MULTI-LABEL CLASSIFICATION SCHEME BASED ON LOCAL REGRESSION FOR RETINAL VESSEL SEGMENTATION

Qi He[1,3], *Beiji Zou*[1,3], *Chengzhang Zhu*[2,3], *Xiyao Liu*[1,3], *Hongpu Fu*[1,3], *Lei Wang*[1,3]

[1] School of Information Science and Engineering, Central South University, Changsha, China
[2] College of Literature and Journalism, Central South University, Changsha, China
[3] Joint Laboratory of Mobile Health, Ministry of Education and China Mobile, Changsha, China

ABSTRACT

The segmentation of small blood vessels whose width is less than 2 pixels in retinal images is a challenging problem. Existed methods rarely focus on the differences between small vessels and big vessels when doing segmentation. Therefore, previous methods are not accurate enough on small blood vessel segmentation. To effectively segment small blood vessels in retinal images including big vessels, we proposed a novel multi-label classification scheme for retinal vessel segmentation. In our proposed scheme, a local de-regression model is designed for multi-labeling and a convolutional neural network is used for multi-label classification. At addition, a local regression method is utilized to transform multi-label into binary label for locating small vessels. The experimental results show that our method achieves prominent performance for automatic retinal vessel segmentation, especially for small blood vessels.

Index Terms— neural network, retinal vessel segmentation, regression, multi-label classification

1. INTRODUCTION

The precise segmentation of retinal vessel serves as an important cue for diagnosis and evaluation of various cardiovascular and ophthalmologic disorders such as diabetes, hypertension and choroidal neovascularization [1]. In addition, retinal vessel is found to be unique for each individual and thus its segmentation result can be used for biometric identification [2]. However, it is tedious and time consuming to segment retinal vessel manually, especially in fundus images. Automatic and reliable segmentation of retinal vessels is a long-standing topic for decades with challenges on variation and diversity of vessel size, shape and surrounding [3].

Existed retinal vessel segmentation techniques can be classified into two main categories: Rule-based or Machine learning-based. The rule-based methods usually involve systems based on vessel tracking [4], adaptive thresholding [5], slice marching [6], and line detectors [7], etc. The machine learning-based methods, which usually require extra hand-labeled images for model learning, outperform most rule-based methods [3], and can be further divided into traditional methods using hand-craft features [8–13] and deep learning methods using deep neural network [14–21]. Recently, the convolutional neural network (CNN) significantly outperform other state-of-the-art methods due to their highly discriminative representations. In 2016, two variants [17, 18] of holistically-nested edge detection method (HED) [16] using the pre-trained parameters [22] are proposed for retinal vessel segmentation. But several methods [19–21], which rely on the strong data augmentation to use the available annotated samples more efficiently, also achieved prominent performances. Most of previous efforts only involved finding better classifier or better network structure. They formulate this task into a binary classification problem, but they rarely consider the differences between small vessels and big vessels as shown in Fig. 1. Therefore, their methods are still not accurate enough for small vessels.

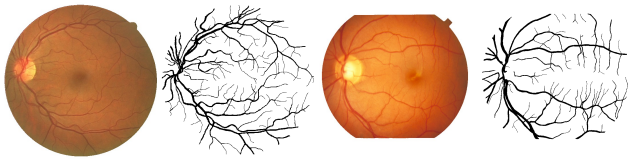

Fig. 1. Example of retinal fundus images and corresponding manual ground truth on DRIVE (left) and STARE (right).

To address this issue, the problem of retinal blood vessel segmentation is reformulated to a multi-label classification task. In this manner, more discriminative representations of small vessel can be captured. To our best knowledge, taking vessel segmentation as multi-label classification rather than binary classification is the first time considered in this paper. Unfortunately, the existed public databases [9, 18] only provide binary classification labels. Therefore, firstly, a local de-regression (LODESS) based model is proposed for labeling. Secondly, a CNN-based classifier is designed for multi-label classification and learning the neighborhood relations of multi-label. But multi-label classification is faced with imbalanced samples. For instance, the area of big ves-

This work was supported by the National Natural Science Foundation of China (No.61573380, No. 61702559) and the Planned Science and Technology Project of Hunan Province, China(No.2017WK2074).

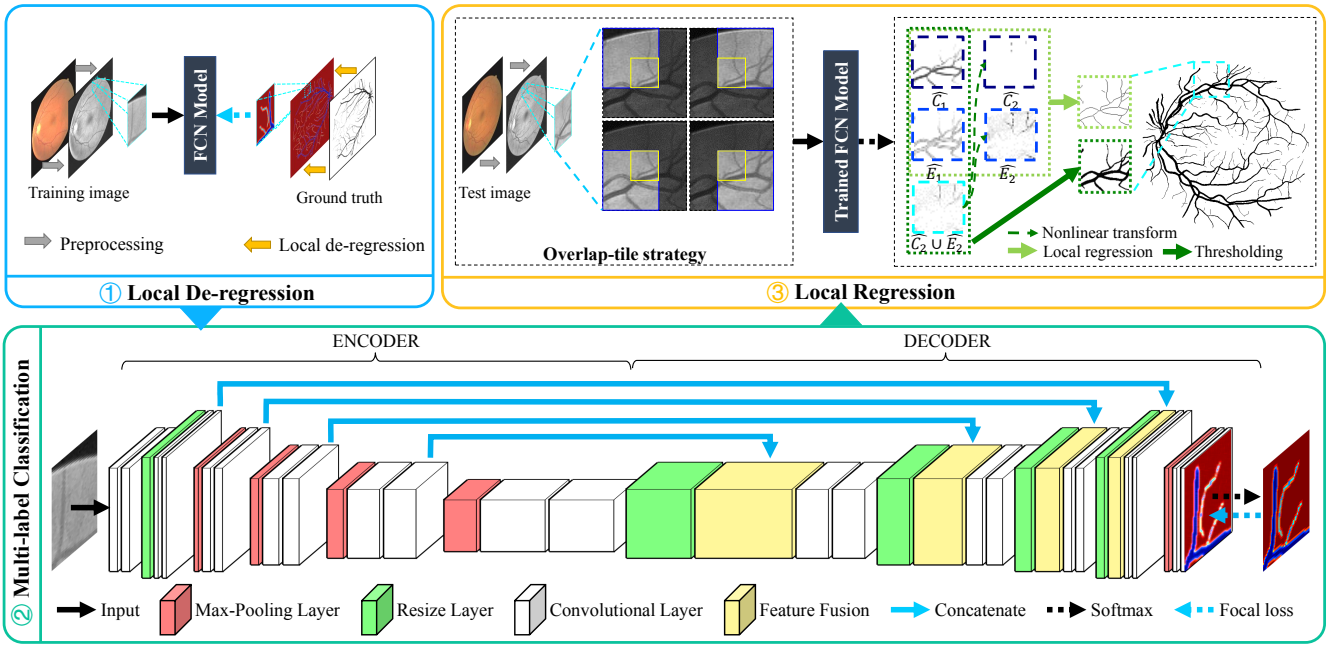

Fig. 2. Overall architecture of our method.

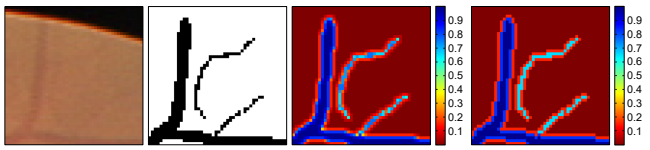

Fig. 3. Example of LODESS (from left to right: original image, ground truth, LODESS, and simplified LODESS).

sels is much larger than the small vessels. Therefore, some dedicated strategies are introduced to solve the problem of imbalanced classification. Finally, the multi-label probability map output by CNN converts into binary label for vessel segmentation. Thus, a local regression (LOESS) based method is utilized for the label simplification.

2. PROPOSED METHOD

Segmentation of small blood vessel for the fundus image is particularly difficult. For such a challenging task, we propose a local regression-based convolutional neural network as shown in Fig. 2, whose three phases are described as follows: *1) local de-regression based model for labeling, 2) convolutional neural network for multi-label classification, 3) local regression based label simplification.*

2.1. Local de-regression for labeling

The local de-regression (LODESS) is designed to address the binary label scenario in which end-to-end method is not accurate enough for small vessel detection. We introduce the LODESS starting from the manual ground truth y of binary classification:

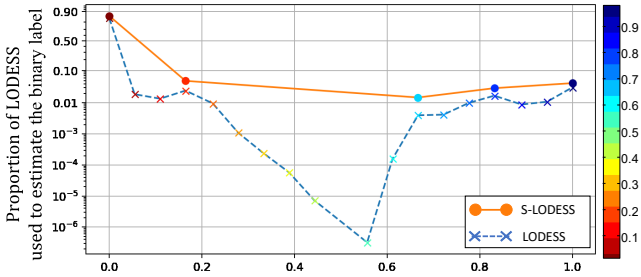

Fig. 4. Distribution of LODESS in the training set (S-LODESS: simplified LODESS labels, LODESS: real LODESS labels).

$$y = \begin{cases} 1 & \text{if } x \text{ is vessel} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In the above $y \in \{0, 1\}$ specifies the ground truth class for notational convenience, we define $P(c \mid x)$:

$$P(c \mid x) = \mathbb{1}(y = c) \tag{2}$$

The LODESS which measures the binary labels' quality of deserving to be believed and trusted is shown in Fig. 3. One notable property of this feature is that pixels with the same binary label may have different LODESS. More formally, we define the LODESS as the local regression of binary label, which is described as:

$$LODESS = \begin{cases} 0.5 + \dfrac{1 + \sum_{z \in N_x} P(1 \mid z)}{2(1 + |N_x|)} & \text{if } x \text{ is vessel} \\ 0.5 - \dfrac{1 + \sum_{z \in N_x} P(0 \mid z)}{2(1 + |N_x|)} & \text{otherwise.} \end{cases} \tag{3}$$

where $z \in N_x$ is the adjacent pixel of x. As shown in Fig. 3, morphological operator can be used for the simplification of LODESS which based on the distribution of LODESS (Fig.

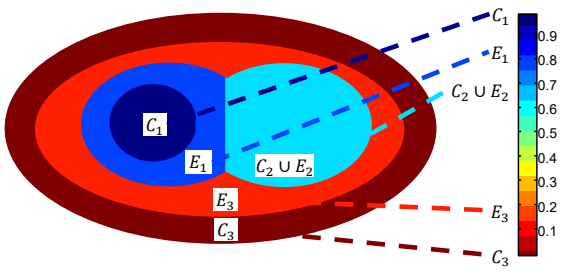

Fig. 5. Simplified mixture model of LODESS.

4) for reducing the number of labels. The statistical data shows that LODESS of background is from 0 to 1/6, LODESS of big vessel is from 5/6 to 1, and LODEESS of the rest which belongs to the small vessel is 2/3.

As a result, five primary classes of LODESS are generated. These labels are referred to the center of big vessel (C_1), the edge of big vessel (E_1), the center and edge of small vessel ($C_2 \bigcup E_2$), the center of background (C_3) and edge of background (E_3) respectively.

2.2. CNN for multi-label classification

The architecture of our proposed CNN for multi-label classification is illustrated in the second part of Fig. 2. This two-stage U-Net-like [19] CNN consisted of an encoder with convolutional and max pooling layers and a decoder with upsampling and convolutional layers. Since the information on the actual class labels of neighboring (Fig. 5) could help making better decision [20], the fully convolutional layer which replace the fully connected layers in traditional CNN is posed for structure prediction. For notational convenience, this end-to-and model is defined as:

$$A = f(X; W) \tag{4}$$

where A is the multi-label output, $f(.)$ is the function set defined by our CNN model, X is the training set, W is the parameter of $f(.)$. Specifically, features learned on the deepest stage of encoder are extremely sparse, thus our network deals well with the optic disc and pathological regions.

Moreover, this segmentation task is faced with a problem of imbalanced classification. To solve this issue, the standard cross entropy loss function is replaced with the focal loss function [23]. The focal loss is aimed to focus on the sample which is difficult to be classified by reducing the weight of easily classified samples:

$$L_{fl} = -\sum_{i=1}^{N}\sum_{j=1}^{k} \mathbb{1}(y^i_{lodess} = j)(1 - A^j_i)^{\gamma} log(A^j_i) \tag{5}$$

where N is the number of pixels in a patch and k is the number of relabeled classes. For the i-th pixel, y^i_{lodess} denotes its label and $(A^1_1, \ldots, A^j_i, \ldots, A^k_N)$ is the output prediction vector. Focusing parameter γ lower the weight of easily classified samples. With the help of this loss function, our model can also be well performed for imbalanced multi-label samples. The object function of this end-to-end model is:

$$W = \arg\min L_{fl} \tag{6}$$

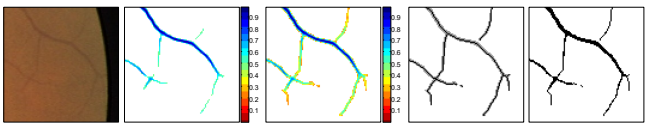

Fig. 6. Each step of refinement (from left to right: original image, LODESS, NLE, LOESS, Proposed).

2.3. Local regression for refinement

As shown in Fig. 6, the holistic flow path for refinement is considered as follows. 1) Prediction of LODESS: a multi-label result output by our CNN is used for the segmentation of big blood vessel in fundus images; 2) Nonlinear Enhancement (NLE): a nonlinear function ($g(x) = x^{0.5}$) is applied to detect small retinal vessels from the last result; 3) LOESS: a LOESS method is utilized for obtain the precise segmentation of small retinal vessels from the NLE images; 4) Proposed: LODESS and LOESS are fused into the proposed segmentation of retinal blood vessels.

Moreover, the overlap-tile strategy [19] is used for accurate prediction as illustrated in the third part of Fig. 2 (areas in the yellow rectangle are the overlapped areas of four blue rectangles). With such a strategy, our method can process images with any size. The nonlinear transform enhanced the $\hat{C}_2 \bigcup \hat{E}_2$ for better detection of small blood vessel in retinal images. After that, to accomplish the LOESS of NLE images, a skeletonization method [24] is designed to obtain the accurate positions of small vessel.

3. EXPERIMENTS

The proposed method is evaluated on the DRIVE [8] and STARE [25] databases (40 and 20 images respectively). As suggested by [20], one-off and leave-one-out strategies are utilized for the division of training set and test set.

3.1. Implementation

Our algorithm is implemented with Keras library and is trained for 20 epochs with standard SGD optimizer. Each training image is pre-processed by enhancement methods, such as graying, normalization, CLAHE and gamma correction and is resampled by 11000 (48×48 pixels) patches for data augmentation. The training time on the augmented training set is about 7 hours using a standard PC with a NVIDIA Titan X Pascal GPU.

3.2. Evaluation of each phase

To evaluate the effectiveness of different phases, we make an ablation experiment on DRIVE using its standard *field of view*

Table 1. Ablation experiment (BC: binary classification, NLE: non-linear enhancement).

Methods	ACC	SEN	SP	F1	ROC
BC	.9552	.8224	.9745	.8239	.9782
+ LODESS	.9546	.7407	.9858	.8059	-
+ NLE	.9504	.8221	.9691	.8085	-
+ LOESS	.9536	.7698	.9804	.8086	-

Table 2. Performance of different methods on DRIVE.

Methods	Year	ACC	SEN	SP	F1	ROC
DRIU [18]	2016	.9528	.8330	.9714	.8262	.9796
CNN [20]	2016	.9517	.8295	.9707	.8221	.9763
HED [16]	2015	.9462	.8009	.9688	.8002	.9699
Proposed	2018	.9519	.7761	.9792	.8129	-

Table 3. Performance of different methods on STARE.

Methods	Year	ACC	SEN	SP	F1	ROC
DRIU [18]	2016	.9669	.8641	.9792	.8489	.9902
CNN [20]	2016	.9738	.8833	.9847	.8791	.9927
HED [16]	2015	.9603	.8398	.9749	.8201	.9860
Proposed	2018	.9704	.8120	.9895	.8553	-

(FOV) [8]. An extra CNN model with the same network architecture as the proposed method is trained by binary labels and is taken as baseline for the evaluation of binary classification (BC) scheme. Each phase is compared with BC by binarizing the soft-map result images at the optimal point of *F1-measure* (or *Dice* [26]). The results of adding different parts progressively is listed in Table 1. The examples in Fig. 7 represent some typical differences in these phases for the qualitative comparison.

The results show that non-liner enhancement (NLE) and LOESS methods perform better than BC on the detection of small retinal vessels. Furthermore, LOESS provides more consist segmentation with respect to the gold standard (more accurate vessel size and shape) than NLE for the small retinal vessels. Moreover, BC is affected by the optic disc and artery (highlighted by red arrows) intuitively. On the contrary, our multi-label classification scheme (LODESS, NLE and LOESS) deals well with these regions.

3.3. Comparison with state-of-the-art methods

As listed in Table 2 and Table 3, the proposed method is compared with the current state-of-the-art for retinal vessel segmentation [18, 20] as well as a prominent edge detection method [16] on DRIVE and STARE respectively. However, each of them has different definition of FOV which might affect our comparison, because the images outside the area of FOV will be removed. To address this issue, the common area of their FOV which is approximately the area of FOV suggested by [20] is used for the quantitative comparison. To compare current state-of-the-art methods on the border of FOV, original masks of FOV are still utilized for the quantitative comparison as shown in Fig. 8.

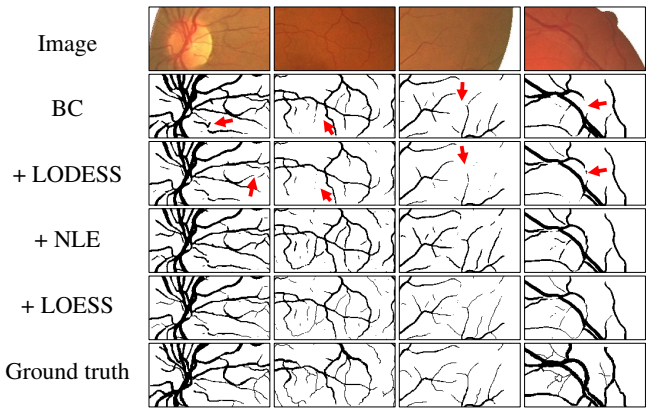

Fig. 7. Qualitative comparison of each phase on DRIVE (black: vessel, white: background).

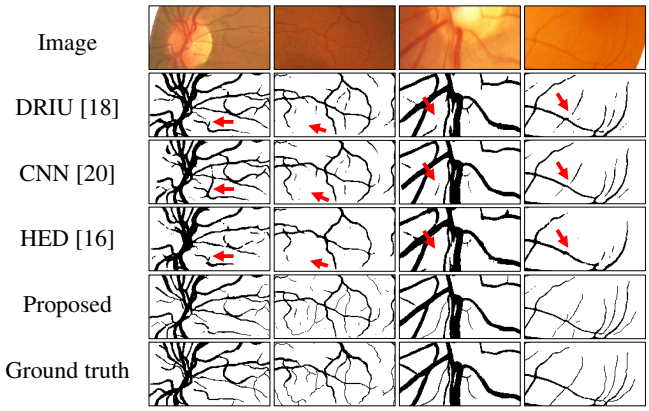

Fig. 8. Qualitative comparison of the current state-of-the-art on DRIVE (the first and second columns) and STARE (the third and fourth columns).

The results show that our method is very effective for blood vessel segmentation in retinal images, and performs better than all methods of comparison on small retinal vessels because of our multi-label classification scheme.

4. CONCLUSION

In this study, a multi-label classification scheme based on local regression is proposed for automatic retinal vessel segmentation. To effectively reformulate this segmentation problem into multi-label classification task, we proposed several dedicated methods, such as local de-regression and local-regression for the conversion between binary labels and multiple labels. We also present a FCN-based deep learning model to solve the imbalance between multi-label classes. Our experimental results have demonstrated that our method achieve prominent accuracy on the segmentation of retinal vessel in fundus images, especially for the segmentation of small blood vessels.

5. REFERENCES

[1] Jack J Kanski and Brad Bowling, *Clinical ophthalmology: a systematic approach*, Elsevier Health Sciences, 2011.

[2] Manjiri B Patwari, Ramesh R Manza, Yogesh M Rajput, Manoj Saswade, and Neha Deshpande, "Personal identification algorithm based on retinal blood vessels bifurcation," in *International Conference on Intelligent Computing Applications (ICICA)*. IEEE, 2014, pp. 203–207.

[3] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, and Sarah A Barman, "Blood vessel segmentation methodologies in retinal images–a survey," *Computer methods and programs in biomedicine*, vol. 108, no. 1, pp. 407–433, 2012.

[4] Yannis A Tolias and Stavros M Panas, "A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering," *IEEE Transactions on Medical Imaging (TMI)*, vol. 17, no. 2, pp. 263–273, 1998.

[5] Xiaoyi Jiang and Daniel Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 25, no. 1, pp. 131–137, 2003.

[6] R Sebbe, B Gosselin, E Coche, and B Macq, "Segmentation of opacified thorax vessels using model-driven active contour," in *Engineering in Medicine and Biology Society, 2005. Ieee-Embs 2005. International Conference of the*, 2005, pp. 2535–2538.

[7] George Azzopardi, Nicola Strisciuglio, Mario Vento, and Nicolai Petkov, "Trainable cosfire filters for vessel delineation with application to retinal images," *Medical image analysis*, vol. 19, no. 1, pp. 46–57, 2015.

[8] Joes Staal, Michael D Abràmoff, Meindert Niemeijer, Max A Viergever, and Bram Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging (TMI)*, vol. 23, no. 4, pp. 501–509, 2004.

[9] João VB Soares, Jorge JG Leandro, Roberto M Cesar, Herbert F Jelinek, and Michael J Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE Transactions on Medical Imaging (TMI)*, vol. 25, no. 9, pp. 1214–1222, 2006.

[10] Elisa Ricci and Renzo Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Transactions on Medical Imaging (TMI)*, vol. 26, no. 10, pp. 1357–1365, 2007.

[11] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, and Sarah A Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering (TBME)*, vol. 59, no. 9, pp. 2538–2548, 2012.

[12] Carlos Becker, Roberto Rigamonti, Vincent Lepetit, and Pascal Fua, "Supervised feature learning for curvilinear structure segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2013, pp. 526–533.

[13] Diego Marín, Arturo Aquino, Manuel Emilio Gegúndez-Arias, and José Manuel Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on Medical Imaging (TMI)*, vol. 30, no. 1, pp. 146–158, 2011.

[14] Yaroslav Ganin and Victor Lempitsky, "N^4-Fields: Neural network nearest neighbor fields for image transforms," in *Asian Conference on Computer Vision (ACCV)*. Springer, 2014, pp. 536–551.

[15] José Ignacio Orlando and Matthew Blaschko, "Learning fully-connected crfs for blood vessel segmentation in retinal images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2014, pp. 634–641.

[16] Saining Xie and Zhuowen Tu, "Holistically-nested edge detection," in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1395–1403.

[17] Huazhu Fu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, and Jiang Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2016, pp. 132–139.

[18] Kevis-Kokitsi Maninis, Jordi Pont-Tuset, Pablo Arbeláez, and Luc Van Gool, "Deep retinal image understanding," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2016, pp. 140–148.

[19] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2015, pp. 234–241.

[20] Paweł Liskowski and Krzysztof Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Transactions on Medical Imaging (TMI)*, vol. 35, no. 11, pp. 2369–2380, 2016.

[21] Debapriya Maji, Anirban Santara, Pabitra Mitra, and Debdoot Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," *arXiv preprint arXiv:1603.04833*, 2016.

[22] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[23] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár, "Focal loss for dense object detection," *arXiv preprint arXiv:1708.02002*, 2017.

[24] Waleed Abu-Ain, Siti Norul Huda Sheikh Abdullah, Bilal Bataineh, Tarik Abu-Ain, and Khairuddin Omar, "Skeletonization algorithm for binary images," *Procedia Technology*, vol. 11, pp. 704–709, 2013.

[25] AD Hoover, Valentina Kouznetsova, and Michael Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, pp. 203–210, 2000.

[26] Jordi Pont-Tuset and Ferran Marques, "Supervised evaluation of image segmentation and object proposal techniques," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 38, no. 7, pp. 1465–1478, 2016.