Retinal Vessel Segmentation Based on the Anam-Net Model

Khursheed Aurangzeb[1, *], Syed Irtaza Haider[2], Musaed Alhussein[1]

[1]*Department of Computer Engineering, College of Computer and Information Sciences, King Saud University,*
Riyadh 11543, Saudi Arabia
[2]*College of Computer and Information Sciences, King Saud University,*
Riyadh 11543, Saudi Arabia
kaurangzeb@ksu.edu.sa

Abstract—Accurate segmentation of retinal blood vessels can help ophthalmologists diagnose eye-related diseases such as diabetes and hypertension. The task of segmentation of the vessels comes with a number of challenges. Some of the challenges are due to haemorrhages and microaneurysms in fundus imaging, while others are due to the central vessel reflex and low contrast. Encoder-decoder networks have recently achieved excellent performance in retinal vascular segmentation at the trade-off of increased computational complexity. In this work, we use the Anam-Net model to accurately segment retinal vessels at a low computational cost. The Anam-Net model consists of a lightweight convolutional neural network (CNN) along with bottleneck layers in the encoder and decoder stages. Compared to the standard U-Net model and the R2U-Net model, the Anam-Net model has 6.9 times and 10.9 times fewer parameters. We evaluated the Anam-Net model on three open-access datasets: DRIVE, STARE, and CHASE_DB. The results show that the Anam-Net model achieves better segmentation accuracy compared to several state-of-the-art methods. For the DRIVE, STARE, and CHASE DB datasets, the model achieved {sensitivity and accuracy} of {0.8601, 0.9660}, {0.8697, 0.9728}, and {0.8553, 0.9746}, respectively. On the DRIVE, STARE, and CHASE_DB datasets, we also conduct cross-training experiments. The outcome of this experiment demonstrates the generalizability and robustness of the Anam-Net model.

Index Terms—Anam-Net; Deep learning; Data augmentation; Retinal vessel segmentation; Semantic segmentation.

I. INTRODUCTION

Partial and sometimes complete blindness is caused due to the longer duration of chronic eye diseases, including Glaucoma, diabetic retinopathy (DR), and cataracts [1]. These diseases slowly deteriorate different parts of the eye, such as retinal vessels, optic cup (OC), and optic disc (OD). Some people experience substantial effects of these chronic diseases mainly due to their ocular weakness or the severity of the chronic condition. It means that there should be easy access to monitor eye health regularly.

The easy and inexpensive availability of regular eye health monitoring procedures will enable us to make a timely diagnosis of the disease, which, in turn, will allow us to avoid the disease or at least delay its impact to a later stage in life. The manual procedures used by physicians/ophthalmologists for the diagnosis of these eye diseases are exhaustive and time-consuming, which sometimes also have difference due to inter- and intra-observer variations. The development of automated methods for the detection of OC, OD, and retinal vessels is of prime importance for the prognosis and diagnosis of Glaucoma and DR [2], [3].

There are two types of glaucoma: closed-angle and open-angle. The first type occurs when some parts of the iris of the eye block the flow of fluid, resulting in increased pressure in the eye. Ocular pain, redness of the eye, significant intraocular pressure, and partial vision loss are all symptoms of this kind of Glaucoma. On the other hand, in its second type, the flow of the fluid is not affected. Patients with this type of Glaucoma cannot feel or notice the symptoms until the disease has progressed to an advanced stage, at which point it is irreversible [4].

In manual procedures, the retinal vessels in the retinal fundus images are segmented by the optometrist and ophthalmologists, which is a tiring task and, at the same time, highly susceptible to doctor-to-doctor observation-variation and may also attribute huge discrepancies due to trainer fatigue. As a result, manual segmentation approaches have an inherent limitation on the quality of the extracted data. On the other hand, diagnostics based on computer aided design (CAD) tools is very efficient and can be utilized with high accuracy for the automated diagnosis of chronic eye diseases in a timely manner. Additionally, these automated diagnostics systems based on CAD tools can be used for population-scale screening programs [5].

In retinal fundus images, semantic segmentation is used to detect OC, OD, and retinal vessels. Semantic segmentation is useful for recognizing these constituent parts of the eye, as well as other abnormalities that are needed to characterize chronic eye diseases such as Glaucoma [6], DR [1], [7], age-related macular degeneration (AMD) [8], vascular occlusions [9], and

chronic systemic hypoxemia [10].

With recent developments and progress in several fields, including high-performance computing (HPC), machine learning, deep learning, image processing, and machine vision, researchers and academicians have started to devise and explore them for problems in different domains of life. Specifically, several researchers devised deep learning models for pixel-wise semantic segmentation for the detection of OC, OD, and retinal vessels. Most of the developed methods so far characterize huge computational load in addition to attributing lower impact/accuracy without pre-processing. Although some of these previous explorations have achieved better accuracy, very little attention has been paid to the computational complexity of the developed model. This necessitates the development of lightweight deep learning models that have a suitable execution time during both training and testing.

For the diagnosis of DR, retinal vascular segmentation in retinal fundus images must be precise and accurate. Maintaining a reasonable computational complexity for a deep learning (DL) model while aiming for higher accuracy is difficult due to several factors such as tortuosity, higher density, vessel shape, diameter, and the formation of various types of lesions such as microaneurysms, hard exudates, soft exudates, and cotton wool in acquired retinal images [11]. With the progression of chronic eye diseases, there are numerous other issues, such as vessel crossing, centreline reflex, vessel branching, and small vessel formation.

Researchers must take equal consideration to each of the issues described above when developing machine learning and deep learning models for retinal vascular segmentation in retinal fundus images. The previously published convolutional neural network (CNN)-based models in the literature have a higher computational complexity; our goal in this paper is to investigate an alternative model for semantic segmentation of retinal fundus images that requires much less trainable parameters compared to the state-of-the-art. The authors in [12] developed the Anam-Net model, which is based on the encoder-decoder architecture. They used Anam-Net for the classification of COVID-19 patients in chest computed tomography (CT) images.

The underlined hypothesis is to investigate and explore a deep learning model, which is deep enough to achieve competitive evaluation metrics compared to best models from the state-of-the-art and at the same time is realizable using normal computing/embedded resources available at the forefront of any healthcare facility. The realization requirement of our hypothesis warrants exploring and developing a deep learning model, which attributes significantly lower trainable parameters and memory overhead.

Hence, the focus of our work is not only on developing deep learning-based automated tool for vessel segmentation, but also to make such tool realizable on ordinary computer systems, which are usually available at forefront in healthcare facility.

The motivation for this work is taken from Anam-Net, which we adapted for vessel detection in retinal fundus images. The fact that it has fewer hyperparameters (resulting in lower computing cost) while preserving competitive

evaluation metrics is our primary rationale for adopting it in our study. These attributes make a significant difference when the overall goal is to use the developed model in real-life scenarios in mobile healthcare or point of care in hospitals, where the heavy computating facility is not available. Hence, the attributes of being lightweight and having higher accuracy make the Anam-Net a better and more reliable option to be employed/used at the point of care or mobile health care facility for large-scale screening programs of the governments.

To the best of our knowledge, this is the first attempt to apply Anam-Net for vessel segmentation in retinal fundus images. The contributions are provided below:

− We used Anam-Net to segment retinal vessels and were able to achieve competitive results in terms of evaluation metrics while keeping the computational complexity to a minimum.

− We chose three publicly available retinal image databases, STARE, DRIVE, and CHASE_DB, to evaluate the performance extensively and to compare the segmentation results to rivals from the state-of-the-art.

− We conducted cross-training experiments on the DRIVE, STARE, and CHASE_DB datasets and obtained results that demonstrate the generalization ability and robustness of Anam-Net.

The remaining manuscript is organized in the following way. The review of the literature for the segmentation of retinal vessels using machine learning and deep learning models is presented in Section II. The description and implementation details of the basic Anam-Net model are provided in Section III. The results and a detailed comparison of applying the Anam-Net model for retinal vessel segmentation are given in Section IV. Finally, Section V provides conclusions and suggestions for future work.

II. RELATED WORKS

Retinal vessel segmentation is an important task for diagnosing eye diseases, which could be done using manual and automated procedures. Recently, great interest has been shown in applying machine learning and deep learning models for accurate and precise segmentation of the retinal vessels. Accurate segmentation of the retinal vessels helps ophthalmologists reliably diagnose different chronic eye diseases such as Glaucoma and DR. Automated retinal vessel segmentation procedures can be divided into two categories, unsupervised and supervised.

In the literature, different approaches have been adapted for retinal vessel segmentation, such as morphological, kernel-based, vessel tracking, multiscale, model-based, and thresholding [13]. The authors in [14] attempted to apply the kernel-based method to the segmentation of the retinal vessels, where they assumed that the width of the vessels remains constant with distance. This is not a realistic assumption, which limits the kernel-based method for real-life implementation, since both the orientation and width of a retinal vessel are continuously changing. The vessel tracking procedures select a set of starting points to trace the ridges of vessels in images of the retinal fundus [15]. The limitation of the vessel tracking procedure is the user's

intervention to follow the ending and starting points. In morphological methods, mathematical equations are used for segmentation of retinal vessels, where mainly top hat operators are applied [16]. The multi-scale detectors were adapted by the authors in [17] for the segmentation of retinal vessels at different orientations and scales. These procedures are efficient, but their performance is degraded for the case of thin vessels. The thresholding-based procedures were implemented by the authors in [18], sometimes leading to an unconnected vascular structure. Model-based approaches perform retinal vessel segmentation, where the basic assumption is to consider the vessels as flexible curves [19]. Model-based approaches are highly sensitive to contrast variations. The interested readers are referred to the extensive review work on unsupervised vessel segmentation performed by the authors in [20].

The authors in [21] developed a model based on a lattice neural network (LNN) for the segmentation of retinal vessels, where the feature extraction was applied as well as reduction. The results were compared to multilayer perceptron (MLP) and support vector machine (SVM) models. They achieved a dice score of 0.66 and 0.69 for the STARE and DRIVE databases.

The authors in [22]–[24] used fully convolutional network (FCN)-based deep neural network (DNN) models for retinal vessel segmentation. However, in the case of [22], there were a large number of false-positive for fundus images having severe pathologies in the optic disc region. On the other hand, the authors in [23] achieved a better segmentation accuracy, but their computational complexity was higher. The authors in [24] also achieved better segmentation accuracy. However, the performance achieved for cross-training results was poor. Specifically, the model trained on the STARE database and tested on images from the DRIVE database produced lower sensitivity, which means lower generalization capability of the developed model.

The authors in [25]–[27] developed and evaluated the different variations of the CNN-based model. The main limitation was the weight gap between the conditional random field (CRF) and CNN, which degraded the overall performance of the developed model. The authors in [28] developed a generative adversarial network (GAN)-based model, where they used a generator along with a patch-based discriminator. The main limitation of the GAN-based models is higher sensitivity to tuning the hyperparameter, higher degree of overfitting, non-convergence, as well as generator gradient vanishing. Due to these different disadvantages, GAN-based models are not suitable for semantic pixel-wise segmentation of retinal fundus images.

Recently, many researchers have targeted retinal vessel segmentation for eye disease classification by developing a different variant of the U-Net model. The authors in [29] developed a Recurrent Residual CNN named "R2U-Net model" for segmentation of retinal vessels. They obtained better dice scores for the images of the CHASE_DB and STARE databases. But they did not validate the generalization capability of their developed model based on the cross-database training and testing strategy. In another attempt, the authors in [30] developed a U-Net-based model,

where they performed cross-database training and testing. However, they achieved lower sensitivity for both the case of cross-database training/testing and experiments performed on images of individual datasets. Furthermore, their model attributed a very high computational complexity based on the possession of 36M trainable parameters.

According to [31], a model based on U-Net and deformable convolutional units was proposed. Their proposed model uses a patch size of 48×48 and replaces the original convolutional layer with a deformable convolutional block. The results indicate that their model achieved low sensitivity compared to several state-of-the-art methods. We noticed that other researchers attempted to segment retinal vessels using different variants of the U-Net model [32]–[39]. In most of these studies, the results indicate that breakage of the small retinal vessels occured in the binarization stage, which requires heavy post-processing operations. The heavy post-processing steps in addition to huge computational complexity of these models limits their capability from considering them for point of care deployment in the front end in hospital, which is the main focus of the current exploration.

Due to the advancements in deep learning, medical image analysis is experiencing a paradigm shift [46]. Despite the rapid development and exploration of machine learning and deep learning models for retinal vessel segmentation, numerous problems remained unresolved, which warrant significant consideration for researchers working in the field of medical image segmentation. Some of the previous studies have given some preference to reduce the computational complexity of the model, but we think that there is still much room for improvement. Specifically, earlier explorations did not pay the due attention to handling the memory overhead. The larger memory overhead in addition to the significantly higher computational complexity of the DNN model limits their application from their deployment equipment for large-scale screening programs, which are possible only monitoring the population during their general check-ups at the point of care in the hospital. Additionally, most of the previous explorations avoided the cross-database training and testing experiment, due to which their generalization capability is limited.

We aim to reduce the computational complexity of the deep learning model by reducing the number of trainable parameters. We attempt to reduce the memory overhead of our previously developed model by employing the concept of skip connections after completing max-pooling jobs to share its indices across the encoder and decoder stages [40]. The idea of introducing skip connections and sharing the max-pooling indices helped to improve feature map resolution, which not only reduced computational complexity, but also improved the various evaluation metrics. However, the total number of parameters achieved was more than 8 million, which could be further reduced by looking at current research into deep learning models in other areas of medical image segmentation. Table I lists the methodologies from the literature that we used to compare our segmentation results with.

TABLE I. LIST OF METHODS FOR COMPARISON.

Reference	Method
Orlando, Prokofyeva, and Blaschko (2017) [22]	Joint FCN-CRF
Mo and Zhang (2017) [23]	FCN with Deep Feature Fusion
Oliveira, Pereira, and Silva (2018) [24]	FCN with Stationary Wavelet Transform
Zhou, Yu, Xu, Gu, and Yang (2017) [25]	Feature Learning with Dense CRF
Feng, Zhuo, Pan, and Tian (2020) [27]	Cross-connected Network
Abbas, Shakeel, Khurshid, and Taj (2019) [28]	GAN
Alom, Yakopcic, Hasan, Taha, and Asari (2019) [29]	R2U-Net
Yan, Yang, and Cheng (2018) [30]	Joint-loss framework based on U-Net
Jin, Meng, Pham, Chen, Wei, and Su (2019) [31]	Deformable U-Net
Wang, Zhao, and Yu (2021) [33]	Nest U-Net and Patch-Learning
Zhang, Zhang, and Xu (2021) [34]	Pyramid U-Net
Tang, Rui, Yan, Li, and Hu (2020) [35]	ResWNet
Lv, Ma, Li, and Liu (2020) [38]	Attention U-Net
Zhuang (2018) [39]	Ladder-Net
Khan, Alhussein, Aurangzeb, Arsalan, Naqvi, and Nawaz (2020) [40]	RCED-Net
Li, Feng, Xie, Liang, Zhang, and Wang (2016) [41]	Cross-modality Approach
Noh, Park, and Lee (2019) [42]	SSANet
Wu et al. (2019) [43]	Vessel-Net
Yang, Li, Guo, and Zhou (2021) [44]	CNN + Cascade Forrest
Uysal and Güraksin (2021) [45]	Patch-based CNN

III. MATERIALS AND METHODS

In this section, first, the fundus images dataset will be discussed along with some pre-processing steps involved in preparing the images for the deep learning model. Next, we will present the Anam-Net architecture and, in the last, we will discuss the implementation details along with the evaluation metrics used to compare the Anam-Net with the state-of-the-art methods.

A. Fundus Image Dataset

The Anam-Net model was tested on three publicly available datasets: DRIVE [47], STARE [14], and CHASE_DB [48]. The DRIVE dataset is made up of 40 fundus images with a resolution of 565×584 pixels that were collected as part of a diabetic retinopathy screening program. The 40 images were separated into two groups, each with 20 images: a training set and a test set. A total of 20 fundus images with a resolution of 605×700 pixels make up the STARE dataset collection. The STARE dataset, unlike the DRIVE dataset, does not include separate training and test data. In this paper, we use a leave-one-out technique for the STARE dataset, in which the model is trained on n-1 samples and evaluated on the remaining one sample. There are 28 images in the CHASE_DB dataset, each with a resolution of 999×960 pixels. A total of 20 images were used to train the network, and the remaining 8 images being used for testing.

The first observer's manual annotations serve as ground truth for our evaluation metrics in all three datasets, which contain two manual segmentations of fundus pictures. Because the image size differs between fundus images in different datasets, we scaled the image to a resolution of 576×576 pixels. We rescale the output probability map to the original size of the test image using bilinear interpolation. The segmentation performance of the Anam-Net is evaluated in the original size of the test image, and thus ensures that the results are not skewed by scale variations.

B. Image Pre-Processing

Pre-processing is essential to suppress noise and improve contrast in the acquired retinal fundus images since they may have non-uniform brightness and intra- and inter-image contrast variations. In this work, we first convert the RGB fundus image to the lightness, channel A, channel B, (LAB) color scheme and apply contrast-limited adaptive histogram equalization (CLAHE) to the lightness channel. The improved L-channel is then combined with the original A and B channels. After that, the image is converted back to the RGB color space, from which the enhanced green channel is retrieved. To further enhance the vessel structure, gamma transformation is used as last pre-processing step.

C. Anam-Net Architecture

The encoder-decoder-based architecture and AD-Block have been combined to make a lightweight CNN-based model. The block diagram of the Anam-Net architecture is shown in Fig. 1. Like the U-Net model, it has a contracting path (encoder) and an expanding path (decoder). With a kernel size of 3 and a filter size of 64, the input fundus image is passed to the convolution layer. A 3×3 convolution is followed by batch normalization and a rectified linear unit (ReLU) layer in each convolution block. A max-pooling layer is incorporated after the convolution block to reduce the dimensionality of the feature by a factor of two. After that, the AD-block is used to learn robust features. The details of the AD-block are shown in Fig. 2. The AD-block comprises of 1×1 depth squeezing convolution, 3×3 feature extraction convolution, and 1×1 depth stretching convolution.

Each encoder and decoder stage in the Anam-Net model has three AD-blocks. In the expansion path, before AD-block, transpose convolution is applied to upsample the feature map to the desired resolution. At the decoder stage, the learnt attributes from the encoder are combined with the layers from the expansion path, enabling the network to learn at multiple levels. For more details on the Anam-Net model, please see the following article [12]. The layer-wise details of the Anam-Net model are shown in Table II.

D. Implementation Details

To train the model, we utilize the Keras deep learning package. With a learning rate of 0.001, a well-known Adam optimizer is utilized. The learning rate is reduced by a factor of 0.1 if the validation loss does not improve after 10 epochs. With a batch size of 4, the model is trained for 150 epochs. To avoid overfitting, we use validation loss as an early stopping criterion. We employ a log dice loss, which focuses on labels that are less accurate [49].
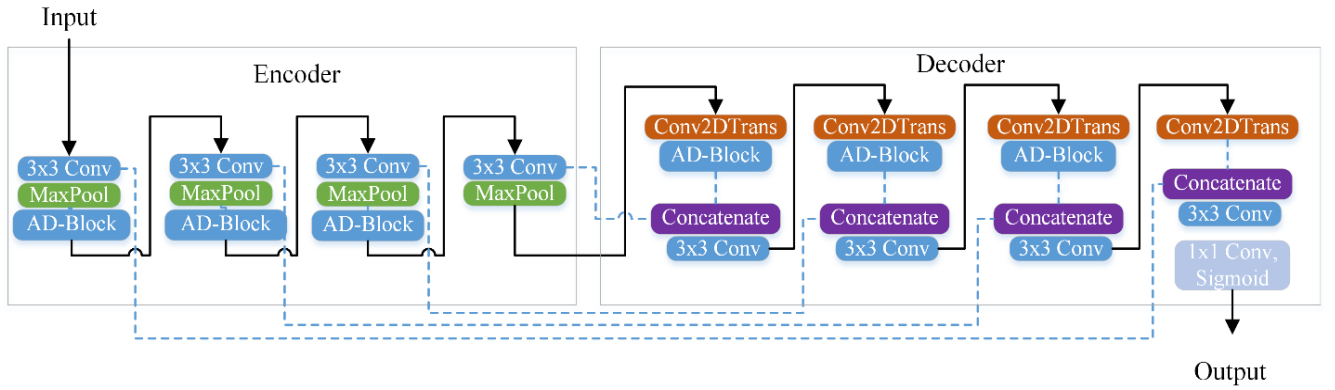
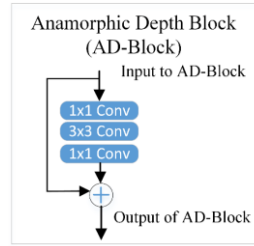Fig. 1. Anam-Net architecture for the segmentation of retinal vessels.

Fig. 2. Architecture of the AD-Block.

TABLE II. ARCHITECTURE DETAILS OF ANAM-NET MODEL.

	Layer	Label	Size	Feature Map Size	Parameter Info
Encoder	1st Convolution Block	EConv1	3×3×64	576×576×64	640
	1st Pooling Layer	Pool-1	2×2	288×288×64	
	1st AD-Block	EAD1Conv1	1×1×16	288×288×16	1040
		EAD1Conv2	3×3×16	288×288×16	2320
		EAD1Conv3	1×1×64	288×288×64	1088
	2nd Convolution Block	EConv2	3×3×128	288×288×128	73856
	2nd Pooling Layer	Pool-2	2×2	144×144×128	
	2nd AD-Block	EAD2Conv1	1×1×32	144×144×32	4128
		EAD2Conv2	3×3×32	144×144×32	9248
		EAD2Conv3	1×1×128	144×144×128	4224
	3rd Convolution Block	EConv3	3×3×256	144×144×256	295168
	3rd Pooling Layer	Pool-3	2×2	72×72×256	
	3rd AD-Block	EAD3Conv1	1×1×64	72×72×64	16448
		EAD3Conv2	3×3×64	72×72×64	36928
		EAD3Conv3	1×1×256	72×72×256	16640
	4th Convolution Block	Econv4	3×3×256	72×72×256	590080
	4th Pooling Layer	Pool-4	2×2	36×36×256	
Decoder	1st Transpose Convolution	DTConv1	3×3×256	72×72×256	262400
	1st AD-Block	DAD1Conv1	1×1×64	72×72×64	16448
		DAD1Conv2	3×3×64	72×72×64	36928
		DAD1Conv3	1×1×256	72×72×256	16640
	1st Concatenate Layer	DConc1		72×72×512	
	1st Conv. Block	DConv1	3×3×256	72×72×256	1179904
	2nd Transpose Convolution	DTConv2	3×3×256	144×144×256	262400
	2nd AD-Block	DAD2Conv1	1×1×64	144×144×64	16448
		DAD2Conv2	3×3×64	144×144×64	36928
		DAD2Conv3	1×1×256	144×144×256	16640
	2nd Concatenate Layer	DConc2		144×144×512	
	2nd Conv. Block	DConv2	3×3×256	144×144×256	1179904
	3rd Transpose Convolution	DTConv3	3×3×128	288×288×128	131200
	3rd AD-Block	DAD3Conv1	1×1×32	288×288×32	4128
		DAD3Conv2	3×3×32	288×288×32	9248
		DAD3Conv3	1×1×128	288×288×128	4224
	3rd Concatenate Layer	DConc3		288×288×256	
	3rd Conv. Block	DConv3	3×3×128	288×288×128	295040
	4th Transpose Convolution	DTConv4	3×3×64	576×576×64	32832
	4th Concatenate Layer	DConc4		576×576×128	
	4th Conv. Block	DConv4	3×3×64	576×576×64	72792
	Final Conv. with Sigmoid	Conv	1×1×1	576×576×1	65

There are only a few training images in each of the three datasets: 20 for DRIVE, 19 for STARE (leave-one-out technique), and 20 for the CHASE_DB. It is very difficult to train a deep learning model with such a short dataset and achieve acceptable segmentation accuracy. We use a variety of data augmentation techniques to improve the robustness and generalization capabilities of the network. Horizontal and vertical flips, random rotations in the range of [0, 360] degrees, random width and height shifts in the range of [0, 0.15], and random magnification in the range of [0.3, 0.12] were among the data augmentation strategies used.

All calculations were performed on IBEX at the King Abdullah University of Sciences and Technology (KAUST) High-Performance Computing (HPC) facility, where we employed a single RTX 2080 Ti GPU for all experiments.

E. Evaluation Metrics

Our proposed model generates a probability prediction map that depicts the likelihood that a pixel belongs to or not to a vessel. For all three datasets, we threshold a probability map with a value of 0.4 to obtain binary segmentation of retinal vessels.

We used standard evaluation metrics that are commonly used in medical image segmentation to evaluate deep learning models. TP, FP, TN, and FN are the abbreviations for true positive, false positive, true negative, and false negative, respectively. Sensitivity (1), specificity (2), accuracy (3), f1-score (5), and Mathew correlation coefficient (MCC) (6) are the evaluation measures used. Below are the formulae for these evaluation metrics:

$$Sensitivity\ (SN) = \frac{TP}{TP + FN}, \quad (1)$$

$$Specificity\ (SP) = \frac{TN}{TN + FP}, \quad (2)$$

$$Accuracy\ (ACC) = \frac{(TP + TN)}{(TP + FP + TN + FN)}, \quad (3)$$

$$Precision = \frac{TP}{(TP + FP)}, \quad (4)$$

$$F1 - score = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity}, \quad (5)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{((TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN))}}. \quad (6)$$

The area under the receiver operating characteristic (ROC) curve (AUC), which ranges from 0 to 1, was used to evaluate image segmentation in addition to the metrics listed above.

IV. RESULTS

In this section, we evaluate the Anam-Net model on the same dataset, i.e., training and test images are both part of the same dataset. Tables III, IV, and V show the results obtained for this experiment. The results highlighted in green, blue, and red indicate the best, second-best, and worst scores of the evaluation metric. As shown in Tables III–V, very few methods have evaluated their proposed model on all evaluation metrics. For this reason, we leave

the evaluation metrics empty if it is not evaluated by the respective state-of-the-art method. For comparison, the evaluation metric scores mentioned in Tables III–V of various methods have been taken from their respective work.

For the DRIVE dataset, the results in Table III show that the Anam-Net model outperforms several methods in terms of sensitivity, accuracy, AUC, and MCC. Our method achieved an f1-score of 0.8152, which is the second-best among all methods listed in Table III. Alom, Yakopcic, Hasan, Taha, and Asari [29] and Wang, Zhao, and Yu [33] achieved the best f1-score and specificity, respectively. However, none of the other evaluation metrics is among the top two, which limits their generalizability. The worst accuracy and sensitivity among all methods is obtained by the models proposed by Zhou, Yu, Xu, Gu, and Yang [25] and Feng, Zhuo, Pan, and Tian [27], respectively.

In the STARE dataset, the Anam-Net model ranked first in terms of sensitivity, f1-score, accuracy, AUC, and MCC, whereas it obtained a competitive score for specificity. Wang, Zhao, and Yu [33] achieved the best specificity and the worst AUC among other methods. Tang, Rui, Yan, Li, and Hu [35] had the second-best specificity and accuracy, whereas the worst sensitivity among the methods listed in Table IV.

TABLE III. COMPARISON WITH METHODS ON DRIVE DATASET.

REF	SN	F1	SP	ACC	AUC	MCC
[22]	0.7897	0.7857	0.9684	-	-	0.7556
[25]	0.8078	0.7942	0.9674	0.9469	-	0.7656
[27]	0.7625	-	0.9809	0.9528	0.9678	-
[29]	0.7792	0.8171	0.9813	0.9556	0.9784	-
[30]	0.7653	-	0.9818	0.9542	0.9752	-
[33]	0.8060	0.7863	0.9869	0.9512	0.9748	-
[34]	0.8213	-	0.9807	0.9615	0.9815	-
[38]	0.7941	-	0.9798	0.9558	0.9847	-
[40]	0.8252	-	0.9787	0.9649	0.9780	-
[42]	0.8354		0.9746	0.9569	0.9820	
[44]	0.8206		0.9728	0.9531	0.9770	
[45]	0.7778	-	0.9784	0.9527	-	-
Proposed	0.8601	0.8152	0.9764	0.9660	0.9863	0.7988

TABLE IV. COMPARISON WITH METHODS ON STARE DATASET.

REF	SN	F1	SP	ACC	AUC	MCC
[22]	0.7680	0.7644	0.9738	-	-	0.7417
[23]	0.8147	-	0.9844	0.9674	0.9885	-
[25]	0.8065	0.8017	0.9761	0.9585	-	0.7830
[27]	0.7709	-	0.9848	0.9633	0.9700	-
[28]	0.7940	-	0.9869	0.9647	0.9885	-
[30]	0.7581	-	0.9846	0.9612	0.9801	-
[31]	0.7595	0.8143	0.9878	0.9641	0.9832	-
[33]	0.8230	0.7947	0.9945	0.9641	0.9620	-
[35]	0.7551	-	0.9903	0.9723	0.9863	-
[38]	0.7598	-	0.9878	0.9640	0.9824	-
[40]	0.8397		0.9792	0.9659	0.9810	-
[45]	0.7558	-	0.9811	0.9589	-	-
Proposed	0.8697	0.8266	0.9813	0.9728	0.9906	0.8145

Regarding the CHASE_DB dataset, the Anam-Net model was ranked first in five out of six evaluation metrics, whereas it ranked second in terms of specificity. In the three datasets, Orlando, Prokofyeva, and Blaschko [22] had the worst f1-score and MCC, while Zhou, Yu, Xu, Gu, and Yang [25] had the worst accuracy.

Tables III–V show the performance of the model in an ideal situation when the model is trained and tested on the same dataset. However, in real world scenario and to test the generalizability of the deep learning model, the model must show robustness for fundus images with high variability, i.e., the acquisition device may belong to different manufacturer or the acquired image came from wide variety of individuals. For this purpose, we conduct cross-training experiments on DRIVE, STARE, and CHASE_DB datasets. The manual annotations by the first human observer were used as ground truth. Tables VI and VII show the best case and worst case segmentation results of the Anam-Net model on the DRIVE, STARE, and CHASE_DB datasets, respectively.

TABLE V. COMPARISON WITH METHODS ON CHASE DATASET.

REF	SN	F1	SP	ACC	AUC	MCC
[22]	0.7277	0.7332	0.9712	-	-	0.7046
[23]	0.7661	-	0.9816	0.9599	0.9812	-
[24]	0.7779	-	0.9864	0.9653	0.9855	-
[25]	0.7553	0.7644	0.9751	0.9520	-	0.7398
[29]	0.7756	0.7928	0.9820	0.9634	0.9815	-
[30]	0.7633	-	0.9809	0.9610	0.9781	-
[31]	0.8155	0.7883	0.9752	0.9610	0.9804	-
[33]	0.8035	-	0.9787	0.9639	0.9832	-
[34]	0.8235	-	0.9711	0.9559	0.9767	-
[38]	0.8176	0.7892	0.9704	0.9608	0.9865	-
[39]	0.7978	0.8031	0.9818	0.9656	0.9818	-
[40]	0.8440	-	0.9810	0.9722	0.9830	-
[43]	0.8132	-	0.9814	0.9661	0.9860	-
Proposed	0.8553	0.8090	0.9826	0.9746	0.9899	0.7968

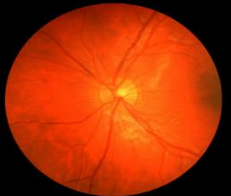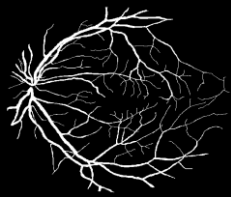TABLE VI. VISUAL REPRESENTATION OF BEST CASE PERFORMANCE OF THE ANAM-NET MODEL.
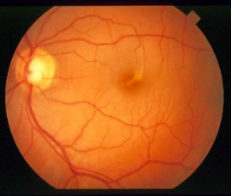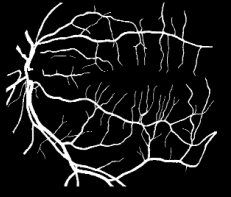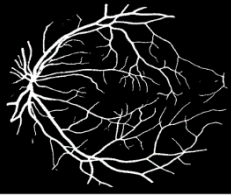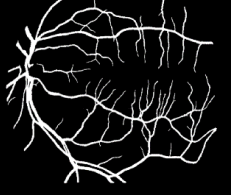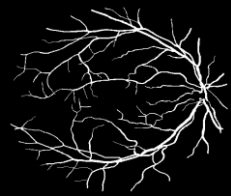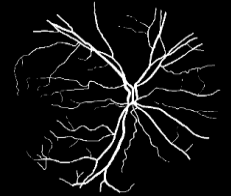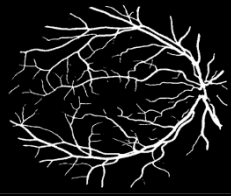
Best Case	DRIVE	STARE	CHASE_DB
Original			
Ground Truth			
Segmented			

TABLE VII. VISUAL REPRESENTATION OF WORST CASE PERFORMANCE OF THE ANAM-NET MODEL.

Worst Case	DRIVE	STARE	CHASE_DB
Original			
Ground Truth			
Segmented			

V. DISCUSSION

In this section, we first perform a cross-database training experiment in which a model trained on one dataset is tested on test images from a second previously unseen dataset. Next, we explore the performance of the Anam-Net model in challenging scenarios such as central vessel reflex and low contrast of fundus images. Finally, we compare the computational complexity of the Anam-Net model with the state-of-the-art methods in terms of the number of trainable parameters.

Table VIII is divided into four sections, where the first section shows the segmentation result on the test images of the STARE dataset when a pre-trained model on the DRIVE dataset is used. The Anam-Net model achieved the best sensitivity, specificity, and accuracy. The method proposed by Mo and Zhang [23] obtained the best f1-score, however, obtained the worst sensitivity among other methods. Oliveira, Pereira, and Silva [24] ranked second in terms of sensitivity, specificity, and accuracy when tested on the STARE dataset. The second section of Table VII shows the segmentation results on the test images of the DRIVE dataset when a pre-trained model on the STARE dataset is used. The Anam-Net ranked first among other methods in three of four evaluation metrics, while ranked second in terms of f1-score. Oliveira, Pereira, and Silva [24] were first in terms of f1-score; however, their method obtained the lowest sensitivity among other methods. When the model is trained on STARE and then tested on DRIVE, we found that it detects fewer thin vessels, resulting in lower sensitivity. On the other hand, the DRIVE database includes substantially more thin vessels annotated than STARE; therefore, sensitivity increased significantly.

To further extend the cross-training experiment, we include the CHASE_DB dataset, which has been reported by very few studies. When the pre-trained model on the DRIVE and STARE datasets is employed, the segmentation results on the test images of the CHASE_DB dataset are shown in sections three and four of Table VIII, respectively.

Several methods have been proposed over the past few years to improve the segmentation accuracy of retinal vessels in retinal fundus images. However, very few studies discuss the performance of their deep learning model on challenging scenarios such as central vessel reflex and low contrast. Table IX shows the exemplar results of the Anam-Net in such challenging scenarios. The first column shows the original fundus image with a yellow bounding box, which represents the region of interest for a specific scenario. The second column shows the cropped image of the original fundus image where the image in the top represents central vessel reflex problem, and the image in the bottom represents low contrast. The third and fourth columns show the manual annotations by the first expert and the predicted probability map by the Anam-Net model, respectively. As shown, Anam-Net predicts the retinal vessels with high probability in the central vessel reflex region. For images with low contrast, the Anam-Net model not only predicts the tiny vessels accurately but also segments additional small vessels that were not annotated by

the first expert. This experiment validates the robustness of the Anam-Net model for such challenging scenarios.

The Anam-Net model is lightweight with only 4.47M trainable parameters. Only two methods have lower computational complexity compared to Anam-Net, as shown in Table X. Zhuang et al. [39] proposed the Ladder-Net model with only 1.38M parameters. On the CHASE_DB dataset, their model achieved a sensitivity of 0.7978 compared to the Anam-Net with a sensitivity of 0.8553. Moreover, the generalizability of their model is not verified by conducting a cross-training experiment. The results show that the Anam-Net model outperforms state-of-the-art methods for segmenting the retinal vasculature.

We thoroughly assessed the performance of our developed deep learning model using healthy and pathological images, where we achieved highly compelling results based on several well-known evaluation metrics, including accuracy, specificity, sensitivity AUC, F1-score, and MCC. We have extensively performed comparison with previous best deep learning models from the state-of-the-art, where our achieved results are not only competitive in terms of evaluation metrics, but also its computation complexity is significantly lower. In our previous work in [40], we achieved better performance with almost 8 million trainable parameters. Compared to our previous work and also other best models from the state-of-the-art, our current model achieved improved segmentation performance with almost 4.5 million trainable parameters. Most importantly, our developed deep learning model attributes significantly lower computational complexity and memory overhead based on fewer trainable parameters. These encouraging attributes of the Anam-Net model advocate for its deployment in normal computing resources available at forefront of hospital, to be used for population-scale eye disease screening programs.

Training the Anam-Net takes approximately 1 hour and 20 minutes on a single NVIDIA RTX2080Ti GPU, while it takes about 0.2 seconds to segment a test image of size 576×576 pixels using the trained model. The authors in [23], [29], and [41] reported an average inference time of 0.5 s, 6 s, and 8.4 s, respectively, for a single test image.

TABLE VIII. CROSS-TRAINING COMPARISON WITH METHODS.

Dataset	REF	SN	F1	SP	ACC
STARE (Train on DRIVE)	[23]	0.7009	0.9843	0.9570	0.9751
	[24]	0.8453	0.9726	0.9597	0.9846
	[25]	0.7525	0.9721	0.9494	-
	[50]	0.7499	0.9798	0.9563	0.9621
	[30]	0.7211	0.9840	0.9569	0.9708
	Proposed	0.8575	0.9727	0.9644	0.9863
DRIVE (Train on STARE)	[23]	0.7412	0.9799	0.9492	0.9653
	[24]	0.6706	0.9916	0.9505	0.9748
	[25]	0.7673	0.9703	0.9443	-
	[50]	0.7217	0.9820	0.9486	0.9327
	[30]	0.7292	0.9815	0.9494	0.9599
	Proposed	0.7864	0.9832	0.9657	0.9767
CHASE (Train on DRIVE)	[23]	0.7003	0.9750	0.9478	0.9671
	[41]	0.7118	0.9791	0.9429	0.9628
	Proposed	0.8652	0.9661	0.9592	0.9797
CHASE (Train on STARE)	[23]	0.7032	0.9794	0.9515	0.9690
	[41]	0.7240	0.9768	0.9417	0.9553
	Proposed	0.7295	0.9776	0.9594	0.9601

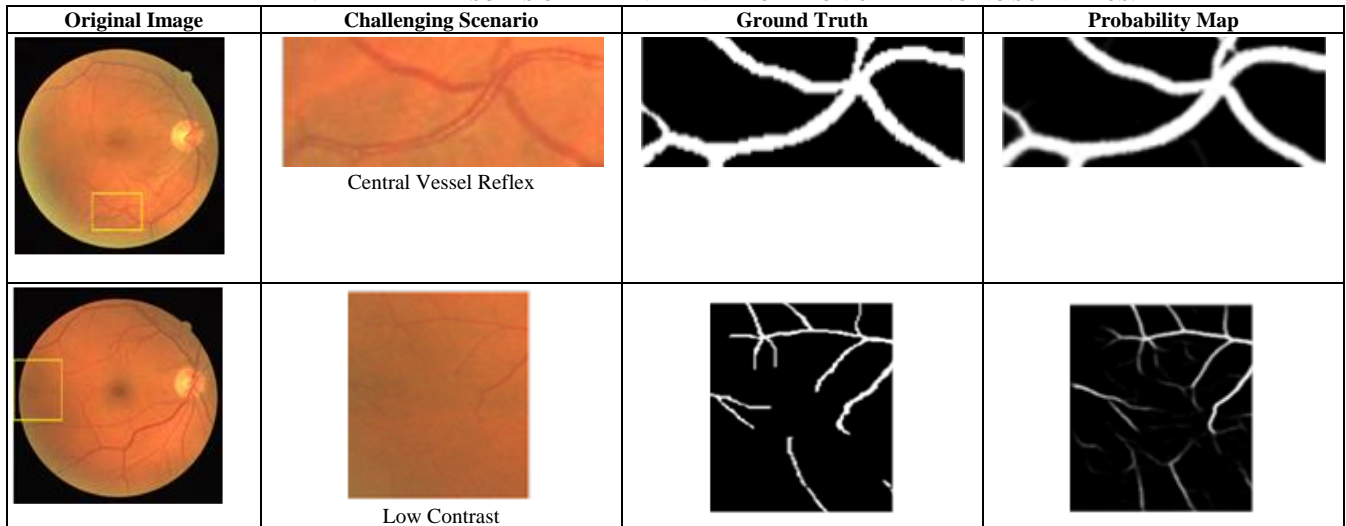TABLE IX. EXEMPLAR RESULTS OF THE ANAM-NET MODEL ON CHALLENGING SCENARIOS.

Original Image	Challenging Scenario	Ground Truth	Probability Map
	Central Vessel Reflex		
	Low Contrast		

TABLE X. TRAINABLE PARAMETER COMPARISON.

Reference	Method	Parameters
Mo and Zhang [23]	FCN with Deep Feature Fusion	8.40M
Alom, Yakopcic, Hasan, Taha, and Asari [29]	R2U-Net	48.90M
Yan, Yang, and Cheng [30]	U-Net with Joint Losses	36.40M
Jin, Meng, Pham, Chen, Wei, and Su [31]	DUNet	0.88M
Wang, Zhao, Ren, Xu, and Yu [32]	Basic U-Net	31.05M
Lv, Ma, Li, and Liu [38]	Attention Guided U-Net	28.25M
Zhuang [39]	Ladder-Net	1.38M
Khan, Alhussein, Aurangzeb, Arsalan, Naqvi, and Nawaz [40]	Basic SegNet	29.50M
Khan, Alhussein, Aurangzeb, Arsalan, Naqvi, and Nawaz [40]	RCED-Net	9.37M
Noh, Park, and Lee [42]	SSANet	25.00M
Anam-Net		4.47M

VI. CONCLUSIONS

We have developed Anam-Net, which is a lightweight CNN-based encoder-decoder architecture with AD-block as bottleneck layers for semantic segmentation of retinal vessels in fundus images. The performance of our developed model has been thoroughly tested using retinal images from the DRIVE, STARE, and CHASE_DB datasets, where we achieved {sensitivity and accuracy} of {0.8601, 0.9660}, {0.8697, 0.9728}, and {0.8553, 0.9746}, respectively. In all three datasets, the results reveal that the Anam-Net model outperforms the state-of-the-art approaches for segmenting the retinal vessels.

Generally, deep learning models achieve improved performance, but at the expense of high-performance computing (GPUs), which are rarely available at the forefront desk of the healthcare facility. This limits the deep learning-based computer aided tools to be applied in facilities having GPU-based computing resources. The basic hypothesis of this study is to explore and develop a deep learning model which not only achieves competitive semantic segmentation performance in terms of evaluation metrics, but is realizable at the mobile/embedded platforms at the point of care, where the computing resources are not supposed to be GPU-based. Such deep learning models should attribute significantly lower trainable parameters, which will lead to reduced computation complexity and will be realizable at ordinary computing/embedded platforms that are readily available at almost every healthcare facility. In this way, the population-scale screening programs for early diagnosis of vision-threatening eye diseases could be effectively managed.

The developed Anam-Net model has better segmentation accuracy and less trainable parameters compared to several other methods, which make it suitable for easy deployment in mobile platforms at the point of care. On the basis of the achieved results for the cross-database training and testing strategy, we may generalize that the Anam-Net model has significantly higher generalization ability in addition to being lightweight (having significantly lower parameters). Due to attributing significantly lower trainable parameters, the Anam-Net model may be used in a variety of medical imaging applications. Hopefully, in the near future, we will be able to achieve competitive classification performance for the diagnosis of COVID-19 disease using the modified Anam-Net model.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Z. Gao, J. Li, J. Guo, Y. Chen, Z. Yi, and J. Zhong, "Diagnosis of diabetic retinopathy using deep neural networks", *IEEE Access*, vol. 7, pp. 3360–3370, 2019. DOI: 10.1109/ACCESS.2018.2888639.
[2] 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 Recognit.*, vol. 44, nos. 10–11, pp. 2314–2324, 2011. DOI: 10.1016/j.patcog.2011.01.007.
[3] M. D. Abràmoff, J. C. Folk, D. P. Han, J. D. Walker, D. F. Williams, S. R. Russell *et al.*, "Automated analysis of retinal images for

detection of referable diabetic retinopathy", *JAMA Ophthalmol.*, vol. 131, no. 3, pp. 351–357, 2013. DOI: 10.1001/jamaophthalmol.2013.1743.

[4] K. Talaat, O. T. Fathi, S. M. Alamoudi, M. G. Alzahrani, R. M. Mukhtar, and M. A. Khan, "Types of glaucoma and associated comorbidities among patients at King Abdulaziz Medical City, Jeddah", *Cureus*, vol. 13, no. 6, p. e15574, 2021. DOI: 10.7759\\cureus.15574.

[5] N. Anuradha and K. Ramani, "Role of optometry school in single day large scale school vision testing", *Oman J. Ophthalmol.*, vol. 8, no. 1, pp. 28–32, 2015. DOI: 10.4103/0974-620x.149861.

[6] K. A. Thakoor, X. Li, E. Tsamis, P. Sajda, and D. C. Hood, "Enhancing the accuracy of glaucoma detection from OCT probability maps using convolutional neural network", in *Proc. of 41st Annu. Int. Conf. of the IEEE Eng. in Med. and Biol. Soc. (EMBC)*, 2019, pp. 2036–2040. DOI: 10.1109/EMBC.2019.8856899.

[7] X. Zeng, H. Chen, Y. Luo, and W. Ye, "Automated diabetic retinopathy detection based on binocular Siamese-like convolutional neural network", *IEEE Access*, vol. 7, pp. 30744–30753, 2019. DOI: 10.1109/ACCESS.2019.2903171.

[8] M. V. Cicinelli, A. Rabiolo, R. Sacconi, A. Carnevali, L. Querques, F. Bandello *et al.*, "Optical coherence tomography angiography in dry age-related macular degeneration", *Surv. Ophthalmol.*, vol. 63, no. 2, pp. 236–244, 2018. DOI: 10.1016/j.survophthal.2017.06.005.

[9] Y. Muraoka, A. Tsujikawa, T. Murakami, K. Ogino, K. Kumagai, K. Miyamoto *et al.*, "Morphologic and functional changes in retinal vessels associated with branch retinal vein occlusion", *Ophthalmology*, vol. 120, no. 1, pp. 91–99, 2013. DOI: 10.1016/j.ophtha.2012.06.054.

[10] S. Traustason, A. S. Jensen, H. S. Arvidsson, I. C. Munch, L. Søndergaard, and M. Larsen, "Retinal oxygen saturation in patients with systemic hypoxemia", *Invest. Ophthalmol. Vis. Sci.*, vol. 52, no. 8, pp. 5064–5067, 2011. DOI: 10.1167/iovs.11-7275.

[11] 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. DOI: 10.1109/ACCESS.2021.3102176.

[12] N. Paluru, A. Dayal, H. B. Jenssen, T. Sakinis, L. R. Cenkeramaddi, J. Prakash *et al.*, "Anam-net: Anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images", *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 932–946, 2021. DOI: 10.1109/TNNLS.2021.3054746.

[13] J. Almotiri, K. Elleithy, and A. Elleithy, "Retinal vessels segmentation techniques and algorithms: A survey", *Appl. Sci.*, vol. 8, no. 2, p. 155, 2018. DOI: 10.3390/app8020155.

[14] A. D. 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, 2000. DOI: 10.1109/42.845178.

[15] J. Zhang, H. Li, Q. Nie, and L. Cheng, "A retinal vessel boundary tracking method based on Bayesian theory and multi-scale line detection", *Comput. Med. Imaging Graph.*, vol. 38, no. 6, pp. 517–525, 2014. DOI: 10.1016/j.compmedimag.2014.05.010.

[16] G. Hassan, N. El-Bendary, A. E. Hassanien, A. Fahmy, A. M. Shoeb, and V. Snasel, "Retinal blood vessel segmentation approach based on mathematical morphology", *Procedia Comput. Sci.*, vol. 65, pp. 612–622, 2015. DOI: 10.1016/j.procs.2015.09.005.

[17] S. A. A. Shah, A. Shahzad, M. A. Khan, C.-K. Lu, and T. B. Tang, "Unsupervised method for retinal vessel segmentation based on Gabor wavelet and multiscale line detector", *IEEE Access*, vol. 7, p. 167221–167228, 2019. DOI: 10.1109/ACCESS.2019.2954314.

[18] X. Jiang and D. Mojon, "Adaptive local thresholding by verification based multi-threshold probing with application to vessel detection in retinal images", *IEEE Trans. Pattern Anal. Mach.*, vol. 25, no. 1, pp. 131–137, 2003. DOI: 10.1109/TPAMI.2003.1159954.

[19] B. S. Y. Lam, Y. Gao, and A. W.-C. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling", *IEEE Trans. Med. Imaging*, vol. 29, no. 7, pp. 1369–1381, 2010. DOI: 10.1109/TMI.2010.2043259.

[20] K. B. Khan, A. A. Khaliq, A. Jalil, M. A. Iftikhar, N. Ullah *et al.*, "A review of retinal blood vessels extraction techniques: Challenges, taxonomy, and future trends", *Pattern Anal. Appl.*, vol. 22, no. 3, pp. 767–802, 2019. DOI: 10.1007/s10044-018-0754-8.

[21] R. Vega, G. Sanchez-Ante, L. E. Falcon-Morales, H. Sossa, and E. Guevara, "Retinal vessel extraction using Lattice Neural Networks", *Comput. Biol. Med.*, vol. 58, pp. 20–30, 2015. DOI: 10.1016/j.compbiomed.2014.12.016.

[22] 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. Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 16–27, 2017. DOI: 10.1109/TBME.2016.2535311.

[23] J. Mo and L. Zhang, "Multi-level deep supervised networks for retinal vessel segmentation", *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 12, pp. 2183–2193, 2017. DOI: 10.1007/s11548-017-1619-0.

[24] A. Oliveira, S. Pereira, and C. A. Silva, "Retinal vessel segmentation based on Fully Convolutional Neural Networks", *Expert Syst. Appl.*, vol. 112, pp. 229–242, 2018. DOI: 10.1016/j.eswa.2018.06.034.

[25] L. Zhou, Q. Yu, X. Xu, Y. Gu, and J. Yang, "Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement", *Comput. Methods Programs Biomed.*, vol. 148, pp. 13–25, 2017. DOI: 10.1016/j.cmpb.2017.06.016.

[26] K. Hu, Z. Zhang, X. Niu, Y. Zhang, C. Cao, F. Xiao, and X. Gao, "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function", *Neurocomputing*, vol. 309, pp. 179–191, 2018. DOI: 10.1016/j.neucom.2018.05.011.

[27] S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "CcNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features", *Neurocomputing*, vol. 392, pp. 268–276, 2020. DOI: 10.1016/j.neucom.2018.10.098.

[28] W. Abbas, M. Shakeel, N. Khurshid, and M. Taj, "Patch-based generative adversarial network towards retinal vessel segmentation", in *Neural Information Processing. ICONIP 2019. Communications in Computer and Information Science*, vol. 1142. Springer, Cham, 2019. DOI: 10.1007/978-3-030-36808-1_6.

[29] Md. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation", *J. Med. Imaging*, vol. 6, no. 1, pp. 1–12, 2019. DOI: 10.1117/1.JMI.6.1.014006.

[30] Z. Yan, X. Yang, and K.-T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation", *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 1912–1923, 2018. DOI: 10.1109/TBME.2018.2828137.

[31] Q. Jin, Z.-P. Meng, T. Pham, Q. Chen, L. Wei, and R. Su, "DUNet: A deformable network for retinal vessel segmentation", *Knowl. Based Syst.*, vol. 178, pp. 149–162, 2019. DOI: 10.1016/j.knosys.2019.04.025.

[32] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, "Dense U-net based on patch-based learning for retinal vessel segmentation", *Entropy*, vol. 21, no. 2, p. 168, 2019. DOI: 10.3390/e21020168.

[33] C. Wang, Z. Zhao, and Y. Yu, "Fine retinal vessel segmentation by combining Nest U-net and patch-learning", *Soft Comput.*, vol. 25, no. 1, pp. 1–14, 2021. DOI: 10.1007/s00500-020-05552-w.

[34] J. Zhang, Y. Zhang, and X. Xu, "Pyramid U-Net for retinal vessel segmentation", in *Proc. of IEEE Int. Conf. on Acoust., Speech and Signal Proc. (ICASSP-2021)*, 2021. arXiv: 2104.02333. DOI: 10.1109/ICASSP39728.2021.9414164.

[35] Y. Tang, Z. Rui, C. Yan, J. Li, and J. Hu, "ResWnet for retinal small vessel segmentation", *IEEE Access*, vol. 8, pp. 198265–198274, 2020. DOI: 10.1109/ACCESS.2020.3032453.

[36] Y. Ma, Z. Zhu, Z. Dong, T. Shen, M. Sun, and W. Kong, "Multichannel retinal blood vessel segmentation based on the combination of matched filter and U-Net network", *BioMed Res. Int.*, vol. 2021, art. ID 5561125, 2021. DOI: 10.1155/2021/5561125.

[37] M. E. Gegundez-Arias, D. Marin-Santos, I. Perez-Borrero, and M. J. Vasallo-Vazquez, "A new deep learning method for blood vessel segmentation in retinal images based on convolutional kernels and modified U-Net model", *Comput. Methods Programs Biomed.*, vol. 205, art. 106081, 2021. DOI: 10.1016/j.cmpb.2021.106081.

[38] Y. Lv, H. Ma, J. Li, and S. Liu, "Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation", *IEEE Access*, vol. 8, pp. 32826–32839, 2020. DOI: 10.1109/ACCESS.2020.2974027.

[39] J. Zhuang, "LadderNet: Multi-path networks based on U-net for medical image segmentation", 2018. arXiv: 1810.07810.

[40] 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", *IEEE Access*, vol. 8, pp. 131257–131272, 2020. DOI: 10.1109/ACCESS.2020.3008899.

[41] 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 Trans. on Med. Imag.*, vol. 35, no. 1, pp. 109–118, 2016. DOI: 10.1109/TMI.2015.2457891.

[42] K. J. Noh, S. J. Park, and S. Lee, "Scale-space approximated convolutional neural networks for retinal vessel segmentation",

Comput. Methods Programs Biomed., vol. 178, pp. 237–246, 2019. DOI: 10.1016/j.cmpb.2019.06.030.

[43] Y. Wu, Y. Xia, Y. Song, D. Zhang, D. Liu, C. Zhang *et al.*, "Vessel-Net: Retinal vessel segmentation under multi-path supervision", in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019. Lecture Notes in Computer Science*, vol. 11764. Springer, Cham, 2019. DOI: 10.1007/978-3-030-32239-7_30.

[44] X. Yang, Z. Li, Y. Guo, and D. Zhou, "Retinal vessel segmentation based on an improved deep forest", *Int. J. of Imag. Sys. Tech.*, vol. 31, no. 4, pp. 1792–1802, 2021. DOI: 10.1002/ima.22610.

[45] E. Uysal and G. E. Güraksin, "Computer-aided retinal vessel segmentation in retinal images: Convolutional neural networks", *Multimed. Tools. Appl.*, vol. 80, pp. 3505–3528, 2021. DOI: 10.1007/s11042-020-09372-w.

[46] F. Altaf, S. M. S. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: Concepts, methods, challenges and future directions", *IEEE Access*, vol. 7, pp. 99540–99572, 2019. DOI:

10.1109/ACCESS.2019.2929365.

[47] 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. on Med. Imaging*, vol. 23, no. 4, pp. 501–509, 2004. DOI: 10.1109/TMI.2004.825627.

[48] 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 Trans. on Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, 2012. DOI: 10.1109/TBME.2012.2205687.

[49] A. Sevastopolsky, "Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network", *Pattern Recognition and Image Analysis*, vol. 27, no. 3, pp. 618–624, 2017. DOI: 10.1134/S1054661817030269.

[50] S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "Ccnet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features", *Neurocomputing*, vol. 392, pp. 268–276, 2020. DOI: 10.1016/j.neucom.2018.10.098.