling technology. Both the proposed networks are shown in Fig. 8 to perform well on images with good contrast and detail. On the contrary, noisy fundus image with poor-quality details of thin and tiny vessels in Fig. 9 prevents the precise localization of vessels, rendering poor detection sensitivity. Based on the highly activated regions (white pixels) of the activation map of this input image in Fig. 10, it is possible that the encoder-decoder layers in U-Net in Fig. 4 learned color, contrast, and edge as among the important features. In

Table 2. Comparison with the recent studies on fundus vessel segmentation performance

Study	Strategy	Average performance metrics			
		ACC	SPEC	SENS	PREC
Girard *et al*. [8]	Contrast-wise global enhancement and local density normalization	0.957	0.98	0.78	-
Ma *et al*. [10]	Spatial activation and Gaussian pixel enhancement	0.95	0.98	0.69	-
Morano *et al*. [11]	Global enhancement and Gaussian filtering	0.96	0.98	0.79	-
Proposed	PSO optimization in network training	0.93 -0.94	0.98	0.52-0.63	0.8-81

contrast, SegNet focuses more on vessel morphology, producing a grainy activation map in subplot (b) of the figure. While both models performed inferiorly on the poor contrast and low illumination images, SegNet was comparatively more effective than the U-Net in locating the vessels under these conditions. The employed SegNet has a comparatively smaller size (~0.67 million learnable), so the *Sgdm* technique was found to help the model generalize better by lowering the convergence speed to reduce the possibility of underfitting. Hence, prompting a larger initial learning rate of 0.1 being chosen during the optimization process to increase the model learning rate. Similarly, the same reasoning applies to U-Net. The *RMSProp* optimizer, which uses adaptive learning rates to update the network, is superior in model training speed. Therefore, a large mini-batch sizes of 81 and a small initial learning rates of $1e^{-4}$ were identified as the best combination for training the large U-Net model.

Although search optimization and augmentation strategies have been adopted here to improve model generalizability, the even better results of the previous studies in Table 2 in locating vessel regions is attributed to two main factors: (1) the use of image contrast enhancement techniques to improve vessel visualization, and (2) the use of rich datasets by combining images from other public sources, namely the AV-DRIVE, INSPIRE-AVR, and high-resolution fundus (HRF) datasets. The improved image quality and quantity (i.e., rich variability in retinography) prevent network overfitting to the background, which is the dataset's most prominent class.

In addition, due to the limitations of the computing memory and resources of our workstation, this study reduced the original image size by a scale of 4 during training to meet the memory requirements. Although it reached a fast inference speed (~0.01 s) during testing, this operation, compounded with augmentation processes, causes geometric distortion and loss of information in the transformed image shown in Fig. 3(b), where circular regions of interest are visibly reduced, and fine capillary vessels appear blurred or disappear, especially after rotational augmentation. These processes reduce image spatial resolution, and obscure image detail information, thus negatively impacting the detection task. Previously Ma *et al.* [10] overcame a similar problem using a novel patch-stitch method by segmenting the original image into several regions to retain the information in the original image without compromising the image quality to fit into the employed computing system. This operation also enriches the training dataset, producing promising accuracy using less computing memory. This strategy may be further explored in training to improve segmentation performance.

The PSO optimization strategy proposed here is a simple, cost-effective, and reproducible alternative to improve segmentation accuracy in fundus vessel detection for better understanding and prognosis of retinal diseases. This automatic process took approximately 1-3 days to complete using our workstation without manual human intervention, making it a potential solution for optimizing the model's learning efficiency. However, several limitations have been identified in the adopted approach. First, despite the larger training data size from physical transformations using image flipping and rotational techniques, the produced images are not loss-free, and most images are degraded in quality through the procedure. Second, even though some of the fundus images are noisy, low-quality in nature, and acquired under varied lighting conditions, this study did not include any image enhancement to exclude factors that could influence the outcome in validating the proposed framework. Third, the rigid ReLU used in the proposed networks hinders the passing of negative values, which, if preserved, could improve the expressiveness of the activation function.

The future scope of this work is to extend this study by further exploring adaptive image enhancement and denoising algorithms and modifying the architecture of the employed networks or improving the network dynamics to enable the model to learn more complex patterns and distinct features in even a small amount of data. The presented results are promising, and the proposed system holds great potential in clinical decision support, especially in examining changes in fundus vessel morphology and physical traits with the progression of retinal vascular diseases

5. Conclusion

Unlike previous studies that focused on improving image segmentation performance using rich datasets and tedious and complex image preprocessing techniques, this study effectively improved the generalization of convolutional-based deep learning networks by using PSO to train the employed models. This cost-effective strategy is practical in real applications when data and computing resources are limited. Its performance tested on the proposed U-Net and SegNet showed variability in their strongest activation maps in fundus vessel detection. This implies that they recognized different features in the image for classification. Although large U-Net has better overall segmentation performance than SegNet, and both models achieved favorable and comparable results to those of conventional methods, inadequate image handling and insufficient preprocessing have been recognized as areas that could be further addressed to improve their recognition accuracy. These future research directions include adopting adaptive image preprocessing and dynamic network representation learning approaches before the system can be deployed for clinical research.

Acknowledgments

Communication of this research is made possible through monetary assistance by Universiti Tun Hussein Onn Malaysia and the UTHM Publisher's Office via Publication Fund E15216.

References

[1] J. Yang and M. Li, "Epidemiology of Noncommunicable Diseases," In: Wang, C., Liu, F. Textbook of Clinical Epidemiology. 1st ed. Singapore: Springer, 2023.

[2] W.S. Wright, R.S. Eshaq, M. Lee, G. Kaur, and N.R. Harris, "Retinal Physiology and Circulation: Effect of Diabetes," Compr. Physiol., vol.10, pp. 933-974, 2020.

[3] T.H. Rim, A.W.J. Teo, H.H.S. Yang, C.Y. Cheung, and T.Y. Wong, "Retinal Vascular Signs and Cerebrovascular Diseases," J. Neuro-Ophthalmol., vol.40, pp. 44-59, 2020.

[4] L.K. Singh, M. Khanna, S. Thawkar, and R. Singh, "Deep-learning based system for effective and automatic blood vessel segmentation from Retinal fundus images," Multimed Tools Appl., vol. 83, pp. 6005–6049, 2024.

[5] J.P.O. Li, H. Liu, D.S.J. Ting, S. Jeon, R.V.P. Chan, J.E. Kim JE, et al. "Digital technology, telemedicine and artificial intelligence in ophthalmology: A global perspective," Prog. Retin. Eye Res., vol. 82, pp. 1-32, 2021.

[6] R.K. Wang and S.L. Jacques, "Innovative Optical Technologies in Ophthalmology and Eye Research," J. Ocul. Pharmacol. Ther., vol. 31, no. 3, pp. 142-142, 2021.

[7] J.X. Ong and A.A. Fawzi, "Perspectives on diabetic retinopathy from advanced retinal vascular imaging," Eye (Lond), vol. 36, pp.

319-327, 2022.

[8] F. Girard, C. Kavalec, and F. Cheriet, “Joint segmentation and classification of retinal arteries/veins from fundus images,” Artif. Intell. Med., vol. 94, pp. 96-109, 2019.

[9] A.Z.M.E. Chowdhury, G. Mann, W.H. Morgan, A. Vukmirovic, A. Mehnert, F. Sohel, “MSGANet-RAV: A multiscale guided attention network for artery-vein segmentation and classification from optic disc and retinal images,” J. Optom., vol. 15, pp. 58-69, 2022.

[10] W. Ma, S. Yu, K. Ma, J. Wang, X. Ding, and Y. Zheng, “Multi-task Neural Networks with Spatial Activation for Retinal Vessel Segmentation and Artery/Vein Classification,” In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Lecture Notes in Computer Science; Springer, 2019.

[11] J. Morano, A.S. Hervella, J. Novo, and J. Rouco, “Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images,” Artif. Intell. Med., vol. 118, pp. 102116, 2021.

[12] D. Shi, Z. Lin, W. Wang, Z. Tan, X. Shang, X. Zhang X, et al., “A Deep Learning System for Fully Automated Retinal Vessel Measurement in High Throughput Image Analysis,” Front Cardiovasc. Med., vol. 22, pp. 823436, 2022.

[13] Setyanto, T.B. Sasongko, M.A. Fikri, and I.K. Kim, “Near-Edge Computing Aware Object Detection: A Review,” IEEE Access, vol. 12, pp. 2989-3011, 2024.

[14] X. Liu, M. Pedersen, and R. Wang, “Survey of natural image enhancement techniques: Classification, evaluation, challenges, and perspectives,” Digit. Signal Process, vol.127, pp. 1-40, 2022.

[15] Huong, K.G. Tay, K.B. Gan, and X. Ngu, “A hierarchical optimisation framework for pigmented lesion diagnosis,” CAAI Trans. Intell. Technol., vol. 7, pp. 34–45, 2022.

[16] Akay, D. Karaboga, and R. Akay, “A comprehensive survey on optimizing deep learning models by metaheuristics,” Artif. Intell. Rev., vol. 55, pp. 829–894, 2022.

[17] Q. Hu, M.D. Abràmoff, and M.K. Garvin, “Automated separation of binary overlapping trees in low-contrast color retinal images,” Med. Image Comput. Assist. Interv., vol. 16, pp. 436-43, 2013.