JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-21, No.-09, September (2026) pp 100-123 ISSN (Print) 0973-8975

A GLOBAL-LOCAL FEATURE LEARNING FOR RETINAL VESSEL SEsGMENTATION USING AN IMPROVED U-NET

Nadeem Akhtar[1], Patil Bhushan[2], Mahire Amit[3], Rahmani Naveed Akhtar[4], Mohammad Shakir[5]

[1, 2, 3] Department of Electrical Engineering, R. C. Patel Institute of Technology, Shirpur, Maharashtra, India.

[4] Department of E&TC, Rajiv Gandhi Institute of Technology, Mumbai, India.

[5] Department of Electrical Engineering, Jamia Institute of Technology, Akkalkuwa, Maharashtra, India.

Email: [1]nadeem.ansari@rcpit.ac.in, [2]bhushan.patil2@rcpit.ac.in, [3]amit.mahire@rcpit.ac.in, [4]rahmani.akhtar@mctrgit.ac.in, [5]shakirshaban@gmail.com

Corresponding Author: **Nadeem Akhtar**

https://doi.org/10.26782/jmcms.2026.09.00006

(Received: March 21, 2026; Revised: August 26, 2026; Accepted: September 10, 2026)

Abstract

The diagnosis of ophthalmic diseases such as glaucoma, diabetic retinopathy, hypertension, and cardiovascular disorders can be recognized through retinal vessel segmentation from fundus images. The low contrast, noise, blurred boundaries, and substantial variations in vessel thickness make the vessel segmentation a challenging task for semantic segmentation of retinal vessels. To mitigate these limitations, an enhanced U-Net variant is proposed. The framework integrates an initial Inception block, a sequential Global-Local (GL) feature engineering module utilizing zooming and shrinking operations, parallel 1×1 and 3×3 skip paths coupled through element-wise addition, and a terminal Atrous Spatial Pyramid Pooling (ASPP) stage. Internal representations across these stages were systematically profiled using Mutual Information (MI) metrics. Quantitative evaluations demonstrate the network's strong performance across standard benchmarks, recording an mIoU of 80.87%, a Mean BF-score of 91.89%, a center-line Dice (clDice) of 80.83%, and an 84.6% vessel classification accuracy alongside a reduced 15.4% false-negative rate. Ablation trials validate the individual architectural contributions: introducing the Inception layer increases IoU from 50.84% to 59.59%, whereas employing element-wise addition over concatenation in the GL module raises clDice to 81.04% (versus 80.72%). Additionally, MI profiling confirms that average pooling and multi-rate ASPP convolutions retain vital spatial entropy and micro-vascular geometry far more effectively than standard operations.

Keywords: Retinal Vessel Segmentation, U-Net, ASPP, CLAHE, Fundus Image.

I. Introduction

The retina is a complex, multi-layered neural tissue lining the back of the eye that consists of specialized photoreceptor cells, rods and cones, which sense the incoming light and transform to the equivalent electrical signals. These transformed signals are sent to the brain for visual perception (Van Gelder et. al., 2022). The structural and functional defects in the retina can disrupt the vision process and may lead to total or partial vision loss. Therefore, early detection and prompt treatment of retinal disorders are crucial in maintaining visual function. Although the retina is highly sensitive, and challenging to detect retinal abnormalities in its blood vessels. This highlights the importance of precise and dependable diagnostic methods (Van Gelder Michael F; Dyer, Michael A; Greenwell, Thomas N; Levin, Leonard A; Wong, Rachel O; Svendsen, Clive N, 2022). Retinal examination is carried out using a non-invasive imaging technique known as funduscopy. This enables the clinician to examine the posterior segment of the eye using a light source and an expert's magnifying lens. The captured fundus image provides a detailed view of important anatomical features including the retina, optic disc, choroid, and retinal blood vessels. Among these structures, the retinal vasculature is significant in early detection of both ocular and systemic diseases.

The precise segmentation of retinal blood vessels is essential for diagnosing and tracking ophthalmic conditions. In clinical practice, ophthalmologists examine the retinal arteries and veins, thickness, and branching patterns to identify retinal diseases in fundus images. The alterations in retinal vessels, such as narrowing, widening, increased twisting, or blockage, are early warning signs of fundus-related diseases (Spaide James G.; Waheed, Nadia K., 2015). The quantitative assessment of these vascular characteristics helps in detecting the abnormalities associated with vascular diseases. Retinal imaging helps in extracting and segmenting the blood vessels in fundus images (Godishala et al., 2025). This technique provides technical insight for diagnosing major eye conditions such as diabetic retinopathy, glaucoma, hypertensive retinopathy, and choroidal neovascularization (Prajna and Nath, 2022). Additionally, structural changes in the retinal blood vessels reflect systemic conditions, including cardiovascular disease and cerebrovascular incidents such as heart attacks and strokes. The important vascular parameters like vessel width, length, branching angles, tortuosity, and inter-vessel distance are analysed for disease progression and diagnosis (Hassan et al., 2015). Funduscopy is a non-invasive, valuable tool for evaluating ocular health, brain condition, and the cardiovascular system (Mardani and Maghooli, 2021). Many studies have shown a strong correlation between retinal vascular structure and systemic conditions such as hypertension, cardiovascular diseases, age-related macular degeneration, and stroke (Ramos-Soto et al., 2021). For instance, increased vessel width and tortuosity are associated with hypertension and retinopathy of prematurity (Toptaş Davut, 2021), while indicators such as arteriovenous nicking and the arteriolar-to-venular width ratio are recognized markers of hypertension and heart disease. In recent years, substantial research efforts have been directed toward improving retinal vessel segmentation techniques using fundus images (Boudegga et al., 2021). According to a recent report by the International Diabetes Federation, the prevalence of diabetes-related retinal conditions has increased by approximately 25% over the past

decade (Kawamura, 2024). This alarming trend highlights the urgent need for effective, accurate, and automated diagnostic tools to support large-scale screening and early intervention. Retinal blood vessel segmentation has been a focal point of extensive research endeavours owing to its paramount significance in the diagnosis and monitoring of a myriad of ophthalmic and systemic pathologies.

II. Related Works

The Early methods of retinal vessel segmentation were based on merging traditional image processing techniques with Convolutional Neural Networks (CNNs) to enhance vessel structure visibility. For example, Uysal and Emre 2020 proposes hybrid framework that integrates image pre-processing, data augmentation, and a CNN model. The pre-processing improves the uneven illumination and augmented vessel-background contrast, while augmentation increases the training dataset size. The proposed method improves the results in terms of a variety of metrics. The advancement from CNN-based approaches, U-Net (Das et al., 2024), becomes an emerging method in segmenting retinal vessel structures that uses the encoder and decoder structure with skip connections (Hossen et al., 2026). The U-Net-based model introduces three specialized pre-processing algorithms for improving the segmentation accuracy (Sanjeewani Arun Kumar et. al., 2024). The experimental results on the DRIVE dataset (Staal Michael D. et. al., 2004) reveal improvements in accuracy, sensitivity, specificity, and F1-score, thereby validating the appropriateness of U-Net variants (Bansal et al., 2026) for capturing intricate vascular structures. Apart from the structural modification in U-Net, the optimization methods play a key role in improving the segmentation performance. For instance, Barnacles Mating Optimizer is employed to enhance deep transfer learning efficacy in malaria detection, which indicates the optimization strategies could be adapted for retinal vessel segmentation to improve convergence speed and model efficiency. Recent research is more focused on refining the U-Net framework by using attention mechanisms and more sophisticated feature modelling approaches. For example, SE-UNet integrates the squeeze-and-excitation module to dynamically adjust channel-wise feature responses, helping to suppress non-vessel regions while emphasizing features relevant to blood vessels (Wan Gaofeng et. al., 2024). Correspondingly, CCS-UNet amalgamates ResNeSt blocks with cross-channel spatial attention to enhance inter-channel information flow, resulting in superior segmentation outcomes across multiple datasets (Zhu Xiang; Zhang, Xue-Dian; Jiang, Min-Shan, 2023). The U-Net performance is improved by using the attention-based mechanism specifically in challenging scenarios such as thin or low-contrast vessels. The influence of kernel size and receptive field design has also been a subject of investigation as critical determinants affecting segmentation performance. A large convolution kernel in U-Net is used by (Boudegga et al., 2021) to capture vessel pixels effectively along with their neighbouring context. Through experimental determination of optimal kernel sizes contingent upon image resolution, the method achieved notable outcomes on both the DRIVE (Staal Michael D. et. al., 2004) and HRF (Odstrcilik Radim et. al., 2013) datasets, significantly surpassing the standard U-Net in terms of accuracy and sensitivity. In (Jin Qi; et. al., 2019), various deep models including modified iterations of AlexNet (Krizhevsky, Sutskever and Hinton, 2017), ZF-Net (Zeiler and Fergus, 2013), VGG (Simonyan and Zisserman, 2015), and

Deformable-ConvNet were tailored for retinal vessel segmentation. By incorporating additional convolutional layers, the models alleviated noise caused by patch interpolation. Assessments conducted on the DRIVE and STARE (Hoover Valentina L.; Goldbaum, Michael H., 2000) datasets illustrated that CNN-based approaches attained elevated accuracy and Area Under the Curve (AUC) values, with Deformable-ConvNet exhibiting particularly robust performance.

The hybrid and multipath derivatives of the U-Net use the multi-scale and structural information. LadderNet (Juntang Zhuang, 2018)uses multiple parallel pathways to augment feature extraction, while IterNet (Li Manisha; Nakashima, Yuta; Nagahara, Hajime; Kawasaki, Ryo, 2020) enhances the segmentation precision by harnessing structural redundancies of vascular structures. Furthermore, the attention-based mechanism refinement based on attention gates (Oktay et al., 2018) and NFN+ cascaded U-Net structures (Wu Yong; Song, Yang; Zhang, Yanning; Cai, Weidong, 2020) are proposed to optimize the spatial and semantic features fusion. This fusion results in improved delineation of vascular structures. A Multichannel U-Net (M-UNet) (Dong Ting; Zhang, Tianyi; Wei, Lifang, 2022) was proposed, which integrates multi-scale equalization sampling to improve the vessel extraction accuracy. M-UNet processes all channels of the DRIVE and STARE image dataset to get more enriched features. The M-UNet achieves enhanced sensitivity and AUC, particularly in the identification of peripheral vessels. A comprehensive comparative analysis is done by (Singh Munish; Thawkar, Shankar; Singh, Rekha, 2023) using U-Net, DenseU-Net, LadderNet, R2U-Net, ATTU-Net, and an augmented R2-ATT U-Net on the STARE dataset. The LadderNet achieves the highest accuracy, while the enhanced R2-ATT U-Net shows commendable performance. MSTP-Net (Wang Xiaobo; Ma, Zhendi, 2025) uses a multi-scale design with three paths: local, global, and fusion, capturing both meticulous details and the principal structure of vessels. The network is tested and found to outperform on publicly available datasets. A U2Net (Chu Jinsong; Zheng, Wenqiang; Xu, Zhengyuan, 2025) based framework augmented with channel and spatial attention along with dilated convolutions achieves high accuracy and robust generalization across clinical datasets. In a broader context beyond fundus imaging, vessel segmentation methodologies have been extended to cover other retinal imaging modalities. SMFDNet (Li Fei et. al., 2025) combines spatial and multi-frequency domain features to improve vessel segmentation accuracy using OCTA images, which indicates robust performance in low-contrast and complex vascular regions across various OCTA datasets. RetinaDA dataset (Guo Shaojia et. al., 2025) integrates images from six distinct public datasets, thereby offering a comprehensive benchmark for assessing the domain adaptation and generalization capabilities of retinal vessel segmentation models. DC Net (Shang et al., 2024) is a lightweight three-layer architecture that leverages dilated convolutions to enlarge the receptive field without excessive pooling, thereby minimizing information loss. The model incorporates several innovative components, including the Positive Sequence Block (PSB) and Reverse Sequence Block (RSB) for contextual and spatial feature extraction, a Nonlinear Feature Extraction Module (NFEM) to enhance shallow feature representation, and a Feature Fusion Module (FFM) for effective multiscale feature integration. LUVS Net (Islam et al., 2023) is a lightweight encoder–decoder architecture based on U-Net, designed specifically for deployment on resource-

constrained devices. The model incorporates several key innovations, including group convolutions to reduce computational complexity, dual-stream information flow for enhanced feature learning, and skip connections to preserve spatial information. Additionally, the architecture minimizes redundant layers and optimizes the number of filters, significantly reducing parameters.

III. Materials and Methods

III.i. Pre-processing of DRIVE Dataset

Initially, the three channels of fundus images are extracted. It is clear from Fig. 1 that most of the information is within the green channel only. It provides superior contrast between blood vessels and the background, while red and blue channels exhibit higher noise levels and lower vessel-background information.

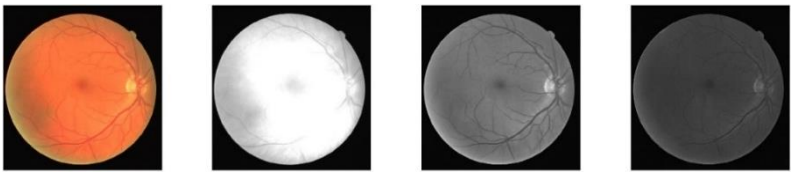

Fig.1. From left to right: Original, Red, Green, and Blue channels of the DRIVE dataset.

CLAHE is applied to the extracted green channel (Jebaseeli, Durai and Peter, 2019) to improve the visibility of vascular structures (Pizer et al., 1987). CLAHE improves the local contrast of retinal vessels and enhances foreground vessels from the background (Jebaseeli, Durai and Peter, 2019). Finally, the gamma correction with a gamma value of 0.5 is applied to further refine image contrast and emphasize the intensity differences around blood vessels. Gamma correction fine-tunes the low- and high-intensity pixel relationships by adjusting the gamma curve. The pre-processed images are shown in Fig. 2.

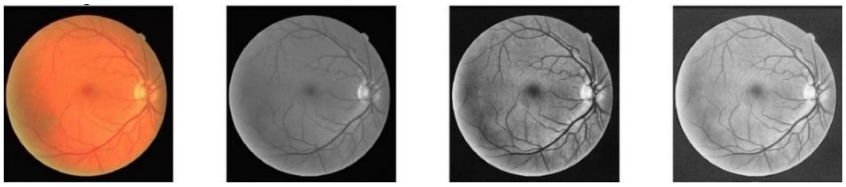

Fig. 2. Pre-processing from left to right: original, green channel, CLAHE, and gamma-corrected.

The training of CNN needs a large number of training images to mitigate overfitting and improve model generalization capability. The limited dataset images of the DRIVE dataset pose a challenge for effective model training. The DRIVE dataset has a total of 40 images; out of these, 20 are used for training and 20 for testing. The image size is large, so we have extracted patches of 256 × 256, but this will extract a limited number of patches; we have used an overlapping patch extraction technique to extract the patches. The half portion of the previous patch is taken as overlapped with the next patch. To further expand the dataset size, we have applied data augmentation using a zooming and shrinking method with a scaling factor of 2 and 0.75 for the training dataset only. The patches are extracted using an overlapped method from the augmented images. For testing images, patches are extracted without overlapping, and no data augmentation is applied. The patches with no information about retinal vessels

are discarded. The augmentation and filtering generate a total of 960, 296, and 80 patches for training, validation, and testing are used for model training.

III.ii. Experimental Model for Segmentation

The proposed U-Net architecture improves the feature representation and effectively recovers spatial details for retinal vessel segmentation. The proposed modification is shown in Fig. 3, which includes an Inception module as an initial feature extractor, a feature processing block, a zooming and shrinking module, and atrous spatial pyramid pooling. The proposed modifications in U-Net capture contextual information at different scales with fine structural characteristics of blood vessels. The inception module is used as an initial feature extractor, which learns features at various spatial scales using different kernel sizes of convolutions. The varying kernel sizes capture the vessels that appear with large variations in width, shape, and orientation.

The zooming and shrinking block after the max pooling learns the global and local features of the retinal vessels. The upsampling in the decoder stage uses bilinear interpolation that smooths feature maps and avoids the checkerboard artifacts that are often introduced by transposed convolution. The encoder features are connected to the decoder path through the feature processing block, which consists of element-wise addition of parallel 3×3 and 1×1 convolutions to improve feature consistency, vessel boundaries, and fine details during reconstruction. An Atrous Spatial Pyramid Pooling (ASPP) is connected at the final step of the network. The ASPP uses convolution at different dilation rates, which allows the network to simultaneously consider local details and global context.

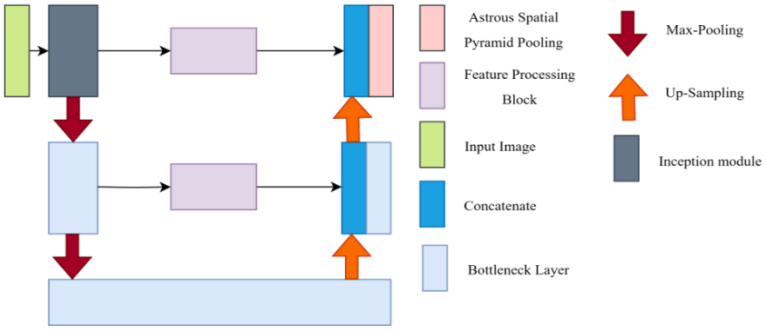

Fig. 3. Proposed architecture for retinal vessel segmentation.

III.ii.a. Inception Module as Initial Feature Extractor

The inception module was proposed by Google researchers in 2014 (Szegedy et al., 2015). The inception module extracts features at a multi-scale level within the single module, which enables learning the discriminative patterns with a smaller number of parameters. The initial features of retinal vessels are important for subsequent feature processing; Inception extracts the features are multiscale level using various convolution kernels and pooling the initial features, as shown in Fig. 4. The Inception module branches have 1×1 convolution layers for dimensionality reduction, 3×3 and 5×5 convolution to capture the spatial features at a multiscale level, pooling pathway that preserves the regional contextual information. The average pooling is added with max pooling paths; the combination of max and average pooling preserves both sharp,

critical edges and general, low-level contextual information is necessary for high accuracy (Zafar et al., 2022; Akhtar and Ragavendran, 2019). The output of average and max pooling is combined and processed by a 1×1 convolution. The output of all these branches is concatenated to produce a unified and rich feature representation for further processing. The ability to capture retinal vessel information at different spatial scales is important since retinal vessels have varying widths and appearance in fundus images. Retinal vessels vary greatly in thickness, orientation, and contrast, with thin capillaries requiring detailed local analysis and larger vessels benefiting from broader contextual awareness. The initial feature extraction at a multiscale level using the Inception module supports better learning in subsequent stages, which leads to improvement.

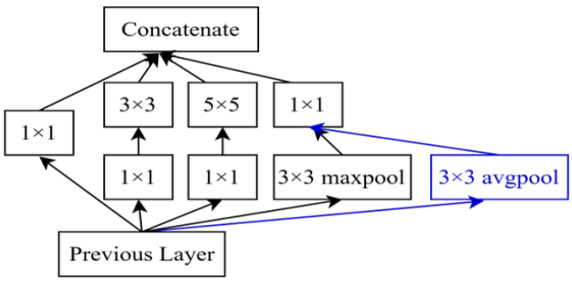

Fig. 4. Modified Inception module as initial feature extractor for retinal vessel segmentation.

III.ii.b. Global Local Feature Extraction Module

The retinal vessels have varying width; it is necessary to learn the global and local features of the vessels. The global and local feature learning module is shown in Fig. 5. This module has zooming and shrinking paths for learning the local and global (shape) features. The zooming and shrinking are achieved through upsampling and downsampling using bilinear interpolation and strided convolution. The zooming and shrinking paths zooms and shrink the image with half and doubling the number of features. The zooming path allows the network to focus on fine structural details, such as thin retinal vessels and subtle boundary information. Similarly, the shrinking path enables the extraction of global-level and shape features by capturing broader contextual patterns and structural relationships within the retinal image. The orange arrow in Fig. 5 indicates upsampling, while downsampling is indicated by the red arrow. The fusion module is the concatenation of zooming and shrinking paths with the original features processed with 1×1 convolution.

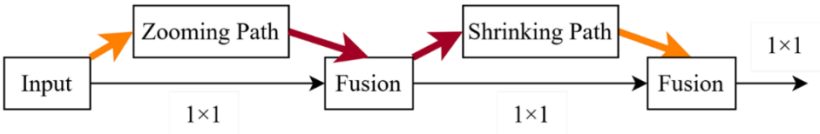

Fig. 5. Global Local Feature extraction module.

A second fusion stage is applied after the shrinking operation to further aggregate multi-scale information. In both fusion stages, addition is used to combine feature maps along the channel dimension. Finally, a 1×1 convolution layer is applied to the fused output to refine channel-wise representations and reduce dimensional redundancy. This

step ensures efficient feature integration while preserving the most discriminative information.

III.ii.c. Feature Processing Block

To enhance the quality of information transferred from the encoder to the decoder, a parallel feature processing block is connected as a skip connection (See Fig. 6). The two convolutions with kernel sizes of 3×3 and 1×1 are used to process the encoder features. The 3×3 convolution captures spatial context and local structural patterns that help in preserving the vessel boundaries and fine details. In contrast, the 1×1 convolution branch performs channel-wise transformation and feature recalibration. It helps to adjust feature dimensionality with representation efficiency improvement. The processed features are fused through element-wise addition rather than concatenation. As demonstrated by He et al. 2016 and Lin et al. 2017, element-wise addition preserves the tensor dimensions without increasing the channel depth.

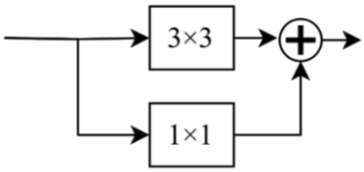

Fig. 6. Feature processing block.

III.ii.d. Atrous Spatial Pyramid Pooling

The ASPP He et al., 2015 illustrated in Fig. 7 processes the input feature with three parallel convolutional branches with dilation rates of 2, 3, and 4. These three dilated convolutions capture the contextual information at varying spatial scales. The dilated convolution enlarges the receptive field without affecting the number of parameters.

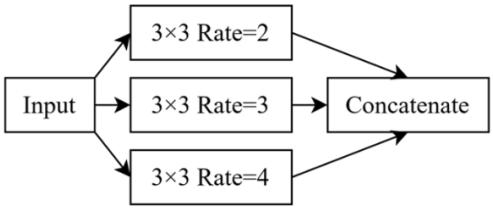

Fig. 7. Atrous Spatial Pyramid Pooling

III.iv. Performance Metrics

Intersection over Union (IoU) measures the overlap between the predicted and labelled regions relative to the total area covered by both regions. In other words, it assesses how much the labelled and predicted regions coincide with each other.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

The mean Intersection over Union (mIoU) is the average of all class IoU which represents the overlap across all the classes, providing an overall measure of segmentation performance.

The mean Boundary F1 (BF) score assesses the accuracy of predicted boundaries with labelled boundaries. It indicates how well the model detects the object edges. The BF score improves when boundary detection increases and segmentation accuracy reaches higher levels.

Centreline Dice (clDice) (Shit et al., 2021) is the metric used to measure the topological preserving similarity for tubular structures in segmentation. It measures the overlap to capture the topological connectivity, which is useful for verifying the branch completeness. It captures structure continuity in thin or branching vessels where small discontinuities can critically impact clinical interpretation (Yazıcı et al., 2026). It is defined as the harmonic mean of topology precision (Tprec) and sensitivity (Tsens) (Bai et al., 2026)

$$\text{clDice} = \frac{2 \times \text{T}_{\text{prec}} \times \text{T}_{\text{sens}}}{\text{T}_{\text{prec}} + \text{T}_{\text{sens}}}$$

This measures how much the centrelines overlap with the volumes. Tprec measures how much of the predicted skeleton overlaps with the actual volume, while Tsens checks how much of the real skeleton is included in the predicted volume (Sobisch et al., 2026).

IV. Results and Discussion

IV.i. Experimental Setup

The model's generalization ability is affected by the selection of hyperparameters. The simulation is carried out with the Adam optimizer, a piecewise learning rate with an initial learning rate of 0.001, and a drop factor of 20 epochs. The batch size we have used is 16, with a maximum of 100 epochs. The early stopping is used with the focal loss function. The proposed framework was implemented using MATLAB 2021.

IV.ii. Comparative Evaluation using Performance Metrics

Table 1 presents the normalized confusion matrix parameters for DC Net (Shang et al., 2024), LUVS Net (Islam et al., 2023), and the proposed network, respectively. DC Net indicates that a significant portion (41.5%) of retinal vessel pixels are misclassified as background, which indicates that it finds difficulty in identifying the thin and low-contrast vessels using this model. LUVS Net improves the vessel classification rate to 67.3%, while the misclassification of vessels as background is reduced to 32.7%. The proposed method shows that vessel classification accuracy is improved to 84.6%, while the misclassification rate of vessels as background is reduced to 15.4%.

Table 1: Normalized Confusion Matrix parameters.

Model	TP	TN	FP	FN
DC Net	58.5%	**99.5%**	**0.5%**	41.5%
LUVS Net	67.3%	98.9%	1.1%	32.7%
Proposed Method	**84.6%**	96.7%	3.3%	**15.4%**

Table provides the quantitative comparison between the proposed model and DC Net, LUVS Net across three evaluation metrics such mIoU, Mean BF Score, and clDice.

These metrics collectively evaluate region-level overlap, boundary fidelity, and vascular topological connectivity. As evaluated across all primary benchmarks, the proposed model demonstrates superior segmentation accuracy compared to both baselines, recording top values of 80.87 mIoU, 91.89 Mean BF Score, and 80.83 clDice.

Table 2: Comparative analysis using mIoU, Mean BF Score, and clDice

Model	mIoU	Mean BF Score	clDice
DC Net	75.38	80.71	67.10
LUVS Net	*78.39*	*87.46*	*73.80*
Proposed	**80.87**	**91.89**	**80.83**

DC Net obtained the lowest performance across the board (75.38% mIoU, 80.71% Mean BF, and 67.10% clDice), pointing to limited spatial overlap and imprecise boundary delineation. Although LUVS Net narrows this gap with higher scores than DC Net, it still trails the proposed framework across every metric. The substantial gains in clDice and Mean BF Score indicate that modifying the skip connections directly strengthens structural continuity and boundary fidelity across complex retinal vascular networks.

Table 3: Quantitative IoU evaluation on selected best-performing test cases.

Performer Model	Test Image	DC Net	LUVS Net	Proposed
Proposed	Image 1	55.76	61.44	72.36
	Image 2	53.61	60.75	70.19
DC Net	Image 3	60.52	58.20	60.63
LUVS Net	Image 4	67.52	72.23	63.21
	Image 5	60.03	66.82	61.63

DC Net obtained the lowest performance across the board (75.38% mIoU, 80.71% Mean BF, and 67.10% clDice), pointing to limited spatial overlap and imprecise boundary delineation. Although LUVS Net narrows this gap with higher scores than DC Net, it still trails the proposed framework across every metric. The substantial gains in clDice and Mean BF Score indicate that modifying the skip connections directly strengthens structural continuity and boundary fidelity across complex retinal vascular networks.

Table presents an image-level comparative analysis between the proposed network, DC Net, and LUVS Net, indicating the comparative analysis of the proposed network with DC Net and LUVS Net for the best-performing test images. In this, we have taken 2 test images (Image 1, Image 2) where the proposed network performs best, 1 image (Image 3) of DC Net, and 2 images (Image 4, Image 5) of LUVS Net are considered. The comparative analysis in graphical form is shown in Fig. 8. The best IoU achieved by the proposed network is about 75.26, while for LUVS Net it is 72.23 and for DC Net it is 67.52. The qualitative analysis of test images listed in DC Net obtained the lowest performance across the board (75.38% mIoU, 80.71% Mean BF, and 67.10% clDice), pointing to limited spatial overlap and imprecise boundary delineation.

Although LUVS Net narrows this gap with higher scores than DC Net, it still trails the proposed framework across every metric. The substantial gains in clDice and Mean BF Score indicate that modifying the skip connections directly strengthens structural continuity and boundary fidelity across complex retinal vascular networks.

Table are shown in Fig. 8. shows the clear differences between the DC Net**,** LUVS Net, and proposed network. In this, the black color represents the True Positive (TP), white is True Negative (TN), green is False Negative (FN), and pink is False Positive (FP). The DC Net exhibits noticeable green (FN) areas indicating the thin vessels are missed, while the proposed network captures finer vascular branches more efficiently. The LUVS Net indicates better continuity than the DC Net, but struggles with some pink pixel (FP) artefacts closer to vessel boundaries and background noise. These test image sets suggest the proposed model provides accurate delineation of vessel continuity, better suppression of background noise, and improved segmentation completeness.

Table 4: Quantitative clDice evaluation on selected best-performing test cases.

Performer	Image Name	DC Net	LUVS Net	Proposed
Proposed	Image 1	76.00	84.46	90.84
	Image 2	81.17	86.18	89.99
DC Net	Image 3	84.52	88.76	89.32
LUVS Net	Image 4	81.54	84.55	87.33
	Image 5	84.52	88.76	89.32

Fig. 8 presents a qualitative visual comparison alongside a pixel-wise error-map evaluation across five representative test images. To ensure a balanced assessment, the dataset features samples where each framework obtained its optimal relative results: the upper two rows contain cases favouring the proposed network, the third row depicts a top-performing instance for DC Net, and the final two rows illustrate cases favouring LUVS Net. The error maps categorize predictions using a four-color scheme: True Positive (black) denotes accurately identified vessel pixels, True Negative (white) denotes correct background pixels, False Negative (green) highlights under-segmented vessel pixels, and False Positive (pink) represents over-segmented background noise.

Proposed Method Top-Performing Cases (Rows 1 and 2): Across the first two test samples, both baseline models demonstrate severe green (FN) fragmentation along peripheral vascular trees. DC Net omits low-contrast terminal vessels almost entirely, whereas LUVS Net exhibits disrupted vessel continuity. Conversely, the proposed framework traces thin capillaries with negligible green omissions, validating that the parallel 1×1 and 3×3 skip-connection module retains fine-grained spatial context and structural topology more effectively than conventional skip pathways.

DC Net Top-Performing Case (Row 3): The third row illustrates a demanding image affected by uneven illumination and low visual contrast. Under these conditions, the proposed model introduces localized pink (FP) artifacts within background regions, revealing minor over-sensitivity to subtle contrast fluctuations. In comparison, DC Net achieves stronger background suppression (pure white regions), albeit at the expense of failing to resolve narrow vascular extensions.

LUVS Net Top-Performing Cases (Rows 4 and 5): The bottom two rows showcase instances where LUVS Net yields sharp boundary definition accompanied by minimal false-positive noise. Even within these samples, DC Net displays persistent under-segmentation (widespread green lines) at vessel endpoints. Although the proposed network generates minor pink (FP) false detections around intricate vascular junctions, it reliably preserves structural continuity and topological accuracy relative to the ground-truth annotations.

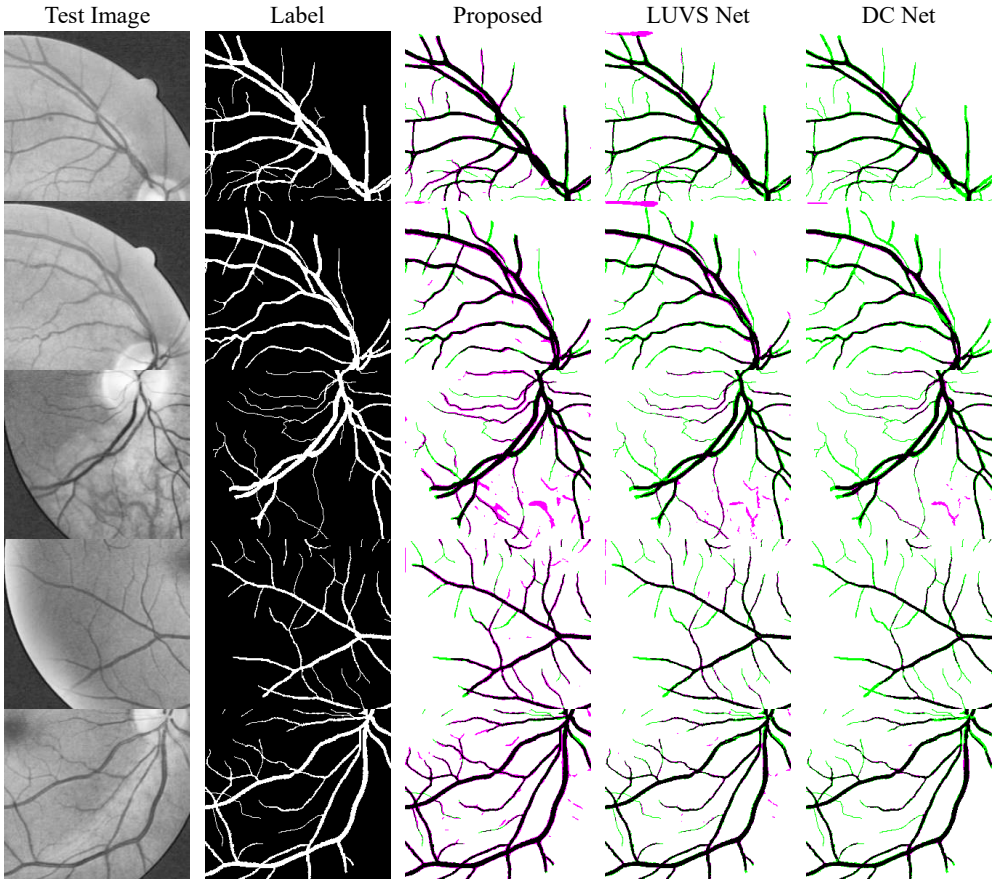

Fig. 8. Visual comparison of vessel segmentation across evaluated models and ground truth labels.

IV.iii. Layer-Wise Information Flow Analysis Using Mutual Information

To investigate how task-relevant information is represented across different stages of the network, a layer-wise information flow analysis was performed using mutual information (MI) (Kim, Choi and Yun, 2025). Mutual information provides a statistical measure of the dependence between two variables and can therefore be used to quantify the degree of correspondence between internal neural representations and the expected semantic information (Issitt et al., 2026). In the context of deep neural networks, this type of analysis provides an interpretable means of examining how representations evolve across different depths of the architecture. A set of representative layers was selected from different stages of the network, including intermediate ReLU layers, pooling layers, and an encoder activation layer. For each selected layer, the activation

maps generated for the input image were extracted. Mutual information was subsequently computed between the resized layer activation map and the binary reference mask. Thus, the obtained MI value characterizes the statistical association between the spatial activation pattern produced by a given network layer and the spatial distribution of the target structure.

The procedure was repeated independently for all selected layers, producing one mutual-information score per layer. The resulting scores were visualized using a bar chart to facilitate direct comparison of the internal representations. A relatively higher MI value indicates a stronger statistical correspondence between a layer's aggregated activation pattern and the target mask, suggesting that its representation contains greater spatial information associated with the target structure. Conversely, a lower MI value indicates weaker statistical dependence between the layer response and the reference mask.

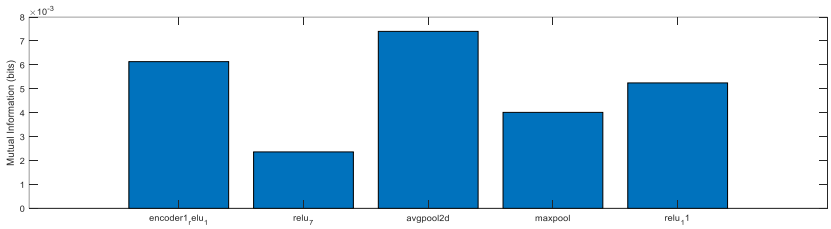

Fig. 9. Layer-wise Mutual Information Analysis of Inception module from left to right (a) Input (b) 1×1 convolution (c) Average Pooling (d) Max pooling (e) 5×5 convolution

Fig. 9 shows the MI across five distinct layers of the Inception module. This reveals that average pooling retains the highest amount of MI, followed closely by the 5×5 convolution operation. The max pool has lower information, while 1×1 convolution exhibits the minimum value among all layers. These quantitative results demonstrate that average pooling retains a substantially higher degree of feature information than max pooling during dimensionality reduction. Consequently, average pooling proves more effective at preserving subtle vascular structures within intermediate activation maps.

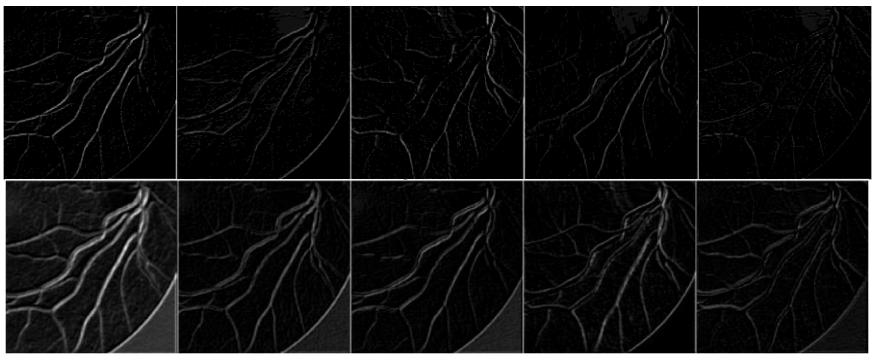

Fig. 10. Feature map at the Input and output of Inception

Fig. 10 shows the top five channels that exhibit the highest information at the input and output of the Inception module. Although the input to the inception module contains illumination variations and background noise alongside the vascular network, the convolutional layers successfully isolate these structural elements across individual

channels. Early feature maps capture primary vascular trunks and sharp vessel boundaries, whereas deeper channels refine these responses by isolating fine capillary networks, suppressing background artifacts, and extracting directional edge patterns.

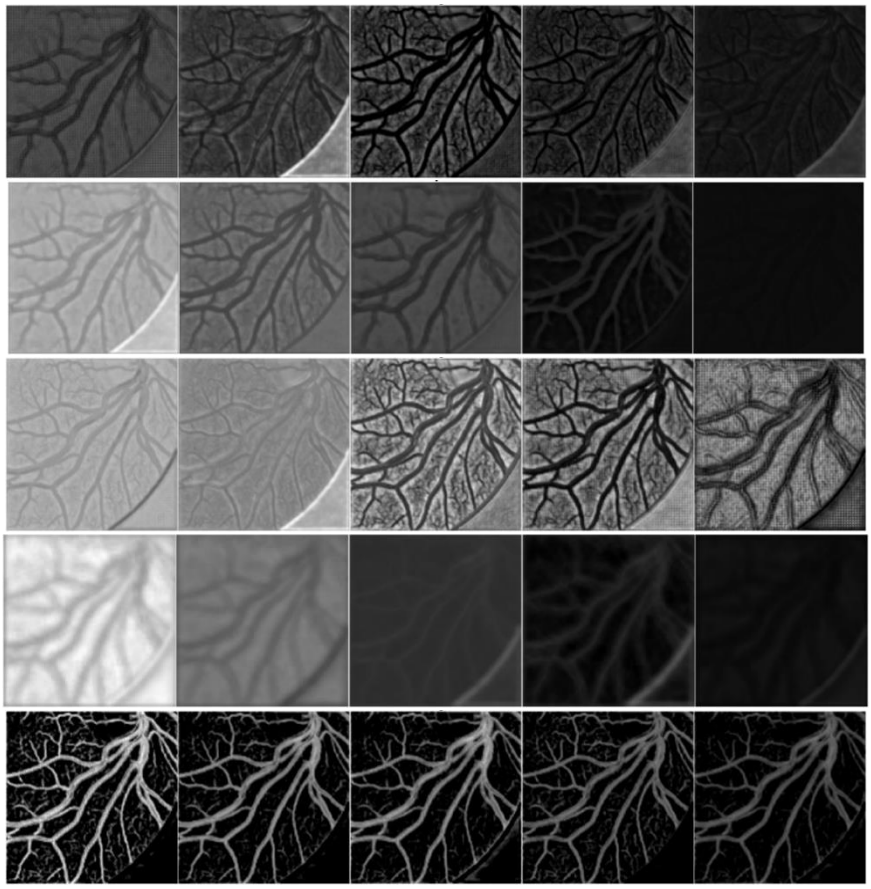

Fig. 11. Visualization of the top five extracted feature maps of the GL feature engineering module from top to bottom: (a) Input, (b) After Zooming, (c) Concatenated, (d) Shrinking, (e) Fused information

Stage-wise visual comparisons in Fig. 11 illustrate how feature channels evolve across the Global-Local (GL) feature engineering module. Although the initial feature maps in row (a) establish basic anatomical context, they remain affected by background illumination shifts and low vessel contrast. Applying the zooming operation in row (b) sharpens local spatial details, producing clearer vessel boundaries. Concatenating these multi-scale channels in row (c) broadens feature diversity, which makes low-contrast microvascular structures far more distinct. In row (d), the shrinking phase filters out non-vessel artifacts and spatial redundancies. As a result, the fused representations in row (e) yield high vessel-to-background contrast, cleanly separating main vascular trunks and fine capillaries before final segmentation.

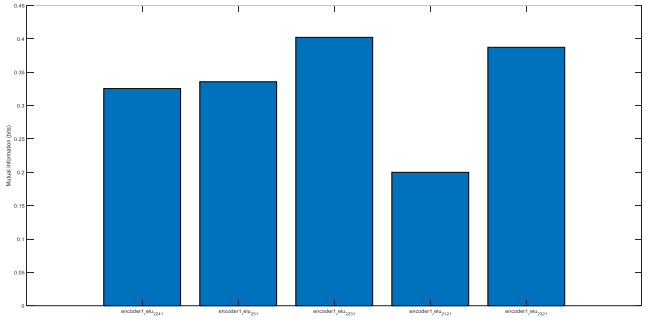

Fig. 12. Layer-wise Mutual Information Analysis of GL feature engineering module from left to right: (a) Input, (b) After zooming, (c) Concatenated, (d) Shrinking, (e) Fused information

Fig. 12 illustrates the quantitative evaluation of Mutual Information (MI) across successive stages of the Global-Local (GL) feature engineering module. While the input (a) and zoomed (b) stages maintain moderate MI values (~0.325 and ~0.335 bits), concatenation (c) maximizes information retention at 0.40 bits. The subsequent shrinking layer (d) deliberately compresses this feature space, reducing MI to 0.20 bits by eliminating redundant spatial context. This intentional downsampling allows the fused output (e) to consolidate target features back to 0.38 bits, preserving critical vascular details within a compact, efficient feature representation.

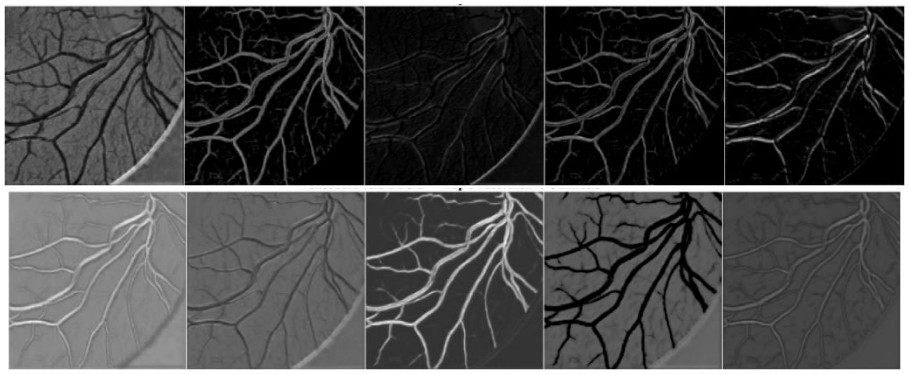

Fig. 13. Input and output at ASPP

Fig. 13 compares incoming feature representations (top row) with post-ASPP channel outputs (bottom row). Input feature maps show initial vessel patterns that suffer from low localized contrast and background noise. Following multi-rate atrous convolutions, the output channels systematically categorize features according to spatial scale: specific branches prioritize thick, high-contrast main vessels, while others isolate delicate capillary networks and directional edges. This multi-scale transformation effectively decouples anatomical structures from background noise, producing refined feature maps optimized for downstream segmentation tasks.

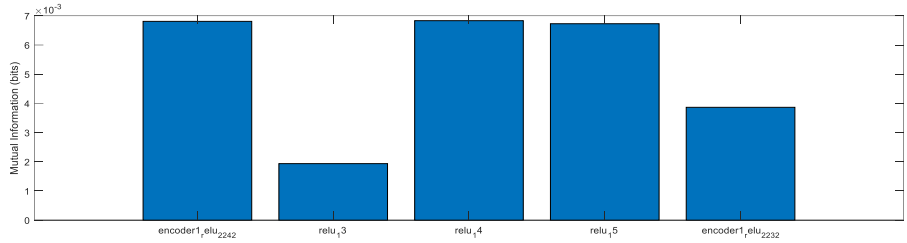

Fig. 14. Layer-wise Mutual Information Analysis of ASPP from left to right: (a) Input, (b) Dilation rate 3, (c) Dilation rate 2, (d) Dilation rate 4, (e) Fused information after convolution

Evaluating Mutual Information across the ASPP module (Fig. 14) demonstrates how specific dilation rates influence feature preservation. The input layer (a) provides an initial MI baseline of 6.8 $\times 10^{-3}$ bits before multi-scale branching. The branch operating at dilation rate 3 (b) experiences a reduction in MI (1.9 $\times 10^{-3}$ bits), while dilation rates 2 (c) and 4 (d) retain dominant spatial features (6.8 $\times 10^{-3}$ and 6.7 $\times 10^{-3}$). Through convolutional feature aggregation, the fused representation (e) stabilizes at 3.9 $\times 10^{-3}$ bits of MI, successfully isolating salient vascular structures across multiple scales.

IV.iv. Ablation Study

The ablation study is conducted to find the contribution of each component and its impact on the performance of retinal vessel segmentation. The improvement in the performance of the baseline model through the addition of each component is studied and evaluated through the performance metrics. The ablation analysis provides insight into how each module contributes to refining retinal vessel segmentation, demonstrating the efficacy of the proposed architecture.

IV.iv.a. Effect of Inception as Initial Feature Extractor

The inception block uses multi-scale kernels in parallel that capture the crucial features of thick and thin blood vessels. Fig. 15 shows the segmented image of U-Net (middle) and U-Net with the Inception model as initial feature extractor (right image). The quantitative analysis (See Table 5) shows that the Inception module as an initial feature extractor has improved the IoU, Dice, and clDice. The Inception module as an initial feature extractor better captures multi-scale vascular features, resulting in more accurate vessel delineation and improved preservation of vessel connectivity. The qualitative analysis shows that U-Net with Inception identifies more continuous vessel structures and recovers a larger number of thin vessels. The baseline U-Net fails to identify several fine branches, mainly in low-contrast regions, that increase the false identifications. The U-Net model with Inception as the initial feature extractor improves the discriminative capability of the encoder during the early stages of learning and reduces the false identification of blood vessels.

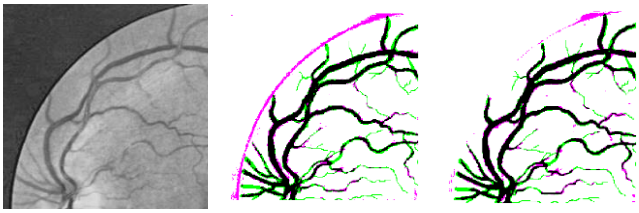

Fig. 15. Effect of Inception: Middle and right are without and with Inception as input layer in U-Net.

Table 5: U-Net and Inception as initial feature extractor U-Net comparison

Model	Vessel IoU	Dice	clDice
U-Net	50.84	66.94	62.76
Inception as Initial layer in U-Net	**59.59**	**74.32**	**72.24**

IV.iv.b. Effect of Global Local Feature Fusion.

Table and Fig. 16 show the quantitative and qualitative comparison of the fusion strategy on GL. The addition feature fusion enhances the IoU, Dice, and clDice. The improved clDice shows the enhanced preservation of the retinal vascular topology, hence effectively integrating the global contextual information with local spatial features

Table 6: Quantitative impact of concatenation versus addition fusion strategies within the GL module

Fusion Strategy	Vessel IoU	Dice	clDice
Concatenation	65.76	78.86	80.72
Addition	**66.54**	**79.69**	**81.04**

Fig. 16 shows the effect of addition and concatenation fusion strategy in Global Local feature extraction. The qualitative analysis shows the GL feature fusion produces more accurate vessel representations, mostly thin vascular branches and vessel bifurcations. The concatenation fusion produces discontinuities in the peripheral regions, while the addition-based fusion strategy preserves the continuity of fine vessels while reducing the number of missed vessel pixels.

Fig.16. Effect of GL fusion, middle and right is the result of concatenation and addition.

IV.iv.c. Effect of Sequential and Parallel GL Feature Fusion without Feature Engineering Block.

The quantitative results showing the effect of series and parallel GL fusion with the FE block are tabulated in Table, and qualitative analysis is shown in Fig. 17. The IoU and Dice for both configurations are almost the same, but series connection exhibits improved clDice. This demonstrates the better preserving capability of series connection, while parallel connection has more missing vessel segments.

Table 7: Quantitative evaluation of series versus parallel GL configurations without the FE block.

GL Connection without FE	Vessel IoU	Dice	clDice
Series	63.75	77.69	**81.24**
Parallel	**63.76**	77.69	77.96

Fig. 17. Effect of series and parallel GL connection in absence of FE Block. Middle and right are parallel and series.

The qualitative analysis shows better continuity along elongated vessel segments in series GL fusion. Thick vessels at the junction of segments are more accurate in series fusion compared to parallel fusion. False negatives (green) surrounding thin vessels are reduced with the series connection. Background noise and isolated false positives (magenta) are marginally lower in the series configuration. Vessel boundaries appear smoother and more coherent when GL modules are connected in series.

IV.iv.d. Effect of Sequential and Parallel GL Feature Fusion with Feature Engineering Block.

The quantitative and qualitative results showing the effect of series and parallel connection with concatenation fusion in the presence of the FE block are shown in Fig. 18 and Table . The IoU and Dice of parallel connection are better than the series connection, but the topological structure preserving capability of series connection is better compared to the parallel connection.

Table 8: GL connection in series and parallel with Feature Engineering

GL Connection with FE	Vessel IoU	Dice	clDice
Series	64.26	78.13	**80.83**
Parallel	**65.18**	**78.77**	80.52

The qualitative analysis demonstrates improvement in clDice, which indicates the enhanced preservation of the connectivity of the retinal vascular network, confirming that sequential integration of global contextual and local spatial information is more effective than parallel fusion for retinal vessel segmentation. The green color vessels segment are broken segments in the segmented image. The segments at the junction are missing in parallel connection as compared to the series connection.

Fig. 18. Effect of GL connection in presence of Feature Engineering Block. Middle and right are series and parallel.

IV.iv.e. Effect of Atrous Spatial Pyramid Pooling.

Table 9 and Fig. 19 show the quantitative and qualitative results showing the effect of ASPP. The IoU and Dice with ASPP are slightly lower by 1.50 and 1.03 percent, while

clDice is improved with ASPP. The improvement in the clDice with ASPP suggests that ASPP enhances the representation of multi-scale contextual information, enabling the network to better recover thin vessels and maintain vessel continuity. The qualitative results in Fig. 19 support this observation, where the model with ASPP produces more continuous vessel segments and recovers finer vascular branches compared with the model without ASPP. Although the enhanced multi-scale context introduces slight pixel-level deviations that marginally reduce IoU and Dice, it yields a more anatomically consistent vascular network, which is particularly important for retinal vessel segmentation.

Table 9: Effect of ASPP

Effect of ASPP	IoU	Dice	clDice
Without ASPP	**65.76**	**79.16**	78.23
With ASPP	64.26	78.13	**80.83**

Fig. 19. Effect of ASPP. Middle and right without and with ASPP

IV.iv.f. Proposed Network

An evaluation of the structural ablation variants of the complete network is detailed in Table 10 and Fig. 20, illustrating the performance gains achieved by combining sequential Global-Local (GL) processing with element-wise addition fusion. Replacing the series GL module with a parallel configuration causes a decline across all quantitative metrics (64.26% Vessel IoU, 78.13% Dice, and 80.83% clDice), showing that parallel feature aggregation is less equipped to model complex vascular spatial hierarchies. Likewise, substituting element-wise addition with concatenation fusion leads to a minor reduction in segmentation accuracy (65.18% Vessel IoU and 78.77% Dice), though it preserves topological structure relatively well (80.52% clDice). Integrating both sequential GL processing and addition fusion within the full architecture yields the top performance, reaching 66.54% Vessel IoU, 79.69% Dice, and 81.04% clDice.

Table 10: Quantitative comparison of the proposed architecture and its structural ablation variants.

Model	Vessel IoU	Dice	clDice
Proposed with Parallel GL connection	64.26	78.13	80.83
Proposed with concatenation	65.18	78.77	80.52
Proposed with series and addition	**66.54**	**79.69**	**81.04**

Qualitative error-map analyses in Fig. 20 support these numerical trends. The parallel GL variant exhibits frequent structural disconnections at intricate bifurcations, whereas the concatenation variant displays localized under-segmentation along narrow distal capillaries. In contrast, the complete framework preserves continuous vascular pathways while producing sharper vessel-to-background contrast in low-contrast regions.

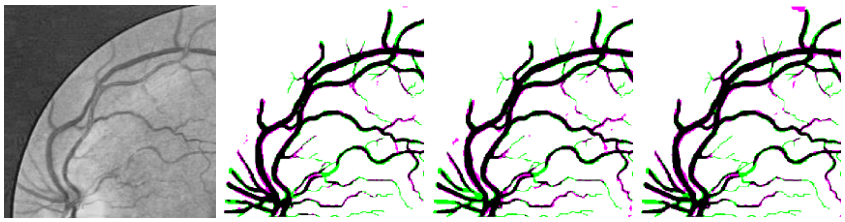

Fig. 20. From left to right: Proposed with series and addition, with concatenation and parallel with concatenation

V. Conclusion

A modified U-Net architecture was developed for retinal vessel segmentation to preserve vascular topology and capture low-contrast capillaries. Integrating an initial multi-scale Inception feature extractor, a sequential GL feature engineering module with element-wise addition, and an ASPP output stage enables the framework to decouple vessel structures from background artifacts. Quantitative evaluations confirm the model outperforms competitive baselines across primary benchmarks, achieving 80.87% mIoU, 91.89% Mean BF Score, 80.83% clDice, and an 84.6% vessel classification rate. Ablation studies validated each component:

- Initializing feature extraction with an Inception layer boosted the early discriminative capacity, raising baseline U-Net IoU from 50.84% to 59.59%.
- Executing element-wise addition outperformed concatenation by restoring micro-vascular bifurcations and resolving peripheral disconnections.
- Arranging GL modules in series connections yielded stronger topological preservation than parallel setups, reaching 81.24% clDice without the feature engineering block.

Layer-wise Mutual Information metrics showed average pooling retains higher spatial entropy than max pooling during compression, while ASPP branches isolate scale-specific features. Visual error maps confirm the framework achieves effective background suppression, accurate boundary delineation, and sustained vascular continuity across varied retinal images.

Conflict of Interest:

There is no competing interest among the authors regarding the publication of the article.

References

I. Akhtar, N. and Ragavendran, U. (2019) 'Interpretation of intelligence in CNN-pooling processes: a methodological survey', Neural Computing and Applications [Preprint]. 10.1007/s00521-019-04296-5.

II. Bai, H. et al. (2026) 'nnLoGoNet: a hybrid local-global network for retinal vessel segmentation with Skeleton Recall Loss', Scientific Reports, 16(1), p. 19297. 10.1038/s41598-026-50020-4.

III. Bansal, A. et al. (2026) 'Retinal Vessel Segmentation: A Comprehensive Review From Classical Methods to Deep Learning Advances (1982–2025)', Advanced Intelligent Systems, n/a(n/a), p. e202501279. Available at: 10.1002/aisy.202501279.

IV. Boudegga, H. et al. (2021) 'Fast and efficient retinal blood vessel segmentation method based on deep learning network', Computerized Medical Imaging and Graphics, 90, p. 101902. Available at: 10.1016/j.compmedimag.2021.101902.

V. Chu Jinsong; Zheng, Wenqiang; Xu, Zhengyuan, B.Z. (2025) '(DA-U) 2Net: double attention U2Net for retinal vessel segmentation.' BMC ophthalmology, 25(1), pp. 86-NA. 10.1186/s12886-025-03908-0.

VI. Das, S. et al. (2024) 'Assessment of retinal blood vessel segmentation using U-Net model: A deep learning approach', Franklin Open, 8, p. 100143. 10.1016/j.fraope.2024.100143.

VII. Dong Ting; Zhang, Tianyi; Wei, Lifang, H.Z. (2022) 'Supervised learning-based retinal vascular segmentation by M-UNet full convolutional neural network', Signal, Image and Video Processing, 16(7), pp. 1755–1761. 10.1007/s11760-022-02132-3.

VIII. Fareed, K. et al. (2025) 'CBAM Attention Gate-Based Lightweight Deep Neural Network Model for Improved Retinal Vessel Segmentation', International Journal of Imaging Systems and Technology, 35(1), p. e70031. 10.1002/ima.70031.

IX. Van Gelder Michael F; Dyer, Michael A; Greenwell, Thomas N; Levin, Leonard A; Wong, Rachel O; Svendsen, Clive N, R.N.C. (2022) 'Regenerative and restorative medicine for eye disease.', Nature medicine, 28(6), pp. 1149–1156. 10.1038/s41591-022-01862-8.

X. Godishala, A. et al. (2025) 'Survey on retinal vessel segmentation', Multimedia Tools and Applications, 84(11), pp. 8361–8411. 10.1007/s11042-024-19075-1.

XI. Guo Shaojia; Ge, Rui; Li, Wenjuan; Lu, Aihong; Feng, Rongzhen; Fang, Wu; Jin, Qiangguo, F.Y. (2025) 'RetinaDA: a diverse dataset for domain adaptation in retinal vessel segmentation', Frontiers of Computer Science, 19(8), p. NA-NA. 10.1007/s11704-024-41114-1.

XII. Hassan, G. et al. (2015) 'Retinal Blood Vessel Segmentation Approach Based on Mathematical Morphology', Procedia Computer Science, 65, pp. 612–622. 10.1016/j.procs.2015.09.005.

XIII. He, K. et al. (2015) 'Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition', IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(9), pp. 1904–1916. 10.1109/TPAMI.2015.2389824.

XIV. He, K. et al. (2016) 'Deep Residual Learning for Image Recognition', in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. 10.1109/CVPR.2016.90.

XV. Hoover Valentina L.; Goldbaum, Michael H., A.D. K. (2000) 'Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response', IEEE transactions on medical imaging, 19(3), pp. 203–210. 10.1109/42.845178.

XVI. Hossen, M.F. et al. (2026) 'A comprehensive review of U-Net architectures for medical image segmentation: emerging trends and federated learning perspectives', Neural Computing and Applications, 38(12), p. 493. 10.1007/s00521-026-12198-6.

XVII. Islam, M.T. et al. (2023) 'LUVS-Net: A Lightweight U-Net Vessel Segmentor for Retinal Vasculature Detection in Fundus Images', Electronics, 12(8). 10.3390/electronics12081786.

XVIII. Issitt, A. et al. (2026) 'Uncovering Neural Learning Dynamics through Latent Mutual Information', Entropy, 28(1). 10.3390/e28010118.

XIX. Jebaseeli, T.J., Durai, C.A.D. and Peter, J.D. (2019) 'Segmentation of retinal blood vessels from ophthalmologic Diabetic Retinopathy images', Computers & Electrical Engineering, 73, pp. 245–258. 10.1016/j.compeleceng.2018.11.024.

XX. Jin Qi; Meng, Zhaopeng; Wang, Bing; Su, Ran, Q.C. (2019) 'Construction of Retinal Vessel Segmentation Models Based on Convolutional Neural Network', Neural Processing Letters, 52(2), pp. 1005–1022. 10.1007/s11063-019-10011-1.

XXI. Juntang Zhuang (2018) 'LadderNet: Multi-path networks based on U-Net for medical image segmentation' CoRR, abs/1810.0.

XXII. Kawamura, T. (2024) 'Annual Report':, Quarterly Journal of Marketing, 44(1), pp. 102–104. 10.7222/marketing.2024.037.

XXIII. Kim, G., Choi, S. and Yun, S.-Y. (2025) 'A Mutual Information-Based Framework for Enhancing Graph Neural Networks on Heterophily', IEEE Access, 13, pp. 199531–199545. 10.1109/ACCESS.2025.3624415.

XXIV. Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) 'ImageNet Classification with Deep Convolutional Neural Networks', Commun. ACM, 60(6), pp. 84–90. 10.1145/3065386.

XXV. Li Fei; Yan, Fen; Meng, Jing; Guo, Yanfei; Liu, Hongjuan; Cheng, Ronghua, S.M. (2025) 'SMFDNet: spatial and multi-frequency domain network for OCT angiography retinal vessel segmentation', The Journal of Supercomputing, 81(4), p. NA-NA. 10.1007/s11227-025-06985-6.

XXVI. Li Manisha; Nakashima, Yuta; Nagahara, Hajime; Kawasaki, Ryo, L.V. (2020) 'IterNet: Retinal Image Segmentation Utilizing Structural Redundancy in Vessel Networks', 2020 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 3645–3654. 10.1109/wacv45572.2020.9093621.

XXVII. Lin, T.-Y. et al. (2017) 'Feature Pyramid Networks for Object Detection', in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 936–944. 10.1109/CVPR.2017.106.

XXVIII. Mardani, K. and Maghooli, K. (2021) 'Enhancing retinal blood vessel segmentation in medical images using combined segmentation modes extracted by DBSCAN and morphological reconstruction', Biomedical Signal Processing and Control, 69, p. 102837. 10.1016/j.bspc.2021.102837.

XXIX. Odstrcilik Radim; Budai, Attila; Hornegger, Joachim; Jan, Jiri; Gazarek, Jiri; Kubena, Tomas; Cernosek, Pavel; Svoboda, Ondrej; Angelopoulou, Elli, J.K. (2013) 'Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database', IET Image Processing, 7(4), pp. 373–383. 10.1049/iet-ipr.2012.0455.

XXX. Oktay, O. et al. (2018) 'Attention U-Net: Learning Where to Look for the Pancreas', CoRR, abs/1804.0, pp. 2–5.

XXXI. Ouyang, J. et al. (2023) 'LEA U-Net: a U-Net-based deep learning framework with local feature enhancement and attention for retinal vessel segmentation', Complex & Intelligent Systems, 9(6), pp. 6753–6766. 10.1007/s40747-023-01095-3.

XXXII. Pizer, S.M. et al. (1987) 'Adaptive histogram equalization and its variations', Computer Vision, Graphics, and Image Processing, 39(3), pp. 355–368. 10.1016/S0734-189X(87)80186-X.

XXXIII. Prajna, Y. and Nath, M.K. (2022) 'Efficient blood vessel segmentation from color fundus image using deep neural network', Journal of Intelligent \& Fuzzy Systems, 42(4), pp. 3477–3489. 10.3233/JIFS-211479.

XXXIV. Ramos-Soto, O. et al. (2021) 'An efficient retinal blood vessel segmentation in eye fundus images by using optimized top-hat and homomorphic filtering', Computer Methods and Programs in Biomedicine, 201, p. 105949. 10.1016/j.cmpb.2021.105949.

XXXV. Sanjeewani Arun Kumar; Akbar, Mohd; Kumar, Mohit; Yadav, Divakar, N.Y. (2024) 'Retinal blood vessel segmentation using a deep learning method based on modified U-NET model', Multimedia Tools and Applications, 83(35), pp. 82659–82678. 10.1007/s11042-024-18696-w.

XXXVI. Shang, Z. et al. (2024) 'DCNet: A lightweight retinal vessel segmentation network', Digital Signal Processing, 153, p. 104651. 10.1016/j.dsp.2024.104651.

XXXVII. Shit, S. et al. (2021) 'clDice - a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation', in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16555–16564. 10.1109/CVPR46437.2021.01629.

XXXVIII. Simonyan, K. and Zisserman, A. (2015) 'Very Deep Convolutional Networks for Large-Scale Image Recognition', in Y. Bengio and Y. LeCun (eds) 3rd International Conference on Learning Representations, {ICLR} 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. http://arxiv.org/abs/1409.1556.

XXXIX. Singh Munish; Thawkar, Shankar; Singh, Rekha, L.K.K. (2023) 'Deep-learning based system for effective and automatic blood vessel segmentation from Retinal fundus images', Multimedia Tools and Applications, 83(2), pp. 6005–6049. 10.1007/s11042-023-15348-3.

XL. Sobisch, J. et al. (2026) 'Toward Well-Connected Retina Segmentation: A Fully Differentiable Endpoint Connectivity Loss (DECL)', IEEE Access, 14, pp. 63326–63341. 10.1109/ACCESS.2026.3686973.

XLI. Spaide James G.; Waheed, Nadia K., R.F.. F. (2015) 'Image Artifacts in Optical Coherence Tomography Angiography', Retina (Philadelphia, Pa.), 35(11), pp. 2163–2180. 10.1097/iae.0000000000000765.

XLII. Staal Michael D.; Niemeijer, Meindert; Viergever, Max A.; van Ginneken, B., J.A. (2004) 'Ridge-based vessel segmentation in color images of the retina', IEEE transactions on medical imaging, 23(4), pp. 501–509. 10.1109/tmi.2004.825627.

XLIII. Szegedy, C. et al. (2015) 'Going deeper with convolutions', in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9. 10.1109/CVPR.2015.7298594.

XLIV. Toptaş Davut, B.H. (2021) 'Retinal blood vessel segmentation using pixel-based feature vector', Biomedical Signal Processing and Control, 70(NA), pp. 103053-NA. 10.1016/j.bspc.2021.103053.

XLV. Uysal Gur Emre, E.G. (2020) 'Computer-aided retinal vessel segmentation in retinal images: convolutional neural networks', Multimedia Tools and Applications, 80(3), pp. 3505–3528. 10.1007/s11042-020-09372-w.

XLVI. Wan Gaofeng; Li, Renxing; Xiang, Yifan; Yin, Dechao; Yang, Minglei; Gong, Deren; Chen, Jiangang, Y.W. (2024) 'Retinal Blood Vessels Segmentation With Improved <scp>SE-UNet</scp> Model', International Journal of Imaging Systems and Technology, 34(4), p. NA-NA. 10.1002/ima.23145.

XLVII. Wang Xiaobo; Ma, Zhendi, J.L. (2025) 'Multi-Scale Three-Path Network (MSTP-Net): A new architecture for retinal vessel segmentation', Measurement, 250(NA), p. 117100. 10.1016/j.measurement.2025.117100.

XLVIII. Wu Yong; Song, Yang; Zhang, Yanning; Cai, Weidong, Y.X. (2020) 'NFN＋: A novel network followed network for retinal vessel segmentation.', Neural networks : the official journal of the International Neural Network Society, 126(NA), pp. 153–162. 10.1016/j.neunet.2020.02.018.

XLIX. Yazıcı, Z.A. et al. (2026) 'VEELA: A Clinically-Constrained Benchmark for Liver Vessel Segmentation in Computed Tomography Angiography'.

L. Zafar, A. et al. (2022) 'A Comparison of Pooling Methods for Convolutional Neural Networks', Applied Sciences, 12(17). 10.3390/app12178643.

LI. Zeiler, M.D. and Fergus, R. (2013) 'Visualizing and Understanding Convolutional Networks', CoRR, abs/1311.2.

LII. Zhu Xiang; Zhang, Xue-Dian; Jiang, Min-Shan, Y.-F.X. (2023) 'CCS-UNet: a cross-channel spatial attention model for accurate retinal vessel segmentation.', Biomedical optics express, 14(9), p. 4739. 10.1364/boe.495766.