High Performance and Low Complexity Retinal Vessel Segmentation Method based on Extended DCNN

Mouna Abdelfattah, Yaroub Elloumi, Rostom Kachouri

High Performance and Low Complexity Retinal Vessel Segmentation Method based on Extended DCNN

Mouna ABDELFATTAH[1], Yaroub ELLOUMI[1,2], Rostom KACHOURI[3]

[1]ISITCom Hammam-Sousse, University of Sousse, Sousse, Tunisia
[2]Medical Technology and Image Processing Laboratory, Faculty of Medicine, University of Monastir, Monastir, Tunisia
[3]LIGM, University Gustave Eiffel, CNRS, ESIEE Paris, Marne-la-Vallée, France

***Abstract*—The segmentation of the retinal vascular tree (RVT) is undoubtedly crucial for visualizing and analyzing vessel morphology, which facilitates the detection and diagnosis of various pathologies that affect the retina. Therefore, segmentation must have high performance with reduced complexity and execution time to meet constraints imposed by the clinical context. In this context, several classic Deep Learning (DL) architectures was proposed for the automatic segmentation of the retinal vascular tree, they generate suboptimal segmentation performance and significant computational complexity.**

In this work, we propose a new method for RVT segmentation. The main contribution consists of suggesting a new DCNN, inspired from Segnet architecture where standard convolution blocks have been substituted by MobileNet convolution block. The proposed method enhance segmentation performance and reduce the computational complexity. It is tested on DRIVE database, and scored an Accuracy of 97.08 %, a sensitivity of 82.38 % and a specificity of 98.55%, within an architecture of 3.4 M parameters.

Keywords—Deep Learning, Automatic Segmentation, Retinal vascular tree, Segnet, MobileNet.

I. INTRODUCTION

The retina is the only part of our body where blood vessels can be observed and photographed directly and easily [1], and in many cases, they reflect the state of the vascular system of the rest of the body. Certain fundus and associated pathologies such as age-related macular degeneration (AMD), diabetic retinopathy (DR) and cardiovascular diseases, affect the retina and his components. Diagnosis and early detection are very important since these pathologies are dangerous and can lead to blindness. The segmentation of the retinal vascular tree presents a major step in detecting these ocular pathologies it helps to analyze the morphology, diameter, area, branching pattern, and tortuosity of retinal blood vessels which are important indicators of these diseases. However, this process is limited by the lack of ophthalmologists and retinal imaging devices. Similarly, manual segmentation of blood vessels in retinal images is both error-prone and time-consuming even for experienced doctors. Since the clinical context expects higher segmentation performance with lower complexity and runtime, various automatic retinal vascular tree segmentation methods have been suggested. The most efficient are the Deep Learning based methods, they have been always belonged to the studies providing higher segmentation performances [24], they have provided higher accuracy rates that exceed 0.95 when performed with the well-used DRIVE database fundus images. In this context, several recent works have proposed segmenting the retinal vascular tree using DL architectures. However, these architectures generate suboptimal segmentation performances and high computational complexity, this complexity generates significant execution times, which in several cases exceed the temporal constraints imposed by the clinical context. These problems are due to the standard and classic convolution blocks on which these architectures are based. Therefore, several works suggest extending these DL architectures, to reduce processing complexity while maintaining high segmentation performance. In this work, we propose a new Deep Learning architecture for the segmentation of the retinal vascular tree; we use this architecture in a new method that we will describe. The idea is to extend the Segnet architecture by replacing classic convolution blocks with new blocks called LightWeight Convolution Module (LCM) of the MobileNet architecture. These new blocks make it possible to extract more efficient features by dividing the convolution into two kernels, while reducing complexity and execution time dedicated to feature extraction. Therefore, our goal for this proposal is to improve segmentation performance while reducing complexity.

The remainder of this paper is organized into four sections. In Section 2, a literature review on the classical and extended DL method for RVT segmentation is presented. Section 3, present the proposed DL architecture. Section 4, describe the segmentation method, Section 5, present the experimental results where the segmentation is evaluated in terms of performance and complexity. The last section is dedicated for conclusion

II. STANDARD AND EXTENDED DEEP LEARNING METHODE

In recent years, Deep Learning has got fast development due to strong context feature expression ability [2]. The advancements in deep learning models are exploited for conducting the analysis and designing the detection methods for accurate segmentation of retinal vessel tree (RVT). Several standards Convolution Neural Networks (CNN) have been used for this purpose of RVT segmentation, such as U-Net network, that was proposed by Ronneberger et al. [6]. It is widely used for medical image segmentation, including retinal images[10]. This network is based on the operation of a fully convolutional network (FCN) [6], its structure is U-shaped composed of two paths encoder and decoder, the blocks of the two paths are connected Through the skip connections. Also, Segnet network proposed by Badrinarayanan et al. [23] has been used for RVT segmentation [11][3]. It is an FCN consists of a network of encoders, which is responsible for capturing the hierarchical characteristics of the input image, a corresponding network of decoders that generates the segmentation map followed by a pixel-by-pixel classification layer. The U-shaped form of Segnet with both encoder and decoder paths, is widely adopted for segmentation and it achieves high performance, moreover, it encourages the model to learn more features by gradually reducing the feature maps and creating connections on both sides. Unlike U-net, SegNet uses a pooling index to perform upsampling through the inverse process of pooling in the decoder part, meaning that the upsampling process does not require learning[9].

These standard Deep Learning models such U-Net and Segnet are used for RVT segmentation[10][11], they have a better performance than traditional Machine Learning based methods, accuracy of U-Net and segnet are respectively 95.54 % and 95.54 %, but they do not achieve the optimal performance required by the clinical context. Furthermore, these architectures are discriminated by a higher computational complexity, proportional to the number of parameters [12], the number of parameters recorded by the U-Net architecture exceeds 31 million, while the segnet architecture recorded a number of parameters of 14.7 million respectively. These problems are due to the standards and classic convolution blocks on which these architectures are based. Indeed, blood vessels in the retina can vary considerably in size and shape, however, standard convolutions may have difficulty capturing these variety of scales effectively and do not allow all the vessels to be modeled adequately taking into account all widths, which can lead to error segmentation, especially for smaller or thinner vessels. Therefore, the performance of the models declines. In addition, in standard convolution operation each layer performs the convolution in a single step. It consists of applying a filter to the entire input image by performing multiplications between the values of the pixels in the image and the weights of the filter; convolution therefore both filters and combines the inputs into a new set of outputs in a single step. Which leads to increased processing time and increased complexity of the entire architecture. These results can go beyond clinical constraints; it generates a significant complexity compared to the performance obtained. The standards networks still unable to reach the trade-off performances, discriminated by an expensive convolution blocks having the effect of increasing computation requirements. Several of these networks have been extended with the aim of performing such clinical needs to maintaining high segmentation performance and reducing the computational complexity. The modifications are usually focused on the convolution layers to make the network wider or on the skip connections to guide the decoder to utilize the most essential features.

In this context, we present some of these extended architectures. Guo et al. [13] proposed SA-UNet network an extension of U-net for retinal vessel segmentation by adding a spatial attention block between the last block of the encoder and the first block of the decoder to acquire the discriminative features about retinal vessels and to guide the decoder to utilize the most essential features. Henda Boudegga et al. [12] proposed an architecture named “RV-Net”. It is a variant of the convolutional neural network U-Net. The main innovation of RV-Net lies in the substitution of large standards convolution layers by low complexity modules named LCMs, each module consist of a 3×3 depth convolution layer followed by a 1×1 convolution layer. This method was proposed to refine the U-Net architecture and reduce the segmentation complexity and execution time To better combine high and low resolution information, K. Ren et al. [14] incorporates the Bi-FPN network into the U-net network to form a new network structure. The classic skip connection is replaced by the Bi-FPN, and the concept of weight is proposed to balance the feature information of various scales better and improve the efficiency of the model. Li et al. [15] propose a combination of U-Net and Dense-Net, the convolution blocks of U-Net are replaced by the Dense module. Each layer of the dense block is directly connected to all the layers before it. Y. Liu et al. [2] propose to extend the classic U-Net network with the incorporation of residual blocks. The classic convolution blocks of U-Net are replaced with new convolutional blocks called ResDO-conv which are inspired by the ResNet architecture, these blocks enable the network to better acquire strong contextual features from retinal images. A pooling fusion block (PFB) is introduced after each Residual DO_Conv block in the encoder to reduce the effect of information loss caused by multiple operations of pooling. The skip connections are replaced by the attention fusion block (AFB) in order to have features at different scales. The M2U-Net network proposed by T. Laibacher et al. [16] has a new encoder-decoder architecture that is inspired by the U-Net. It adds pretrained components of MobileNetV2 in the encoder part and novel contractive bottleneck blocks in the decoder part that combined with bilinear upsampling, in order to have a fast and efficient model for use in embedded and mobile environments. V.Sathananthavathi et al. [17] propose a new deep neural network, named as ‘Enhanced Encoder Atrous Unet’ (EEA Unet) for retinal vessels segmentation. In order to increase the receptive field, the classic convolution blocks of U-Net are replaced as the atrous convolution. Renyuan liu et al. [18] propose the Dual Attention Res2UNet (DA-Res2UNet) network. This model is based on an encoder-decoder U-Net structure, replaces the convolutional layers in U-Net with Res2block [8] and Dropblock. Res2block effectively obtains the multiscale information, and Dropblock avoids overfitting. In addition, Spatial Attention and Dual Attention blocks [10] are added between the encoder and the decoder, thus this model can get global dependency information with less amount computation. The work of Y. Liu et al. [19] proposes a new lightweight segmentation network for precise retinal vessel segmentation, which is called as Wave-Net model. The simple skip connections of original U-Net are replaced by the block (DED) to improve the segmentation precision on thin vessels, in addition, a multi-scale feature fusion block (MFF) is proposed to fuse cross-scale contexts. Zhang et al. [20] introduce Pyramid-scale aggregation blocks (PSABs) in U-Net to obtain higher features extraction. Inception-Like U-Net (ILU-Net) is proposed by Z. Zhu et al. [21] to segment vessels. The simple convolution blocks are replaced by Down-sampling Inception Blocks (DIB) in encoder and up-sampling

Inception Blocks (UIB) in decoder, in addition, a novel skip connection scheme is provided to better connect low-level features to high-level features. To obtain refined features of retinal blood vessels, Z. CUI et al. [22] combine three cascade connected U-Net networks. Skip connections in each Unet are replaced by residual paths, an atrous space pyramid pooling (ASPP) module is employed between each encodor and decoder, a novel loss function is proposed to minimize the loss between predicted images and the ground truth. T. Khan et al. [11] introduce RCED-Net a variant of the CNN network Segnet. The key idea in this architecture is a residual connection introduced between first block of the encoder and final block of the decoder. The purpose of this connection is to include the vessel information directly into the final predictions. Table 1. compare these works according to the performance metrics Accuracy(Acc), Sensitivity(Se) and Specificity(Sp). In addition, to the number of parameters.

III. Proposal of a new Deep learning architecture

A. Principe of our proposed architecture

An efficient segmentation of retinal vascular tree is indispensable for a reliable diagnosis of ocular pathologies, ensure high segmentation performance while saving computational complexity is indispensable to be suitable for clinical practice. Our main idea consists in proposing a new DL network which performs retinal vessel segmentation using LCMs. The suggested network is an extension of the Segnet network, had a U-shaped form since it achieves high performance. As we said in the previous section, that the convolution blocks of Segnet are mainly based on standard convolution layers which does not suitable for multi échelle segmentation of rétinal vessel that leads to non-optimal performance and an augmented processing complexity, we suggest to replace the classic blocks by a new blocks. Recently, another form of convolution formed by lighter convolution modules called Lightweight Convolution Modules (LCMs) was proposed by Howard et al.[4], who used them in the MobileNet architecture[4]. These modules are based on depthwise separable convolution, that provides the same convolution as the standard convolution, but with a different technique, which consists of dividing the convolution operation into two separate successive operations, one for filtering and one for combining. Indeed, Vessel has different informations among the channels, extract features from each channel separately rather than from all image channel since, you can capture channel-specific details that might be missed when considering all channels together. These extracted features can be better highlighted, thus segmentation accuracy will be improved. Also, such modification reduces significantly the complexity, feature extraction from all channels simultaneously may lead to high-dimensional feature vectors, which can increase computational complexity and require more resources for processing. Extracting features separately from each channel can reduce the dimensionality of the feature space, making segmentation algorithms more computationally efficient while still capturing relevant information.

Thus, our main contribution consists of extending the Segnet network by using the u-shaped form, which facilitates training and allows high segmentation performances to be achieved. And by replacing the standard convolution layers with the Lightweight Convolution Modules (LCMs) to reduce the complexity. A 3x3 depthwise convolution layer followed by a 1x1 Pointwise convolution layer forms the module LCM.

TABLE I. Performances results of state of the art works

Works	**Year**	**Acc(%)**	**Se(%)**	**Sp(%)**	**Parameters (M)**
Y.Liu et al. [2]	2023	95.61	79.85	97.91	_
R.Liu et al. [18]	2023	97.04	81.50	98.56	0.845
Y.Liu et al. [19]	2023	95.61	81.64	97.64	_
K. Ren et al. [14]	2022	95.51	80.64	97.67	_
Z. CUI et al. [22]	2022	97.00	80.13	98.42	_
Z. Zhu et al. [21]	2022	97.33	79.19	98.95	2.925
Li et al. [15]	2021	96.98	79.31	98.96	_
H.boudegga et al.[12]	2021	98.19	84.48	99.00	10.87
V.Sathananthavathi et al. [17]	2021	95.77	79.18	97.08	_
J.Zhang et al. [20]	2021	96.15	82.13	98.07	_
C.Guo et al. [13]	2020	96.98	82.12	98.40	_
T.KHAN et al. [11]	2020	96.49	82.52	97.87	9.7
T.Khan et al.[11] Segnet	2020	95.79	79.49	97.38	14.7
Q. Jin et al.[10]U-Net	2019	95.54	78.49	98.04	31.03
T.Laibacher et al.[16]	2019	96.30	_	_	0.55

The depthwise convolution layer takes as input M feature maps of size L × H, where L and H represent respectively the width and height of feature maps. Each map is extracted separately to iteratively apply a convolution kernel of size K × K × 1, where K represents the width and height of the kernel. Consequently, M feature maps are provided through the first layer, having the dimension of L × H × 1. These cards are transferred to the second Pointwise convolution layer on which a 1 × 1 × N kernel is applied iteratively to create linear combinations of output channels from the depthwise convolution to generate N feature maps.

B. Architecture description

The proposed architecture is formed by an encoder and a decoder. The downsampling path of the suggested network is composed of five blocks, which are represented with green frames in Fig.1. we have an input image of size 128×128×3 towards the first block, that is composed by a convolution layer using a 3 × 3 kernels to produce 64 feature maps. Then, an LCM is applied where his first layer has convolved 64 feature maps separately, using 3 × 3 kernels, folloxed by a convolution layer with 1 × 1 kernels, to produce 128 feature maps. Those three convolution layers are characterized by a stride equal to 1 and an activation function RELU. This module is followed by a “max pooling” layer parameterized with a 2 × 2 kernels and a stride of 2. The second block is composed of two LCM layers followed by a max pooling layer to produce 256 feature maps. The third block, is composed by three LCM layers followed by a 2×2 maxPooling layer. The fourth and fifth blocks, each one is composed of 3 successive LCM layers, followed by a "max pooling" layer parameterized by a 2 × 2 kernels and a stride equal to 2, to produce 512 feature maps. In the encoder, each convolution block doubles the feature map size, and each maxPooling layer reduces the image size to half, until we obtain an image size of 8×8×512 at the end of the encoder.

Similarly, the upsampling path is composed of five blocks. which are represented with red frames in Fig.1. Each block contains an upsampling layer parameterized with a 2 × 2 kernel and a stride of 2. For the first and the second block, the upsampling layers are followed by three LCMs, their first depthwise convolution layers are parameterized with a 3 × 3 kernel, where their second convolution layers are parameterized with 1 × 1 kernels. These convolution layers are parameterized with an activation function RELU and a stride equal to 1 with number of filters fixed at 512. The upsampling layer of the third block is followed by three LCMs with 256 karnels. The upsampling layers in the fourth and the last blocks, are followed each one by two LCMs. Contrary to the downsampling path, the number of output feature map is reduced in each block by half and the output size of the feature map is doubled on the length, as well as on the width.

After the final decoder, the output is sent to a 1x1 convolution layer with a number of filters equal to 2. This layer is followed by a "Sigmoïde" activation function which will give the final prediction and which will classify the pixels into vessels or non-vessels. At the end, we have an image with size 128x128x2.

IV. Proposed method for retinal vessel segmentation

The retinal vessel segmentation method essentially relies on the training model based on the proposed architecture that we describe in the previous section. The entire method illustrated in Fig.2. Composed of two stage training and test. Each stage composed of successive steps to have at the end a segmented image. Retinal images are taken with a digital fundus camera, despite the controlled conditions, many retinal images suffer from non-uniform illumination given by several factors: the curved surface of the retina, presence of diseases[1]. Similarly, retinal blood vessels have thin, dark, elongated structures with variation in thickness, therefore, the color of fine vessels can be close to the color of the fundus. This explains the need to preprocess retina images in order to increase image quality, compensate for lighting variations and improve contrast. We first apply a normalization based on Histogram Equalization (HE) [7] which will apply a local normalization of each pixel to zero mean and unit variance, aims to compensate for lighting variations and improve the local contrast. Subsequently, we apply Contrast Limited Adaptive Histogram Equalization (CLAHE) [5] in order to increase the contrast. At the end, we add a gamma correction, which allows the control of the brightness of an image. Subsequently, and given that the initial number of retinal images is too small, a data augmentation method is applied in order to increase the size and diversity of the training set, we applied image transformations like rotation and flipping. First, we will flip the fundus images vertically and then we will do a horizontal flip for the original images and the vertically flipped images. In a second step, we will apply rotations to the images according to the following angles: 30°, 60°, 120°and 150°. These transformations are performed for all fundus images and their corresponding ground truth images in the same dataset, to avoid distorting training. Retinal fundus images are characterized by a significant number of blood vessels that are depicted with different thicknesses, orientations and tortuosities, which makes segmentation of an entire image difficult and gives inaccurate results thus a step of patch extraction is then carried out, it involves taking each time a patch of size "µ" from the image and segmenting it separately. In the testing phase, the preprocessing and patch extraction steps are applied to the test set, these patches will then be segmented by the new architecture that we have proposed. The set of segmented patches obtained will then go through the post-processing stage where this set will be merged to obtain the final result of the entire segmented image.

V. Experimental validation

A. *Experiment setup*

The training of the proposed network is done using a set of parameters chosen experimentally. The experimentation consists at varying the parameter value and evaluated the results. The parameter providing the higher performances are, the ADAM optimizer algorithm with a learning rate value of 0.001, a batch_size of 16. To minimize the loss between predicted images and the ground truth, we use "BinaryCrossnetropy" function. The entire code executed according to these parameters for 20 epochs. These parameters are provided in Table 3. The validation of our model is done with the public retinal image database Digital Retinal Images for Vessel Extraction "DRIVE". This database contains 40 images having sizes of 565 × 584, each retinal image is joined with its manual blood vessel segmentation. We divided the dataset into two subsets, 30 images for training, while 10 images were reserved for testing. The training images were then augmented using various combinations of image processing schemes as we said in the previous section. After images augmentation and cropping, we were able to generate 192 images for every

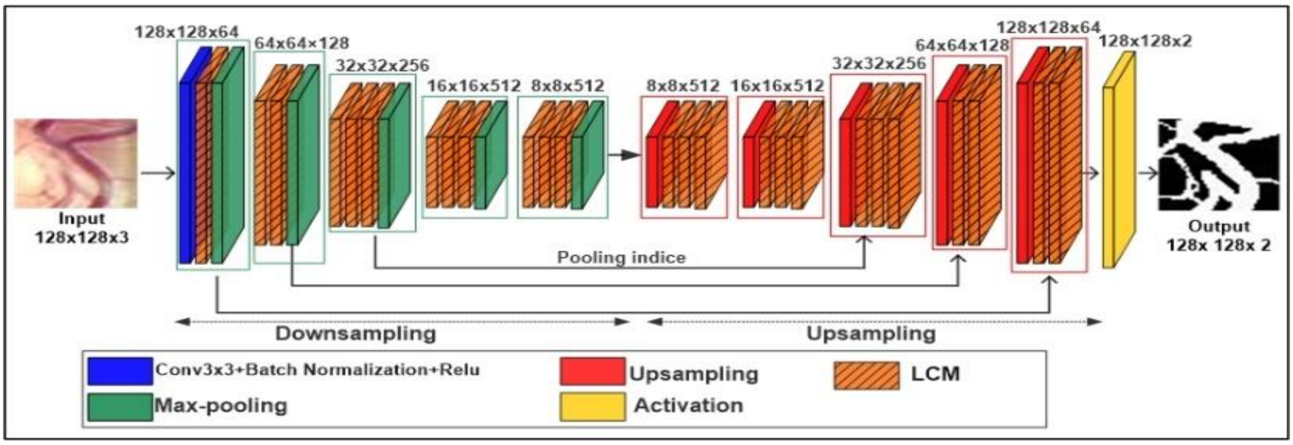

Fig 1. Proposed DL network for retinal blood vessel segmentation

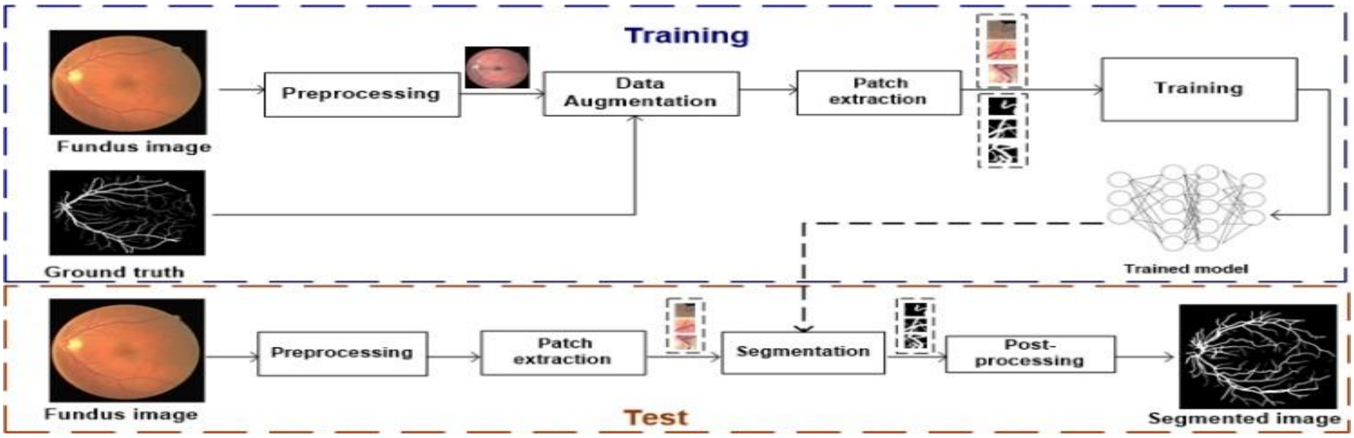

Fig 2 The process of proposed retinal blood vessel segmentation method.

available image, a total of 5760 images for training. Out of these, 3840 images were used to train the networks with 1920 images reserved for validation. Table 2 summarizes the numbers of images used in each set.

B. Evaluation metrics

Evaluation of the segmentation results is necessary to analyze the robustness of the architecture. It is based on the classification of each pixel as true positive (TP), true negative (TN), false positive (FP) and false negative (FN). The TP and TN are respectively the number of pixels that are correctly classified, while the FP and FN are respectively the number of misclassified pixels. Segmentation performance metrics are Accuracy (ACC) that describes classification performance. Sensitivity (Se) and Specificity (Sp) which reflect the ability to correctly classify pixels. The DICE coefficient that represents the similarity between the segmented image and the groundtruth image. Precision that indicates the proportion of classify pixels correctly. Table 5. describes the performance metrics for the evaluation of retinal vessel segmentation.

C. Segmentation Performance

The obtained results during segmentation with the new architecture with DRIVE dataset based on ACC, SE, SP, Dice and Precision, in addition, to the number of parameters, are computed by referring to the manual segmentation provided by the expert. Our model shows a high performance, where ACC, Se, Sp, Dice and Precision are respectively in the order of 97.08 %, 82.38 %, 98.55 %, 81.12 %, 84.25%. The total number of parameters needed in this model is 3.4 M. These results are shown in Table 4. Examples of segmentation result are illustrated in Fig 3.

TABLE III. Training parameters

Parameter	**Value**
Learning rate	0.001
Optimizer	ADAM
Loss function	BinaryCrossnetropy
Epoch number	20
Batch_size	16
Patch_size	128

D. Comparison of segmentation performances for DRIVE

Subsequently to evaluate the suggested method, we propose to compare our segmentation result with some methods like Segnet, U-net, RCED-net an extension of Segnet and RV-net an extension of U-Net with LCMs module in convolution blocks, for the DRIVE database, according to the following metrics Acc, Se and Sp and the number of parameters. The comparison results are presented in Table 6. We notice that our model trained by the new convolution blocks improves the segmentation performance of Segnet network. Indeed, the accuracy value is increased from 95.79 % to 97.08 %, likewise the sensitivity and specificity values which reflect the ability to correctly classify and to have a true prediction, are higher in our model. In addition, our method achieves better accuracy than RCED-Net and U-net. RV_net have a higher performance, an accuracy of 98.19 %. The number of parameters provided by our model is 3.4 M. We reduce the complexity better then all methods in table 5. The complexity is reduced more than four times from 14.7M to 3.4 M compared to Segnet.

TABLE IV. Performance and complexity result of our model

Metrics					***Parameters***
Acc(%)	***Se(%)***	***Sp(%)***	***Dice(%)***	***Precision(%)***	
97.08	82.38	98.55	81.12	84.25	3.4 M

TABLE II. Information of the database used in our experiment.

Database	**Dataset size**	**Training set size**		**Validation set size**		**Test set size**
		Before augmentation and cropping	***After augmentation and cropping***	***Before augmentation and cropping***	***After augmentation and cropping***	
Drive	40	20	3840	10	1920	10

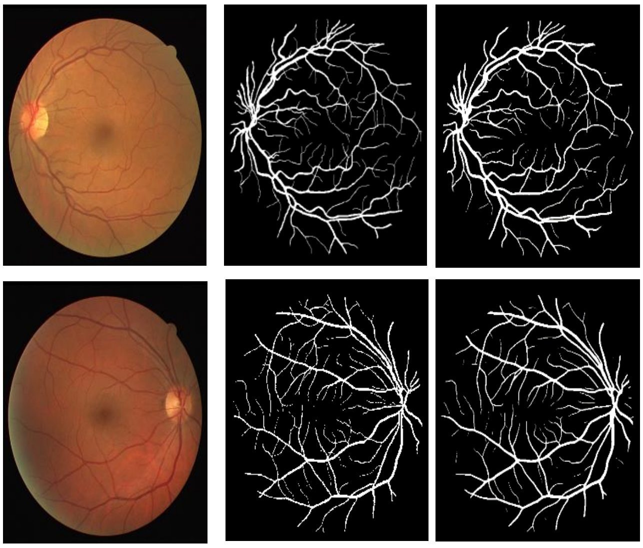

Fig 3. Segmentation results: first column: retinal image, second column: Ground truth, third column: segmentation result

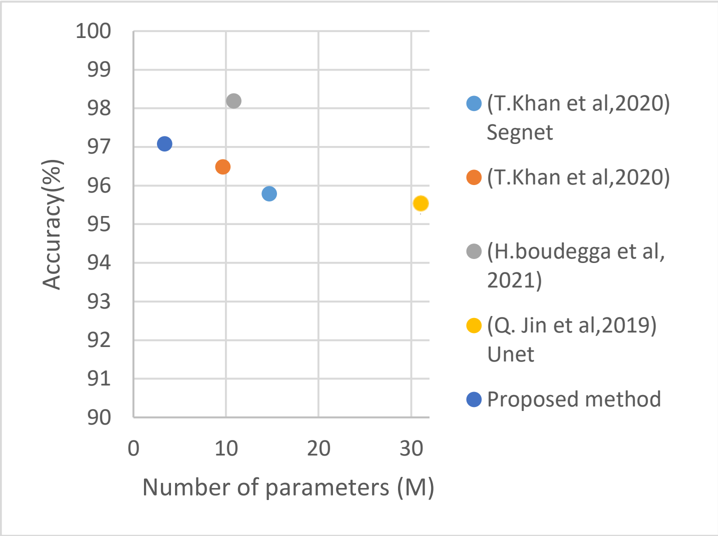

Fig 4. Tradeoff between segmentation accuracy and execution time.

TABLE V. Performance metrics for the evaluation of retinal vessel segmentation.

Metrics	Description
Accuracy (Acc)	TP+TN/(TP+TN+FP+FN)
Sensitivity (Se)	TP/(TP+FN)
Specificity(SP)	TN/(TN+FP)
Precision	TP/(TP+FP)
Dice	(2×TP)/(2×TP+FP+FN)

TABLE VI. Comparison of segmentation performances for DRIVE database

Works	**Metrics**			**Parameters(M)**
	Acc(%)	***Se(%)***	***Sp(%)***	
T.Khan et al. [11] Segnet	95.79	79.49	97.38	14.7
T.Khan et al. [11] RCED-Net	96.49	82.52	97.87	9.7
Q. Jin et al. [10] U-Net	95.54	78.49	98.04	31.03
H.boudegga et al.. [12] RV-Net	**98.19**	**84.48**	**99.00**	10.87
Proposed method	**97.08**	**82.38**	**98.55**	**3.4**

Additionally, we suggest studying the trade-off between the segmentation performance and the complexity of retinal segmentation methods. For that purpose, we propose to investigate about the evolution of accuracy according to the complexity. For that, the retinal vessel segmentation methods are depicted into a 2D Cartesian space where horizontal and vertical coordinates correspond to the complexity and the accuracy rate, respectively, as shown in Fig 4. Our method reaches the less complexity with a higher accuracy rate among those methods, as represented by the blue point. For other work, the growth of the accuracy rate has been always associated with the rise in complexity.

VI. Conclusion

Several ocular and associated pathologies affect the retina such as hypertensive retinopathy, diabetic retinopathy and AMD. These pathologies can cause blindness so, early detection is important to prevent advanced stages. The diagnosis is made by segmentation of the retinal vascular tree (RVT) since these diseases have effects on the vascular network and the analysis of this network facilitates the diagnosis. Several Deep Learning architectures have been proposed for the automatic segmentation of RVT, respecting the constraints, the performances that must be improved while reducing the complexity. In this context, the objective of this work is to propose a new architecture and apply it in a new segmentation method, to maintain high performance while reducing complexity. Therefore, we proposed an architecture inspired by the segnet architecture by replacing the standard convolutional blocks with the reduced complexity convolutional blocks of the MobileNet architecture. The evaluation of this architecture is done with the DRIVE database, based on these metrics Accuracy, Sensitivity, Specificity, Precision and Dice. The obtained results for these metrics are respectively 97.08%, 82.38%, 98.55%, 81.12% and 84.25%. The total number of parameters is 3.4M. As a result, the segmentation performance is improved, the accuracy is increased from 95.79% to 97.08%, and the complexity is reduced by more than 4 times since the total number of parameters is reduced from 14.7M to 3.4M. We can therefore say that we have achieved our objectives of improving performance and reducing complexity.

References

[1] A. Oliveira, S. Pereira et C. A. Silva. "Retinal vessel segmentation based on Fully Convolutional Neural Networks". In : Expert Systems with Applications 112 (2018), p. 229-242. url https://www.sciencedirect.com/science/article/pii/S0957417418303816.

[2] Y. Liu et al. "ResDO-UNet : A deep residual network for accurate retinal vessel segmentation from fundus images". In : Biomedical Signal Processing and Control 79 (jan. 2023), p. 104087

[3] Xu et al. “Retinal Vessel Automatic Segmentation Using SegNet”. In : Computational and Mathematical Methods in Medicine 2022 (mar. 2022), p. 1-11.

[4] A. G. Howard et al. MobileNets : Efficient Convolutional Neural Networks for Mobile Vision Applications. cite arxiv :1704.04861. 2017. url : http : / / arxiv . org/abs/1704.04861.

[5] In : Graphics gems. 1994. url : https://api.semanticscholar.org/CorpusID:62707267.

[6] O. Ronneberger, P. Fischer et T. Brox. “U-Net : Convolutional Networks for Biomedical Image Segmentation”. In : t. 9351. Oct. 2015, p. 234-241.

[7] H. Cheng et X. Shi. “A simple and effective histogram equalization approach to image enhancement”. In : Digital Signal Processing 14.2 (2004), p. 158-170. url : https://www.sciencedirect.com/science/article/pii/S105120040300037X.

[8] U. Ozgunalp, R. Fan et A. Serener. “Semantic Segmentation of Retinal Vessels Using SegNet”. In : 2020 28th Signal Processing and Communications Applications Conference (SIU). 2020, p. 1-4.

[9] Q. Qin et Y. Chen. “A review of retinal vessel segmentation for fundus image ana- lysis”. In : Engineering Applications of Artificial Intelligence 128 (2024), p. 107454. url : https://www.sciencedirect.com/science/article/pii/S095219762301638X.

[10] Q. Jin et al. “DUNet : A deformable network for retinal vessel segmentation”. In : Knowledge-Based Systems 178 (mai 2019).

[11] T. Khan et al.“Residual Connection-Based Encoder Decoder Network (RCED-Net) for Retinal Vessel Segmentation”. In : IEEE Access (juil. 2020).

[12] H. Boudegga et al. “Fast and efficient retinal blood vessel segmentation method based on deep learning network”. In : Computerized Medical Imaging and Graphics 90 (mar. 2021), p. 101902.

[13] C. Guo et al. “SA-UNet : Spatial Attention U-Net for Retinal Vessel Segmentation”. In : (avr. 2020).

[14] Ren et al. “An improved U-net based retinal vessel image segmentation method”. In : Heliyon 8.10 (2022), e11187. url : https://www.sciencedirect.com/science/article/pii/S2405844022024756.

[15] Li et al. “Blood Vessel Segmentation of Retinal Image Based on Dense-U-Net Network”. In : Micromachines 12 (nov. 2021), p. 1478.

[16] T. Laibacher, T. Weyde et S. Jalali. “M2U-net : Effective and efficient retinal vessel segmentation for real-world applications”. In : © 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. IEEE, avr. 2020, p. 115-124. url : https://openaccess.city.ac. uk/id/eprint/24152/.

[17] V. Sathananthavathi et I. Ganesan. “Encoder Enhanced Atrous (EEA) Unet architecture for Retinal Blood vessel segmentation”. In : Cognitive Systems Research 67 (jan. 2021).

[18] R. Liu et al. “DA-Res2UNet : Explainable blood vessel segmentation from fundus images”. In : Alexandria Engineering Journal 68 (2023), p. 539-549. url : https://www.sciencedirect.com/science/article/pii/S111001682300073X.

[19] Y. Liu et al. “Wave-Net : A lightweight deep network for retinal vessel segmentation from fundus images”. In : Computers in Biology and Medicine (2023). url : https://www.sciencedirect.com/science/article/pii/S0010482522010496.

[20] J. Zhang, Y. Zhang et X. Xu. “Pyramid U-Net for Retinal Vessel Segmentation”.In : ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2021), p. 1125-1129. url : https://api.semanticscholar.org/CorpusID:233033718.

[21] Z. Zhu et al. “ILU-Net : Inception-Like U-Net for retinal vessel segmentation”. In : Optik - International Journal for Light and Electron Optics 260 (avr. 2022), p. 169012.

[22] Z. CUI, S. SONG et J. QI. “MF2ResU-Net : a multi-feature fusion deep learning architecture for retinal blood vessel segmentation”. In : Digital Chinese Medicine 5.4 (2022), p. 406-418. url : https://www.sciencedirect.com/science/article/pii/S2589377722000787.

[23] V. Badrinarayanan, A. Kendall et R. Cipolla. “SegNet : A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation”. In : IEEE Transactions on Pattern Analysis and Machine Intelligence 39.12 (2017), p. 2481-2495.

[24] S. Minaee et al. “Image Segmentation Using Deep Learning : A Survey”. In : IEEE transactions on pattern analysis and machine intelligence PP (f´ev. 2021).