



Arterioles and Venules Classification in Retinal Images Using Fully Convolutional Deep Neural Network

Sufian AlBadawi^(✉) and M. M. Fraz^(✉)

School of Electrical Engineering and Computer Science,
National University of Science and Technology, Islamabad, Pakistan
{sufian.badawi, moazam.fraz}@seecs.edu.pk

Abstract. The abnormalities in size, shape and other morphological attributes of retinal vasculature have been prospectively associated as a physio-marker and predictor of many microvascular, systemic and ophthalmic diseases. The progression of retinopathy has a very different evolution in venules and arterioles with some biomarkers associated with only one type of vessel. The robust classification of retinal vasculature into arteriole/venule (AV) is the first step in the development of automated system for analyzing the vasculature biomarker association with disease prognosis.

This paper presents an encoder-decoder based fully convolutional deep neural network for pixel level classification of retinal vasculature into arterioles and venules. The feature learning and inference will be done directly from the image without the requiring the segmented vasculature as a preliminary step. The complex patterns are automatically learned from the retinal image without requiring the handcrafted features. The methodology is trained and evaluated on a subset of the images collection obtained from a population-based study in the UK (EPIC Norfolk), producing 93.5% detection rate. This proposed technique will be optimized further and may replace the AV classification module in the QUARTZ software which is developed earlier by our research group.

Keywords: Computer aided diagnosis · Deep learning
Arteriole/venule classification · Fully convolutional neural networks
Semantic segmentation

1 Introduction

The analysis of shape, appearance, tortuosity and other morphological attributes of human retinal blood vessels can be the important diagnostic indicator of various ophthalmic and system disease which includes diabetic retinopathy, hypertensive retinopathy, arteriolar narrowing, arteriosclerosis and age related macular degeneration [1]. The association of abnormalities in retinal vasculature with cardiovascular disease has been reported in the studies [2]. The effect of onset of systemic and ophthalmic disease on arterioles and venules is very much different. For instance, generalized arteriolar narrowing is one among the early signatures of hypertensive retinopathy. The decrease

in Arteriole to Venule Ratio (AVR) is a well-known predictor of stroke and other cardiovascular disease in later life. Moreover, Arterio-Venous (AV) nicking is associated with long term risk of hypertension [2].

The advancement in retinal image acquisition and the availability of retinal fundus images make it possible to run the large population based screening programs to examine the early biomarkers of these diseases. Besides improving the diagnostic efficiency, the computerized retinal image analysis can help in reducing the workload of ophthalmologists. Therefore an efficient algorithm for classification of retinal vasculature into the constituent venules and arterioles is an essential part of automated diagnostic retinal image analysis system.

The arterioles/venules in the retinal images look like much similar to each other with only very few known discriminating features [3]. The venules appear to be a little bit wider than the arterioles particularly at the place closer to the optic disc. The arterioles exhibit clearer and wider center light reflex as compared to the venules. The venules appear to be a bit darker in color than arterioles. Moreover, generally the arterioles do not cross other arterioles and venules do not cross other venules within the retinal vasculature tree. The intra/inter image variability in color, contrast and illumination are further added challenges in developing automated AV classification system. The width as well as the color of retinal vessels change across their length as they originated from optic disc and spread in the retinal. The color change is due to the variability in oxygenation level.

Deep learning [4] is gaining importance in the last few years due to the ability to efficiently solve complex nonlinear classification problems. The main advantage of deep learning is the automated feature learning from the raw data. The convolutional neural network (CNN) [5] architectures have been used for variety of image classification and detection tasks with the human level performance. The CNNs have been used to detect diabetic retinopathy in retinal images in recent Kaggle competition with very encouraging results. The promising results of CNN based architectures in retinal image analysis motivates us to investigate the application of deep learning for pixel level classification and labeling.

In this paper, we have modeled the vessel classification task as semantic segmentation. Semantic segmentation [6] refers to the pixel level understanding of an image, and each pixel in the image will be assigned to a particular object class. For vessel classification, the aim is to assign every pixel in the retinal image to either of the three classes i.e. the arteriole, the venule or the background. We have presented a CNN based architecture for pixel level classification of retinal blood vessels into arterioles and venules. The proposed methodology can perform end-to-end vessel classification directly on the retinal image without the need of separately segmenting the blood vessels, or delineating the vessel centerlines as proposed in other algorithms. To the limit of our knowledge, the deep learning-based pixel level semantic segmentation has been utilized for the first time for classifying retinal blood vessels into arterioles/venules. The proposed AV classification algorithm will replace the current AV classification module in the QUARTZ retinal image analysis software tool [7], which is developed by our research

group for quantification of retinal vessel morphology, with the aim to help epidemiologists analyze the association of retinal vessel morphometric properties with the prognosis of various systemic/ophthalmic disease biomarkers.

The rest of this paper is arranged as follows: a review of the related techniques is reported in Sect. 2. The following section provides the detailed description of proposed methodology. In Sect. 4, the experimental results are presented. The discussions and conclusion are illustrated in Sect. 5.

2 Related Work

A number of techniques are reported in the literature for arteriole/venule classification in retinal images [8]. These approaches may be categorized into two major groups; the graph based approaches and the feature based approaches.

The feature based approaches prepare set of features for each pixel that eventually used as input to a classification algorithm. The first step in the majority of the approaches is segmenting the vasculature tree, followed by vessel skeletonization. The next step is the identification of bifurcations and crossovers. The complete vasculature is divided into vessel segments by removing the pixels at the crossover/bifurcation point in the vessel centerline images. The features are computed from these vessel segments which are further classified by a suitable classifier to be arteriole or venule. The graph based approaches usually represent the vasculature tree into a graph planner. The contextual information in the graph is utilized in making local decisions for a pixel whether it belongs to the arteriole or the venule.

Li [9] introduced a Gaussian filter model designed to detect vessel's center light reflex and used Minimum Mahalanobis distance classifier. However, the classification accuracy is mentioned at artery/vein level, and not at pixel level. Grisan [10] proposed the dividing the retinal image into four quadrants assuming that each of the divided regions has at least one arteriole/venule and afterwards applied fuzzy clustering. Saez [11] and Vazquez [12] improved the quadrant based approach, computed the pixel level features from RGB and HSL color spaces and utilized K-Mean clustering for AV classification. Kondarmann [13] proposed background normalization followed by computing the features of vessel centerline pixels in the 40-pixel square neighborhood and use Neural Network classifier for AV classification. Niemijar et al. [14] have computed a 27 dimensional feature vector for each pixel and classify the vasculature segments using linear discriminate classifier. Fraz [15] introduced features at different levels (pixel, segment, profile based) and use ensemble classifier for pixel level classification. Relan [16] computed the features set from the circular neighborhood around the current pixel within a specific radius and used the least square SVM classifier. Xu [17] built an innovative feature set from first and second order texture based and pass it to KNN classified for pixel classification.

Rothaus et al. [18] and Dashtbuzorg et al. [19] have built the planner graph from the vessel centerlines such that the branches and crossovers in the vascular network represent the nodes in the graph and the vessel segments represent the link between the graph nodes. The contextual information i.e. the link orientation across nodes, and the count

of links associated with each graph node is used to identify the node type. After identification of all the nodes on the graph, the links corresponding to the vessel segment can be identified as arteriole or venule. Rothaus et al. [18] also created vessel graph, initialize few vessel segments manually and employ a rule based algorithm to propagate the vessel labels across the graph. Dashtbozorg et al. [19] combine the supervised pixel classification approach with graph based methodology to obtain pixel level classification. A color information based 30-D feature vector is computed for every centerline pixel followed by linear discriminant analysis classifier. The classification results are combined with graph labeling to attain excellent results. Estrada [20] applied global likelihood model to assign the a/v label to the links.

The feature based and graph based approaches can struggle in case the vascular tree is not correctly segmented. Moreover, these approaches are heavily relying on the hand crafted features. Welikala et al. [21] have employed the deep learning for the first time in the context of AV classification, and used a six layers convolutional neural network for feature learning from the retinal vasculature. The methodology achieves significant results in terms of accuracy, but it also relies on the accurate segmentation of vessels in the retinal image. We have proposed an end-to-end pixel-level AV classification techniques based on encoder-decoder based fully convolutional neural network. The proposed technique does not rely on the segmented vasculature, rather it learns and classify the pixels directly from the image.

3 The Methodology

In this work, we have presented a fully convolutional encoder-decoder based deep neural network architecture for pixel-wise segmentation of retinal vasculature and classification of arterioles and venules simultaneously. The proposed network architecture takes inspiration from SegNet [22] and perform the semantic segmentation of retinal images by associating each pixel of an image with a class label, i.e. background, arteriole or venule, without performing retinal vasculature segmentation separately, which usually had been a preliminary step in the traditional computer vision based AV classification approaches.

The network is composed of convolutional layers without any fully-connected layers which are usually found at the end of the traditional CNN. The encoder-decoder based fully convolutional neural networks take the input of arbitrary size and produce correspondingly-sized output. The feature learning and inferencing is performed as a whole-image-at-a-time basis by dense feedforward computation and backpropagation.

The encoder part of the network takes an input image and generates a high-dimensional by learning the features at multiple abstractions and aggregating the features at multiple levels. The decoder part of the network takes a high dimensional feature vector and generates a semantic segmentation mask. The building blocks of the network are convolutional layers, down-sampling and up-sampling. The learning is performed within subsampled layers using stride convolutions and max pooling. The up sampling layers in the network enable pixel wise prediction by applying unpooling and deconvolutions.

3.1 The Deep Network Architecture

The architecture consists of a sequence of encoder-decoder pairs which are used to create feature maps followed by pixel wise classification. The encoder-decoder architecture is illustrated in Fig. 1. The complete network consists of three layers of encodes-decoder blocks as shown Fig. 1(b). The input encoder block and output decoder block is presented in Fig. 1(a) and Fig. 1(c) respectively.

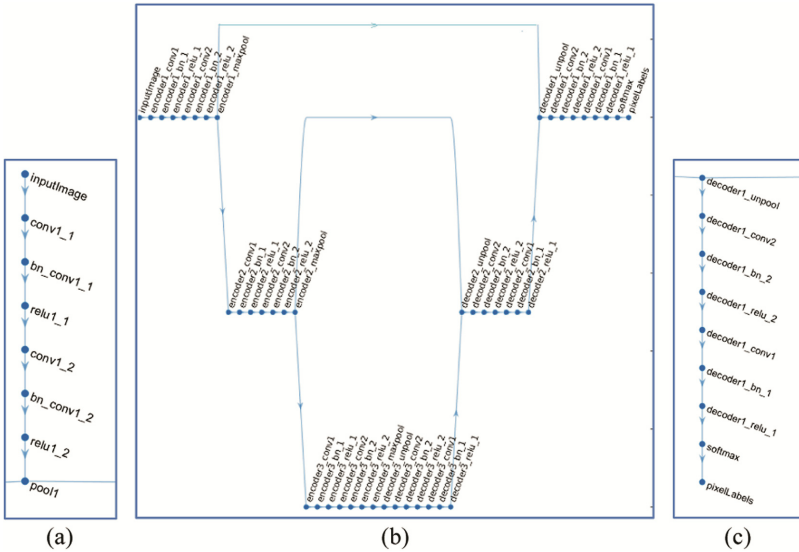


Fig. 1. The network architecture; (a) The input block of encoder part; (b) The complete network diagram; (c) The output block of decoder part.

The encoder part of the network closely resembles the VGG16 [5] architecture with the difference that only convolution layers are retained while the fully connected layers are excluded which makes it smaller and easier to train. A set of feature maps is produced by performing convolutions with filter bank. The feature map is batch normalized and element wise Rectified Linear Unit (RELU) activation is performed. Afterwards, a 2×2 max pooling with a non-overlapping stride of 2 units is applied. We have modified the architecture by reducing the number of layers to seven; hence the number of trainable parameters is also reduced.

The decoder part is comprised of nonlinear up sampling and convolution layers. The feature map is up sampled by the decoder network by utilizing the maxpooling indices which have been computed from the corresponding encoder phase. The pooling indices of encoder and decoder parts are connected with each other hence incorporating the capability to retain high frequency details in the neural network. As a result of up sampling, sparse feature maps are produced. Afterwards, dense feature-maps are generated by convolving previously generated sparse feature maps with a trainable filter bank. Softmax classifier is applied after the restoration of feature maps to the original

resolution. The softmax performs independent classification of each pixel as arteriole, venule or background and produces the final multiclass segmentation.

3.2 Learning Details

The methodology is evaluated on a dataset of 100 images, such that 90 images are used for training and 10 images are used for testing. The available pertained models which include AlexNet, VGG and ResNet are trained on PASCAL VOC [23] or ImageNet [24]. These datasets are very much different than that of retinal images therefore the pre-trained weights are not used. The Stochastic Gradient Descent (SGD) is used to train all the network. The learning rate fixed at 0.1 and a mini-batch of 12 images is used for training.

4 Experimental Evaluation

4.1 Materials

The methodology is evaluated on a sub set of images from EPIC Norfolk study [25]. The study was started as a large multi-center cohort with the aim to investigate the relationship among diet, lifestyle factors and cancer/other disease prognosis. The subset is comprised of 100 images acquired from 50 middle aged participants using Topcon non-mydratic fundus cameras having a size of 3000×2002 pixels. The images are captured from both of the left and right eyes. The other biomarkers are also recorded which includes weight, BMI and family history of diabetes and hypertension. The vessels are manually labeled by two experts using image labeler application available with Matlab R2017b. The labels are verified by the ophthalmologists at St Georges University of London UK.

4.2 Performance Measures

The performance measures used to quantitatively evaluate the algorithm performance are summarized in Table 1.

Table 1. Performance metrics for vessel classification

Metric	Description
Global accuracy	It measures the percentage of pixels correctly classified in the dataset. $Global Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)}$
Mean accuracy	The fraction of correctly classified pixels averaged over the classes
Mean Intersection over Union (mIoU)	Mean IoU is used to compare the similarity between the predicted class and the actual ground truth class
Weighted Intersection over Union (wIoU)	Performs IoU average weighted by the count of pixels in each class
Mean BF score	The measure depicts that how well predicted objects boundaries match ground truth boundaries

*TP, FP, and TN and FN refers to true positive, false positive, true negative, false negative respectively

4.3 Experimental Results

The attained performance measures by the proposed methodology are summarized in Table 2.

Table 2. Performance measures of AV Classification

Global accuracy	Mean accuracy	Mean IoU	Weighted IoU	Mean BF score
0.9352	0.8228	0.7905	0.9457	0.8201

The comparison of the algorithm accuracy with previously published algorithms is shown in Table 3.

Table 3. Performance comparison with different vessel classification methods

S. No	Algorithm	Accuracy
1	Grisan [14]	87.58%
2	Kondarmann [13]	95.32%
3	Niemijar [14]	93.5%
4	Vazquez [12]	87.68%
5	Fraz [15]	83.27%
6	Relan [16]	89.4%
7	Xu [17]	92.3%
8	Dashtbuzorg [19]	87.4%
9	Estrada [20]	93.5%
10	Welikala [21]	91.97%
11	Proposed Methodology	93.5%

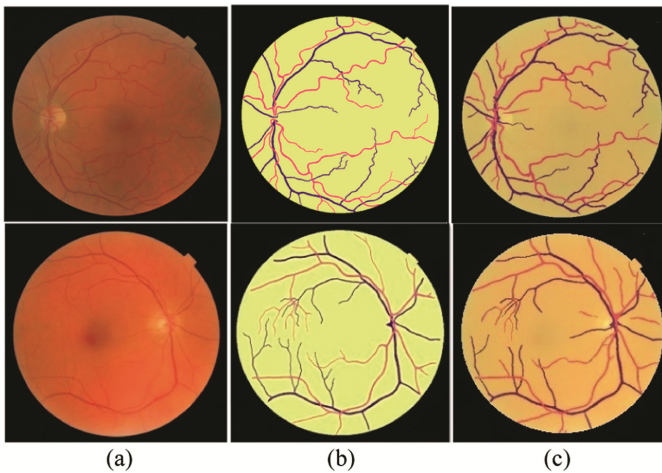


Fig. 2. Vessel Classification Results: (a) Original images (b) labels (c) segmentation results (Color figure online)

Figure 2 shows the classification results of the proposed methodology. The first column is the retinal image, the ground truth and the classification results are shown in 2nd the 3rd column respectively. The background is marked with yellow color and the arterioles and venules are marked with red and blue color respectively.

5 Discussion and Conclusion

In this paper a novel deep learning based methodology for AV classification of retinal blood vessels is presented. An encoder-decoder based deep convolutional neural network is proposed for pixel level classification of retinal vessels into the arterioles and venules. The methodology does not rely on prior segmentation of retinal blood vessels, which have been the preliminary step for approximately all of the AV classification techniques available in the literature. The proposed network architecture has taken inspiration from SegNet, which is used in semantic segmentation paradigm but to the best of our knowledge, has been utilized for the first time in the context of automated AV classification.

The major contribution of this paper is the application of novel encoder-decoder based fully convolutional deep network for robust AV classification. In future we aim to extend this methodology such that it will be used in place of current AV classification module in the QUARTZ software [7], which is developed by our research group for automated quantification the retinal vessel morphometry, with the aim study associations between vessel change and systemic/ophthalmic disease prognosis. Furthermore, we aim to use the proposed methodology as a preliminary step in the development the modules in QUARTZ for identification of venous beading and measurement of arterio-venous nicking.

Acknowledgements. We thankfully appreciate the support of NVIDIA Corporation with the donation of the TitanX GPU used for this research.

References

1. Kanski, J.J., Bowling, B.: *Clinical Ophthalmology: A Systematic Approach*, 8th edn. Elsevier Health Sciences (UK), London (2015)
2. Wong, T.Y., Klein, R., Sharrett, A.R., Duncan, B.B., Couper, D.J., Tielsch, J.M., et al.: Retinal arteriolar narrowing and risk of coronary heart disease in men and women. *JAMA, J. Am. Med. Assoc.* **287**, 1153–1159 (2002). <http://jama.ama-assn.org/content/287/9/1153.abstract>
3. Fraz, M.M., Barman, S.A.: Computer vision algorithms applied to retinal vessel segmentation and quantification of vessel caliber. In: *Image Analysis and Modeling in Ophthalmology*, pp. 49–84. CRC Press (2014)
4. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436 (2015)
5. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. *CoRR*, vol. abs/1409.1556 (2014)
6. Garcia-Garcia, A., Orts, S., Oprea, S., Villena-Martinez, V., Rodríguez, J.G.: A Review on Deep Learning Techniques Applied to Semantic Segmentation. *CoRR*, vol. abs/1704.06857 (2017)

7. Fraz, M.M., Welikala, R.A., Rudnicka, A.R., Owen, C.G., Strachan, D.P., Barman, S.A.: QUARTZ: quantitative analysis of retinal vessel topology and size – an automated system for quantification of retinal vessels morphology. *Expert Syst. Appl.* **42**, 7221–7234 (2015)
8. Miri, M., Amini, Z., Rabbani, H., Kafieh, R.: A comprehensive study of retinal vessel classification methods in fundus images. *J. Med. Sig. Sens.* **7**, 59–70 (2017)
9. Li, H., Hsu, W., Lee, M.L., Wang, H.: A piecewise gaussian model for profiling and differentiating retinal vessels. In: *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, vol. 1, pp. I-1069-72 (2003)
10. Grisan, E., Ruggeri, A.: A divide et impera strategy for automatic classification of retinal vessels into arteries and veins. In: *Engineering in Medicine and Biology Society (2003) Proceedings of the 25th Annual International Conference of the IEEE*, vol. 1, pp. 890–893 (2003)
11. Saez, M., González-Vázquez, S., González-Penedo, M., Barceló, M.A., Pena-Seijo, M., Coll de Tuero, G., et al.: Development of an automated system to classify retinal vessels into arteries and veins. *Comput. Methods Programs Biomed.* **108**, 367–376 (2012)
12. Vázquez, S.G., Cancela, B., Barreira, N., Penedo, M.G., Rodríguez-Blanco, M., Seijo, M.P., et al.: Improving retinal artery and vein classification by means of a minimal path approach. *Mach. Vis. Appl.* **24**, 919–930 (2013)
13. Kondermann, D., Kondermann, C., Yan, M.: Blood vessel classification into arteries and veins in retinal images. In: *Medical Imaging*, p. 651247 (2007)
14. Niemeijer, M., Xu, X., Dumitrescu, A.V., Gupta, P., Ginneken, B.V., Folk, J.C., et al.: Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. *IEEE Trans. Med. Imaging* **30**, 1941–1950 (2011)
15. Fraz, M.M., Rudnicka, A.R., Owen, C.G., Strachan, D.P., Barman, S.A.: Automated arteriole and venule recognition in retinal images using Ensemble classification. In: *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, pp. 194–202 (2014)
16. Relan, D., MacGillivray, T., Ballerini, L., Trucco, E.: Automatic retinal vessel classification using a least square-support vector machine in VAMPIRE. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 142–145 (2014)
17. Xu, X., Ding, W., Abràmoff, M.D., Cao, R.: An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image. *Comput. Methods Programs Biomed.* **141**, 3–9 (2017)
18. Rothaus, K., Jiang, X., Rhiem, P.: Separation of the retinal vascular graph in arteries and veins based upon structural knowledge. *Image Vis. Comput.* **27**, 864–875 (2009)
19. Dashtbozorg, B., Mendonça, A.M., Campilho, A.: An automatic graph-based approach for artery/vein classification in retinal images. *IEEE Trans. Image Process.* **23**, 1073–1083 (2014)
20. Estrada, R., Allingham, M.J., Mettu, P.S., Cousins, S.W., Tomasi, C., Farsiu, S.: Retinal artery-vein classification via topology estimation. *IEEE Trans. Med. Imaging* **34**, 2518–2534 (2015)
21. Welikala, R.A., Foster, P.J., Whincup, P.H., Rudnicka, A.R., Owen, C.G., Strachan, D.P., et al.: Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort. *Comput. Biol. Med.* **90**, 23–32 (2017)
22. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 2481–2495 (2017)
23. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. *Int. J. Comput. Vis.* **88**, 303–338 (2010)

24. Shuhan, C., Ben, W., Jindong, L., Xuelong, H.: Semantic image segmentation using region-based object detector. In: 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), pp. 505–510 (2017)
25. EPIC-Norfolk.: European Prospective Investigation of Cancer (EPIC), February 2013. <http://www.srl.cam.ac.uk/epic/>