Retinal blood vessel segmentation from high resolution fundus image using deep learning architecture

Henda Boudegga, Yaroub Elloumi, Asma Ben Abdallah, Rostom Kachouri, Nesrine Abroug, Mohamed Hedi Bedoui

Retinal blood vessel segmentation from high resolution fundus image using deep learning architecture

Henda Boudegga[1,3], Yaroub Elloumi[1,3], Asma Ben Abdallah[1], Rostom Kachouri[2], Nesrine Abroug[4], Mohamed Hedi Bedoui[1]

[1]Medical Technology and Image Processing Laboratory, Univ. of Monastir, Tunisia.
[2]LIGM, Univ. Gustave Eiffel, CNRS, ESIEE Paris, F-77454 Marne-la-Vallée France
[3]ISITCom Hammam-Sousse, University of Sousse, Tunisia.
[4]Department of Ophthalmology, Fattouma Bourguiba University Hospital, Faculty of Medicine, University of Monastir, Tunisia

Abstract. The Retinal Vascular Tree (RVT) segmentation is required to diagnose various ocular pathologies. Recently, fundus images are acquired with higher resolution, which allows representing a large range of vessel thickness. However, standard Deep Learning (DL) architectures with static and small convolution size have failed to achieve higher segmentation performance. In this paper, we propose a novel DL architecture for RVT segmentation dedicated for high resolution fundus images. The idea consists at extending the U-net architecture by increasing (e.g. decreasing) convolution kernel size through convolution blocs, in correlation with downscale (e.g. upscale) of feature map dimensions. The proposed architecture is validated on HRF database, where average sensitivity is increased from 56% to 84%.

1 Introduction

The retinal blood vessels are a main component responsible of the visual acuity. Several ocular pathologies alter the anatomy of the vascular tree, such as the proliferative stage of the diabetic retinopathy [1] [2], the Age-related Macular Generation (AMD) and the Retinopathy Hypertensive (RH) [3]. Hence, several process of clinical diagnosis is based on segmenting the retinal vascular tree. For this purpose, several methods of retinal vessel tree segmentation have been put forward. With respect to the complicated vessel morphology, Deep Convolutional Neural Networks (DCNNs) methods have achieved great successes in retinal vessel segmentation task with respect to convolutional approach. The DCCNs are based on convolution layers, which are responsible for feature extraction task.

The current imaging technology allows capturing fundus images with higher resolution, which are illustrated with a large gap of multi vessels thickness. Thus, the static setting of convolution layers with reduced kernel size cannot guarantee properly feature extraction whatever the vessel thickness is. Even several have achieved higher segmentation quality in lower resolution fundus, a lower detection rates are deduced when segmenting higher resolution images [4].

For this context, we propose a new DCNN that allows extracting the RVT representation into high-resolution fundus image. The main contribution consists of changing the size of convolutional filters into an increased (e.g. decreased) way over convolution blocs correlated with downscale (e.g. upscale) of fundus image size. Existing CNN architecture has been used as a backbone network, giving rise to an extended architecture for the retinal vascular segmentation from high-resolution fundus image. The remainder of the paper is organized in five sections. In section 2, a brief review of multi-scale segmentation approaches is described. Thereafter, the proposed network is detailed. The experimental results are presented in section 4 where proposed architecture is evaluated on high resolution databases and compared to multi-scale segmentation approaches. The conclusion is given in the last section.

2 Related works

In this section we suggest reviewing some vessel segmentation approach based deep learning networks. Several methods suggested extending several networks to be adaptable for the detection of the multi vessel thickness throw substituting standard convolution layer or adjusting their kernel forms or sizes or making additional connection between layers or managing multi-networks. For instance, wang and al. [5] introduce dense connections into U-Net architecture to make full use of the multi-scale features for vessels segmentation. Jin et al. [6] have replaced the regular kernel by deformable kernel into the convolution layers of U-Net architecture. Additional, the work [4] and [7] suggest switching the standard convolution layers within U-net blocks respectively by dilated convolution layers parameterized by a kernel size in the order of 3x3 dilated by a factor of 2 and depthwise convolution modules. Even, several have optimized the size of convolution kernel. The work [8] propose enlarging the size of kernel of the U-net convolution layers with respect to vessel representation into fundus images. By way of contrast, other methods have employed several networks, where each one ensured segmenting vessels with a specific thickness. Tian and al. [9] present a multi-path network composed by low and high frequency paths for segmenting thick and thin vessels respectively, where high frequency path contain dilated convolutional blocks. Furthermore, the work depicted in [10] suggested three stages of deep learning networks for the segmentation of the retinal vessel tree. Two networks were proposed to respectively segment thick and thin vessels and another was suggested for merging the thicker and thinner segmentation map to generate the final segmentation map.

Those proposed contributions have guaranteed effective segmentation results to address the limitations of standard versions. However, their potential results are generated on lower resolution databases such as DRIVE and STARE, where sensitivity rates on DRIVE fundus images of [6] and [4] are in the range of 80% respectively. In contrast, their sensitivity rates on HRF fundus images are only about 74 and 65%, respectively. Thus, those proposed contributions are inappropriate to tolerate the retinal vascular tree from high resolution fundus images.

3 Deep Learning architecture for the segmentation of retinal blood vessel on high resolution fundus images

3.1. Problem analysis

The retinal vascular tree spread from the OD center with a larger width and expends into the whole retina with a curvilinear or slightly tortuous shape, providing vessels descends having decreased width ranging approximately from 200µm to some micro. Recently, fundus images are captured with retinographs characterized by high resolution allowing the representation of fundus images in the order of 6µm/pixel as example. Hence, a large gap of vessel thickness variation is modeled. The segmentation throw a Convolution Neural Networks (CNNs) consists of extracting features from several features maps having different sizes, basing on convolution processing. The convolution processing consists of sliding a convolution kernel “K” with small matrix “nxn” of weights “W”, managing for each slid an element-wise multiplication between the weight values “W” and the kernel-sized patch of the input feature map “I”. Then, the results are summed to obtain an output value in the corresponding position (i,j) of the output feature map “O”, as donated in equation (1). Figure 1, illustrate the convolution processing.

$$O(i,j)=\sum_{p=0}^{nxn} W(Kpx,Kpy)*I(i+Kpx,j+Kpy) \quad \textbf{(1)}$$

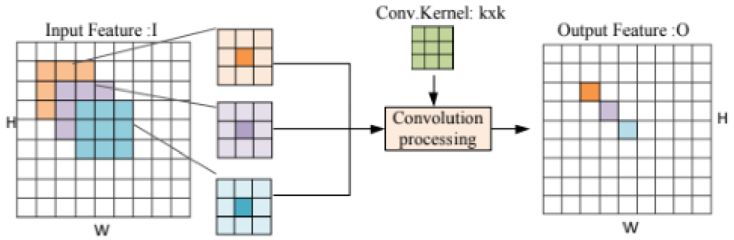

Fig. 1**:** Convolution processing

In most CNN models, the convolution layers have similar settings over the network with a fixed kernel size in the order of “nxn=3x3”. Even, this convolution setting is insufficient to cover vessel pixels sharing common features, leading to low segmentation quality.

As described in section 2, the idea of [8] consists of enlarging the convolution kernel size in order to make the kernel look at more surrounding neighbors while computing convolution. This purpose has been validated on high resolution fundus images setting the kernel sizes to 7x7, resulting an improved result with respect to the standard network. Although, as vessels are represented differently in a high-resolution fundus image, scaling in a large thickness gap, as well as features extracted from the most anterior layers will be used in the deepest layer. Hence, applying a large convolution kernel in the order of 7x7 for all convolution layers in the network may not extract vessel details represented on a few pixels. Therefore, to extract vessels from a high-resolution fundus image, the earliest layer of the network must be configured with the smallest kernel size correlating with the smallest vessel representation. The aim is to extract and record local details to be used in the deeper layers. Additionally, most of CNN networks process several downscale operations, resulting different sizes of features maps over blocs. Thus, local features details are hierarchically dropped over blocs and high level features are retained. Therefore, the convolution kernel sizes should be hierarchically increased to “mxm” as illustrated in Figure 2 over blocs in correlation with downscale processing. The aim is to make the kernel look at more surrounding neighbors while computing convolution processing to extract more complex features over blocs to be used in the innermost blocs. The following Figure shows the idea of increasing the convolution kernel over blocs.

In this context, within the aim of guaranteeing an efficient segmentation of blood vessels from high resolution fundus image, the main contribution of this work is to propose a deep learning architecture with multi-scale convolution kernel size for the segmentation of retinal vascular tree from high resolution fundus image.

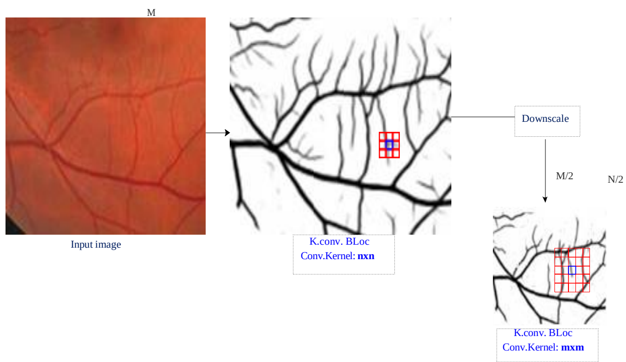

Fig. 2: Application of increased convolution kernel size for the segmentation of blood vessels from high resolution fundus image.

3.2. Proposed network

The main contribution of this work consists of proposing a deep learning network for the segmentation of the retinal vascular tree from high resolution fundus images. Thus, we propose extending the well-known architecture U-net through configuring their convolution layers over blocs by multi-scale kernel sizes. The objective is to guarantee efficient segmentation quality by adopting the U-form and making the network gradually capturing feature complexity and covering local features details as well as global features.

The U-net architecture is structured on Downsampling and Upsampling paths composed respectively by five and four blocks, each of which contains two convolution layers. Those blocs are separated respectively by 2x2 maxpooling and 2x2 Upsampling layers. The proposed configuration consists of setting the convolution kernel sizes of the earliest Downsampling blocs to the smallest vessel scales ranging in the order of “3x3”. The aim is to collect as much as possible features and to simulate the simple representation of vessels and to capture more detailed features. Thereafter, the bellow convolution blocs are parameterized by an increased order of convolution kernel size “mxm” correlated with the downscaled features map sizes. Thus, those convolution blocs are configured by kernel sizes respectively in the order of “5x5”, “7x7”, “9x9” and “11x11”. The aim is to capture hierarchically features and to represent more global and high-level representative information. Similarly to the standard U-net architecture, the Upsampling path configuration is mirrored with Dowsampling configuration path, where blocks with the same level are parameterized with the same convolution kernel size. Thus, the Upsampling convolution blocs are configured by kernel sizes respectively in the order of “9x9”, “7x7”, “5x5” and “3x3”. For the last bloc, an additional convolution layer is proposed as illustrated with yellow box parameterized with a kernel size in the order of 1x1 and activated with softmax function to make vessel detection decision. The proposed U-net model for the segmentation of high resolution fundus image is illustrated in Figure 3.

The proposed network is undergoes for training process to adjust weight nodes and achieve a precise model. The process involves utilizing a specific set of training parameters. The Xavier initialization technique was adopted to initialize the weights and biases. In addition, Adam optimizer with a learning rate in the order of 0.001 was used. In addition, the cross entropy loss is applied to minimize gap between the prediction and ground truth.

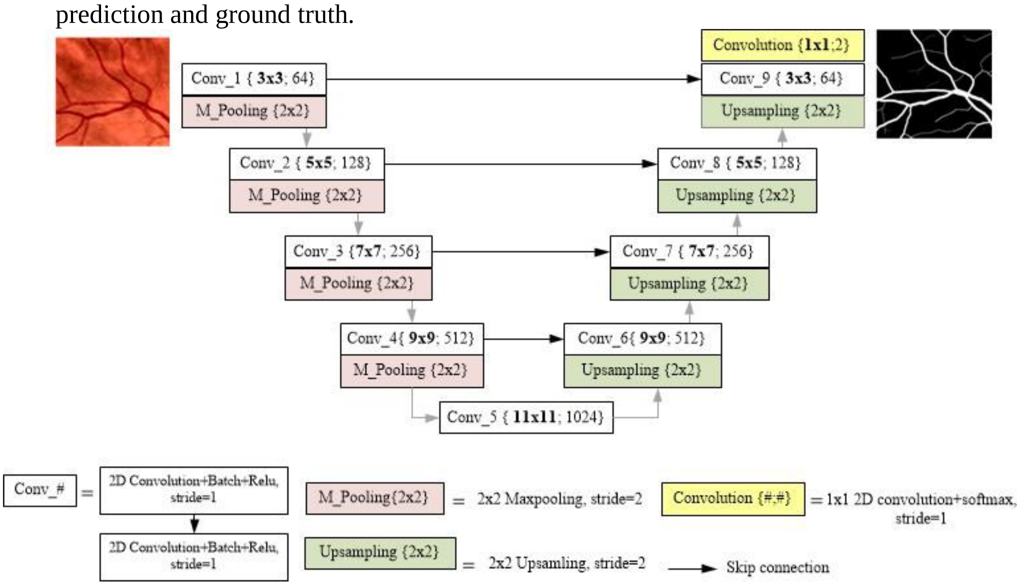

Fig. 3: Proposed network for retinal blood vessel segmentation into high resolution fundus image.

In order to evaluate the proposed network's ability to detect blood vessels from high resolution fundus image, we employed our approach, proposed in [7]. A preprocessing step is firstly applied to enhance image quality. Following this, a cropping process is employed to generate multiple sub-images with dimensions around 192x192, passing through the proposed network. A post-processing step is implemented to merge the segmented sub-images and generate a complete segmented image of the RVT.

4 Experiments

4.1. Database and evaluation criteria and experiment setup

The validation of the proposed model is ensured by the deployment of the high-resolution retinal database HRF. The database is composed by 45 retinal images having the size of 3504x2336 captured with a resolution in the order of 6μm/pixel. Various evaluation metrics are applied for pixel classification results basing on, including Accuracy (Acc), Sensibility (Sens), Specificity (Spec) and DICE. All those evaluation metrics are computed respectively as equation in the following Table 1. The implementation is conducted on Intel core i7 with a 3.67 GHZ frequency processor, 8Go RAM and a NVIDIA GTX 980 GPU using CUDA 9.0 with CUDNN 7.6.3.

Metrics	Description
Sensitivity (SE)	TP / (TP+FN)
Specificity (SP)	TN / (TN+FP)
Accuracy (ACC)	TN+TP/TP+FP+TN+FN
Precision	2*TP/ 2 * TP + F N + FP

Table 1: Performance metrics for the evaluation of retinal vessel segmentation.

1.1. Experimental results

In this section, we assess the effectiveness of the proposed method for segmenting the RVT from the high resolution fundus images of the HRF database, as illustrated in Fig.4. The segmentation results in terms of ACC, SE, SP, and DICE, are respectively in the order of 96.75%, 84.5%, 97.86% and 76.03%. Figure 4 show segmentation results of HRF fundus image.

In addition, we narrowed our investigation by comparing their performance with the baseline U-net and others state-of-the-art methods. From Table 2, their segmentation accuracy rates are closed, due to the unbalanced vessel and background pixels count. Hence, for RVT segmentation tasks the sensitivity is an important metric as it highlights the model's ability to identify the subtle vessel details.

Basing on the sensitivity rates of Table 2, the segmentation quality of the segmented RVT has improved with respect the U-Net model [17], where sensitivity is improved from 56% to 84%. Additionally, we can conclude that the proposed method's sensitivity rate outperforms DL-based methods such as [11] , [12], [13], [14], [15], [6], [4] and [16] exhibiting a notable difference of approximately 14%. Consequently, the proposed network demonstrates a robust capability to detect vessel pixels. This significant performance can be attributed to the integration of multi-scale convolutional kernel sizes into the U-Net model.

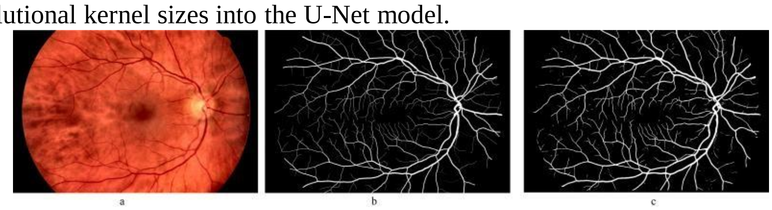

Fig.4: Segmentation results of HRF database; (a) Retinal images, (b) Ground truth, (c) Segmented results.

Methods	[11]	[12]	[13]	[14]	[15]	[6]	[16]	[4]	U-net [17]	**Proposed model**
ACC(%)	95.57	96.98	94.37	96.5	96.54	96.51	96.37	92.44	95.77	**96.75**
SE(%)	70.54	80.76	78.81	80.10	78.03	74.64	80.37	65.77	56.29	**84.5**
SP(%)	98.34	98.31	95.92	80.10	98.43	98.74	97.96	97.99	97.25	**97.86**

Table 3: Comparison of segmentation performances on HRF database.

5. Conclusion

In this paper, we have proposed a novel DL architecture to achieve efficient segmentation from high resolution fundus image. The contribution consists of extending the well-known architecture U-net throw expanding an increased (e.g. decreased) convolution kernel size through convolution blocs, in correlation with downscale (e.g. upscale) of feature map dimensions. The proposed architecture is validated on HRF database reaching a higher accuracy of 96.75 with a sensitivity of 84%. With the permanent increase in image resolution, the convolution kernel size can be adjusted following the same principle, in order to perform an accurate feature extraction. The same DL extension can be adopted for the segmentation of artery and vein, which is crucial to detect several pathologies such hypertensive retinopathy.

Knowledge

This work is carried out as part of the MAGHREB 2023 PHC project (Project Code: 48672RG).

References

[1] S. S. Kar and S. P. Maity, „Detection of neovascularization in retinal images using mutual information maximization", *Computers & Electrical Engineering*, vol. 62, pp. 194–208, Aug. 2017.

[2] S. Yu, D. Xiao, and Y. Kanagasingam, „Automatic detection of neovascularization on optic disk region with feature extraction and support vector machine", in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2016, pp. 1324–1327.

[3] S. Khitran, M. U. Akram, A. Usman, and U. Yasin, „Automated system for the detection of hypertensive retinopathy", in *2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oct. 2014, pp. 1–6.

[4] S. V. and I. G., „Encoder Enhanced Atrous (EEA) Unet architecture for Retinal Blood vessel segmentation", *Cognitive Systems Research*, vol. 67, pp. 84–95, Jun. 2021.

[5] 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, p. 168, Feb. 2019, doi: 10.3390/e21020168.

[6] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, „DUNet: A deformable network for retinal vessel segmentation", *Knowledge-Based Systems*, vol. 178, pp. 149–162, Aug. 2019.

[7] H. Boudegga, Y. Elloumi, M. Akil, M. Hedi Bedoui, R. Kachouri, and A. B. Abdallah, „Fast and efficient retinal blood vessel segmentation method based on deep learning network", *Computerized Medical Imaging and Graphics*, vol. 90, p. 101902, Jun. 2021.

[8] H. Boudegga, Y. Elloumi, R. Kachouri, A. Ben Abdallah, and M. H. Bedoui, „Extended U-net for Retinal Vessel Segmentation", in *Advances in Computational Collective Intelligence*, C. Bădică, J. Treur, D. Benslimane, B. Hnatkowska, and M. Krótkiewicz, Eds., in Communications in Computer and Information Science. Cham: Springer International Publishing, 2022, pp. 564–576.

[9] C. Tian, T. Fang, Y. Fan, and W. Wu, „Multi-path convolutional neural network in fundus segmentation of blood vessels", *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 583–595, Apr. 2020.

[10] Z. Yan, X. Yang, and K.-T. Cheng, „A Three-Stage Deep Learning Model for Accurate Retinal Vessel Segmentation", *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1427–1436, Jul. 2019.

[11] P. M. Samuel and T. Veeramalai, „VSSC Net: Vessel Specific Skip chain Convolutional Network for blood vessel segmentation", *Computer Methods and Programs in Biomedicine*, vol. 198, p. 105769, Jan. 2021.

[12] Z. Fan *et al.*, „Accurate Retinal Vessel Segmentation via Octave Convolution Neural Network". arXiv, Sep. 22, 2020.

[13] Z. Yan, X. Yang, and K.-T. Cheng, „Joint Segment-Level and Pixel-Wise Losses for Deep Learning Based Retinal Vessel Segmentation", *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 1912–1923, Sep. 2018.

[14] Z. Jiang, H. Zhang, Y. Wang, and S.-B. Ko, „Retinal blood vessel segmentation using fully convolutional network with transfer learning", *Computerized Medical Imaging and Graphics*, vol. 68, pp. 1–15, Sep. 2018.

[15] D. Wang, A. Haytham, J. Pottenburgh, O. Saeedi, and Y. Tao, „Hard Attention Net for Automatic Retinal Vessel Segmentation", *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 12,

pp. 3384–3396, Dec. 2020.

[16] D. Li and S. Rahardja, „BSEResU-Net: An attention-based before-activation residual U-Net for retinal vessel segmentation", *Computer Methods and Programs in Biomedicine*, vol. 205, p. 106070, Jun. 2021.

[17] J. Han, Y. Wang, and H. Gong, „Fundus Retinal Vessels Image Segmentation Method Based on Improved U-Net", *IRBM*, vol. 43, no. 6, pp. 628–639, Dec. 2022.