ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

# 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], 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.

**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

Retinal blood vessels play a crucial role in visual acuity. Various ocular pathologies, including diabetic retinopathy, age-related macular degeneration, and hypertensive retinopathy, can affect the vascular tree's anatomy. Clinical diagnosis often relies on segmenting the retinal vascular tree, leading to the development of numerous retinal vessel segmentation methods. Deep Convolutional Neural Networks (DCNNs) have shown significant success in this task due to their ability to handle the complex vessel morphology through convolutional layers responsible for feature extraction. Advancements in imaging technology enable the capture of high-resolution fundus images, showcasing a wide range of vessel thicknesses. The conventional setup of convolution layers with reduced kernel size might not effectively extract features across all vessel thicknesses. While some methods achieved better segmentation quality in lower resolution fundus images, lower detection rates were observed when dealing with higher resolution images [1].

In this context, we introduce a new DCNN for extracting the RVT representation from high-resolution fundus images. The contribution involves altering the size of convolutional filters in correlation with the upscale or downscale of the fundus image size. We build upon an existing CNN architecture, extending it for retinal vascular segmentation. The paper is structured into five sections. Section 2 provides a brief review of multi-scale segmentation approaches, followed by a detailed explanation of our proposed network. Section 4 presents experimental results, evaluating our architecture on high-resolution databases and comparing it with other multi-scale segmentation approaches. The paper concludes with a summary in the last section.

## 2 Related works

In this section, we review vessel segmentation approaches based on DL networks. Several methods extend existing networks to adapt to detecting vessels with varying thicknesses, by substituting standard convolution layers, either adjusting kernel forms or sizes or making additional connections between layers. Wang et al. [2] introduced dense connections into the U-Net architecture to leverage multi-scale features for vessel segmentation. Jin et al. [3] replaced the regular kernel with deformable kernels in U-Net convolution layers. Additionally, works like [1] and [4] suggest switching standard convolution layers within U-Net blocks with dilated convolution layers parameterized by a 3x3 kernel dilated by a factor of 2 and depthwise convolution modules. Even, several have updated the convolution kernel size respect to vessel representation in fundus images, such as the work described in [5].

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 [3] and [1] 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 RVT originates from the center of the OD and gradually expands into the entire retina with a curvilinear or slightly tortuous shape. The vessels have varying widths, decreasing from around 200µm to microscopic dimensions. Fundus images captured nowadays have high resolution, represented at approximately 6µm/pixel, leading to a significant range of vessel thickness variation. To perform segmentation through Convolutional Neural Networks (CNNs), features are extracted from multiple feature maps with different sizes using convolution processing. This process involves sliding a convolution kernel "K" with a small matrix "nxn" of weights "W" over the input feature map "I". At each position (i,j) of the output feature map "O," an element-wise multiplication between the weight values "W" and the kernel-sized patch of the input feature map "I" is performed, and the results are summed to obtain the output value (as shown in equation (1)). In most CNN models, the convolution layers have typically fixed settings with a kernel size of "nxn=3x3." However, this configuration is insufficient to model all neighbor features, resulting a low segmentation quality.

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

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

As described in section 2, the approach proposed in [5] and the work [6] involves enlarging the convolution kernel size to capture more surrounding neighbors during computation, improving results on high-resolution fundus images when using a 7x7 kernel size. However, applying a large kernel size throughout the entire network may miss vessel details represented on a few pixels and may mix features from different vessel thicknesses. To address this, the earliest layer of the network should have the smallest kernel size, extracting local details for use in deeper layers.

In addition, CNN networks often involve downscale operations, resulting in different sizes of feature maps across blocks. This hierarchical processing drops local feature details and retains high-level features. To extract complex features for innermost blocks, the convolution kernel sizes should be hierarchically increased to "m x m" in correlation with the downscale processing, as shown in Figure 1. This allows the kernel considering more surrounding neighbors during computation and extract more complex features over blocks. In order to efficiently segmenting blood vessels from high-resolution fundus images, the main contribution of this work is the proposal of a deep learning architecture with a multi-scale convolution kernel size.

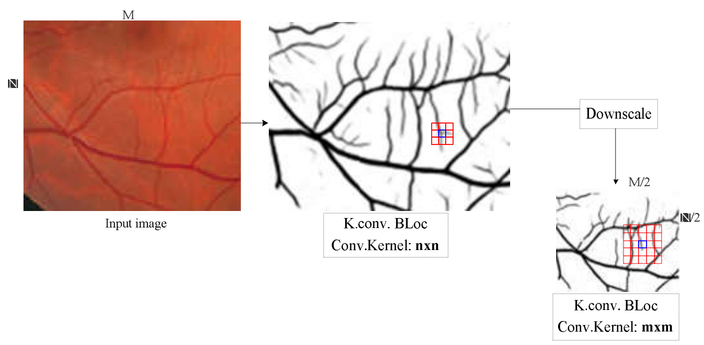


Fig. 1: Increased convolution kernel size for blood vessel segmentation from high-resolution fundus image.

## 3.2. Proposed network

This work introduces a deep learning network for retinal vascular tree segmentation from high-resolution fundus images. The proposed method extends the U-net architecture by configuring convolution layers in blocks with multi-scale kernel sizes. This approach ensures efficient segmentation quality by gradually capturing feature complexity, covering both local and global features.

The U-net architecture is structured on Downsampling and Upsampling paths with five and four blocks, respectively. Each block consists of two convolution layers and separated by 2x2 maxpooling for the downsampling blocks and 2x2 Upsampling layers for the Upsampling blocks. The configuration consists of setting the convolution kernels size in the earliest Downsampling blocks to the smallest vessel scale (3x3), in order to capture basic vessel representations and detailed features.

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

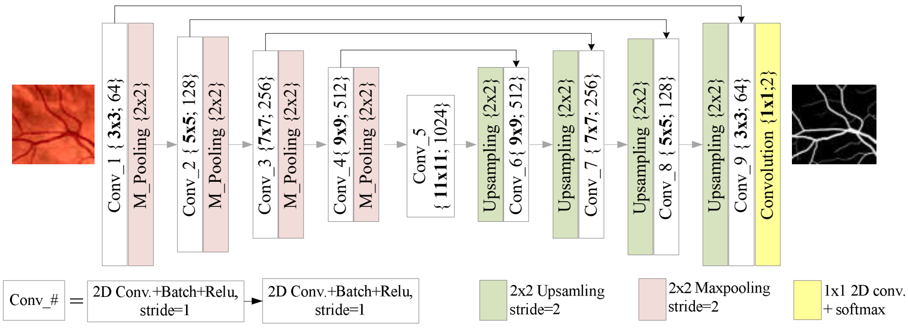


Fig. 2: Proposed network for retinal blood vessel segmentation into high resolution fundus image

Subsequent convolution blocks use increasing kernel sizes (5x5, 7x7, 9x9, and 11x11) to capture hierarchical and global features. The Upsampling path mirrors the Downsampling path with the same kernel size configurations (9x9, 7x7, 5x5 and 3x3). The final block includes an additional 1x1 convolution layer as shown with yellow box activated with softmax for vessel detection. The proposed U-net model for high-resolution fundus image segmentation is shown in Figure 2.

The proposed network is undergoing for training to adjust weight nodes and achieve precise model. Thus, the Xavier technique is used to initialize weights and biases. The Adam optimizer with a learning rate of 0.001 and cross-entropy loss are employed to minimize the gap between predictions and ground truth. In order to evaluate the proposed network, we employed our approach proposed in [4]. A preprocessing step is firstly applied to enhance image quality. Then, a cropping is preformed to generate multiple sub-images with dimensions around 192x192, before providing then to the proposed network. A post-processing step is implemented to merge the segmented sub-images and generate a complete segmented image.

# 4 Experiments

## 4.1. Database and evaluation criteria and experiment setup

The proposed model is validated using the high-resolution retinal database HRF, comprising 45 retinal images sized at 3504x2336 with a resolution of approximately 6µm/pixel. Various evaluation metrics, including Accuracy (Acc), Sensibility (Sens), Specificity (Spec), and DICE, are employed for pixel classification results calculated as equations (2-5). The implementation is performed on an Intel Core i7 processor with a frequency of 3.67 GHz, 8GB RAM, and a NVIDIA GTX 980 GPU using CUDA 9.0 with CUDNN 7.6.3.

$$\text{Accuracy (Acc)} = TN+TP/(TP+FP+FN+TN) \quad (2)$$

$$\text{Sensitivity (Sens)} = TP/(TP + FN) \quad (3)$$

$$\text{Specificity (Spec)} = TN/(TN + FP) \quad (4)$$

$$\text{DICE}= 2*TP/ 2 * TP + F N + FP \quad (5)$$

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

### 4.2. 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.3. The segmentation results in terms of Acc, Sens, Spec, and DICE, are respectively in the order of 96.75%, 84.5%, 97.86% and 76.03%.

In addition, we narrowed our investigation by comparing their performance with the baseline U-net and others state-of-the-art methods. From Table 1, 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 1, the segmentation quality of the segmented RVT has improved with respect the U-Net model [7], where sensitivity is improved from 56% to 84%. This significant enhancement is attributed to the integration of multi-scale convolutional kernel sizes into the U-Net model. Additionally, we can conclude that the proposed method's sensitivity rate outperforms DL-based methods such as [8], [9], [10], [3], [1], [5] and [11] exhibiting a notable difference of approximately 14%. Consequently, the proposed network demonstrates a robust capability to detect vessel pixels.

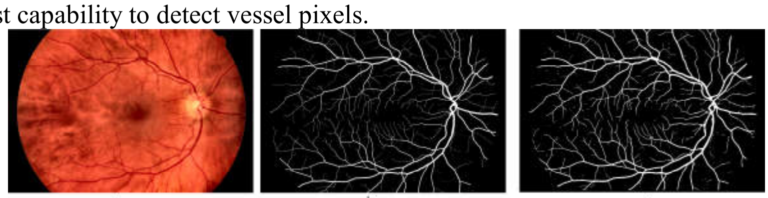


Fig.3: (a) Retinal images, (b) Ground truth, (c) Segmented results.

| Methods | [8] | [9] | [10] | [3] | [11] | [1] | [5] | U-net [7] | our model |
|---|---|---|---|---|---|---|---|---|---|
| Acc(%) | 95.57 | 96.5 | 96.54 | 96.51 | 96.37 | 92.44 | 96.4 | 95.77 | 96.75 |
| Sens(%) | 70.54 | 80.10 | 78.03 | 74.64 | 80.37 | 65.77 | 80.3 | 56.29 | 84.5 |
| Spec(%) | 98.34 | 80.10 | 98.43 | 98.74 | 97.96 | 97.99 | 96.37 | 97.25 | 97.86 |

Table 1: 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 difference in image resolution, the proposal of adjusting the convolution kernel size may be applied on other databases following the same principle. This approach offers the possibility of applying the proposed technique across various datasets, 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 retinal components [12] or pathologies such hypertensive retinopathy [13], cataract [14], diabetic retinopathy [15] and aged macular degeneration [16].

## References

[1] 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

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] X. Ding, X. Zhang, Y. Zhou, J. Han, G. Ding, and J. Sun, 'Scaling Up Your Kernels to 31x31: Revisiting Large Kernel Design in CNNs'. arXiv, Apr. 02, 2022.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] S. Ben Sayadia, Y. Elloumi, M. Akil, M. H. Bedoui, "Computational Efficiency of Optic Disk Detection on Fundus Image: A survey", SPIE Proceeding on Real-Time Image and Video Processing 2018, May 2018, Orlando, Florida (United States).

[13] M. Akil, Y. Elloumi, R. Kachouri, "Detection Of Retinal Abnormalities In Fundus Image Using Cnn Deep Learning Networks", Book Title: State Of The Art In Neural Networks, Volume Number: Vol. 1, Publisher: Elsevier, 2020.

[14] Y. Elloumi, "Cataract Grading Method Based on Deep Convolutional Neural Networks and Stacking Ensemble Learning", Int. Journal of Imaging Systems and Technology, Tome 32, N° 3, Mai 2022, Pages 798-814.

[15] Y. Elloumi, M. Ben Mbarek, R. Boukadida, M. Akil, M. H. Bedoui, "Fast and accurate mobile-aided screening system of moderate diabetic retinopathy", 13th conference on machine vision, 116050U, 2021.

[16] S. Ben Sayadia, Y. Elloumi, R. Kachouri, M. Akil, A. Ben Abdallah, M. H. Bedoui, "Automated Method For Real-Time Amd Screening Of Fundus Images Dedicated For Mobile Devices", Medical & Biological Engineering & Computing, 60, 1449–1479 (2022).

[17] I. Assali, A. Ghazi Blaiech, A. Ben Abdallah, K. Ben Khalifa, M. Carrère, Mohamed Hédi Bedoui, "CNN-based classification of epileptic states for seizure prediction using combined temporal and spectral features", Volume 82, 2023.

[18] A. Bouden, A. Ghazi Blaiech, K. Ben Khalifa, A. Ben Abdallah, M. Hédi Bedoui, A novel deep learning model for COVID-19 detection from combined heterogeneous X-ray and CT chest images, Artificial Intelligence in Medicine: 19th International Conference on Artificial Intelligence in Medicine, AIME 2021, Virtual Event, June 15–18, 2021, pp. 378-383.