Deep Learning-Driven Retinal Vessel Segmentation Using a Transformer-Based Architecture: A Performance Evaluation

Yves Byiringiro
Martin J. Whitman School of Management
Syracuse University
Syracuse, NY, 13244, United States
ybyiring@syr.edu

Jean De Dieu Niyonteze
Goizueta Business School
Emory University
Atlanta, GA 30322, United States
jniyont@emory.edu

Liliane Tuyishime
Martin J. Whitman School of Management
Syracuse University
Syracuse, NY, 13244, United States
Ituyishi@syr.edu

Mediatrice Dusenge
African Centre of Excellence in Data Science
University of Rwanda
Kigali, Rwanda
dusengemeddy@gmail.com

Josue Uzigusenga
College of Sciences and Technology
University of Rwanda
Kigali, Rwanda
uzigusengajosue@gmail.com

Benny Uhoranishema
Goizueta Business School
Emory University
Atlanta, GA 30322, United States
buhorani@alumni.cmu.edu

***Abstract*—Retinal vessel segmentation is essential for automated ophthalmic diagnosis and remains challenging due to variations in image quality, illumination, and the presence of fine, low-contrast vessels. While transformer-based models have recently emerged as powerful alternatives to convolutional neural networks, their effectiveness across diverse retinal imaging datasets has not been systematically assessed. To our knowledge, this study presents a multi-dataset evaluation of the SegFormer transformer architecture for retinal vessel segmentation using eight widely adopted fundus datasets. A five-fold cross-validation strategy was employed to assess generalization across heterogeneous imaging conditions, after which the best-performing dataset, LES-AV, was used for final model training. The resulting model was directly applied to all external datasets without fine-tuning, enabling a rigorous assessment of cross-dataset robustness. SegFormer achieved consistently strong outcomes, obtaining Dice scores up to 0.82 and accuracy above 0.95 across all datasets. The model demonstrated reliable extraction of both major vessels and thin structures, highlighting the advantage of transformer-based representations for retinal image analysis. This work provides a unified multi-dataset benchmark for transformer-based retinal vessel segmentation and demonstrates the clinical potential of SegFormer as a high-performance, generalizable, and scalable framework for next-generation ophthalmic imaging applications.**

Keywords*—*Deep learning, Medical image analysis, Retinal vessel segmentation, SegFormer, Transformer models

I. Introduction

Retinal vessel segmentation refers to the automated delineation of blood vessels within retinal fundus images. This process separates vascular structures from the background to generate an accurate representation of the retinal vascular network, which serves as a critical biomarker for the detection and assessment of various ocular and systemic pathologies, including diabetic retinopathy, glaucoma, hypertension, and cardiovascular diseases [1]. However, retinal vessel segmentation remains challenging due to the difficulty of detecting small vessels, low contrast between vessels and background, uneven illumination, image noise, and variability across imaging conditions [1], [2].

Convolutional neural network (CNN)–based approaches such as U-Net, ResU-Net, Dense U-Net, and Attention U-Net have achieved remarkable success in retinal vessel segmentation by effectively capturing local spatial features and hierarchical representations [3]. Nonetheless, their limited receptive field and lack of explicit global context modeling can hinder the accurate representation of complex vascular structures, particularly thin and tortuous vessels [4].

More recently, transformer-based models such as TransUNet, Swin Transformer, AttUNet, and SwinUNet have been proposed to overcome these limitations by leveraging self-attention mechanisms to capture long-range dependencies and global contextual information. This capability enhances the preservation of vessel continuity and improves the detection of fine, low-contrast structures in retinal images [5].

In this study, we evaluate the efficacy of a transformer-based deep learning architecture, SegFormer, for retinal vessel segmentation. Our analysis focuses on the model's ability to capture global contextual information, preserve vascular continuity, and accurately delineate thin vessel structures across datasets of varying resolutions and imaging characteristics. The remainder of this paper is organized as follows. Section II reviews related work on CNN- and transformer-based retinal vessel segmentation methods. Section III describes the datasets, preprocessing pipeline, SegFormer architecture, and evaluation metrics. Section IV presents cross-validation results and cross-

979-8-3315-5723-2/26/$31.00 ©2026 IEEE

dataset prediction experiments. Section V concludes the paper and outlines directions for future research.

II. Related Work

Diabetic retinopathy is a complex and progressive ocular disease that can lead to irreversible vision loss if not detected and treated at an early stage [6]. Accurate retinal vessel segmentation is essential for its diagnosis and monitoring since vascular abnormalities, including caliber changes, occlusions, and microaneurysms, serve as early indicators of disease progression [7]. Deep learning techniques have shown strong potential in extracting detailed vascular structures from retinal fundus images and have consequently become an important component of modern ophthalmic image analysis.

Since 2020, convolutional neural network (CNN)-based models have been widely used for retinal vessel segmentation. Architectures such as Spatial Attention U Net (SA U Net), Dense U Net, and other lightweight variants have demonstrated improvements in segmentation accuracy and model efficiency. SA U Net has reported Dice similarity coefficients above 0.80 and accuracies above 0.95 on benchmark datasets including DRIVE, CHASE_DB1, STARE, and HRF [8]. Lightweight U Net architectures have also achieved competitive performance, with Dice scores of 0.7871, 0.7946, and 0.6902 on the DRIVE, CHASE_DB1, and HRF datasets, respectively, and corresponding accuracies of 0.9113, 0.9718, and 0.8437. Although enhanced loss functions and attention mechanisms improved results on DRIVE and CHASE_DB1, performance gains were limited on HRF due to significant image variability in contrast and illumination.

In recent years, transformer-based architectures have demonstrated state-of-the-art performance in retinal vessel segmentation. These models address key limitations of CNNs by capturing long-range dependencies and global contextual relationships through self-attention mechanisms. This capability enhances the detection of thin vascular branches, improves vessel continuity, and supports better discrimination between vessels and background structures [9]. SegFormer has shown strong performance in dense prediction tasks and has achieved results such as 51.8 percent mIoU on ADE20K, 84.0 percent mIoU on Cityscapes, and 46.7 percent mIoU on COCO Stuff while being significantly smaller and faster than earlier transformer models like SETR [10]. Its lightweight version, SegFormer B0, further demonstrated efficiency by achieving 37.4 percent mIoU on ADE20K and 76.2 percent mIoU on Cityscapes at 50 FPS with only 3.8 million parameters.

Additional transformer-based architectures, including stimulus-guided transformer networks and hybrid CNN-transformer frameworks, have further strengthened the field by improving generalization in heterogeneous retinal images and enhancing performance on low-contrast vessel regions [11]. Collectively, these advancements highlight the potential of transformer-driven models in retinal vessel segmentation and motivate the present study to investigate SegFormer within this context.

III. Materials and Methods

A. Datasets and Preprocessing Steps

This study evaluated the performance of the SegFormer transformer-based architecture using eight public retinal fundus datasets: DRIVE [12], STARE [13], CHASE_DB1 [14], HRF [15], RETA Benchmark [16], AV-DRIVE [17], LES-AV [18], and IOSTAR [19]. These datasets differ in resolution, illumination, and annotation quality, providing a diverse benchmark for retinal vessel segmentation. Representative images and masks from each dataset are shown in Fig. 1.

All retinal images were resized to 512 × 512 pixels and normalized to a standard intensity range. To enhance vessel visibility, CLAHE was applied consistently across all datasets. For AV-DRIVE, artery and vein annotations were merged into a single binary vessel mask to maintain uniform labeling. Ground truth masks from all datasets were converted into binary vessel maps before training and evaluation. This unified preprocessing pipeline ensured consistency across datasets and enabled fair performance comparison of the SegFormer model.

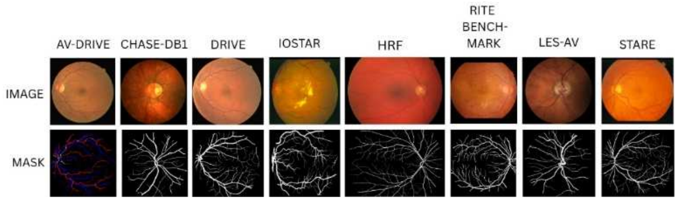


Fig. 1. Representative color fundus images and their corresponding vessel segmentation masks from the eight public retinal datasets used in this study.

B. SegFormer Architecture and Training Strategy

SegFormer was selected for this study because of its efficiency and strong performance in dense prediction tasks. The model uses the Mix Transformer (MiT) encoder, which captures long-range dependencies through self-attention while preserving fine spatial details by means of overlapping patch embeddings. This hierarchical design enables the model to detect both large vessels and thin peripheral branches more effectively than conventional CNN architectures. A lightweight MLP decoder aggregates multi-scale encoder features without relying on computationally heavy upsampling blocks, making the architecture suitable for medical image segmentation.

All experiments used the standardized preprocessing pipeline described in Section III-A (512×512 resizing, normalization, and CLAHE). The model was optimized with the Adam optimizer at an initial learning rate of 1×10^{-4} and trained for a maximum of 100 epochs on an NVIDIA GeForce RTX 4070 GPU. A cosine annealing learning rate schedule was applied to stabilize convergence, and the batch size was set to 4 due to memory considerations.

To ensure robustness, a five-fold cross-validation procedure was performed independently on each dataset. For every fold, the checkpoint achieving the best validation Dice score was retained for evaluation. This approach provided reliable performance estimates across datasets of different resolutions, illumination levels, and labeling characteristics. The cross-validation results also motivated the selection of LES-AV as the final training dataset for cross-dataset prediction experiments presented later in this study.

The model was trained using a combined Binary Cross-Entropy (BCE) and Dice loss function, computed on vessel versus background pixel classes.

C. Metrics, Evaluation Procedures, and Validation Strategy

Segmentation performance was evaluated using standard metrics commonly applied in retinal vessel segmentation. Dice score and accuracy served as the primary measures. Dice score quantifies the spatial overlap between predicted vessels and ground truth, while accuracy measures the proportion of correctly classified pixels. Specificity, precision, recall, and F1-score were also computed to characterize model behavior on vessel and background classes.

For each dataset, metrics were computed on a per-image basis and averaged to obtain representative values. A five-fold cross-validation procedure was used to evaluate generalization, where four folds were used for training and one for validation. This approach ensured stable performance estimates across datasets with varying illumination, resolution, and annotation characteristics. After completing cross-validation, the dataset with the highest overall performance was selected as the best-performing dataset.

This best-performing dataset was then used for a separate final training stage, where the model was retrained using a conventional 70 percent training and 30 percent testing split. The resulting model was subsequently applied to all remaining datasets without fine-tuning. All evaluations followed the definitions of accuracy (Eq. 1), Dice score in both pixel-wise and set-based forms (Eqs. 2–3), Specificity (Eq. 4), Precision (Eq. 5), Recall (Eq. 6), and F1-score (Eq. 7). In the following equations, TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. X represents the set of predicted vessel pixels, and Y denotes the ground-truth vessel pixel set used for overlap-based evaluation.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Dice\ Score = \frac{2TP}{2TP + FP + FN} \tag{2}$$

$$Dice\ Score = \frac{2|X \cap Y|}{|X| + |Y|} \tag{3}$$

$$Specificity\ = \frac{TN}{TN + FP} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall\ = \frac{TP}{TP + FN} \tag{6}$$

$$F1 - score = \frac{2\ \times\ Precision\ \times\ Recall}{Precision + Recall} \tag{7}$$

IV. Results and Discussion

A. Cross-Validation Performance Across Multiple Datasets

Table I presents the SegFormer model's segmentation performance across the eight public retinal vessel segmentation datasets. The mean Dice and accuracy values obtained through five-fold cross-validation are supported by the training and validation curves in Figs. 2–7. Overall, the model showed stable behavior across datasets with varying resolutions and illumination conditions.

DRIVE produced the lowest Dice performance, likely due to its small size and strong background–vessel imbalance. In contrast, STARE and CHASE_DB1 achieved higher and more consistent results, reflecting clearer vascular boundaries and reliable annotations. LES-AV showed the highest Dice and accuracy in Table I, confirming its suitability for transformer-based segmentation. The RETA Benchmark also performed well, benefiting from modern imaging quality and more consistent labeling.

HRF and IOSTAR achieved moderate Dice scores, influenced by illumination variation, limited sample size, and annotation inconsistencies. These differences are observable in Figs. 2–7, where datasets with greater variation display larger fluctuations in the validation curves. Despite these challenges, SegFormer maintained stable performance across all datasets.

Overall, the cross-validation results in Table I indicate that SegFormer generalizes effectively across heterogeneous fundus datasets, with LES-AV providing the strongest and most consistent performance. For each dataset, segmentation performance was computed independently across the five folds, and the results reported in Table I represent the mean and standard deviation aggregated over all folds. However, a limitation of this work is that although cross-validation mitigates overfitting, the LES-AV dataset contains only 22 images; therefore, the 70/30 split used during the final training stage is based on a relatively small sample size.

While prior CNN-based approaches such as SA U-Net and lightweight U-Net report Dice scores of 0.79–0.80 on DRIVE, 0.79–0.97 on CHASE_DB1, and 0.69–0.84 on HRF [8], these results typically rely on dataset-specific preprocessing, tuning, and training protocols. Under our strictly unified preprocessing pipeline and five-fold cross-validation applied uniformly across eight datasets, SegFormer achieves competitive Dice scores (0.70–0.82) while maintaining consistently high accuracy above 0.95. This highlights the strength of transformer-based models in delivering reliable and stable performance across heterogeneous imaging conditions rather than excelling only on a single dataset.

TABLE I. PERFORMANCE EVALUATION OF THE SEGFORMER MODEL USING FIVE-FOLD CROSS-VALIDATION ACROSS MULTIPLE RETINAL VESSEL SEGMENTATION DATASETS

Dataset	**# Images / Modality / Resolution in Pixels**	**Segmentation results of the K-fold cross-validation of the SegFormer model (k = 5).**		**Comments**
		Average: Mean± Std		
		Dice	Accu-racy	
DRIVE	40 / Color fundus images / Fundus photography / 768×584	0.742 ± 0.02	0.952 ± 0.008	Limited generalization due to small dataset and strong background–vessel imbalance.
STARE	20 / Fundus/ 605×700	0.770 ± 0.012	0.964 ± 0.005	Excellent generalization; well-annotated dataset.
CHASE _DB1	28 / Fundus/ 1280×960	0.764 ± 0.005	0.964 ± 0.005	Stable performance despite limited sample size.
HRF	45 / Color Fundus / 3504×2336	0.698 ± 0.015	0.950 ± 0.000	Reduced Dice due to illumination variation and limited training data.
IOSTAR	30 / Color fundus photographs / 1024×1024 masks	0.752 ± 0.044	0.951 ± 0.005	Performance was affected by image variability and small dataset size.
RETA Benchmark	81 / Color fundus photographs / 2896 × 1944	0.764 ± 0.008	0.960 ± 0.000	Consistent performance; high-quality modern benchmark.
AV-DRIVE	40 / Fundus / 565 × 584	0.768 ± 0.011	0.958 ± 0.005	Reliable reference for artery–vein segmentation tasks.
LES-AV	22 / Fundus / 1620 × 1444 and 1958×2196	0.780 ± 0.007	0.970 ± 0.000	Highest overall performance; strong annotation consistency.

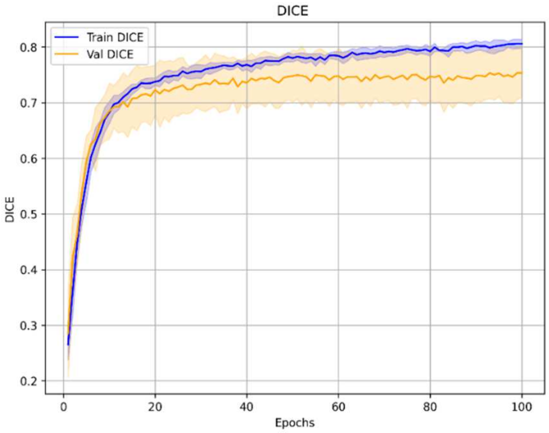


Fig. 2. Training and validation dice curves of the SegFormer model on the LES-AV dataset obtained from five-fold cross-validation.

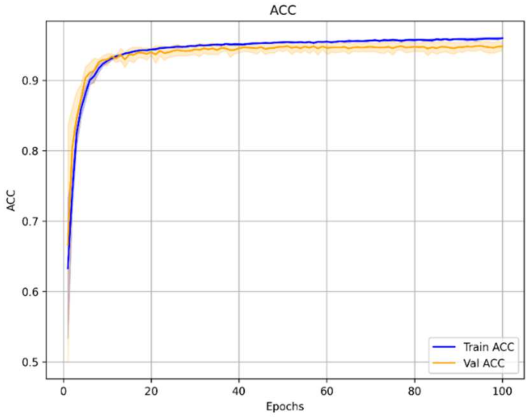


Fig. 3. Training and validation accuracy curves of the SegFormer model on the LES-AV dataset obtained from five-fold cross-validation.

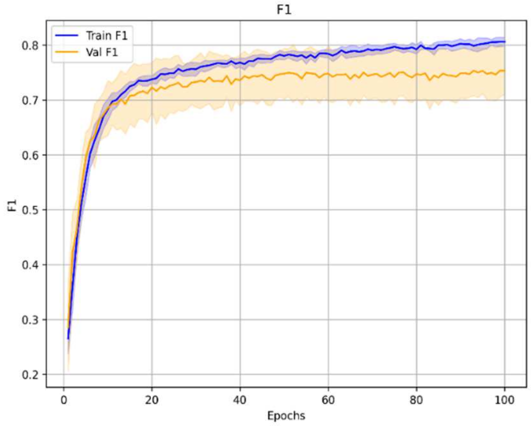


Fig. 4. Training and validation F1-score curves of the SegFormer model on the LES-AV dataset obtained from five-fold cross-validation.

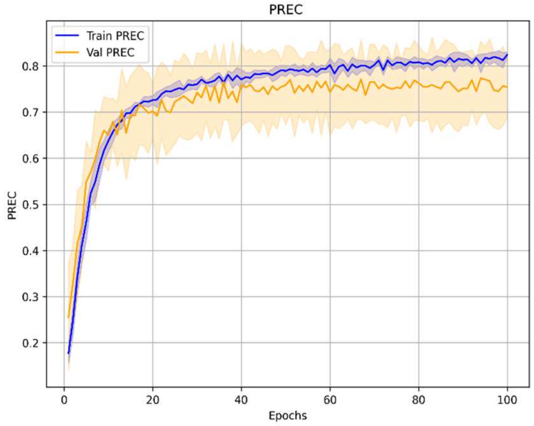


Fig. 5. Training and validation precision curves of the SegFormer model on the LES-AV dataset obtained from five-fold cross-validation.

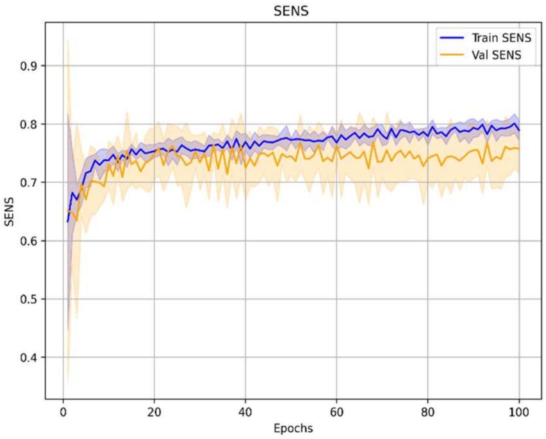


Fig. 6. Training and validation sensitivity curves of the SegFormer model on the LES-AV dataset obtained from five-fold cross-validation.

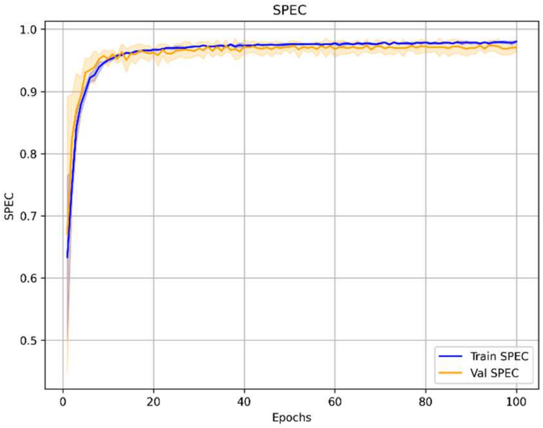


Fig. 7. Training and validation specificity curves of the SegFormer model on the LES-AV dataset obtained from five-fold cross-validation.

B. *Final Model Training on LES-AV and Cross-Dataset Prediction*

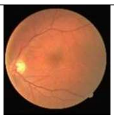
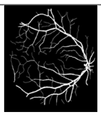
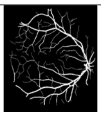
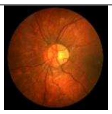
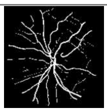
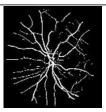
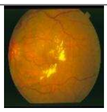
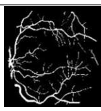
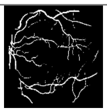
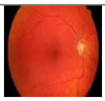
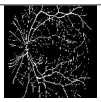
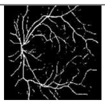
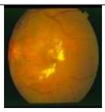
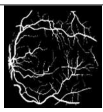
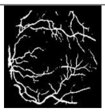
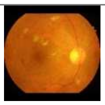
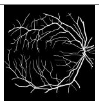
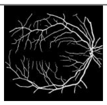
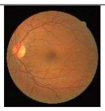
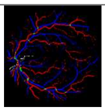
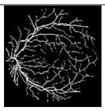
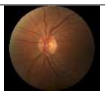
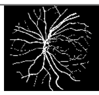
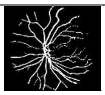
After completing the cross-validation phase, the LES-AV dataset was identified as the best-performing dataset in Table I and was therefore used for the final training stage. The SegFormer model was retrained on LES-AV using a conventional 70 percent training and 30 percent testing split, applying the same preprocessing steps used throughout the study. The training and validation behavior of this final model is illustrated in Figs. 8 and 9.

The trained LES-AV model was then applied to all remaining datasets without additional fine-tuning, using the same preprocessing steps applied during training to assess cross-dataset generalization. This evaluation reflects a realistic scenario in which a model trained on a single high-quality dataset is deployed on external images captured under different conditions. Representative qualitative predictions and their corresponding Dice and accuracy values are summarized in Table II. The results show that the model preserved major vessel structures across all datasets, achieving Dice scores between 0.70 and 0.82 and accuracies above 0.95.

Performance remained highest on LES-AV, confirming its superior annotation consistency and resolution. The model also performed competitively on DRIVE, STARE, and CHASE_DB1, demonstrating good recovery of both large vessels and fine branches. Slightly lower Dice values on HRF and IOSTAR reflect the impact of illumination variability, reduced sample size, and heterogeneous labeling, yet the overall vascular topology was still maintained. These findings show that the SegFormer architecture trained on a single high-quality dataset can generalize across multiple external datasets with diverse imaging characteristics.

It is important to note that the last two columns of Table II report the dataset-level mean Dice and accuracy values obtained when applying the LES-AV-trained SegFormer model to each external dataset.

TABLE II. QUALITATIVE SEGMENTATION EXAMPLES FROM THE LES-AV–TRAINED SEGFORMER MODEL ACROSS ALL DATASETS, INCLUDING ORIGINAL IMAGES, GROUND-TRUTH MASKS, PREDICTED MASKS, DICE SCORES, AND ACCURACY.

Datasets	**Original Image**	**Ground Truth**	**Predicted Mask**	**Dice Score**	**Accuracy**
DRIVE				0.70	0.96
CHASE_DB1				0.75	0.96
STARE				0.75	0.96
HRF				0.72	0.96
IOSTAR				0.75	0.95
RETA Benchmark				0.76	0.96
AV-DRIVE				0.78	0.96
LES-AV				0.82	0.96

V. CONCLUSION AND FUTURE WORK

This study evaluated a transformer-based architecture for retinal vessel segmentation using eight publicly available fundus image datasets. Through five-fold cross-validation, the SegFormer model demonstrated stable and competitive performance across datasets with varying resolutions, illumination conditions, and annotation quality. LES-AV yielded the highest Dice and accuracy values, confirming its consistency and suitability for training transformer-based models. Retraining the model on LES-AV and applying it to all remaining datasets further demonstrated strong cross-dataset generalization, with Dice scores ranging from 0.70 to 0.82 and accuracies exceeding 0.95. These results highlight the capability of transformer models to capture both major vascular structures and finer branches across heterogeneous imaging sources.

Future work will focus on improving the transformer-based framework in several directions. SegFormer will be further fine-tuned to enhance its segmentation performance, particularly on thin and low-contrast vessels. Additional state-of-the-art transformer-based models, such as Swin Transformer V2, Mask2Former, TransUNet, and UNETR, will also be evaluated to determine whether more advanced architectures provide superior generalization or computational benefits. In addition, inference time will be analyzed more extensively across datasets to assess the feasibility of real-time or near-real-time deployment in clinical environments.

To further advance this research direction, future studies will incorporate knowledge from recent literature to guide deeper investigation into performance evaluation, inference efficiency, and the practical considerations necessary for clinical implementation. The overarching goal is to develop robust, scalable, and computationally efficient transformer-based segmentation models that support reliable retinal image analysis in real-world ophthalmic workflows.

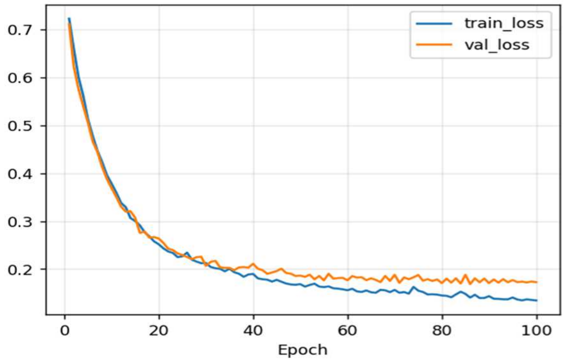


Fig. 8. Training and validation loss curves of the SegFormer model on the LES-AV dataset.

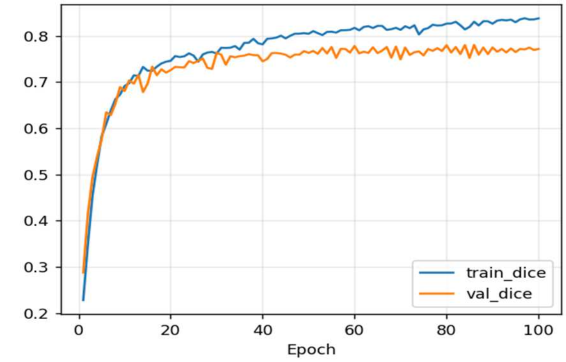


Fig. 9. Training and validation dice curves of the SegFormer model on the LES-AV dataset.

References

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

[2] S. S. Alqahtani *et al.*, "Impact of retinal vessel image coherence on retinal blood vessel segmentation," *Electronics*, vol. 12, no. 2, Art. no. 396, Jan. 2023.

[3] Y. B. Rong *et al.*, "Segmentation of retinal vessels in fundus images based on U-Net with self-calibrated convolutions and spatial attention modules," *Med. Biol. Eng. Comput.*, vol. 61, no. 7, pp. 1745–1755, 2023.

[4] Z. Liu *et al.*, "Deep learning for retinal vessel segmentation: A systematic review of techniques and applications," *Med. Biol. Eng. Comput.*, vol. 63, no. 8, pp. 2191–2208, 2025.

[5] M. T. Sadrabadi and H. Agahi, "Comparative analysis of retinal vessel segmentation utilising convolutional and transformer-based architectures," engrXiv preprint, 2024.

[6] J. Zhou and B. Chen, "Retinal cell damage in diabetic retinopathy," Cells, vol. 12, no. 9, Art. no. 1342, 2023.

[7] A. Khandouzi, A. Ariafar, Z. Mashayekhpour, M. Pazira, and Y. Baleghi, "Retinal vessel segmentation: A review of classic and deep methods," Ann. Biomed. Eng., vol. 50, no. 10, pp. 1292–1314, 2022.

[8] Z. Zeng et al., "Efficient retinal vessel segmentation with 78K parameters," J. Imaging, vol. 11, no. 9, Art. no. 306, 2025.

[9] H. J. Kim, H. Eesaar, and K. T. Chong, "Transformer-enhanced retinal vessel segmentation for diabetic retinopathy detection using attention mechanisms and multi-scale fusion," Appl. Sci., vol. 14, no. 22, 2024.

[10] E. Xie *et al.*, "SegFormer: Simple and efficient design for semantic segmentation with transformers," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 12077–12090, 2021.

[11] J. Lin, X. Huang, H. Zhou, Y. Wang, and Q. Zhang, "Stimulus-guided adaptive transformer network for retinal blood vessel segmentation in fundus images," Med. Image Anal., vol. 89, Art. no. 102929, 2023.

[12] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," IEEE Trans. Med. Imaging, vol. 23, no. 4, pp. 501–509, Apr. 2004.

[13] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," IEEE Trans. Med. Imaging, vol. 19, no. 3, pp. 203–210, Mar. 2000.

[14] M. M. Fraz *et al.*, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, 2012.

[15] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," Int. J. Biomed. Imaging, vol. 2013, Art. ID 154860, 2013.

[16] X. Lyu, L. Cheng, and S. Zhang, "The RETA benchmark for retinal vascular tree analysis," Sci. Data, vol. 9, no. 1, Art. no. 775, 2022.

[17] T. A. Qureshi, M. Habib, A. Hunter, and B. Al-Diri, "A manually-labeled, artery/vein classified benchmark for the DRIVE dataset," in Proc. 26th IEEE Int. Symp. Computer-Based Medical Systems (CBMS), Porto, Portugal, Jun. 2013, pp. 485–488.

[18] J. I. Orlando, J. Barbosa Breda, K. Van Keer, M. B. Blaschko, P. J. Blanco, and C. A. Bulant, "Towards a glaucoma risk index based on simulated hemodynamics from fundus images," in Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI), ser. Lecture Notes in Computer Science, vol. 11071, pp. 65–73, 2018.

[19] J. Zhang, B. Dashtbozorg, E. Bekkers, J. Pluim, R. P. W. Duits, and B. M. ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation," IEEE Trans. Med. Imaging, vol. 35, no. 12, pp. 2631–2644, Dec. 2016.