



# Deep Convolutional Artery/Vein Classification of Retinal Vessels

Maria Ines Meyer<sup>1</sup>(✉), Adrian Galdran<sup>1</sup>, Pedro Costa<sup>1</sup>,  
Ana Maria Mendonça<sup>1,2</sup>, and Aurélio Campilho<sup>1,2</sup>

<sup>1</sup> INESC-TEC - Institute for Systems and Computer Engineering,  
Technology and Science, Porto, Portugal

{maria.i.meyer,adrian.galdran}@inesctec.pt

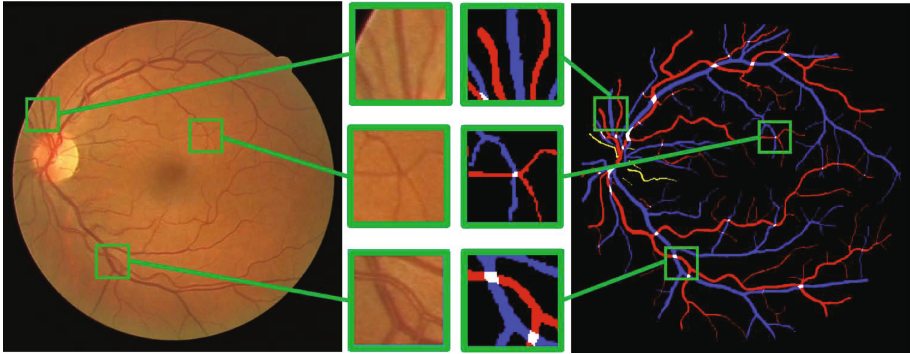
<sup>2</sup> Faculdade de Engenharia da Universidade do Porto, Porto, Portugal  
{amendon,campilho}@fe.up.pt

**Abstract.** The classification of retinal vessels into arteries and veins in eye fundus images is a relevant task for the automatic assessment of vascular changes. This paper presents a new approach to solve this problem by means of a Fully-Connected Convolutional Neural Network that is specifically adapted for artery/vein classification. For this, a loss function that focuses only on pixels belonging to the retinal vessel tree is built. The relevance of providing the model with different chromatic components of the source images is also analyzed. The performance of the proposed method is evaluated on the RITE dataset of retinal images, achieving promising results, with an accuracy of 96% on large caliber vessels, and an overall accuracy of 84%.

## 1 Introduction

The retina is routinely examined in medical settings as a means of diagnosis for several different pathologies. In particular, several types of alterations in the vasculature are known to be indicative of disease. For instance, decreased arteriolar caliber may indicate coronary artery disease, while increased venular caliber is associated with diabetic retinopathy and risk of stroke [11]. To assess these alterations, a commonly employed biomarker is the ratio between arteriolar and venular diameters (AVR) [2]. Abnormal AVR has been correlated with high cholesterol levels or high blood pressure [7]. In addition, a decreased AVR is an early sign of Diabetic Retinopathy. This is caused by a pathological widening of the veins produced by retinal hypoxia, arising after retinal microvascular degradation due to high blood sugar levels [5]. The early detection of such symptoms can increase the chance of disease detection and enable early adoption of preventive treatments to avoid vision loss.

To examine the retinal vessels, ophthalmologists often rely on images of the retina acquired by fundus cameras. However, objective visual analysis of the retinal vasculature on these images is a time consuming task that requires expert knowledge. For this reason, in recent years a large number of techniques have



**Fig. 1.** Left: Retinal image. Right: Expert A/V annotations, encoded as: blue: veins; red: arteries, white: crossings; yellow: uncertain labels. Center: Zoomed regions illustrating the main visual features and challenges related to the A/V problem. Top Row: Arteries show brighter intensities. Center Row: Fine veins and arteries are hardly distinguishable. Bottom Row: Central reflex effect. Reflex size is slightly larger for arteries. (Color figure online)

been proposed to automate several retinal image analysis tasks, *e.g.*, vessel segmentation [10], optic disk localization [3], or detection of markers for disease [1]. The classification of retinal vessels into arteries or veins is another relevant challenge, since this differentiation is necessary for measuring AVR. An automated approach to artery/vein (A/V) classification could lead to a systematic approach for AVR estimation, which would be a very useful tool for ophthalmologists (Fig. 1).

Existing techniques can be grouped into graph-based and dense methods. Graph-based methods reduce the segmented vasculature to a linear skeleton, further divided into separate edges, considering junction and ending points as nodes, and building an undirected graph on which geometry and connectivity can be more easily analyzed [4, 5, 14]. Dense methods assign an artery or vein label to each pixel in a pre-existing segmentation of the retinal vessel tree. This classification is achieved by supplying to a machine learning model representations of artery and vein pixels in terms of visual descriptors that model differences between both classes. These descriptors attempt to capture relevant color, size, or geometrical information [9, 12, 17, 18]. A notable exception are recent deep-learning based approaches, which automatically learn the most useful representation [16].

In this work, we propose to solve the A/V classification task by means of a Deep Convolutional Neural Network. This allows us to avoid generating an initial skeletonized vessel map representation, and directly classify every pixel in a vessel. We bypass the implementation of cumbersome tracking mechanisms to propagate predictions along vessels, and other type of post-processing. We also provide an experimental analysis of the influence of each particular color

component on the resulting model, which allows us to discard non-relevant information and leads to a more efficient training.

## 2 A/V Classification Method

The proposed method builds on a Convolutional Neural Network specifically designed to deal with the classification of pixels lying on one-dimensional structures (vessels), which are embedded in a two-dimensional space (the background of the retinal image). Below we detail the selected architecture and a suitable modification of a standard pixel-wise loss function to take into account this scenario.

### 2.1 Adapting CNNs for A/V Classification

Over the past few years Convolutional Neural Networks (CNNs) have achieved remarkable success in medical image analysis tasks. When applied for the task of image segmentation, we are interested in assigning a prediction to each pixel on the image. CNNs can be reformulated to perform image segmentation based on the notions of *Fully-Convolutional Neural Networks* (F-CNN) and *skip connections*, which allow to link coarse and fine layers of a CNN.

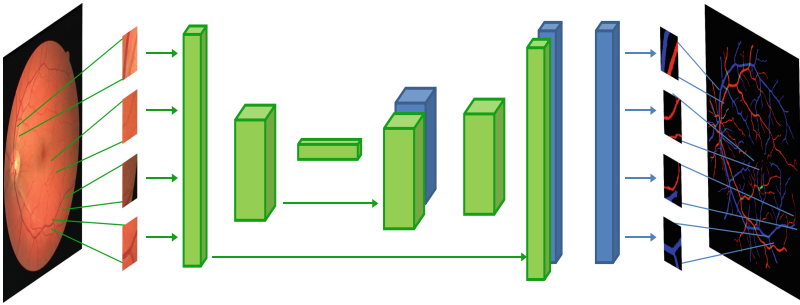
A popular F-CNN architecture is U-net, introduced in [13] for the task of segmenting biological tissue images. In this architecture, the upsampling of the feature map is done symmetrically to the regular contracting CNN section. Specifically, the output feature map of the last layer of the contracting path is upsampled so that it has the same dimension as the second last layer. The resulting feature map is fused with the feature map from the corresponding layer in the contracting section. This procedure is repeated until the output of the upsampling operation has the same dimension as the input of the network. This results in a U-shaped network, where the output feature maps of the layers in the contracting section are fused with the output from the upsampling operations.

The problem of classifying retinal vessels can also be formulated as a pixel-wise classification problem, suitable for the application of a U-net-like network. Nevertheless, solving a three-class segmentation in the context of retinal images is a complex challenge that remains unsolved [5]. As such, in this work we focus on the simpler problem of classifying pixels belonging to retinal vessels, while ignoring the task of separating the vasculature from the background.

In its common formulation the convolution operation is not suitable for the task of vessel pixel classification, since the input of a CNN must be a rectangular image region and all of its pixels must be classified. Accordingly, a U-net-like model returns a prediction for each pixel, and a third class accounting for background pixels needs to be considered. The solution adopted in this paper for solving the two class A/V problem consists on weighting the penalization of pixels belonging to different classes. For that, we propose to modify the cross-entropy loss function, typically used for classification problems, as follows:

$$L_i = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^3 w_j y_{ij} \log(\hat{y}_{ij}), \quad w_j = \begin{cases} 0, & j = 1 \\ 1, & j = 2, 3 \end{cases} \quad (1)$$

where  $\hat{y}_{ij}$  is the prediction associated to a pixel  $i$  with ground-truth label given by  $y_{ij}$ . In this case,  $j \in \{1, 2, 3\}$  models each of the possible classes occurring in a retinal image, *i.e.* background ( $j = 1$ ), artery ( $j = 2$ ), and vein ( $j = 3$ ), and  $w_n$  is the weight attributed to each class. In order to obtain classifications only for the vessel tree, the background class is assigned a weight  $w = 0$  at training time. With this formulation, Eq. (1) returns a score penalizing only incorrect classification of vessel pixels, and the error is backpropagated through the network (Fig. 2).



**Fig. 2.** Proposed U-Net-like architecture adapted for A/V classification

**Implementation and Training.** In our implementation of the U-net architecture, each convolution layer uses a stride of two, and is followed by Rectified Linear Unit activations, batch normalization and Dropout. The feature maps between the shallow contracting layers and the deep expanding layers are fused by a concatenation operation. The final layer is a convolution followed by a softmax activation in order to map each feature vector onto the classes of interest.

The model is optimized by minimizing the pixel mis-classification error, according to the cross-entropy function given by Eq. (1). The minimization follows stochastic mini-batch gradient descent, with the gradients computed by standard backpropagation. The Adam algorithm [8] is used for optimization, with learning rate  $1 \times 10^{-2}$ . The loss was monitored on the validation set, and training was stopped when it did not decrease anymore. Training took approximately 20h on a NVIDIA GeForce GTX 1080 GPU. After training, the model can generate a prediction for 90  $64 \times 64$  pixels patches (the minimum number of patches needed to create a reconstruction of the original image) in approximately 10.4s.

### 3 Experimental Evaluation

#### 3.1 Data and Evaluation

The method was trained and evaluated on the RITE dataset [6], which contains A/V ground-truth of the retinal images from the publicly available DRIVE dataset [15]. DRIVE contains 40 images of size  $565 \times 584$ , divided into 20 training and 20 test images. Before training, 4 images were randomly selected for validation.

Images in RITE contain labels for artery, vein, crossing and uncertain. Pixels labeled as crossing or uncertain were ignored during training. The images were divided into patches of  $64 \times 64$  pixels centered on randomly selected pixels of both classes. At test time predictions were performed patch-wise, and the vessel tree was reconstructed by combining these predictions. For this, when dividing the input image into patches the coordinates corresponding to the position of the upper left corner of each patch in the original image were recorded. These patches were fed to the model, and the output probability maps were combined into a full-size image. Whenever patches overlapped, the pixel-wise probabilities were simply averaged.

When evaluating the models only pixels corresponding to the vessel tree were considered. Predictions from the same vessel segments were averaged to reach a consistent result. ROC analysis was performed, and the optimal threshold maximizing the Youden Index was selected. This threshold was used to compute Accuracy, Sensitivity and Specificity.

#### 3.2 Chromatic Component Analysis for the A/V Problem

When dealing with color images, RGB is the typical choice for most computer-aided diagnosis tasks. However, it is well-known that, in the case of retinal images, the blue channel contains large quantities of noise, whereas the green channel is the component of choice for vessel segmentation methods due to its good contrast between background and vessels. On the other hand, the red component seems to be visually relevant for the discrimination of vessel types.

In order to analyze the contribution of each color channel for classifying vessels, we take advantage of CNNs being agnostic feature extractors: the most relevant information is extracted by the model driven only by the misclassification loss, thus bypassing manual feature engineering. We conduct a simple experiment: we train seven equal models, considering different inputs for each of them. We provide to each model different combinations of chromatic components, spanning from single channels to every possible pair, and finally to the three color channels. All models are trained for 150 epochs with the same hyperparameter configuration. The results of these experiments are shown in Table 1.

As expected, when considering individual channels, the performance of the model trained on the Red component only was higher than when trained on

**Table 1.** Classification performance when considering different chromatic components.

	R	G	B	RG	RB	BG	RGB
Accuracy	0.82	0.79	0.74	0.83	<b>0.84</b>	0.82	<b>0.84</b>
AUC	0.89	0.87	0.81	0.91	<b>0.92</b>	0.89	0.91

Green, and training with the Blue channel obtained the worst performance. However, when considering combinations of two different channels, we observed that the Red channel was slightly better complemented by the Blue component than by the Green channel, although the difference was marginal. The combination of chromatic components that ignored the Red channel obtained the worst result, indicating that Red is an essential component to consider when solving the A/V classification problem. In addition, the combination of the three components did not lead to any improvement in performance when compared to employing either the Red and Green components or Red and Blue. Since considering less input channels leads to removing redundant information and a more efficient memory usage, we discarded the option of training on the entire RGB information for the next section.

### 3.3 Quantitative Comparison with Previous Approaches

Since the difference in performance between using RG and RB chromatic component combinations was not significant, we decided to select both input types and train two models until full convergence. In order to test our approach against other A/V classification techniques, we present in Table 2 a comparison of our method and the methods introduced in [4, 12] for the same dataset. Both methods originally reported performance when classifying only major vessels with a caliber greater than 3 pixels. For a fair comparison, we also computed the performance of the technique introduced in this paper on these vessels.

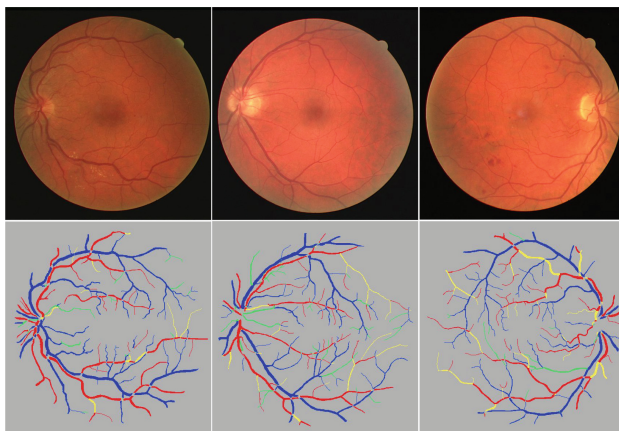
## 4 Discussion and Future Work

The proposed model achieves a competitive performance in terms of accuracy when compared with other approaches. It is also interesting to note that the model trained on Red-Blue components achieved similar AUC, but greater specificity than the one trained on Red-Green channels, indicating that the information in each channel may be complementary. It must be stressed that this performance was computed only for predictions on thick vessels. This is a useful

**Table 2.** Performance of the proposed models on thick vessels.

	Sensitivity	Specificity	Accuracy	AUC
Dashtbozorg et al. [4]	0.90	0.84	0.874	—
Niemejer et al. [12]	0.80	0.80	—	0.88
Proposed approach – <b>RG</b>	<b>0.96</b>	0.93	0.95	<b>0.99</b>
Proposed approach – <b>RB</b>	0.95	<b>0.96</b>	<b>0.96</b>	<b>0.99</b>

scenario, since for AVR measurements only large arcades are typically considered. However, when predictions are computed on the entire vessel tree, the performance drops significantly, as shown in Table 1. In Fig. 3 we show the predictions for all vessel pixels. The displayed predictions correspond to the best, the average, and the worst results. We can observe how, while the classification of the main arcades is mostly correct, thin vessels are wrongly predicted in a substantial amount. The performance of other recent methods, when predicting the entire vessel tree, is greater than the proposed model [5, 16], indicating that there is a wide margin for improvement. Accordingly, the next steps will consist of designing domain-related regularizing mechanisms for the presented deep CNN, mainly taking into account geometrical constraints on the distribution of arteries and veins within the retina.



**Fig. 3.** Left: Best classification result (RB); Middle: Average classification result; Right: Worse classification result. Color code: Red, correctly predicted arteries. Blue, correctly predicted veins. Green, wrongly predicted veins. Yellow, wrongly predicted arteries. (Color figure online)

## 5 Conclusions

We have introduced a novel approach for classification of retinal vessels into arteries and veins. The method builds on a Deep Neural Network, and the loss function that drives the model learning is suitably modified to account for vessel pixels only, while ignoring the background. The method achieves promising results in the task of artery/vein classification, with high performance on thick vessels.

**Acknowledgments.** This work is funded by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and the European Regional Development Fund (ERDF), within the project “NanoSTIMA:



Macro-to-Nano Human Sensing: Towards Integrated Multimodal Health Monitoring and Analytics/NORTE-01-0145-FEDER-000016”.

## References

1. Costa, P., Campilho, A.: Convolutional bag of words for diabetic retinopathy detection from eye fundus images. *IPSN Trans. Comput. Vis. Appl.* **9**(1), 10 (2017)
2. Dashtbozorg, B., Mendonça, A.M., Campilho, A.: An automatic method for the estimation of arteriolar-to-venular ratio in retinal images. In: *IEEE International Symposium on Computer-Based Medical Systems*, pp. 512–513, June 2013
3. Dashtbozorg, B., Mendonça, A.M., Campilho, A.: Optic disc segmentation using the sliding band filter. *Comput. Biol. Med.* **56**, 1–12 (2015)
4. Dashtbozorg, B., Mendonça, A.M., Campilho, A.: An automatic graph-based approach for artery/vein classification in retinal images. *IEEE Trans. Image Process.* **23**(3), 1073–1083 (2014)
5. 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**(12), 2518–2534 (2015)
6. Hu, Q., Abramoff, M.D., Garvin, M.K.: Automated separation of binary overlapping trees in low-contrast color retinal images. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *MICCAI 2013. LNCS*, vol. 8150, pp. 436–443. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-40763-5\\_54](https://doi.org/10.1007/978-3-642-40763-5_54)
7. Ikram, M.K., de Jong, F.J., Vingerling, J.R., Witteman, J.C.M., Hofman, A., Breteler, M.M.B., de Jong, P.T.V.M.: Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? The rotterdam study. *Invest. Ophthalmol. Vis. Sci.* **45**(7), 2129–2134 (2004)
8. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: *International Conference on Learning Representations*, pp. 1–13 (2014)
9. Kondermann, C., Kondermann, D., Yan, M.: Blood vessel classification into arteries and veins in retinal images. In: *Image Processing*, p. 651247 (2007)
10. Meyer, M.I., Costa, P., Galdran, A., Mendonça, A.M., Campilho, A.: A deep neural network for vessel segmentation of scanning laser ophthalmoscopy images. In: Karray, F., Campilho, A., Cheriet, F. (eds.) *ICIAR 2017. LNCS*, vol. 10317, pp. 507–515. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-59876-5\\_56](https://doi.org/10.1007/978-3-319-59876-5_56)
11. Nguyen, T.T., Wong, T.Y.: Retinal vascular changes and diabetic retinopathy. *Curr. Diabetes Rep.* **9**(4), 277–283 (2009)
12. Niemeijer, M., van Ginneken, B., Abramoff, M.D.: Automatic classification of retinal vessels into arteries and veins. In: *SPIE medical imaging. International Society for Optics and Photonics*, November 2016, vol. 7260, pp. 1–8 (2009)
13. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
14. Rothaus, K., Jiang, X., Rhiem, P.: Separation of the retinal vascular graph in arteries and veins based upon structural knowledge. *Image Vis. Comput.* **27**(7), 864–875 (2009)
15. Staal, J.J., Abramoff, M.D., Niemeijer, M., Viergever, M.A., Ginneken, B.V.: Ridge based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* **23**(4), 501–509 (2004)



16. Welikala, R., Foster, P., Whincup, P., Rudnicka, A., Owen, C., Strachan, D., Barman, S.: Automated arteriole and venule classification using deep learning for retinal images from the UK biobank cohort. *Comput. Biol. Med.* **90**, 23–32 (2017)
17. Yu, H., Barriga, S., Agurto, C., Nemeth, S., Bauman, W., Soliz, P.: Automated Retinal Vessel Type Classification in Color Fundus Images, vol. 8670 (2013)
18. Zamperini, A., Giachetti, A., Trucco, E., Chin, K.S.: Effective features for artery-vein classification in digital fundus images. In: *Proceedings - IEEE Symposium on Computer-Based Medical Systems* (2012)