Towards Explainable Retinal Vessel Segmentation: A Deep Learning Approach

Sounak Sadhukhan
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, UP 201310, India
Email: sounaks.cse@gmail.com

Susmita Das
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, UP 201310, India
Email: susmitad900@gmail.com

Arunabha Tarafdar
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, UP 201310, India
Email: arunabhatarafdar@gmail.com

Aparna Singh
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, UP 201310, India
Email: aparnasingh2211@gmail.com

2025 IEEE 6th India Council International Subsections Conference (INDISCON) | 979-8-3315-1504-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/INDISCON66021.2025.11254539

***Abstract*—Retinal fundus image analysis has been broadly used to identify numerous cardiovascular and eye-related diseases. In ophthalmology, fundus image segmentation is a crucial step in the automated identification and description of blood vessels. As the encoder-decoder architecture based on convolutional neural networks (U-Net) was created for pixel-wise segmentation problems and has already been applied in several fields, it has been used here for retinal blood vessel segmentation. It successfully blends efficient learning, spatial information preservation, and feature extraction. The U-Net's encoder path captures semantic features, and its decoder refines those features for precise localization. This method has been assessed on two different datasets, and encouraging outcomes were found. Because our model generates higher validation accuracy, it produces findings on these datasets that are equivalent to the state-of-the-art approaches. Gradient-weighted class activation mapping is also used to increase this model's explainability. The mapping tool assists us in identifying the areas of the images that have the most influence on the segmentation process and in comprehending the decision-making process.**

***Index Terms*—Fundus Images, Retinal Vessels Segmentation, CNN, U-Net, Gradient-weighted Class Activation Map**

I. Introduction

A fundus image contains the core surface of the eye that provides a visual representation of the retina. It regulates how light is transformed into electrical impulses that are sent straight to the brain, and is sensitive to light. The intricate system of blood vessels known as the retinal vasculature provides the retina with oxygen and nutrients. Therefore, modifications in vessels' morphology, like branching pattern, diameter, and tortuosity, indicate underlying physiological processes and systemic health.

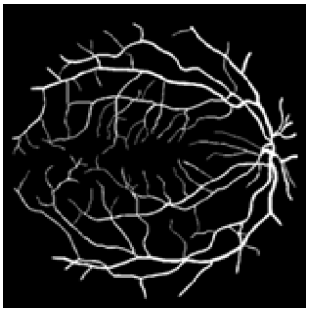
Retinal fundus image analysis has been broadly used to identify numerous cardiovascular [1] and eye-related diseases. Among several features of fundus images, structural characteristics of retinal vasculature are widely utilized as a biomarker for identifying hypertension [2], arteriosclerosis [3], diabetic retinopathy [4], and stroke [5]. Therefore, automatic detection and analysis of these vasculatures are important to avoid visual impairment.

In ophthalmology, fundus image (Fig. 1a) segmentation is one of the crucial steps that consists of automated identification and description of retinal blood vessels (Fig. 1b). Either human or automated methods can be used for this segmentation. However, manual processes are time-consuming, expensive, and require high expertise. It makes the automated segmentation process a highly desirable one as it is quick and effective in pathological cases. However, it is not widely accepted due to structural pathology of retinal vessels like: 1) irregularity in structure, sometimes leaky, or haemorrhages, and drusen; 2) diameter of vessels is either thin or thick; 3) vessels are low in contrast at the background; 4) vessel bifurcation and crossing. There are two types of automated vascular segmentation techniques: supervised and unsupervised.

Supervised algorithms always depend on manual annotation for the labelled actual truth images that are used as references in classification. In order to segment the annotated images, it learns the mapping function from the vessel features. These techniques are further divided into deep neural network (DNN) based methodologies and traditional machine learning algorithms. In the case of classical machine learning, k-nearest neighbour [6] and support vector machine (SVM) [7] are mainly used for segmentation. Ricci et al. [8] used SVM for extracting important features by line detector and segmented vasculatures, while Lupascu et al. [9] used an AdaBoost classifier and a customized feature vector of local intensity, spatial properties, and geometry. On the other hand, Soares et al. [10] applied a Bayesian classifier and the 2-D Gabor wavelet transformation with the pixel intensity value as a feature vector for retinal vessel segmentation. Martin et al. [11] illustrated a different method based on custom 7-D feature vectors and neural networks. Fraz et al. [12] proposed

979-8-3315-1504-1/25/$31.00 ©2025 IEEE

(a) Retinal fundus image sample from the DRIVE dataset in RGB format



(b) Corresponding ground truth image that is manually drawn from sample

Fig. 1: Sample Image from DRIVE Dataset

a method based on ensemble algorithms with bagging and boosting decision trees on custom-created vectors with the help of gradient, morphology, line strength, and Gabor filter response.

Scientists have also come up with some DNN-based approaches [13] [14] [15] [16] that showed promising outcomes as they automatically extract features from the images and learn from them. Fu et al. [17] [18] developed a fully convolutional DNN (CDNN) that is combined with fully connected conditional random fields. Orlando et al. [19] trained fully connected conditional random fields for solving this problem. Li et al. [20] transformed the retinal vessel segmentation into a cross-modality data transformation and then solved it using deep learning. Dasgupta et al. [21] illustrated a deep learning-based iterative process to classify fundus image pixels.

As per the recent trends, CDNN is highly effective for segmenting blood vessels in fundus images [22]. This paper uses a CNN-based U-Net framework for retinal vessel segmentation, as it has successfully applied segmentation in several research areas. Moreover, it can effectively learn from limited-sized datasets and accurately localize objects in the images due to its skip connection, down, and up-sampling. Testing this approach on two datasets, STARE and DRIVE, has produced positive results. It produced results comparable to the state-of-the-art techniques and outperformed other approaches in terms of validation accuracy on these datasets. The gradient-weighted class activation (Grad-CAM) mapping is another tool used to improve this model's explainability. To understand the decision-making process and determine which regions of the pictures have the biggest impact on the segmentation process, we use the Grad-CAM. This tactic increases segmentation accuracy while also providing valuable insights into the behavior of the model, fostering more transparency and confidence in therapeutic procedures.

The structure of this paper is as follows: Sect. II focuses on the solution framework; Sect. III describes the performance of this method and compares it with other models; and Sect. IV concludes this paper.

II. Proposed Framework

As U-Net performed well in a variety of research domains, including satellite imagery and medical imaging, it is utilized in this study to separate the retinal blood vessels from fundus pictures. Besides that, it has some advantages. It can learn effectively from a limited training dataset due to its skip connections and efficient use of feature maps. U-Net accurately localizes objects in the images due to its down- and up-sampling with skip connections. In the next sub-section (Sec. A), our focus will be on U-Net architecture and its characteristics. Our discussion also includes the technique (Grad-CAM) used for models' performance explainability (Sec. B). Moreover, the datasets (STARE and DRIVE) will be described in Sec. C on which the experiment was performed.

A. U-Net

The primary application for U-Net, a CNN-based DNN, is pixel-based segmentation jobs. There are two pathways in it: the expanding path (decoder) and the contracting path (encoder). It looks like a U-shape, which is the main reason for its name. After downsampling the input images using a cascaded convolutional layer (3x3) with a ReLU activation function, the encoder route employs a max-pooling (2x2) with a stride of 2. The feature maps' spatial dimensions are decreased using the max-pooling layer by choosing the maximum value within the immediate region, whereas convolution layers extract complicated properties from the images. This down-sampling helps the network to gain knowledge about more abstract and robust features while reducing computational complexity. At every down-sampling step, the number of channels is doubled, which helps the network to capture complex information. The encoder path learns hierarchical features like capturing object boundaries, textures, etc. The bottleneck serves as a link between the encoder and decoder and is the deepest layer of the U-Net. It contains two convolution layers (3x3) and a ReLU activation function only (no max-pooling is applied). Bottleneck extracts high-level abstract features while maintaining the reduced spatial resolution.

On the other hand, the decoder uses deconvolution techniques to upsample the feature maps to improve spatial resolution. U-Net's skip connections are its most crucial feature. It helps to create a link feature map from the corresponding levels of the encoder to the decoder. Feature maps from the relevant encoders are concatenated with the upsampled features. These facilitate the decoder's access to fine-grained data from the previous phase, which is essential for accurate localization. Along with low-level geographical features, this process also yields high-level semantic information. After concatenation, it uses convolutional layers to refine the segmentation map and further process the combined information. The decoder path learns to reconstruct the input image by preserving the spatial information. A 1x1 convolution layer in the network's output layer gives each pixel in the picture a probability map indicating whether it is part of the background or blood vessels.

TABLE I: The U-Net layers, output form, and quantity of trainable parameters.

Block	Layer Type	Output Shape (H, W, C)	Trainable Parameters	Connected To
-	Input Layer	(512, 512, 3)	0	-
-	Lambda	(512, 512, 3)	0	Input Layer
C1.0	Con.Layer(1.0)	(512, 512, 16)	448	Lambda
	Dropout(1.0)	(512, 512, 16)	0	Con.Layer(1.0)
	Con.Layer(1.1)	(512, 512, 16)	2,320	Dropout(1.0)
-	Max-pooling(1.0)	(256, 256, 16)	0	Con.Layer(1.1)
C2.0	Con.Layer(2.0)	(256, 256, 32)	4,640	Max-pooling(1.0)
	Dropout(2.0)	(256, 256, 32)	0	Con.Layer(2.0)
	Con.Layer(2.1)	(256, 256, 32)	9,248	Dropout(2.0)
-	Max-pooling(2.0)	(128, 128, 32)	0	Con.Layer(2.1)
C3.0	Con.Layer(3.0)	(128, 128, 64)	4,640	Max-pooling(2.0)
	Dropout(3.0)	(128, 128, 64)	0	Con.Layer(3.0)
	Con.Layer(3.1)	(128, 128, 64)	9,248	Dropout(3.0)
-	Max-pooling(3.0)	(64, 64, 64)	0	Con.Layer(3.1)
C4.0	Con.Layer(4.0)	(64, 64, 128)	73,156	Max-pooling(3.0)
	Dropout(4.0)	(64, 64, 128)	0	Con.Layer(4.0)
	Con.Layer(4.1)	(64, 64, 128)	1,47,584	Dropout(4.0)
-	Max-pooling(4.0)	(32, 32, 128)	0	Con.Layer(4.1)
C5.0	Con.Layer(5.0)	(32, 32, 256)	2,95,168	Max-pooling(4.0)
	Dropout(5.0)	(32, 32, 256)	0	Con.Layer(5.0)
	Con.Layer(5.1)	(32, 32, 256)	5,90,080	Dropout(5.0)
-	Con.Transpose(5.0)	(64, 64, 128)	1,31,200	Con.Layer(5.1)
C6.0	Concatenate(5.0)	(64, 64, 256)	0	Con.Transpose(5.0), Con.Layer(4.1)
	Con.Layer(6.0)	(64, 64, 128)	2,95,040	Concatenate(5.0)
	Dropout(6.0)	(64, 64, 128)	0	Con.Layer(6.0)
	Con.Layer(6.1)	(64, 64, 128)	1,47,584	Dropout(6.0)
-	Con.Transpose(6.0)	(128, 128, 64)	32,832	Con.Layer(6.1)
C7.0	Concatenate(6.0)	(128, 128, 128)	0	Con.Transpose(6.0), Con.Layer(3.1)
	Con.Layer(7.0)	(128, 128, 64)	73,792	Concatenate(6.0)
	Dropout(7.0)	(128, 128, 64)	0	Con.Layer(7.0)
	Con.Layer(7.1)	(128, 128, 64)	36,928	Dropout(7.0)
-	Con.Transpose(7.0)	(256, 256, 32)	8,224	Con.Layer(7.1)
C8.0	Concatenate(7.0)	(256, 256, 64)	0	Con.Transpose(7.0), Con.Layer(2.1)
	Con.Layer(8.0)	(256, 256, 32)	18,464	Concatenate(7.0)
	Dropout(8.0)	(256, 256, 32)	0	Con.Layer(8.0)
	Con.Layer(8.1)	(256, 256, 32)	9,248	Dropout(8.0)
-	Con.Transpose(8.0)	(512, 512, 16)	2064	Con.Layer(8.1)
C9.0	Concatenate(8.0)	(512, 512, 32)	0	Con.Transpose(8.0), Con.Layer(1.1)
	Con.Layer(9.0)	(512, 512, 16)	4,624	Concatenate(8.0)
	Dropout(9.0)	(512, 512, 16)	0	Con.Layer(9.0)
	Con.Layer(9.1)	(512, 512, 16)	2,320	Dropout(9.0)
C10.0	Con.Layer(10.0)	(512, 512, 1)	17	Con.Layer(9.1)

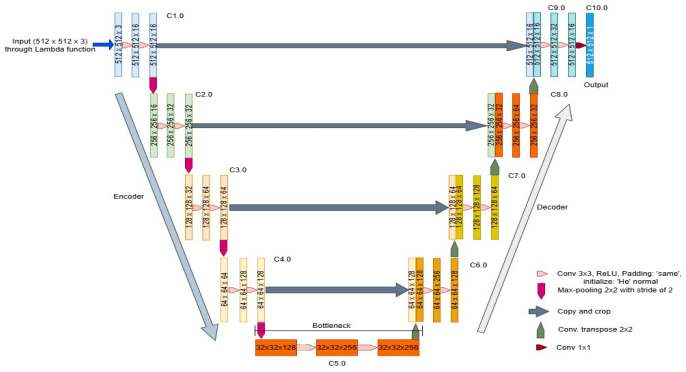


Fig. 2: Diagram of U-Net architecture

A detailed description of the structure of U-Net is provided in Table I and Fig. 2. In our model, we have 19,41,105 trainable parameters in total.

B. Model's Explainability

This paper incorporates explainability techniques (Grad-CAM) not only to improve segmentation accuracy but also to improve the transparency and trustworthiness of the model. The trustworthiness of our model will help the ophthalmologist with its clinical adoption.

Grad-CAM was used for visualizing the most important regions of the image for the model's prediction. It generates a heatmap that identifies the most contributing areas of the image for the model's decision-making to classify a pixel as a vessel or background. The red or yellow portions represent the highly influential areas of the image that contribute to the prediction. If the red or yellow region aligns closely with an actual object in the original image, then it indicates that the prediction of the model is meaningful. If certain regions of misalignment indicate overfitting.

C. Datasets

The performance of our model was recorded on two benchmarked datasets: STARE and DRIVE. These datasets are specifically designed for retina vessel segmentation research.

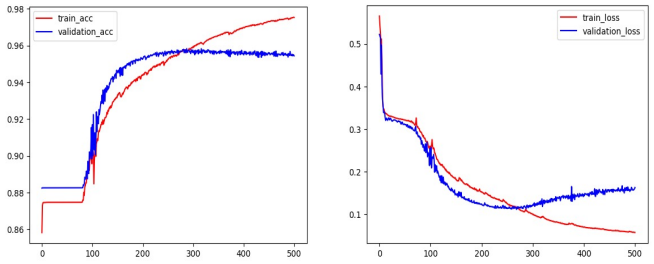
DRIVE contains 40 fundus images in RGB format that include 7 abnormal pathology cases (early-stage diabetic retinopathy). These 40 images are equally divided into a training set and a testing set. Every image has a resolution of 584x565 pixels with 8 bits per color channel. The images were acquired using a Canon CR5 non-mydriatic 3CCD camera with a 45-degree field of view (FOV). It also provides manually segmented vessel images as ground truth. These ground truth images are binary masks, where pixels belonging to the blood vessels are represented as 1, otherwise labelled as 0.

The STARE dataset contains only 20 colored fundus images with a resolution of 700x605. This dataset contains a variety of retinal pathologies like macular degeneration, hypertensive retinopathy, and diabetic retinopathy. Additionally, it offers two distinct sets of hand labeled ground truth photos that were produced by two people. These dual-labelled ground truths allow researchers to identify ambiguous regions and assessment of variability. The ground truth images are binary masks. Blood vessel-related pixels are represented by 1; otherwise, they are represented by 0.

III. Results and Discussions

Two publicly accessible datasets (DRIVE and STARE) are used to assess the proposed methodology. At first, all fundus images are converted to greyscale and shrunk to 512x512 resolution to simplify the training process. Datasets are split into a 90% training and a 10% validation set at random for each epoch. For both datasets, the model is iterated 500 times. Here, we only discuss the execution of the model on the DRIVE, as execution processes on STARE are almost similar.

Initially, the training as well as validation losses are high, but gradually, they decrease with the epochs. At the midway point of the execution (around 250 epochs), the validation loss was minimal. After that, the loss increases, while the training loss still decreases up to 500 epochs. In contrast, training and validation accuracy were around 0.8 (approx.) initially. It went up gradually as every epoch was executed. After 250 epochs (approx.), validation accuracy was flat, though training accuracy still shows an upward movement



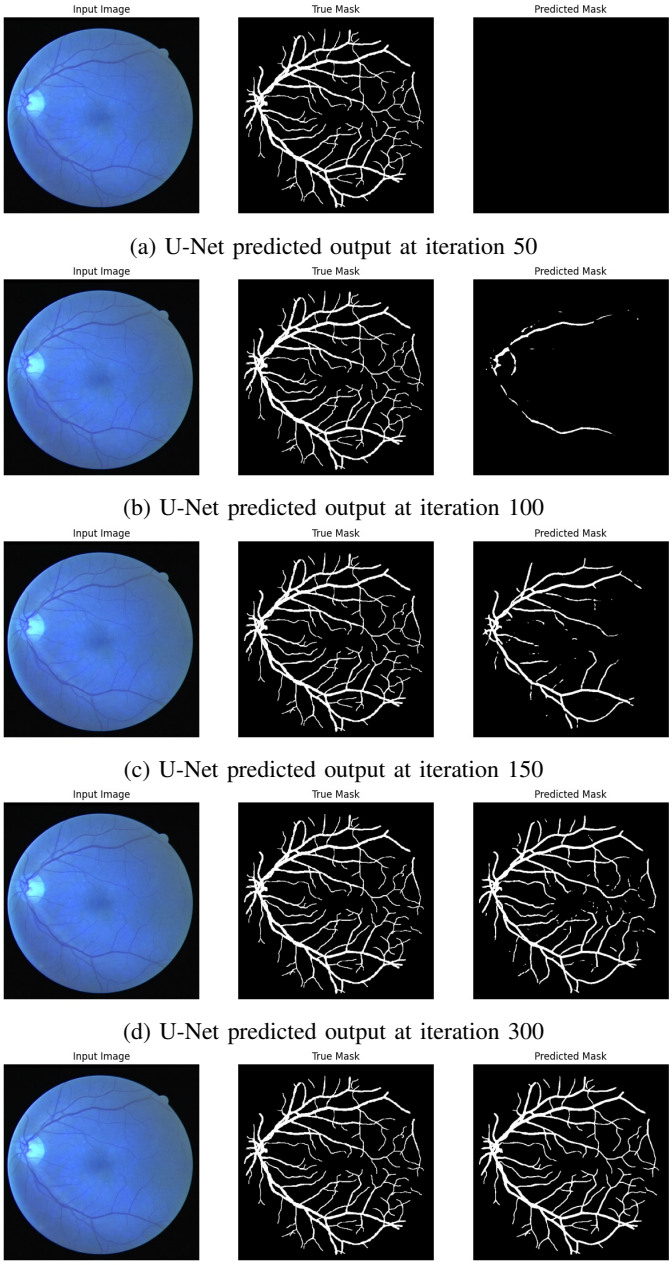
(a) U-Net model accuracy on training and validation up to 500 iterations

(b) Training and validation loss up to 500 iterations in the DRIVE dataset

Fig. 3: Results Obtained from DRIVE Dataset

till 500 epochs. Therefore, we can decide from these two plots depicted in Fig. 3 ((3a) and (3b)) that, training till 250 epochs was justified. Anything beyond it overfits the data. The intermediate results for DRIVE are shown in Fig. 4 ((4a)-(4e)). The model's potential is demonstrated by the fact that its output almost matches the real truth value. Additionally, several of the current supervised state-of-the-art models have been compared (Table II). Our model has demonstrated high validation accuracy for both datasets. The validation accuracy was 0.9578 (DRIVE) and 0.9565 (STARE), respectively, outperforming some of the most advanced models. It was difficult to recognize the blood vessels accurately because of the significant imbalance between the vascular and non-vascular pixels. Even though our model has demonstrated a high likelihood of properly identifying it.

After testing the model, we intend to improve the transparency and credibility of the model. The credibility of our model will help the physician with its clinical adoption. Grad-CAM is applied to understand the model's predictions (Fig. 5). First, Grad-CAM calculates the class score's gradient about the final convolution layer's feature mappings. Then it identifies the important region of the image that contributes extensively to the prediction by weighting the feature map with the gradients. The output picture of the second-to-last convolution layer in our experiment served as the Grad-CAM's input, preserving the image's semantic and spatial information. The Grad-CAM output is displayed in Fig. 5(c). The regions in red or yellow colors represent a high-value zone that indicates a highly influential part of the image during decision-making. This part of the image likely belongs to the retinal blood vessels and our model learns to recognize the structure successfully. In contrast, the blue portion represents the non-blood vessel areas or the background of the image. The bright central spot of the image represents the optic disc is an irrelevant object for this task, though the model showed an over-attention to it. Overall, Grad-CAM indicates that the proposed model strongly depends on the structural characteristics of blood vessels, such as their branching and cylindrical shapes.



(a) U-Net predicted output at iteration 50

(b) U-Net predicted output at iteration 100

(c) U-Net predicted output at iteration 150

(d) U-Net predicted output at iteration 300

(e) U-Net predicted output at iteration 500

Fig. 4: Retinal fundus image sample from the DRIVE dataset, its ground truth and the U-Net predicted output

TABLE II: Comparisons of state-of-the-art methods with proposed method

Article	DRIVE(Accuracy)	STARE(Accuracy)
Staal et al. [6]	0.9441	0.9516
Soares et al. [10]	0.9461	0.9480
Ricci and Perfetti [8]	0.9595	0.9646
Lupascu et al. [9]	0.9597	-
Marín et al. [11]	0.9452	0.9526
Fraz et al. [12]	0.9480	0.9534
Proposed Method	**0.9578**	**0.9565**

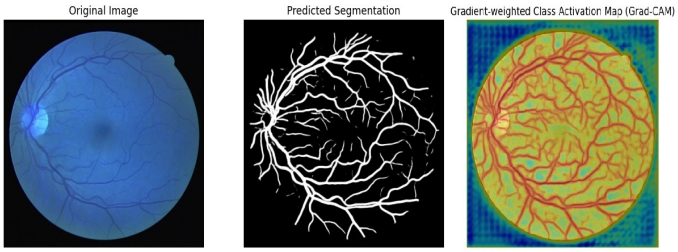


Fig. 5: (a) Retinal fundus image, (b) model's output, (c) outcome of Grad-CAM

IV. Conclusions

Here, U-Net is used for retinal vessel segmentation. It can effectively learn from limited-sized datasets and accurately localize objects in the images. On the STARE and DRIVE, the suggested approach yielded results that were on par with the state-of-the-art approach. Grad-CAM was also used to increase the model's explainability. The mapping tool aids in our comprehension of the decision-making process and identifies the areas of the photos that have the greatest influence on the segmentation procedure. This technology increases confidence and transparency in healthcare procedures by improving segmentation accuracy and providing useful information about the model's behavior. Overall, the experimental results reflect good performance on normal and pathological images. Hyperparameter tuning of this model and its application to a large-scale dataset constitute key components of our future research plan.

Acknowledgement

We are sincerely thankful to Dr. Siddhartha Ghosh, M.B.B.S., M.S. (ophthalmology), Director, Global Eye Hospital, Kolkata, West Bengal 700098 for his valuable suggestions and inputs.

References

[1] S. Chikumba, Y. Hu, and J. Luo, "Deep learning-based fundus image analysis for cardiovascular disease: a review," *Therapeutic Advances in Chronic Disease*, vol. 14, p. 20406223231209895, 2023.

[2] G. Dai, W. He, L. Xu, E. E. Pazo, T. Lin, S. Liu, and C. Zhang, "Exploring the effect of hypertension on retinal microvasculature using deep learning on east asian population," *PloS one*, vol. 15, no. 3, p. e0230111, 2020.

[3] R. G. Barriada and D. Masip, "An overview of deep-learning-based methods for cardiovascular risk assessment with retinal images," *Diagnostics*, vol. 13, no. 1, p. 68, 2022.

[4] J. Sahlsten, J. Jaskari, J. Kivinen, L. Turunen, E. Jaanio, K. Hietala, and K. Kaski, "Deep learning fundus image analysis for diabetic retinopathy and macular edema grading," *Scientific reports*, vol. 9, no. 1, p. 10750, 2019.

[5] Z. Girach, A. Sarian, C. Maldonado-García, N. Ravikumar, P. I. Sergouniotis, P. M. Rothwell, A. F. Frangi, and T. H. Julian, "Retinal imaging for the assessment of stroke risk: a systematic review," *Journal of Neurology*, vol. 271, no. 5, pp. 2285–2297, 2024.

[6] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[7] X. You, Q. Peng, Y. Yuan, Y.-m. Cheung, and J. Lei, "Segmentation of retinal blood vessels using the radial projection and semi-supervised approach," *Pattern recognition*, vol. 44, no. 10-11, pp. 2314–2324, 2011.

[8] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE transactions on medical imaging*, vol. 26, no. 10, pp. 1357–1365, 2007.

[9] C. A. Lupascu, D. Tegolo, and E. Trucco, "Fabc: retinal vessel segmentation using adaboost," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 5, pp. 1267–1274, 2010.

[10] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE Transactions on medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.

[11] D. Marín, A. Aquino, M. E. Gegúndez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on medical imaging*, vol. 30, no. 1, pp. 146–158, 2010.

[12] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[14] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[15] L.-C. Chen, "Semantic image segmentation with deep convolutional nets and fully connected crfs," *arXiv preprint arXiv:1412.7062*, 2014.

[16] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*. Springer, 2016, pp. 140–148.

[17] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in *2016 IEEE 13th international symposium on biomedical imaging (ISBI)*. IEEE, 2016, pp. 698–701.

[18] H. Fu, Y. Xu, S. Lin, D. W. Kee Wong, and J. Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*. Springer, 2016, pp. 132–139.

[19] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 16–27, 2016.

[20] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE transactions on medical imaging*, vol. 35, no. 1, pp. 109–118, 2015.

[21] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*. IEEE, 2017, pp. 248–251.

[22] A. E. Ilesanmi, T. Ilesanmi, and G. A. Gbotoso, "A systematic review of retinal fundus image segmentation and classification methods using convolutional neural networks," *Healthcare Analytics*, vol. 4, p. 100261, 2023.