






AV-Net: deep learning for fully automated artery-vein classification in optical coherence tomography angiography

MINHAJ ALAM,^{1,3}  DAVID LE,^{1,3}  TAEYOON SON,¹ JENNIFER I. LIM,^{1,2} AND XINCHENG YAO^{1,2,*} 

¹*Department of Bioengineering, University of Illinois at Chicago, Chicago, IL 60607, USA*

²*Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL 60612, USA*

³*These authors contributed equally to this work*

**xcy@uic.edu*

Abstract: This study is to demonstrate deep learning for automated artery-vein (AV) classification in optical coherence tomography angiography (OCTA). The AV-Net, a fully convolutional network (FCN) based on modified U-shaped CNN architecture, incorporates enface OCT and OCTA to differentiate arteries and veins. For the multi-modal training process, the enface OCT works as a near infrared fundus image to provide vessel intensity profiles, and the OCTA contains blood flow strength and vessel geometry features. A transfer learning process is also integrated to compensate for the limitation of available dataset size of OCTA, which is a relatively new imaging modality. By providing an average accuracy of 86.75%, the AV-Net promises a fully automated platform to foster clinical deployment of differential AV analysis in OCTA.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Early disease diagnosis and effective treatment assessment are essential to prevent vision loss. Differential artery-vein (AV) analysis can provide valuable information for disease detection and classification. It has been demonstrated to be valuable for evaluating diabetes, hypertension, stroke and cardiovascular diseases [1–3] along with common retinopathies [4,5]. Several clinical studies have evaluated AV abnormalities in different diseases. However, clinical deployment of the AV analysis for routine management of eye diseases is challenging. Most of the clinical studies relied on manual or semi-automated approaches to identify arteries and veins, which is ineffective in a clinical setting. Therefore, a fully automated platform for AV classification is important.

To date, automated AV classification has been primarily used in color fundus images acquired with traditional fundus photography [6–15], which provide limited resolution and sensitivity to reveal microvascular abnormalities associated with eye conditions [16]. Microvascular anomalies that occur at early stages of eye diseases, cannot be reliably identified in traditional fundus photography [17–19]. An alternative to traditional color fundus imaging is optical coherence tomography (OCT) and OCT angiography (OCTA). OCT and OCTA can provide depth-resolved visualization of individual retinal layers with capillary level resolution. Especially, OCTA is sensitive to identify subtle microvascular changes, and thus has been extensively explored for quantitative analysis and objective classification of retinal diseases [20–24]. Using quantitative feature analysis, we have recently demonstrated the potential of differentiating arteries and veins in OCTA [4,5,25,26]. Differential AV analysis showed improved OCTA performance to identify abnormal changes in diabetic retinopathy (DR) and sickle cell retinopathy (SCR) eyes [4,5,26]. However, clinical deployment of the AV analysis in OCTA requires an automated, simple, but robust method. A potential solution is to employ deep machine learning i.e., convolutional neural

networks (CNNs) for AV classification automatically. A fully convolutional network (FCN) can be trained with a ground truth dataset for a specific task and can be implemented on validation or testing dataset. A fully automated method is a key factor for clinical deployment of artificial intelligence (AI) based screening, diagnosis, and treatment evaluation.

In this study, we develop and validate AV-Net, an FCN based on a modified U-shaped CNN architecture, for deep learning AV classification in OCTA. A multi-modal training process involves both enface OCT and OCTA, which provide intensity and geometric profiles, respectively, for AV classification. Transfer learning is employed to compensate for the limitation of available dataset size of OCTA which is a relatively new imaging modality. By incorporating transfer learning and multi-modal training approaches, fully automated AV classification is demonstrated. The AV-Net performance is validated with manual AV ground truth maps using accuracy and intersection over union (IOU) metrics.

2. Methods

This study is in adherence to the ethical standards present in the Declaration of Helsinki and was approved by the institutional review board of the University of Illinois at Chicago (UIC).

2.1. Data acquisition

Spectral domain (SD) -enface OCT and OCTA data were acquired using an Angiovue SD-OCT device (Optovue, Fremont, CA, USA). The OCT device had a 70,000 Hz A-scan rate, $\sim 5 \mu\text{m}$ axial and $\sim 15 \mu\text{m}$ lateral resolutions. All enface OCT/OCTA images used for this study were $6 \text{ mm} \times 6 \text{ mm}$ scans; only superficial OCTA images were used. The enface OCT was generated as a maximum intensity 3D projection of the retinal slabs from internal limiting membrane to outer plexiform layer. After image reconstruction, both enface OCT and OCTA were exported from Revue software interface (Optovue) for further processing.

2.2. Model implementation

In this paper, we present for the first time 'AV-Net', an FCN based on a modified U-Net architecture. Recent studies using UNet have demonstrated AV classification using fundus photographs. To the best of our knowledge, our study is the first to demonstrate AV classification in OCTA. The input of the AV-Net is a 2-channel system to combine grayscale enface OCT and OCTA. Enface OCT is a near infrared (NIR) image, which is equivalent to a fundus image, to provide vessel intensity profiles. On the other hand, OCTA contains the information of blood flow strength and vessel geometry features. The output of AV-Net is an RGB (red-green-blue) image. The R and B channels correspond to artery and vein systems, respectively, and the G channel represents the background.

The overall design of the AV-Net follows an encoder-decoder architecture (Fig. 1(a)). The encoder of the AV-Net is a combination of dense, convolutional and transition blocks (Fig. 1), making the network deeper compared to UNet which incorporated a shallower 'VGG16' architecture. The encoder, also known as the contracting path, extracts the context of the image. The decoder, also termed as the expanding path, identify image features. The addition of bridging between the encoder and decoder is to enable precise localization and mapping of feature maps to produce the output image [27]. The convolution blocks are similar to the identity block in ResNet, except for the use of concatenation instead of summation operations [28,29]. The dense block is composed of convolution blocks, with each subsequent block connected to the previous blocks by skip-connections.

Skip-connections is to alleviate the vanishing-gradient problem in deep learning [30]. Following each dense block, a transition block is used to reduce the dimensions of the output feature maps. In the decoder, we employ upsampling operations and use decoder blocks. The decoder block concatenates the outputs of the upsampling operation and the output of the convolution from the

Table 1. Comparative classification performance of AV-Net.

Comparison	Cross Validation	Accuracy	F1	IOU
UNet Dice + Focal Loss Pre-trained Weights	Artery	88.054 ± 0.343	77.743 ± 0.673	63.711 ± 0.879
	Vein	88.653 ± 0.704	78.978 ± 0.722	65.354 ± 0.953
	Average	88.353 ± 0.500	78.360 ± 0.675	64.533 ± 0.884
AV-Net Dice Loss Pre-trained Weights	Artery	83.570 ± 0.734	61.769 ± 1.480	44.871 ± 1.495
	Vein	80.125 ± 1.587	63.697 ± 1.401	47.017 ± 1.388
	Average	81.848 ± 1.020	62.733 ± 1.423	45.944 ± 1.422
AV-Net Focal Loss Pre-trained Weights	Artery	81.007 ± 0.404	46.031 ± 1.236	29.980 ± 1.045
	Vein	81.566 ± 0.249	46.137 ± 0.942	30.066 ± 0.801
	Average	81.287 ± 0.307	46.084 ± 0.964	30.023 ± 0.816
AV-Net Dice + Focal Loss Random Weight	Artery	85.957 ± 0.842	73.243 ± 1.377	58.018 ± 1.656
	Vein	85.501 ± 0.582	73.631 ± 1.258	58.412 ± 1.544
	Average	85.729 ± 0.572	73.437 ± 1.304	58.215 ± 1.584
AV-Net Dice + Focal Loss Pre-trained Weights	Artery	86.705 ± 1.087	82.761 ± 1.677	70.658 ± 2.404
	Vein	86.798 ± 1.174	82.850 ± 1.666	70.781 ± 2.399
	Average	86.751 ± 1.126	82.805 ± 1.670	70.719 ± 2.399

classification accuracy on the ImageNet validation dataset. To employ the pre-trained encoder into a FCN, the fully connected layer was removed. The intermediate outputs of the encoder network are subsequently connected to the decoder network. This procedure is initialized with random weights using the glorot uniform distribution, as corresponding inputs to the appropriate layers (Fig. 1). Our newly constructed FCN, AV-Net, using transfer learning is then trained on OCTA images for the task of image segmentation. This method of employing transfer learning for image segmentation, is repeated in the comparative study with the state-of-the-art networks, i.e. UNet. FCN training procedure utilized the Adam optimizer with a learning rate of 0.0001, a dice loss function, and a minibatch size of 8. Moreover, regularization procedures including data augmentation and cross-validation were used to prevent overfitting. Training was performed on a Windows 10 computer using NVIDIA Quadro RTX 5000 Graphics Processing Unit (GPU). The FCN was trained and evaluated on Python (v3.7.1) using Keras (2.2.4) with Tensorflow (v1.31.1) backend. In this study, the OCTA dataset comprised of 50 images. To evaluate our network, a 5-fold cross validation method, with each fold following an 80/20 train/test split procedure, was employed. Due to a limited dataset, data augmentation, i.e., random flips, rotation, zooming, and image shifting, was implemented during the training process. Therefore, in each fold the network was trained with 3,000 images, and testing evaluation was performed on the 8 original images of each fold. Average accuracy, intersection-over-union (IOU) and F1-score was used as an evaluation metric for AV classification, by comparing with manually labelled ground truths for each cross-validation folder. In the revision, we have clarified that (page 4): For evaluation of average accuracy of artery identification, we conducted one vs all pixel-wise classification (artery pixels vs vein + background pixels), and measured the average accuracy from prediction performance of both labels. Similarly, for evaluating vein accuracy, the one vs all classification labels were vein vs artery + background pixels. The average performance accuracy is the mean of artery and vein accuracies (Table 1). Both IOU and F1 score are standard metrics for segmentation and pixel-wise classification tasks. IOU measures the similarity between a predicted region (AV) and the ground truth region in an image and can be defined as the size of intersection, divided by the union of two regions [31]. IOU was measured separately for artery and vein, by comparing predicted pixels for each category to the pixels from ground truth.

Average was calculated by taking the mean of artery and vein IOU. F1 score is also a robust metric for pixel-wise classification and can be defined as a harmonic mean of precision and recall [32].

2.3. Loss functions

In this study, the AV-Net was trained using a compound loss function derived from dice loss [33] and focal loss [34] and was defined as Eq. (1):

$$L = L_{dice} + L_{focal} \quad (1)$$

Where L_{dice} is the dice loss (Eq. (2)) and L_{focal} is the focal loss (Eq. (3)). Recent studies have found the combination of multiple losses improves image segmentation tasks with class imbalances [35,36]. Dice score measures the degree of overlap between the prediction and ground truth and is therefore suited for image segmentation (pixel-wise classification) tasks. The dice loss can be written as

$$L_{dice} = 1 - \frac{2 \sum_{x \in \omega} p_I(x)g_I(x)}{\sum_{x \in \omega} p_I^2(x) + \sum_{x \in \omega} g_I^2(x)} \quad (2)$$

The focal loss function is used to help mitigate the imbalance between foreground and background classes during training. The focal loss is derived from the cross entropy (CE) loss and introduces a focusing parameter γ that helps increase the importance of correcting misclassified examples [34]. L_{focal} can be written as

$$L_{focal} = - \sum_{x \in \omega} (\alpha(1 - p_I(x))^\gamma g_I(x) \log p_I(x) + (1 - \alpha)p_I^\gamma(x)(1 - g_I(x)) \log(1 - p_I(x))) \quad (3)$$

Where the weighting factor $\alpha \in [0, 1]$, focusing parameter $\gamma \geq 0$, $g_I(x)$ and $p_I(x)$ are label and estimated probability vectors, respectively. In our experimental designs, $\alpha = 0.25$ and $\gamma = 2$ works best in practice [34].

3. Results

3.1. Patient demographics

Our dataset comprised of images from 50 patients (20 control eyes and 30 DR eyes). Subjects and diabetic patients with and without DR were recruited from the UIC retina clinic. The patients present in this study are representative of a university population of diabetic patients who require clinical diagnosis and management of DR. Two board-certified retina specialists classified the patients based on the severity of DR according to the Early Treatment Diabetic Retinopathy Study (ETDRS) staging system. All patients underwent complete anterior and dilated posterior segment examination. All control OCTA images were obtained from healthy volunteers that provided informed consent for OCT/OCTA imaging. All subjects underwent OCT and OCTA imaging of both eyes (OD and OS). The images used in this study did not include eyes with other ocular diseases or any other pathological features in their retina such as epiretinal membranes and macular edema. Additional exclusion criteria included eyes with prior history of intravitreal injections, vitreoretinal surgery or significant (greater than a typical blot hemorrhage) macular hemorrhages. Validation dataset comprised of healthy volunteers that provided informed consent for OCT/OCTA imaging.

3.2. Classification evaluation

The AV-Net achieved an average accuracy of 86.75% (86.71% and 86.80% respectively for artery and vein) on the test data and a mean IOU was 70.72%, and F1-score of 82.81%. The accuracy metric considers segmentation of artery, vein and background pixels, and takes an average of the three parameters for final accuracy value. We observed that, the classifier is extremely robust for

background prediction, i.e., it is very good for segmenting the blood vessels (average accuracy ~97%). To demonstrate a more robust measure for AV classification performance, we utilize IOU which compares the AV-Net generated AV map pixel to pixel with the ground truth.

A comparative analysis of AV-Net performance is summarized in Table 1. The optimal AV-Net was trained with pre-trained 'imagenet' weights and a training loss function that integrated both Dice and Focal loss. In Table 1, we demonstrated AV classification performance using UNet ('Imagenet' weights, Dice + Focal loss); AV-Net with Dice loss; AV-Net with Focal loss; AV-Net with both Dice and Focal loss but without transfer learning 'imagenet' weights.

For comparative analysis, the results of AV-Net implementation with and without transfer learning are reported. It is observed that transfer learning improves the performance of AV-net compared to random weight initialization. Despite the high dissimilarity between ImageNet and OCTA dataset, there may be certain features that are transferable, such as simple features, morphology or intensity-based features, in early layers of the encoder.

Additionally, the combination of multiple losses has been demonstrated to improve the performance for image segmentation tasks. The most used loss function for segmentation tasks is the dice loss function. On the other hand, recent studies have revealed that the focal loss function mitigates class imbalances between the foreground and background. Therefore, our hypothesis is that the combination of both dice and focal loss can improve the performance of AV-Net. To test our hypothesis, we performed a comparative study by training AV-Net with dice and focal loss separately. The results from training with the dice and focal loss function separately, revealed that individually each loss function had adequate performance, with the focal loss function having the worst performance. However, when combining both the dice and focal losses improved the performance of AV-Net.

We further compared the performance of AV-Net with the state of the art UNet model. For comparative analysis, both architectures were trained using transfer learning and the combined dice and focal loss functions. As shown in Fig. 2, the AV-Net demonstrates improved performance compared to UNet. Interestingly, it is observed that the UNet showed slightly better accuracy values compared to AV-Net (88.25% vs 86.71%). Since UNet is a comparatively shallower network, it is less prone to overfitting, providing a better semantic segmentation performance. Since accuracy metric considers prediction of all pixels (i.e., artery, vein, and background), the overall accuracy is increased. However, this does not necessarily mean better performance of identifying artery and vein pixels. That is why similarity metrics, i.e., F1 and IOU scores of UNet are comparatively low. As a more complex network, AV-Net is much better for identifying artery and vein pixels, as represented by the F1 and IOU scores which reflect the comparison of artery and vein pixels with the ground truth. While AV-Net was inspired by the UNet architecture, the incorporation of short and long skip connections and the increased depth of the network improved the performance.

To check the AV classification performance on diseased eyes, we further tested the AV-Net performance on only the OCT/OCTA data from DR cohort. The accuracies of predicting arteries and veins were 85.94% and 85.85%, respectively. The mean IOU scores for artery and vein classification were 68.28% and 68.65%, respectively. The mean F1 scores for artery and vein classification were 81.12% and 81.4%, respectively.

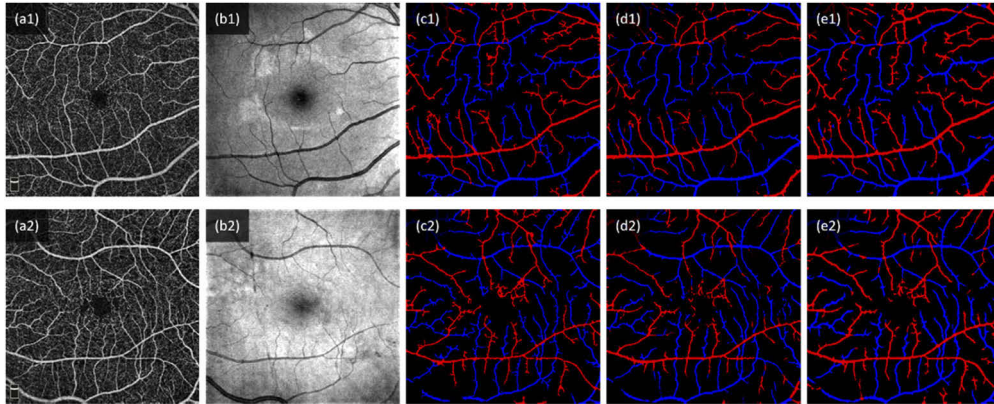


Fig. 2. Examples of control and DR (top and bottom, respectively) (a) input OCTA, (b) enface OCT, (c) the ground truth, (d) UNet predicted AV-maps, and (e) AV-Net predicted AV-maps.

4. Discussion

In summary, we have demonstrated the AV-Net for fully automated AV classification in OCTA. The AV-Net achieved an average accuracy of 86.75% (86.71% and 86.80% respectively for artery and vein) on the test data and a mean IOU was 70.72%, and F1-score of 82.81%.

Differential AV analysis is known to be valuable for quantifying subtle microvascular changes and distortions due to retinopathies. Incorporating AV classification capability into the clinical imaging devices would enhance the diagnostic ability and quantitative power of OCTA. Previous studies exploring the use of deep learning for AV classification have been primarily focused on traditional fundus photography. Xu et al. adapted a UNet for AV classification using publicly available fundus datasets, such as DRIVE and INSPIRE, and achieved high accuracy [37]. Similarly, Meyer et al. employed deep learning using a patch-wise prediction strategy and included regularization techniques such as dropout and batch normalization [38]. To our knowledge, this is the first study to employ deep learning for AV classification in OCTA.

In this study, we employed an FCN, based on the UNet architectures. In a previous study, Ronneberger et al. [27] have shown the use of long skip connections, that can help the network localize high resolution features, thereby a more precise output. In AV-Net, we employ dense blocks that utilize short skip connections. These short skip connections encourage the network to reuse features, making the model more compact. In comparison to other networks such as VGG16, AV-Net is a 5 times deeper network (having more convolutional layers) but the number of parameters is significantly smaller (approximately 17 times less). Having deeper network enables more learning capability, whereas smaller number of parameters means less computational burden. By leveraging both long and short skip connections, we are able to train our AV-Net for robust AV classification. From comparative analysis shown in Table 1, AV-Net performance is improved compared to AV classification performance using a standard UNet architecture. Furthermore, a comparison of AV-Net trained with ‘Dice’ loss, ‘Focal’ loss and ‘Dice + Focal’ loss showed that incorporating both Dice and Focal loss improved segmentation since Dice loss compares similarity between AV map and ground truth, and Focal loss compensates for class imbalance between AV pixels and background pixels.

The input of the AV-Net consists of both enface OCT and OCTA. While OCTA does provide highly detailed vasculature maps, the arteries and veins are indistinguishable from each other by OCTA information itself. On the other hand, OCT retains reflectance information to differentiate artery and vein [25]. By combining both images, the FCN can learn the intensity information

from the OCT and the highly detailed vasculature from the OCTA. Employing both OCT and OCTA is also convenient since they are from same OCT data volume and OCTA is reconstructed based on OCT processing. Therefore, using enface OCT and OCTA as 2-channel input of the AV-Net requires no pre-processing and image registration.

The results of the cross-validation study revealed an adequate IOU and F1 score. Qualitatively AV-Net has good vessel segmentation and AV classification performance. However, the predicted AV maps do appear more dilated compared to the ground truths. There are notable areas of misclassification, i.e., at vessel cross points. Future improvements to AV-Net could include developing a dataset with ground truth for vessel crossings. Additional validation with enlarged datasets from different OCTA devices will be required to pursue clinical deployments of the AV-Net for differential AV analysis.

5. Conclusion

The AV-Net has been demonstrated for fully automated AV classification in OCTA. The AV-Net is based on one FCN with modified U-shaped CNN architecture. A multi-modal training process was involved to include both enface OCT and OCTA for robust AV classification, and a transfer learning procedure was integrated to compensate for the limited size of OCTA dataset. By incorporating transfer learning and multi-modal training, the AV-Net achieved an accuracy of 86.75% for robust AV classification.

Funding

National Eye Institute (P30 EY001792, R01 EY023522, R01 EY030101, R01EY029673, R01EY030842); Research to Prevent Blindness; Richard and Loan Hill Foundation (Endowment); Illinois society to prevent blindness (ISPB_Minhaj Alam).

Disclosures

No competing interest exists for any author.

References

1. Y. Hatanaka, T. Nakagawa, A. Aoyama, X. Zhou, T. Hara, H. Fujita, M. Kakogawa, Y. Hayashi, Y. Mizukusa, and A. Fujita, "Automated detection algorithm for arteriolar narrowing on fundus images," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, (IEEE, 2006), 286–289.
2. M. K. Ikram, J. A. Janssen, A. M. Roos, I. Rietveld, J. C. Witteman, M. M. Breteler, A. Hofman, C. M. Van Duijn, and P. T. de Jong, "Retinal vessel diameters and risk of impaired fasting glucose or diabetes: the Rotterdam study," *Diabetes* **55**(2), 506–510 (2006).
3. M. Alam, T. Son, D. Toslak, J. Lim, X. Yao, and T. V. S. Tech, "Combining optical density ratio and blood vessel tracking for automated artery-vein classification and quantitative analysis in color fundus images," *Trans. Vis. Sci. Tech.* **7**(2), 23 (2018).
4. M. Alam, J. I. Lim, D. Toslak, and X. Yao, "Differential Artery–Vein Analysis Improves the Performance of OCTA Staging of Sickle Cell Retinopathy," *Trans. Vis. Sci. Tech.* **8**(2), 3 (2019).
5. M. Alam, D. Toslak, J. I. Lim, and X. Yao, "Color fundus image guided artery-vein differentiation in optical coherence tomography angiography," *Invest. Ophthalmol. Visual Sci.* **59**(12), 4953–4962 (2018).
6. W. Aguilar, M. E. Martinez-Perez, Y. Frauel, F. Escolano, M. A. Lozano, and A. Espinosa-Romero, "Graph-based methods for retinal mosaicing and vascular characterization," *Lecture Notes in Computer Science* **4538**, 25–36 (2007).
7. R. Chrástek, M. Wolf, K. Donath, H. Niemann, G. Michelson, and M. V. Appl, "Automated Calculation of Retinal Arteriovenous Ratio for Detection and Monitoring of Cerebrovascular Disease Based on Assessment of Morphological Changes of Retinal Vascular System," in *MVA*, 2002), 240–243.
8. E. Grisan and A. Ruggeri, "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*, (IEEE, 2003), 890–893.
9. H. Jelinek, C. Depardieu, C. Lucas, D. Cornforth, W. Huang, and M. Cree, "Towards vessel characterization in the vicinity of the optic disc in digital retinal images," in *Image Vis Comput Conf*, 2005), 2–7.

10. H. Li, W. Hsu, M.-L. Lee, and H. Wang, "A piecewise Gaussian model for profiling and differentiating retinal vessels," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, (IEEE, 2003), 1–1069.
11. M. Niemeijer, B. van Ginneken, and M. D. Abràmoff, "Automatic classification of retinal vessels into arteries and veins," *Med. Imaging* **7260**, 72601F (2009).
12. K. Rothaus, X. Jiang, and P. Rhiem, "Separation of the retinal vascular graph in arteries and veins based upon structural knowledge," *Image Vis. Comput.* **27**(7), 864–875 (2009).
13. A. Simó and E. de Ves, "Segmentation of macular fluorescein angiographies. A statistical approach," *Pattern Recognit.* **34**(4), 795–809 (2001).
14. S. Vázquez, N. Barreira, M. Penedo, M. Penas, and A. Pose-Reino, "Automatic classification of retinal vessels into arteries and veins," in *7th international conference biomedical engineering (BioMED 2010)*, 2010), 230–236.
15. S. Vázquez, B. Cancela, N. Barreira, M. G. Penedo, M. Rodríguez-Blanco, M. P. Seijo, G. C. de Tuero, M. A. Barceló, and M. Saez, "Improving retinal artery and vein classification by means of a minimal path approach," *Mach. Vis. Appl.* **24**(5), 919–930 (2013).
16. S. Zahid, R. Dolz-Marco, K. B. Freund, C. Balaratnasingam, K. Dansingani, F. Gilani, N. Mehta, E. Young, M. R. Klifto, B. Chae, L. A. Yannuzzi, and J. A. Young, "Fractal Dimensional Analysis of Optical Coherence Tomography Angiography in Eyes With Diabetic Retinopathy," *Invest. Ophthalmol. Visual Sci.* **57**(11), 4940–4947 (2016).
17. B. I. Gramatikov, "Modern technologies for retinal scanning and imaging: an introduction for the biomedical engineer," *BioMed. Eng. OnLine* **13**(1), 52 (2014).
18. K. R. Mendis, C. Balaratnasingam, P. Yu, C. J. Barry, I. L. McAllister, S. J. Cringle, and D.-Y. Yu, "Correlation of histologic and clinical images to determine the diagnostic value of fluorescein angiography for studying retinal capillary detail," *Invest. Ophthalmol. Visual Sci.* **51**(11), 5864–5869 (2010).
19. S.-C. Cheng and Y.-M. Huang, "A novel approach to diagnose diabetes based on the fractal characteristics of retinal images," *IEEE Trans. Inform. Technol. Biomed.* **7**(3), 163–170 (2003).
20. A. Y. Kim, Z. Chu, A. Shahidzadeh, R. K. Wang, C. A. Puliafito, and A. H. Kashani, "Quantifying Microvascular Density and Morphology in Diabetic Retinopathy Using Spectral-Domain Optical Coherence Tomography Angiography," *Invest. Ophthalmol. Visual Sci.* **57**(9), OCT362 (2016).
21. N. V. Palejwala, Y. Jia, S. S. Gao, L. Liu, C. J. Flaxel, T. S. Hwang, A. K. Lauer, D. J. Wilson, D. Huang, and S. T. Bailey, "Detection of non-exudative choroidal neovascularization in age-related macular degeneration with optical coherence tomography angiography," *Retina* **35**(11), 2204–2211 (2015).
22. G. Holló, "Vessel density calculated from OCT angiography in 3 peripapillary sectors in normal, ocular hypertensive, and glaucoma eyes," *Eur. J. Ophthalmol.* **26**(3), e42–e45 (2016).
23. M. Alam, D. Thapa, J. I. Lim, D. Cao, and X. Yao, "Quantitative characteristics of sickle cell retinopathy in optical coherence tomography angiography," *Biomed. Opt. Express* **8**(3), 1741–1753 (2017).
24. M. Alam, D. Thapa, J. I. Lim, D. Cao, and X. Yao, "Computer-aided classification of sickle cell retinopathy using quantitative features in optical coherence tomography angiography," *Biomed. Opt. Express* **8**(9), 4206–4216 (2017).
25. M. Alam, D. Toslak, J. I. Lim, and X. Yao, "OCT feature analysis guided artery-vein differentiation in OCTA," *Biomed. Opt. Express* **10**(4), 2055–2066 (2019).
26. T. Son, M. Alam, T.-H. Kim, C. Liu, D. Toslak, and X. Yao, "Near infrared oximetry-guided artery-vein classification in optical coherence tomography angiography," *Exp. Biol. Med.* **244**(10), 813–818 (2019).
27. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, (Springer, 2015), 234–241.
28. Q. Ji, J. Huang, W. He, and Y. Sun, "Optimized Deep Convolutional Neural Networks for Identification of Macular Diseases from Optical Coherence Tomography Images," *Algorithms* **12**(3), 51 (2019).
29. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017), 4700–4708.
30. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016), 770–778.
31. M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *International Symposium on Visual Computing*, (Springer, 2016), 234–244.
32. Y. Sasaki, "The truth of the f-measure. 2007," (2007).
33. F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, (IEEE, 2016), 565–571.
34. T.-Y. L. P. G. Ross and G. K. H. P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017).
35. W. Zhu, Y. Huang, L. Zeng, X. Chen, Y. Liu, Z. Qian, N. Du, W. Fan, and X. Xie, "AnatomyNet: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy," *Med. Phys.* **46**(2), 576–589 (2019).
36. M. Chen, L. Fang, and H. Liu, "FR-NET: Focal loss constrained deep residual networks for segmentation of cardiac MRI," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, (IEEE, 2019), 764–767.
37. X. Xu, R. Wang, P. Lv, B. Gao, C. Li, Z. Tian, T. Tan, and F. Xu, "Simultaneous arteriole and venule segmentation with domain-specific loss function on a new public database," *Biomed. Opt. Express* **9**(7), 3153–3166 (2018).
38. M. I. Meyer, A. Galdran, P. Costa, A. M. Mendonça, and A. Campilho, "Deep convolutional artery/vein classification of retinal vessels," in *International Conference Image Analysis and Recognition*, (Springer, 2018), 622–630.