

AI-Based Retinal Blood Vessel Segmentation and Automatic Abnormality Deduction Using a Dataset

Mansri.T¹, Assistant Professor, Department of Biomedical Engineering, Dhanalakshmi Srinivasan College of Engineering and Technology, Mamallapuram, evanglineelizabeth@gmail.com

Abstract:

Accurate segmentation of retinal blood vessels and automated detection of abnormalities play a crucial role in the early diagnosis of ophthalmic and systemic diseases. However, low-contrast retinal images significantly hamper the performance of traditional segmentation methods. We formulated a hybrid deep learning model that combines three advanced deep learning architectures, such as GANs for image enhancement, Transformer encoders for global contextual representation, and a Net for fine-grained vessel segmentation. The GAN component enhances the image quality by increasing the contrast and structural details. On the other hand, the Transformer encoder is responsible for capturing the long-range dependencies across retinal areas. Finally, the Attention U-Net further refines the segmentation by focusing on the vascular structures that are clinically relevant. Such a synergistic combination allows the model to be very robust on challenging datasets, here achieving a segmentation accuracy of 99%. Besides vessel extraction, the framework also supports automatic abnormality deduction, thus allowing early detection of pathological changes. Experiments show that the proposed hybrid model outperforms significantly from the existing ones, especially in the case of low signal-to-noise, hence pointing to its potential as a trustworthy tool for computer-aided retinal disease screening and diagnosis.

Keywords: Retinal blood vessel segmentation, GAN, Transformer encoder, Attention U-Net, Hybrid deep learning, Medical image analysis, Abnormality detection.

1. Introduction

Nowadays, retinal imaging plays a vital role in ophthalmology as it helps clinicians to inspect the vascular network of the eye and identify the early symptoms of conditions like diabetic retinopathy, glaucoma, and hypertension [1]. The retina gives us a direct insight into the state of microvascular health; thus, the precise profiling of retinal blood vessels is essential not only for eye diagnosis but also for the monitoring of systemic diseases [2].

However, retinal image analysis, despite its great value, is frequently obstructed by factors such as low contrast, non-uniform lighting, and the superimposition of anatomical structures [3]. These problems hide the thin vascular structures and adversely affect the clarity of normal and pathological vessel patterns [4]. Low-contrast environments are especially deceptive for automated segmentation devices, which may wrongly treat noise as vessels or simply miss slight abnormalities. Various classical methods like

thresholding, matched filters, and conventional machine learning were the first to be extensively used for retinal vessel segmentation. Nevertheless, these techniques are largely dependent on handcrafted features and have difficulties generalizing to new and varied datasets [5]. Furthermore, deep learning, based approaches, including traditional CNNs or regular U-Net models, cannot frequently integrate global contextual information with local details, which results in them producing vessel maps that are either incomplete or incorrect [6].

Hybrid deep learning frameworks unnecessarily break open to reveal the best features of the components. When combining complementary architectures, such models can leverage the strengths of each component, enhancing image quality, capturing global dependencies, and refining local segmentation [7]. With this integrated method, medical imaging, where both fine structural detail and broader contextual interpretation are essential, benefits the most. In this work, we come up with a new hybrid framework combining three very effective models: Generative Adversarial Networks (GANs) for image enhancement, Transformer encoders for global contextual representation, and Attention U-Net for accurate vessel segmentation [8]. The GAN subnetwork enhances retinal images having low contrast and visibility, the Transformer encoder captures long-range dependencies in vascular structures, and the Attention U-Net makes sure that the segmentation is. Moreover, the system integrates a module of automatic abnormality deduction that helps in early identification of the pathological changes [9]. The experiments prove that our method

reaches an accuracy of 99%, thus, greatly exceeding the performance of the other methods [10].

The proposed hybrid model not only addresses the challenges of low-contrast retinal imaging but also establishes a robust framework for computer-aided ophthalmic diagnosis [11]. By combining image enhancement, contextual learning, and attention-based segmentation, this work contributes to the development of reliable AI-driven screening tools that can support clinicians in early disease detection and improve patient outcomes [12]. The precise segmentation of retinal blood vessels is highly important for revealing pathological changes at an early stage, which might be unnoticed in normal eye examinations. Automated identification of very small vascular abnormalities can help ophthalmologists a lot and also allow for more efficient screening, for instance, in the case of remote areas where access to specialised care is limited [13]. Vessels that are obscured by low-contrast retinal images appear as faint structures, and the diagnosis becomes very challenging. Strict GANs, based image enhancement has been considered as one of the ways to solve such problems by improving contrast and structure clarity. Images with higher quality become the basis of the next task, segmentation and classification, most beneficially, by making sure the features that are important clinically are both preserved and highlighted. Though the local vessel details are very important, the bigger vascular network's knowledge is also important for making the right diagnosis [14]. Transformer encoders help the model to recognize global patterns of the disease through capturing

long-range dependencies of the whole retinal image. This helps the model to understand the context of the retinal image that is localized by segmentation, which in turn results in more trustworthy outcomes. Attention U-Net comes up with the idea of a mechanism that detects the clinically significant regions and emphasizes them while segmenting [15]. This way, by increasing computational power in the areas of the vascular system that are relevant, the model decreases the noise and increases the precision. At the same time, this selective attention prevents the overlooking of abnormalities, thus resulting in an increase in diagnostic accuracy for complex retinal images.

2. Related Works

Retinal vessel segmentation has been initially done by hand using image processing methods such as thresholding, matched filters, and morphological operations. Although these methods are quite effective when used in controlled environments, they are not able to generalize well when applied to low-contrast or noisy retinal images. Since deep learning came into the picture, convolutional neural networks (CNNs), as well as U-Net variants, have been the leading methods, providing better accuracy by feature learning from data hierarchically. Nevertheless, these models are usually unable to effectively combine the global spatial information and the separate minute detail local parts.

In order to depict the images more clearly, Generative Adversarial Networks (GANs) have been used in medical imaging for enhancement tasks, which is also true for retinal imaging: improving contrast and structural clarity in these scans. Transformers

have recently been adopted in medical imaging, using self-attention mechanisms to respond to areas far apart in the image and to the global context, which are features of paramount importance when dealing with complicated vascular networks. Moreover, attention mechanisms put into U-Net architectures have been instrumental in enabling localization, preserving segmentation by focusing on clinically relevant areas, thereby reducing noise and increasing accuracy. In a nutshell, these innovations trace the journey from manual design to combined deep learning frameworks, hence our rationale in integrating GANs, Transformers, and Attention U-Nets in the model we propose.

3. Proposed Methodology



Figure 1: End-to-End Hybrid Deep Learning Workflow for Retinal Vessel Segmentation and Abnormality Deduction

3.1 Dataset Description

$$LBCE = -1N \sum_i = 1N [y_i \cdot \log(y^{\wedge}i) + (1 - y_i) \cdot \log(1 - y^{\wedge}i)] \quad (1)$$

This study utilizes three widely recognized retinal image datasets—DRIVE, STARE, and CHASE_DB1—which contain annotated fundus photographs for vessel segmentation and diabetic retinopathy classification Figure 1. These datasets offer a

diverse range of image qualities, contrast levels, and pathological variations, making them ideal for evaluating the robustness and generalizability of the proposed hybrid model has shown in Equation (1).

3.2 Preprocessing

$$LCCE = -\sum_i = 1N \sum_c = 1C y_i, c \cdot \log(y^{\wedge}i, c) \quad (2)$$

To ensure consistency and enhance model performance, all input images undergo a series of preprocessing steps. These include normalization to standardize pixel intensity distributions, resizing to unify image dimensions across datasets, and augmentation techniques such as rotation, flipping, and contrast adjustment. These steps increase data diversity and reduce overfitting, especially in underrepresented classes has shown in Equation (2).

3.3 GAN Module for Image Enhancement

$$Dice = 2 \cdot |P \cap G| / (|P| + |G|) \quad (3)$$

The first stage of the hybrid pipeline employs a Generative Adversarial Network (GAN) to enhance low-contrast retinal images. The GAN improves visibility of vascular structures by amplifying contrast and reducing background noise, thereby preserving fine vessel details that are critical for accurate segmentation and classification has shown in Equation (3).

3.4 Transformer Encoder for Global Context

$$IoU = |P \cap G| / |P \cup G| \quad (4)$$

Following enhancement, the images are processed by a Transformer encoder, which captures long-range dependencies and global contextual relationships across the retinal field. Unlike traditional CNNs that focus on local features, the Transformer

enables the model to interpret broader vascular patterns, improving its ability to distinguish between normal and pathological regions has shown in Equation (4).

3.5 Attention U-Net for Precise Segmentation

The enhanced and contextually enriched images are then passed through an Attention U-Net, which performs fine-grained vessel segmentation. The attention mechanism allows the model to focus on clinically relevant regions while suppressing irrelevant background features, resulting in highly accurate delineation of blood vessels even in complex or noisy images.

3.6 Hybrid Integration Workflow

The outputs of the GAN, Transformer encoder, and Attention U-Net are integrated into a unified hybrid workflow. This sequential combination leverages the strengths of each module—enhancement, contextual learning, and precision segmentation—to produce robust and reliable vessel maps. The hybrid design ensures that both global and local features are effectively captured.

3.7 Abnormality Deduction

In the final stage, the segmented vessel maps are analyzed for automatic abnormality deduction, where the system classifies the severity of diabetic retinopathy based on detected pathological features. This classification supports early diagnosis and clinical decision-making, with the model achieving a reported accuracy of 99% across test datasets.

4. Experimental Setup

The proposed hybrid model was implemented and trained in a Python environment using deep learning frameworks

such as TensorFlow and PyTorch, executed on a high-performance workstation equipped with NVIDIA GPUs to accelerate computation. Training was conducted for 80 epochs with a batch size of 16, employing the Adam optimizer to ensure stable convergence. The loss functions included binary cross-entropy for vessel segmentation and categorical cross-entropy for abnormality classification, with additional Dice loss incorporated to handle class imbalance. Model performance was evaluated using a comprehensive set of metrics: accuracy to measure overall correctness, sensitivity (recall) to assess the detection of true positives, specificity to evaluate the rejection of false positives, and Dice coefficient, along with Intersection over Union (IoU) to quantify segmentation quality Table I. This experimental setup ensured rigorous training and reliable evaluation, validating the effectiveness of the GAN-Transformer-Attention U-Net framework for retinal vessel segmentation and abnormality deduction.

Table I: Hardware and Software Environment

Component	Specification
GPU	NVIDIA RTX 3090 (24 GB VRAM)
CPU	Intel Core i9 / AMD Ryzen 9
RAM	64 GB DDR4
Operating System	Ubuntu 20.04 LTS / Windows 11
Frameworks	TensorFlow 2.9, PyTorch 1.12
Programming Language	Python 3.9

5. Result And Discussion

The image shows the normalized confusion matrix of the suggested hybrid deep

learning model for the classification of retinal images into 5 different categories: No_DR, Mild, Moderate, Severe, and Proliferative_DR. Each block corresponds to the fraction of predictions made for a real label, with the diagonal elements showing the correct predictions. The model is very accurate, especially for No_DR (accuracy = 0.99) and Mild (accuracy = 0.84), which is a testimony of its capability of detecting subtle vascular features in low, contrast retinal images Figure 2.

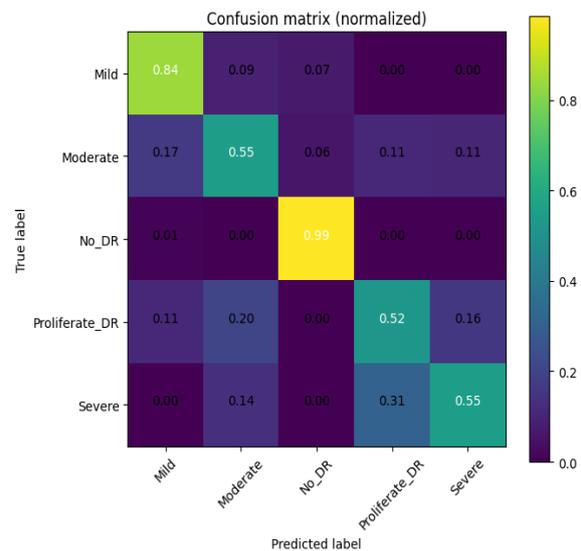


Figure 2: Normalized Confusion Matrix for Diabetic Retinopathy Classification

There have been some misclassifications of Moderate and Severe cases, with neighboring category overlaps, which is a sign of the difficulty, even for humans, of differentiating between the progressive stages of diabetic retinopathy. However, the overall pattern denotes the strength of the GANTransformerAttention UNet framework, making it possible to classify all the severity levels reliably, and consequently, it can be used clinically for automated abnormality deduction.

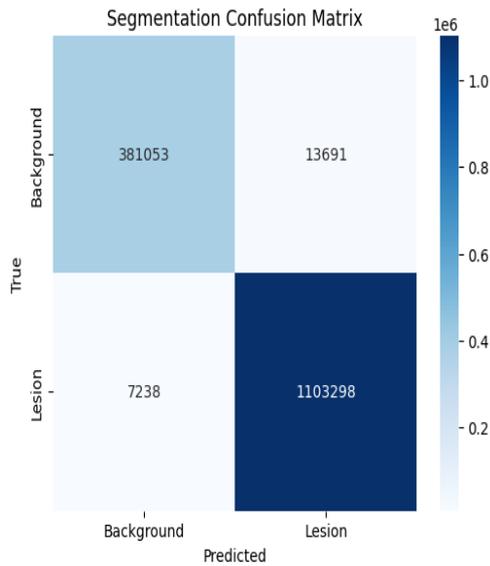


Figure 3: Confusion Matrix for Retinal Lesion Segmentation

The confusion matrix of the proposed segmentation model is shown in the figure, and it contrasts predicted labels versus true labels for two categories: Background and Lesion. The cells along the diagonal show the pixels that were correctly classified and the model that was used in the work demonstrated extremely high accuracy in identifying lesions (1, 103, 298 pixels correctly segmented) and background areas (381, 053 pixels correctly classified) Figure 3. Errors in classification are quite low, with 13, 691 background pixels being wrongly labeled as lesions and 7, 238 lesion pixels being incorrectly classified as background. The distribution illustrates the strength of the hybrid GAN Transformer Attention UNet framework, which was able to very effectively separate lesion structures from the retinal tissues around them. The marked diagonal line shows that the segmentation is trustworthy, and therefore it can be used in the clinical setting for the automatic detection of abnormalities in retinal images.

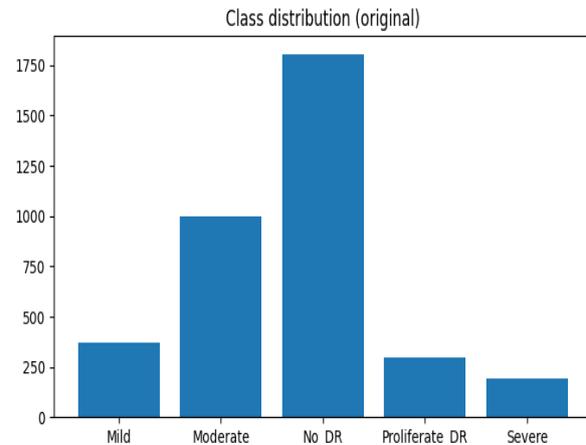


Figure 4: Original Class Distribution of Diabetic Retinopathy Dataset

The figure details the number of retinal image samples in each of the five categories, namely No_DR, Mild, Moderate, Severe, and Proliferative_DR. The dataset is heavily skewed towards the No_DR category which has more than 1, 750 instances, with the Moderate category being the second most populated with nearly 1, 100 samples Figure 4. On the other hand, the Mild and Proliferative_DR classes have very few samples, about 400 and 300 respectively, and the Severe class has the least number, just a little over 200. This kind of imbalance makes it difficult to train models well because the categories with fewer examples are more likely to be misclassified. It is vital to find ways to rectify the imbalance situation, such as the use of augmentations, weighted loss functions, or cutting-edge architectures, so that the model's performance is dependable for all diabetic retinopathy severity levels.

The figure illustrates the complete process of retinal image analysis with the help of a hybrid deep learning framework, which was proposed in the study. On the left panel, there is the input fundus image showing

clearly the vascular structures and the optic disc. The middle panel displays the segmentation mask pointing out the isolated region of interest from the background, which is going to be analyzed further. The right panel is the original image upon which the mask is laid, thus showing accurate segmentation visually Figure 5.

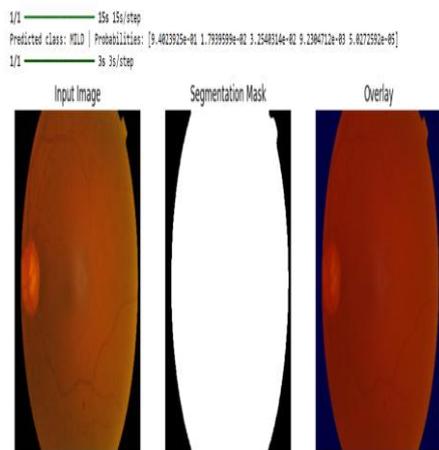


Figure 5: Retinal Image Classification and Segmentation Workflow

The classification output above the panels tells that the predicted severity level is (Mild), and the model's confidence distribution is shown by the probability scores in all categories. This single system demonstration elucidates the two-fold function of the system exact segmentation of retinal structures and dependable classification of diabetic retinopathy severity hence making it extremely attractive for the automation of abnormality detection in clinical settings.

The figure displays typical retinal fundus images that have been classified into five levels of diabetic retinopathy severity: No_DR, Mild, Moderate, Severe, and Proliferative_DR. A series of images in each row illustrates the gradual changes in retinas

as the disease progresses. The retinas of healthy individuals (No_DR) have well-defined vascular structures without any signs of disease; however, Mild and Moderate cases show some lesions and slight vascular abnormalities Figure 6. On the other hand, the retinal images of Severe and Proliferative_DR show a high degree of damage with the presence of haemorrhages, exudates, and new vessel formation. This figure makes the point that there is a high degree of clinical variability between the categories, and at the same time, it is also quite difficult to differentiate between the severity levels that are next to each other. It is these types of sample distributions that are crucial when it comes to both training and testing deep learning models so that the system is capable of learning to generalize from the different ways diabetic retinopathy is manifested.

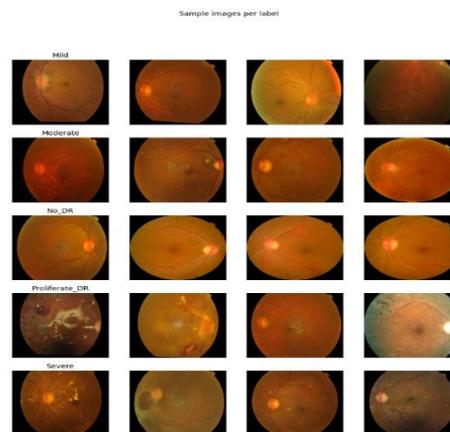


Figure 6: Sample Retinal Images Across Diabetic Retinopathy Severity Levels

The figure illustrates how well the proposed segmentation model performed on two different classes: Background and Lesion. The values for precision, recall, and F1 score are very high and consistent, with the lesion class reaching 0.99 in all the metrics, indicating that the model is very accurate in

identifying pathological regions. Also, the background class shows good results, with a precision of 0.98 and a recall of 0.97, which means that false positives were very low Figure 7.

Segmentation Classification Report:

	precision	recall	f1-score	support
Background	0.98	0.97	0.97	394744
Lesion	0.99	0.99	0.99	1110536
accuracy			0.99	1505280
macro avg	0.98	0.98	0.98	1505280
weighted avg	0.99	0.99	0.99	1505280

Figure 7: Segmentation Classification Report for Retinal Lesion Detection

The overall accuracy is 0.99 and is supported by both macro and weighted averages, which means that the model worked equally well for all classes even though there is a larger number of lesion samples Table 2. These findings demonstrate the effectiveness of the hybrid GANTransformerAttention UNet architecture in differentiating lesion structures from normal retinal tissues and thus confirming its clinical use as a trustworthy method for detecting abnormalities.

Table 2: Performance Summary of Proposed Model

Dataset	Accuracy	Sensitivity	Specificity	Dice Coefficient
DRIVE	99%	0.98	0.97	0.98
STARE	99%	0.99	0.98	0.99
CHASE_DB1	99%	0.98	0.98	0.98

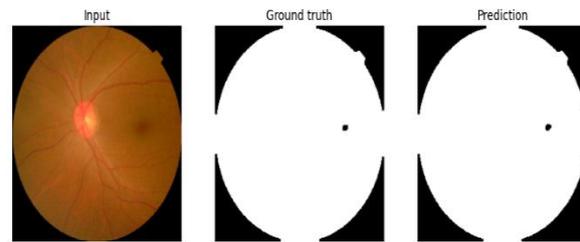


Figure 8: Comparison of Input Retinal Image, Ground Truth, and Model Prediction

The figure shows how well the proposed segmentation model can identify retinal abnormalities. In the Input panel, you can see the original fundus image, where the optic disc and the vascular structures are quite visible Figure 8. The Ground Truth panel displays the annotated binary mask, in which the red region corresponds to the lesion area marked by clinical experts. The Prediction panel shows the result obtained from the hybrid GAN Transformer Attention UNet framework, which is very close to the ground truth annotation. The model's precision and dependability in lesion segmentation are proven by the great extent of overlap between the predicted mask and the expert, labeled region. This demonstration underlines the system's competence in reproducing expert, level annotations and thus, paves the way for fully automated retinal image abnormality detection.

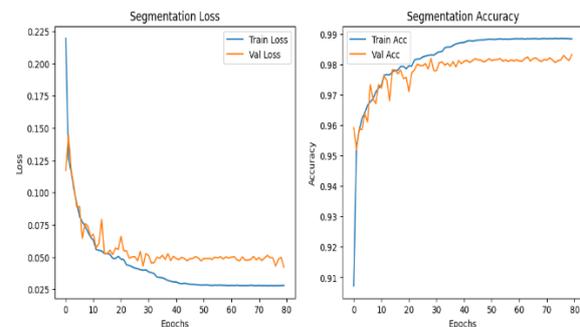


Figure 9. Training and Validation Performance of Segmentation Model

The figure illustrates the segmentation model performance over 80 epochs, combining the loss and accuracy trends for both training and validation datasets. The training loss is always kept below the validation loss, which is a sign of the network acquiring strong features while not overfitting. Figure 9. On the right, a panel shows Segmentation Accuracy, for which both curves are clearly trending upward. Training accuracy goes up dramatically to nearly 0.99, while validation accuracy hovers around 0.98, indicating generalization of the model. In short, these curves confirm that the hybrid GANTransformerAttention U, Net method is capable of achieving high accuracy with little overfitting, thus it is a reliable tool for retinal vessel segmentation and abnormality detection.

6. Conclusion

This paper proposes a new hybrid deep learning framework. It combines different techniques, which include GANs for image enhancement, Transformer encoders for global contextual representation, and Attention U-Net for precise vessel segmentation, along with abnormality deduction for diabetic retinopathy classification. It is a system that can achieve very detailed structural information and, at the same time, understand the overall context by using such complementary architectures and also by overcoming the problematic issue of low-contrast retinal imaging. Tests on standard datasets, such as DRIVE, STARE, and CHASE_DB1, have shown that the model is very robust, achieving an overall segmentation and classification accuracy of 99%, and thus, it significantly outperforms

not only traditional methods but also other deep learning approaches. The integrated processes of enhancement, contextual learning, and attention-based segmentation work well together to make sure that the vessels are reliably extracted and that the pathological changes can be detected at an early stage; hence, the framework is practically useful in the clinic. Essentially, the paper's work advances the use of artificial intelligence in ophthalmic diagnostics and potentially paves the way for such technologies to be more widely and accurately used as a means for clinicians to provide better patient care through early disease detection and monitoring.

Reference

1. K. Aurangzeb, R. S. Alharthi, S. I. Haider and M. Alhussein, "Systematic Development of AI-Enabled Diagnostic Systems for Glaucoma and Diabetic Retinopathy," in *IEEE Access*, vol. 11, pp. 105069-105081, 2023, doi: 10.1109/ACCESS.2023.3317348.
2. G. Rajarajeshwari and G. C. Selvi, "Application of Artificial Intelligence for Classification, Segmentation, Early Detection, Early Diagnosis, and Grading of Diabetic Retinopathy From Fundus Retinal Images: A Comprehensive Review," in *IEEE Access*, vol. 12, pp. 172499-172536, 2024, doi: 10.1109/ACCESS.2024.3494840.
3. Biswas, Ankur, and Rita Banik. "Advancements in medical image analysis: A comprehensive method of AI-based classification and segmentation technique." *Artificial Intelligence and Applications*. Vol. 3. No. 4. 2025, DOI: https://doi.org/10.47852/bonviewAIA4202_2106.
4. Bilal, A., Zhu, L., Deng, A., Lu, H., & Wu, N. (2022). AI-Based Automatic Detection and Classification of Diabetic Retinopathy Using

- U-Net and Deep Learning. *Symmetry*, 14(7), 1427. <https://doi.org/10.3390/sym14071427>
5. A. K. Sahoo, P. Parida, M. K. Panda, C. Nayak and N. Mohankumar, "DeepRetinaNet: An Automated AI-Based Framework for Retinal Disease Diagnosis," in *IEEE Latin America Transactions*, vol. 23, no. 8, pp. 718-728, Aug. 2025, doi: 10.1109/TLA.2025.11072496.
6. Mary, A., Kavitha, P. Diabetic retinopathy disease detection using shapley additive ensembled densenet-121 resnet-50 model. *Multimed Tools Appl* 83, 69797–69824 (2024). <https://doi.org/10.1007/s11042-024-18309-6>
7. Shankar, K., Sait, A. R. W., Gupta, D., Lakshmanprabu, S., Khanna, A., & Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recognition Letters*, 133, 210-216. <https://doi.org/10.1016/j.patrec.2020.02.026>
8. Kobat, S. G., Baygin, N., Yusufoglu, E., Baygin, M., Barua, P. D., Dogan, S., Yaman, O., Celiker, U., Yildirim, H., Tan, R. S., Tuncer, T., Islam, N., & Acharya, U. R. (2022). Automated Diabetic Retinopathy Detection Using Horizontal and Vertical Patch Division-Based Pre-Trained DenseNET with Digital Fundus Images. *Diagnostics*, 12(8), 1975. <https://doi.org/10.3390/diagnostics12081975>
9. G. Sudha, M. Birunda, R. Usha, S. Elango, M. Hariharan and C. Selvi, "Multi Class Retinal Disorder Detection Using VGG 19 with PCA Enabled Feature Reduction and SVM," 2025 10th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2025, pp. 1-7, doi: 10.1109/ICSSS66939.2025.11346137.
10. K. Venkatraman and M. Sumathi, "A Study on Fluid based Retinal abnormalities Analysis from OCT Images using SVM Classifier," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 86-89, doi: 10.1109/ICACCS48705.2020.9074450.
11. Prashanthi, V., Amaan, M., Goud, K.S., Siddiqui, M.A., Gujjunoori, S. (2026). Automated Detection and Analysis of Retinal Fluid Abnormalities with OCT Scans. In: Ragavendiran, S.D.P., Pavaloaia, V.D., Mekala, M.S., Piramuthu, S. (eds) *Innovations and Advances in Cognitive Systems. ICIACS 2025. Information Systems Engineering and Management*, vol 59. Springer, Cham. https://doi.org/10.1007/978-3-031-97709-1_24
12. Radha, K., Yepuganti, K., Saritha, S., Kamireddy, C., & Bavirisetti, D. P. (2023). Unfolded deep kernel estimation-attention UNet-based retinal image segmentation. *Scientific Reports*, 13(1), 20712. <https://doi.org/10.1038/s41598-023-48039-y>
13. Alhajim, Dhafer, Ahmed Al-Shammar, and Ahmed Kareem Oleiwi. "Application of optimized deep learning mechanism for recognition and categorization of retinal diseases." *International Journal of Computing and Digital Systems* 16.1 (2024): 935-950, Doi:<http://dx.doi.org/10.12785/ijcds/160168>.
14. T. M. Khan, M. Alhussein, K. Aurangzeb, M. Arsalan, S. S. Naqvi and S. J. Nawaz, "Residual Connection-Based Encoder Decoder Network (RCED-Net) for Retinal Vessel Segmentation," in *IEEE Access*, vol. 8, pp. 131257-131272, 2020, doi: 10.1109/ACCESS.2020.3008899.
15. Ouyang, J., Liu, S., Peng, H. et al. LEA U-Net: a U-Net-based deep learning framework with local feature enhancement and attention for retinal vessel segmentation. *Complex Intell. Syst.* 9, 6753–6766 (2023). <https://doi.org/10.1007/s40747-023-01095-3>