Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images

José Morano [*], Álvaro S. Hervella, Jorge Novo, José Rouco

Centro de Investigación CITIC, Universidade da Coruña, A Coruña, Spain
VARPA Research Group, Instituto de Investigación Biomédica de A Coruña (INIBIC), Universidade da Coruña, A Coruña, Spain

ARTICLE INFO

ABSTRACT

Background and objectives: The study of the retinal vasculature represents a fundamental stage in the screening and diagnosis of many high-incidence diseases, both systemic and ophthalmic. A complete retinal vascular analysis requires the segmentation of the vascular tree along with the classification of the blood vessels into arteries and veins. Early automatic methods approach these complementary segmentation and classification tasks in two sequential stages. However, currently, these two tasks are approached as a joint semantic segmentation, because the classification results highly depend on the effectiveness of the vessel segmentation. In that regard, we propose a novel approach for the simultaneous segmentation and classification of the retinal arteries and veins from eye fundus images.

Methods: We propose a novel method that, unlike previous approaches, and thanks to the proposal of a novel loss, decomposes the joint task into three segmentation problems targeting arteries, veins and the whole vascular tree. This configuration allows to handle vessel crossings intuitively and directly provides accurate segmentation masks of the different target vascular trees.

Results: The provided ablation study on the public Retinal Images vessel Tree Extraction (RITE) dataset demonstrates that the proposed method provides a satisfactory performance, particularly in the segmentation of the different structures. Furthermore, the comparison with the state of the art shows that our method achieves highly competitive results in the artery/vein classification, while significantly improving the vascular segmentation.

Conclusions: The proposed multi-segmentation method allows to detect more vessels and better segment the different structures, while achieving a competitive classification performance. Also, in these terms, our approach outperforms the approaches of various reference works. Moreover, in contrast with previous approaches, the proposed method allows to directly detect the vessel crossings, as well as preserving the continuity of both arteries and veins at these complex locations.

1. Introduction

The analysis of the anatomical structures of the retina represents a key step in the diagnosis and screening of many relevant diseases [1,2]. This is due to the significant changes that some of these structures manifest when a certain disease is present. Thus, for example, ophthalmic and systemic diseases like Age-Related Macular Degeneration (AMD), glaucoma, and diabetes provoke significant alterations in the retinal vessels, the optic disc or the macula. The abnormal growth of new blood vessels within or under the macula is representative of the AMD disease [3]. Microaneurysms, retinal haemorrhages, exudates and neovascularization are representative of Diabetic Retinopathy (DR) [1],

as well as particular changes in the diameters of small retinal vessels can be a sign of diabetes [4]. The increased venous diameter and venous oxygen saturation is associated with glaucomatous damage [5,6]. The generalized retinal arteriolar narrowing is significantly associated with the optic nerve damage caused by open-angle glaucoma [7,8]. Many of these disorders are of a great severity, and they can lead to the partial or total loss of vision [2,9–11]. Nevertheless, an early diagnosis and treatment can avoid or mitigate some of the most harmful consequences. In this regard, color retinography, an affordable and widely available eye fundus imaging technique, is commonly used for early diagnosis and pathological screening. Thereby, performing an accurate analysis of this type of images is crucial.

* Corresponding author at: Centro de Investigación CITIC, Universidade da Coruña, A Coruña, Spain

  E-mail addresses: j.morano@udc.es (J. Morano), a.suarezh@udc.es (Á.S. Hervella), jnovo@udc.es (J. Novo), jrouco@udc.es (J. Rouco).

One of the most commonly analyzed structures of the retina is the vascular tree, since it provides particularly meaningful information for numerous diseases [1,3,12–14]. Thus, a complete analysis of the retinal vascular tree requires the segmentation of the retinal vessels from eye fundus images as well as its classification into arteries and veins. The retinal vasculature segmentation allows the measurement of different features of the blood vessels (such as width, length and tortuosity) with proven relevance in evaluating and monitoring some of the aforementioned diseases (e.g. DR). Complementarily, vasculature classification allows measuring some arteries/veins specific features that supplement the previous data, such as the arteriolar and venular diameters [15]. Among other utilities, the measurement of these diameters allows the calculation of the Arteriolar-to-Venular diameter Ratio (AVR), that has been found useful for the diagnosis of multiple diseases, such as hypertension or diabetes [16–18]. In clinical practice, however, the application of these analyses is limited due to two main reasons: first, they are challenging, repetitive and time-consuming for the clinical specialists, and second, the results that are obtained by distinct experts commonly present significant differences, especially for the small vessels. In this context, automatic vasculature segmentation and classification methods emerged as an auspicious alternative.

Regarding the segmentation of the retinal vasculature, the early works were mainly based on either ad hoc image processing techniques [12,19–21] or the use of hand-engineered image features along with traditional supervised learning methods, such as Artificial Neural Networks [22,23]. Later, advances in deep learning motivated vasculature segmentation methods based on the patch-wise application of Convolutional Neural Networks (CNN) [24]. Currently, the state of the art in vascular segmentation is dominated by Fully Convolutional Neural Networks (FCNN), of varying architectures and training approaches, which allow to directly obtain a vascular segmentation map from a full size image [25–30].

On the other hand, the classification of the retinal vessels into arteries and veins (A/V) followed a similar state of the art progression to vasculature segmentation, since the latter has always been conceived as a prior step to A/V classification. Methodologically, A/V classification works have evolved from ad hoc and traditional supervised learning methods [31–34] to CNN-based methods [35]. However, the overall approach of these works consisted on sequentially performing the vasculature segmentation followed by the A/V classification. Despite its wide adoption, the main drawback of this overall approach is that the errors in the vasculature segmentation are propagated to the classification stage. To avoid this issue, several recent works proposed to simultaneously address both tasks as a multiclass semantic segmentation problem with FCNNs [36–41]. Some of these works formulate the problem with three classes—background, artery and vein—[36,37,40,41]. Others, instead, add a fourth class "uncertain" [38,39]. This class, as specified in several annotated datasets [38,42–44], comprises all the vessels which experts cannot determine whether they are arteries or veins. The fourth class presented in [38] agrees with this definition. However, in [39], the vessel crossings are also included in the uncertain class. The crossings class comprises all the areas where a vein overlaps an artery or vice versa [45], and it is also commonly identified in A/V datasets [42–44]. Along the article, in line with our approach, we will use the same nomenclature as the one originally specified in the datasets, which does not include the vessel crossings in the uncertain class. In that regard, previous alternatives [38,39] handle vessels crossings and uncertain vessels in a counter-intuitive way, and give raise to incomplete segmentation maps for both arteries and veins, as the network is forced to decide to classify a vessel as one of the aforementioned classes.

In this work, we propose a novel approach for the simultaneous segmentation and classification of the retinal arteries and veins from eye fundus images using FCNNs. This approach decomposes these tasks into three related segmentation problems: the segmentation of arteries, veins, and vessels. To train the networks using this approach, we propose a novel loss named "Binary Cross-Entropy by 3" (BCE3) that combines three independent segmentation losses, one for each class of interest: arteries, veins and vessels. This approach manages uncertain vessels and vessel crossings in an intuitive way. On the one hand, uncertain vessels are discarded for the computation of the arteries and veins segmentation losses, but included for the computation of the vessels segmentation loss. In this way, the network can freely assign uncertain vessels to any of the artery and vein classes, while still receiving feedback for uncertain vessels in the vessels segmentation subtask. On the other hand, vessel crossings are included for the computation of the three independent losses. Thus, the networks are able to effectuate complete segmentation maps for the three classes: arteries, veins, and vessels. Moreover, the proposed setting allows the networks to detect the crossings between the vessels of different type, despite not being specifically trained for it, through the intersection (product) of the predicted arteries and veins segmentation masks. To evaluate the potential and advantages of this approach, we performed an exhaustive comparative study with the current state-of-the-art approaches, which commonly use multi-class Cross-Entropy loss, for the A/V segmentation and classification tasks in a multi-class semantic segmentation setting. For this comparison, we employ the publicly available Retinal Images vessel Tree Extraction (RITE) dataset, which is considered the reference standard for the classification of retinal arteries and veins.

1.1. State of the art

In the literature, several works have approached the segmentation of the retinal vasculature and, to a lesser extent, the classification of vessels into arteries and veins. Additionally, only in recent times, some works have approached both tasks simultaneously.

Regarding the vasculature segmentation, the initially proposed methods were chiefly based on ad hoc image processing techniques. Adaptive local thresholding [19], deformable models [20], vessel tracking [21] and ridge detection [12] are some representative examples. Moreover, traditional supervised learning methods, such as Artificial Neural Networks (ANNs) [22,23] or k-nearest neighbors [12], were also applied. Later, the use of deep learning, and specifically Convolutional Neural Networks (CNNs), meant an important advance [24,25]. These networks allow to automatically learn features during the training, contrary to classical approaches, where the feature extraction had to be manually designed. Thus, these networks can be trained end-to-end, from the raw data to the target decisions, and commonly produce better results [46,47]. Additionally, the use of Fully Convolutional Neural Networks (FCNNs) [48,49], that are only composed of local operators, allowed to directly handle image inputs of arbitrary sizes, without needing to iterate over them using patches. Currently, the state-of-the-art approaches on retinal vasculature segmentation utilize FCNNs [26–30].

The automatic A/V classification has been less explored than the vasculature segmentation. Also, until recently, it was conceived as a two-stage process, in which the vessel classification followed the vasculature segmentation. In this way, the initial A/V classification works either assumed the presence of vasculature segmentation masks, or some vessel pixels that were manually annotated, [31–33,50] or proposed their own automatic vasculature segmentation method as an independent first stage [34,35]. Hence, the classification stage was only focused on classifying the pixels previously segmented as vessels. Furthermore, early works typically restricted A/V classification to certain regions of interest, normally the region around the optic disc (OD) [50,31,32].

Zamperini et al. [50] employed supervised classifiers with color, contrast and position features for classifying a reduced set of central vessel points. This points were previously selected by experts and were located near the OD. Saez et al. [51] proposed an unsupervised method based on k-means clustering for the classification of previously

segmented vessels around the OD. Relan et al. [31] proposed a semi-automatic unsupervised method based on Gaussian Mixture Model-Expectation Maximization (GMM-EM) clustering, instead of k-means, due to its less dependency on the initialization. This method makes use of color features and, similarly to [50], it needs a set of manually marked vessel points to work. Later, the same authors [32] also proposed a supervised method based on Least Square-Support Vector Machine classifier. Their approach assumed the presence of previously extracted vascular masks near the OD. Dashtbozorg et al. [33] introduced a classification method that follows a graph-based approach. They were one of the first to propose a method for classifying the whole vascular tree. However, their method used previously segmented vessel masks for extracting the vascular graph. Estrada et al., in [34], presented a semi-automatic graph-based framework that incorporates a global likelihood model for A/V classification. Like [33], their method classifies the whole vascular tree, but it implies the participation of experts to manually correct the errors of the primarily extracted graph. Afterward, Welikala et al. [35] were the first to propose the use of a CNN for the classification step. They employed a patch-wise approach with patches centered on the centerline vessel pixels. For the extraction of these pixels, the authors utilized an unsupervised vessel segmentation approach based on a multi-scale line detector.

The two-stage approach followed by early works, however, presents an important drawback. Since the classification is done only for the pixels previously segmented as vessels, the final results heavily depend on the efficacy of the segmentation method, and the errors in it are propagated to the classification step. In recent times, to avoid this problem, several works have approached the simultaneous segmentation and classification of the retinal vessels as a semantic pixel-level classification task [36–41]. To that end, all these works employ FCNNs, as they currently represent the state of the art for semantic segmentation [52]. Furthermore, all of them use some type of image preprocessing to facilitate the training of the networks. Beyond this common characteristics, the works differ in many points. One of the most relevant is the formulation of the problem. Xu et al. [36], Girard et al. [37], Ma et al. [40] and Kang et al. [41] formulated the problem as a three-class semantic segmentation task—background, artery, vein—, while Hemelings et al. [38] and Galdran et al. [39] added a fourth class "uncertain". In [38], this class only comprises the vessels of uncertain class, coinciding with the definition provided in several A/V segmentation datasets [38,42–44]. However, in [39], the "uncertain" class also includes vessel crossings (areas where a vein overlaps an artery or vice versa). In [38], the crossing regions are labelled with the class of the upper vessel, so they are not handled in a special way. From these works [38,39], the approach proposed by Galdran et al. [39] is the only one that does not ignore the pixels from the "uncertain" class in the computation of the loss.

Another important point in related works is the strategy that was used by each of them to alleviate the scarcity of annotated data. This problem, despite not being exclusive of medical imaging, is particularly notable in this field, since data annotation is usually a challenging process that must be performed by clinical experts. Beyond the typical data augmentation operations (i.e. affine and elastic transformations and color and intensity variations), more or less common to all of them, the works implement two different types of mechanisms: random patch extraction and transfer learning. Girad et al. [37] train the network using small patches randomly extracted from areas close to the vessels. Hemelings et al. [38] train the network using large patches randomly extracted. Kang et al. [41] use an encoder petrained on ImageNet based on the GoogLeNet deep network architecture [53]. Ma et al. [40] use small random patches and an encoder pretrained on ImageNet and based on the ResNet CNN architecture [54]. Finally, some of these works also propose specific mechanisms to mitigate the manifest misclassification of certain sections of vessels, e.g. the classification of a segment halfway down a vein as "artery". Girard et al. [37] propose a postprocessing method based on graph propagation, Ma et al. [40] include an spatial

activation mechanism, and Kang et al. [41] incorporate a category-attention weighted fusion (CWF) module and a graph based vascular structure reconstruction algorithm for postprocessing the result images.

The rest of the manuscript is organized as follows. In Section 2, we present the methodology for the A/V segmentation and classification task, including the description of the different approaches to compare, the network architecture, the data, and the training details. Then, in Section 3, we present a comprehensive evaluation and discussion of the presented approaches. Also, we expose the comparisons with the state-of-the-art works for both the vasculature segmentation and the A/V classification. Finally, in Section 4, we present the main conclusions derived from the results and the potential future work.

2. Materials and methods

The simultaneous segmentation and classification of the retinal arteries and veins (SSCAV) requires both to identify the blood vessels in the image and to differentiate them between arteries and veins. In this work, we approach this integrated vascular analysis using FCNNs that are trained to provide pixel-wise image segmentations. The proposed methods decomposes the SSCAV into the three semantic segmentations targeting arteries, veins, and vessels, as described in Section 2.2. In order to evaluate the performance of this approach and quantify and demonstrate its advantages, we compare it with the traditional semantic segmentation approach followed by the state-of-the-art works. The details of this approach are explained in Section 2.1.

The visual differentiation between arteries and veins can be difficult in some cases, like small vessels with poor contrast and not clearly connected to a specific arterial or venular tree. These cases can be identified with certainty as vessels, but their classification into arteries and veins is uncertain, even for an expert. Additionally, it is common to find crossings between arteries and veins in the retina. Thus, although for these positions either the artery or the vein is above the other, these pixels can be regarded as simultaneously belonging to both types of vessels. This allows to account for continuous arterial and venular trees regardless of crossings below the other one. The identification of these two special situations (uncertain vessels, and vessel crossings) is common in manually annotated retinal vessel classification datasets [38,42–44]. Fig. 1 shows an example of a ground truth image from the RITE dataset labelled this way, along with its decomposition into arteries, veins, crossings and uncertain vessels.

2.1. Traditional approach

The traditional approach, depicted in Fig. 2, addresses the SSCAV as a single multi-class semantic segmentation task. Thus, for each pixel, an FCNN predicts the likelihoods of N mutually exclusive classes. To that end, the network output is composed of N channels that are tied with each other using a softmax activation. In this approach, Cross-Entropy (CE) is used as loss function, similarly to most state-of-the-art works in semantic segmentation. Formally, CE for joint vasculature segmentation and classification is defined as:where $f(r)_c$ denotes the network output for a given input retinography r and class c, s_c denotes its corresponding ground truth, ω_c denotes the weight of the class c, N denotes the number of classes and Ω denotes a Region of Interest (ROI) mask, with the set of valid pixels in the image.

In this case, we consider $N = 4$ classes, as depicted in Fig. 2 (for this reason, we name this loss CE4). These classes correspond to "background", "artery", "vein", and "uncertain or crossing". Thus, as in [39], the fourth class includes both the vessels of uncertain class and the vessel crossings.

Following prior works [37–39], we set the ω_c weights for the "background", "artery" and "vein" target classes to 1, so that all pixels of these classes count equally for the loss. Conversely, the weights for the "uncertain or crossing" class are set to 0. Thus, neither the vessel crossings, which simultaneously belong to the artery and vein classes,

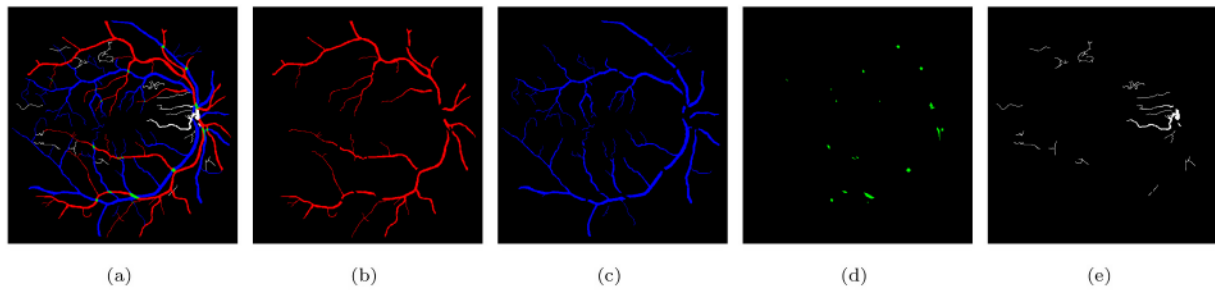(a)   (b)   (c)   (d)   (e)

Fig. 1. Example RITE ground truth image (a) and its decomposition into (b) arteries (red), (c) veins (blue), (d) crossings (green) and (e) uncertain (white) classes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
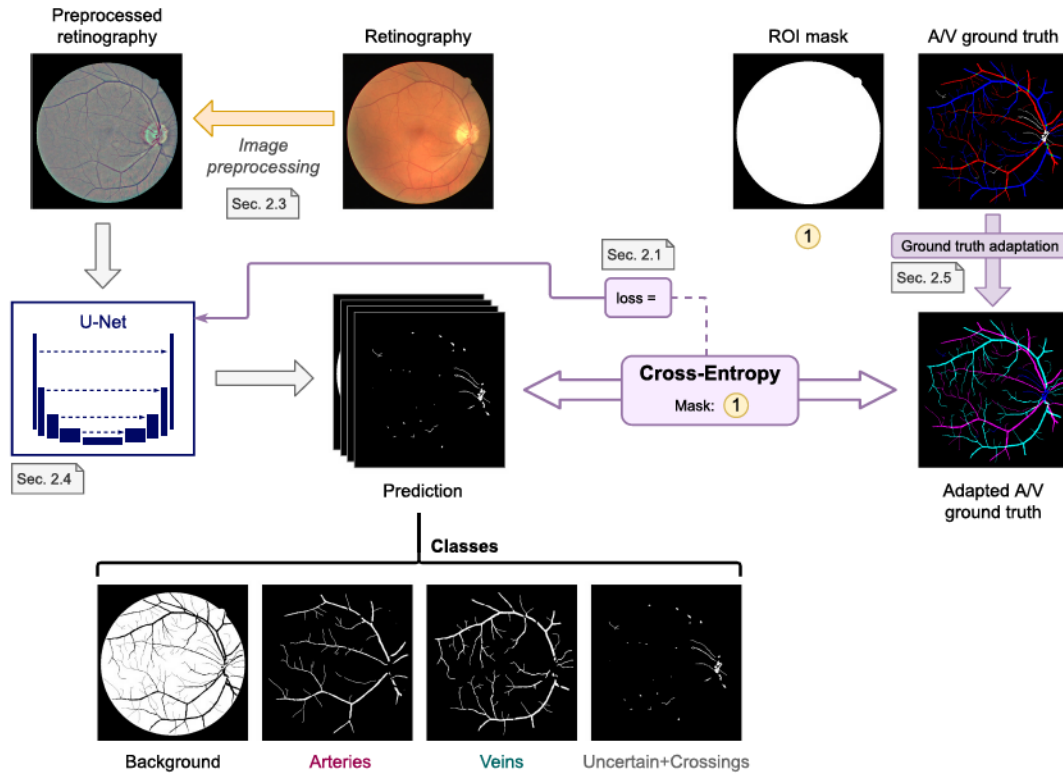

Fig. 2. Traditional A/V segmentation and classification approach. The gray notes indicate the section of the article where the method is described.

$$\mathscr{L}_{CE}(f(r),s) = -\sum_{\Omega}\sum_{c=1}^{N} \omega_c \cdot s_c \cdot log(f(r)_c), \tag{1}$$

nor the uncertain pixels, provide learning feedback in this approach. This is the most commonly used approach in previous works [36–38,40,41].

Using this classification approach, it is possible to obtain the vascular segmentation map by performing a pixel-wise addition of the predicted probability maps for arteries, veins and "uncertain or crossings", or using the inverse of the background class. However, it is not likely that the arterial and venular tree segmentation maps derived from the artery and vein classes are continuous, not even considering the "uncertain or crossings" class to complete them. The reason is that the network must decide to classify the vessel crossings as either of the above classes, in case of detecting them. Due to the fact that either the artery or the vein will be above the other, and the conditions for uncertain vessel classification are not commonly met at regular vessels, it is not probable that the "uncertain or crossings" class will be used. Thus the vascular tree for the vessel below will be discontinuous.

One possible solution to fix the vessel crossing discontinuities is to also use a weight $\omega_c = 1$ for the "uncertain or crossings" class in Eq. (1), as in [39]. However, while this solution forces the network to mark the crossings as this specific class, there is no distinction among these regions and the uncertain vessels. Thus, an additional split into two classes, with a very low number of learning samples for each one, would be further necessary. In this work, we provide a more straightforward and balanced solution, as explained in the next section, that addresses this issue.

2.2. Proposed multi-segmentation approach

Our proposed approach for SSCAV is depicted in Fig. 3. To both segment and classify the retinal vessels, we decompose the SSCAV problem into three different tasks: arteries segmentation, veins segmentation and vessels segmentation. This multi-segmentation (MS)
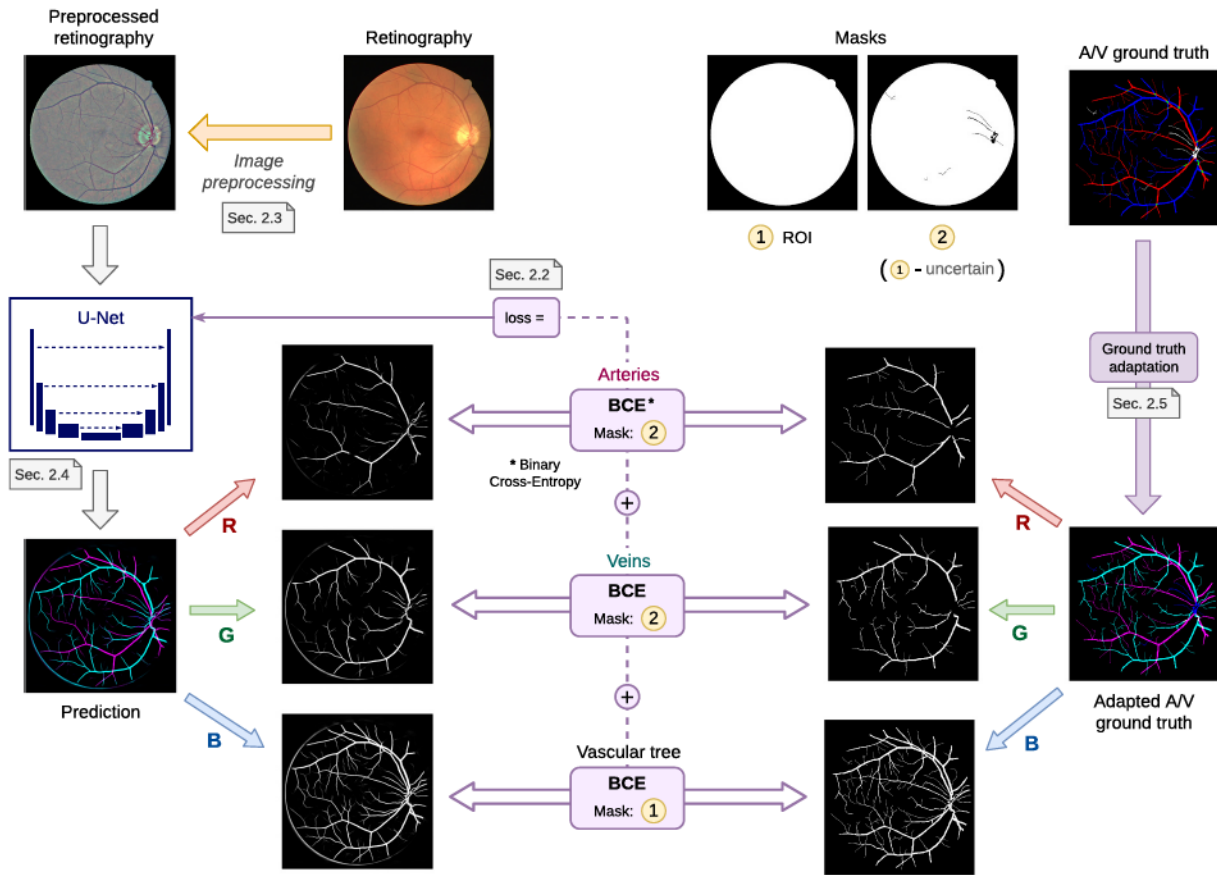
Fig. 3. Proposed A/V segmentation and classification approach. In each case, the corresponding Binary Cross-Entropy (BCE) is calculated only for the pixels that match the region delimited by the mask. For vessels, the delimited region includes all the pixels of the ROI, while for arteries and veins, this region does not include "uncertain" vessels. The gray notes indicate the section of the article where the method is described.

approach is in line with the dual nature of the retinal vascular tree, where arteries and veins individually spread throughout the retina and intersect frequently at different points. Moreover, following this approach, the three continuous vascular tree segmentation maps (arteries, veins and all the vessels) are also directly provided by the neural network.

In this case, the network output consists of 3 independent channels (see Fig. 3). Each channel contains the predicted probability map for one of the structures of interest: arteries, veins and vessels. The total loss is computed as the sum of the individual losses for each structure, each being calculated as the Binary Cross-Entropy (BCE) between the manually annotated segmentation mask of the structure and the corresponding predicted probability map. Mathematically, the "Binary Cross-Entropy by 3" (BCE3) loss is defined as

$$\mathcal{L}_{BCE3}(f(r),s) = -\sum_{c=1}^{3}\sum_{\Omega_c}\left[s_c{\cdot}log\big(f(r)_c\big)+(1-s_c){\cdot}log\big(1-f(r)_c\big)\right], \quad (2)$$

where $\mathbf{r}$ is the input retinography, $f(r)_c$ and s_c the network output and the ground truth for class c and Ω_c is the set of all the pixels within a class-specific ROI mask. The class-specific ROI mask for the vessels class is the retinal ROI, and it is equivalent to the Ω used in Eq. (1). Differently, for the artery and vein classes, the ROI mask Ω_c is computed as the retinal ROI minus the vessel pixels of uncertain class, as labelled by the experts in the ground truth. Examples of these ROI masks are represented in Fig. 3.

Notice that avoiding the uncertain pixels in the artery and vein segmentation losses (i.e. to mask them for the computation of the loss), and not in the vascular segmentation one, allows the system two things: first, to freely assign those pixels to artery or vein classes (as in the case

of the traditional approach), and second, to simultaneously receive feedback for those pixels in the vascular segmentation subtask. Consequently, the network receives feedback for every pixel inside the ROI.

In addition to this feature, it is also relevant to note that, in this case, the "uncertain" class does not include the vessel crossings, unlike the "uncertain or crossing" class defined in the traditional approach. Instead, in this approach, the vessel crossings are encoded as the superposition of the artery and vein segmentation maps, which should be both indicating positive class for these pixels. In order to do that, the ground truth s_c for each class c should be adapted to the output of the network and these restrictions. This can be achieved by considering the vessel crossing regions as positives in both artery and vein segmentation maps, and combining these two maps with the uncertain vessel regions to obtain the vascular tree segmentation map. An example of such an adapted ground truth is shown in Fig. 3. Additionally, this setting allows to detect vessel crossings through the intersection (product) of the predicted arteries and veins segmentation masks.

2.3. Preprocessing

Retinography datasets usually present a significant illumination and contrast variability among the images. This variability is due to different factors that affect the image acquisition. For example, distinct pupil sizes or exposure levels. In addition, retinographies also present illumination and contrast variability within the image due to the curvature of the retina or uneven lighting during the acquisition.

To correct this variability and therefore facilitate the training of the networks, most of the state-of-the-art-works implement an image preprocessing technique [36–40]. The ablation studies that are provided in

these works clearly show the beneficial effect of the image preprocessing. However, the employed technique varies among the works, and there is not an accepted standard. In this context, and given the adequate results, we include the method proposed in the reference work of Girard et al. [37]. In particular, this technique performs a channel-wise global contrast enhancement and a local intensity normalization as follows:

$$I_{\text{norm}}^C = \sigma_0 \frac{I^C - I_l^C}{\sigma_{I^C - I_l^C}},\tag{3}$$

where I_{norm}^C is the normalized channel C, I^C is the channel C of the original retinography, I_l^C is the low-pass filtered channel C and $\sigma_{I^C - I_l^C}$ is the global standard deviation of the channel resulting from the subtraction of the low-pass filtered channel I_l^C to the original channel I^C.

In our case, $\sigma_0 = 1$. Also, as low-pass filter, a gaussian filter with zero mean and standard deviation $\sigma = 10$ is used.

In Fig. 4, an example of a retinography before and after applying the preprocessing method is provided, along with the density histograms of vessel pixels and background pixels for each image. The histograms have been computed from the images converted to grayscale. It can be seen at first glance that the enhanced image, i.e. the one that has been preprocessed, has a greater contrast between the vessels and the background and it presents a more uniform illumination. This enhancement can also be seen in the histograms of the images.

2.4. Network architecture

For performing the SSCAV with both approaches, we use the fully convolutional neural network architecture U-Net [49]. This architecture was initially proposed for biomedical image segmentation, and its valuable results [55,56], particularly in several tasks related to this field [57–59], have made it to be considered a reference in computer vision. In fact, most of the state-of-the-art works in artery and vein classification are based on this architecture [36–40]. Furthermore, U-Net was successfully applied in our previous works for vasculature segmentation [60,61], fovea localization, optic disc localization and segmentation [61], and multimodal reconstruction [62].

In Fig. 5, a scheme of the U-Net architecture is shown. As can be seen in the figure, the network is composed of two main paths almost symmetrical: a contracting path (encoder), and an expansive path (decoder). This encoder-decoder shape is complemented with skip connections via concatenation between both paths.

The contracting path, or encoder, consists of 4 downsampling blocks. These blocks comprise two 3×3 convolutions, each followed by a ReLU, and a max-pooling of stride 2. Thus, after each downsampling block, the spatial resolution is halved. Moreover, the number of feature maps is duplicated.

The expansive path, or decoder, is composed of 4 upsampling blocks and, as said, it is almost symmetrical to the encoder. Each upsampling
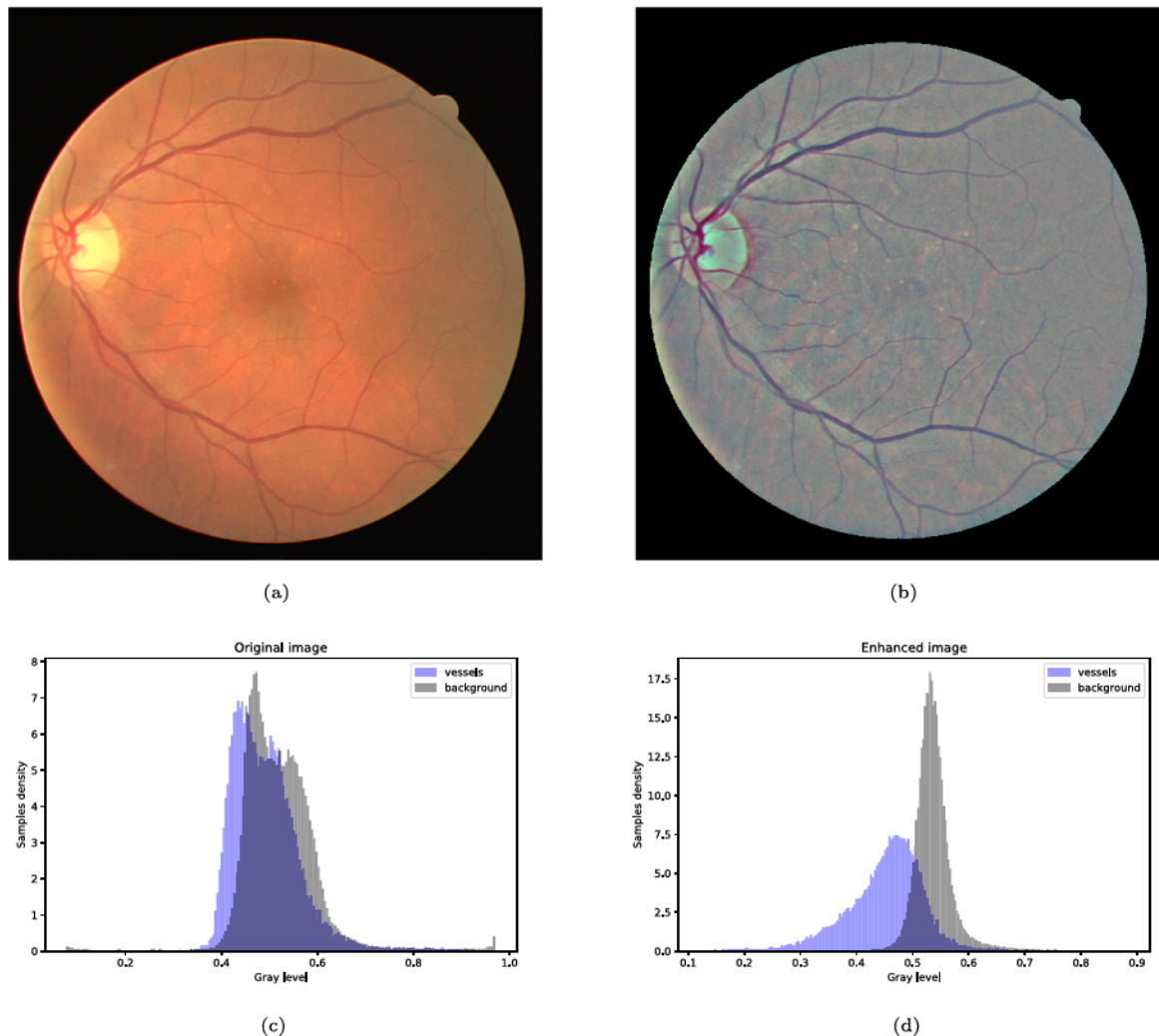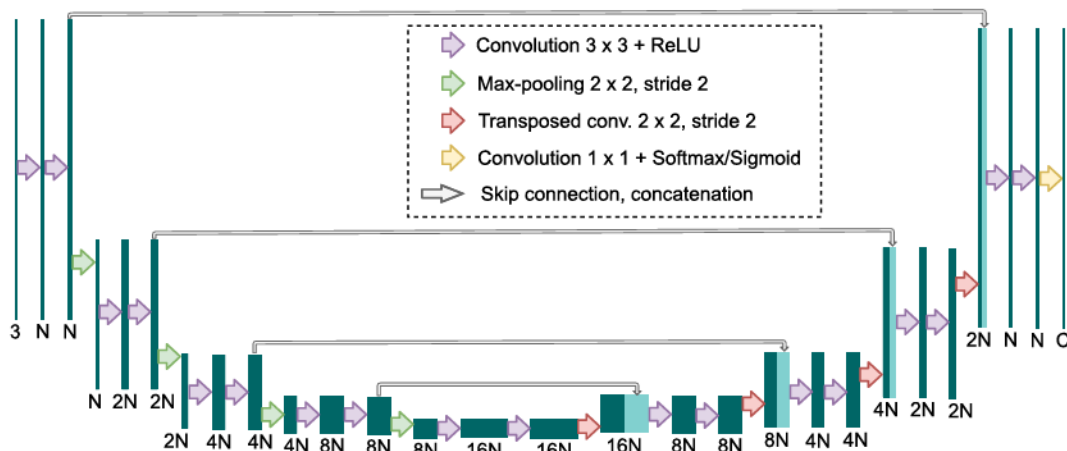

Fig. 4. Example of a RITE retinography before and after applying the preprocessing method, along with the gray level density histograms of vessel pixels and background pixels for each case. (a) Original retinography. (b) Preprocessed retinography. (c) Density histogram of the original retinography. (d) Density histogram of the preprocessed retinography.

Fig. 5. U-Net architecture. N represents the number of base channels. In our case, $N = 64$. C stands for the number of channels of the output. In the MS approach, the sigmoid activation function is applied at the last convolutional layer of the network. In the traditional approach, this layer is followed by a softmax layer.

block consists of two 3×3 convolutions, each followed by a ReLU, and a 2×2 transposed convolution. These transposed convolutions produce the opposite effect to the max-pooling operations, duplicating the spatial resolution. Also, the number of feature maps is reduced by half. The result of each transposed convolution is then concatenated with the feature maps of the same spatial resolution from the contracting path through a skip connection. In this way, the features from the contracting path are transferred to the expanding path at different resolution levels, leading to more detailed results in the network output.

In the final part of the network, two 3×3 convolutions followed by a ReLU are applied to the result of the latter upsampling block, and then, a last convolution of 1×1 is used to reduce the number of feature maps as desired (e.g. to a certain number of classes). In the MS approach, a sigmoid activation function is applied at the last convolutional layer, while in the traditional approach, the last layer is followed by a softmax layer.

2.5. Data

For the experiments in this work, we employed the publicly available RITE dataset [42], which is a reference standard for A/V classification. This dataset is based on the widely used Digital Retinal Images for Vessel Extraction (DRIVE) dataset [12], which contains color retinography images and is likewise considered a reference standard for vascular segmentation. Both the DRIVE and RITE datasets are divided into the same training and test subsets, including 20 color fundus photographs each one. From these 40 images, 7 are from patients with diabetic retinopathy (mild early stage) and 33 are from healthy patients. All the images are 768×584 pixels with a circular Region Of Interest (ROI).

For all the images, the manual segmentation of the retinal vessels and their classification is available. The DRIVE dataset provides a gold standard manual vasculature segmentation map for each of the training and test images by an expert (first expert). Additionally, the test set images are annotated by another independent observer (second expert). The RITE dataset uses these manual annotations as reference to label the vessels. Specifically, the training set annotations are based on the first expert annotations from DRIVE, while the test set annotations are based on those of the second one. In the RITE annotations, all the vessels pixels were classified into one of four classes: *artery, vein, crossing* or *uncertain*. The *crossing* class is used for pixels in the regions where arteries and veins overlap, while the *uncertain* class is used for the pixels that the clinical experts have been able to identify as vessels but not to discriminate as arteries or veins. The distribution of data samples in these classes is shown in Table 1. Also, an example of a retinography and its corresponding vasculature segmentation and A/V classification

Table 1
Distribution of data samples.

Class	Samples	%
Background	3,828,898	87.52
Vessel	546,029	12.48
- Artery	227,210	5.19
- Vein	278,576	6.37
- Crossing	14,003	0.32
- Uncertain	26,240	0.60

ground truths is depicted in Fig. 6.

For training the different networks, we use the 20 images of RITE-train with a random split of 15 and 5 images for training and validation, respectively. For testing, we use the entire RITE-test set.

In order to train the networks following the traditional and MS approaches, the ground truth images from the RITE dataset are adapted, in each case, to the output of the network. In the traditional approach, each pixel of the ground truth is assigned to either background, artery, vein or "uncertain or crossing" class, which comprises both uncertain vessels and crossings. The resulting ground truth is a grayscale image in which each pixel has a numeric label of the class to which it belongs. On the other hand, in the MS approach, each pixel is assigned to any number of the following classes: artery, vein and vessel. In this case, crossings belong to the three classes, and uncertain vessels belong only to the "vessel" class. Background pixels are not assigned to any class. The resulting ground truth is an RGB image in which each channel contains a manual segmentation mask of one of the aforementioned structures. An example of a RITE ground truth adapted to both the traditional and the MS approaches can be found in Fig. 7.

Data processing for the evaluation of the different methods is described in detail in Section 2.7.

2.6. Training details

To train the different models, the Adam optimization algorithm is used [63]. The parameters of the algorithm were set as follows: initial learning rate $\alpha = 1 \times 10^{-4}$, decay rates $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The training is performed at a constant learning rate, and early stopping is applied when the validation loss does not improve for 200 epochs. Both β_1 and β_2 values were taken from the original Adam paper [63], while the initial learning rate and the patience for early stopping were empirically established based on the learning curves for both training and validation data. To avoid an exhaustive exploration of the parameter space, we relied on prior evidence of the most suitable values that can be found in the literature [38,39,61,62,64].
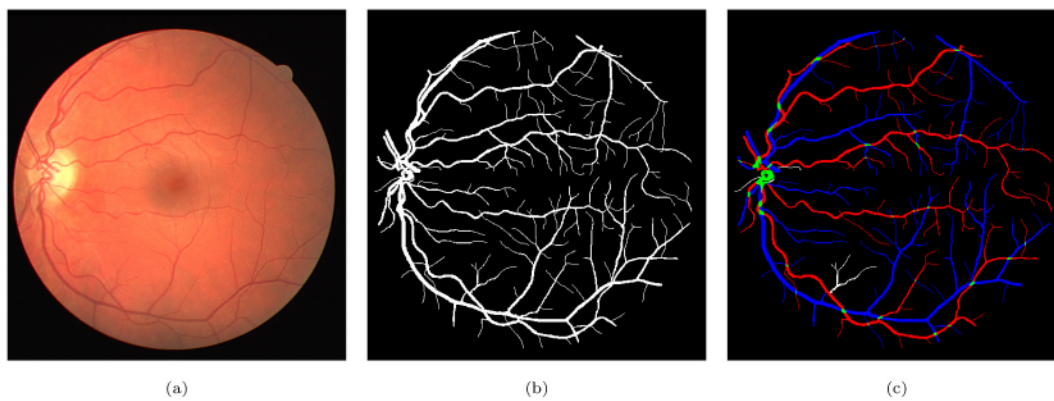
(a) 　　　　　　　　　　　　　(b) 　　　　　　　　　　　　　(c)

Fig. 6. Example of RITE retinography and its ground truths. (a) Retinography. (b) Vasculature segmentation ground truth. (c) A/V classification ground truth. In the A/V ground truth, arteries, veins, crossings and "uncertain" vessels are represented by the colors red, blue, green and white, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
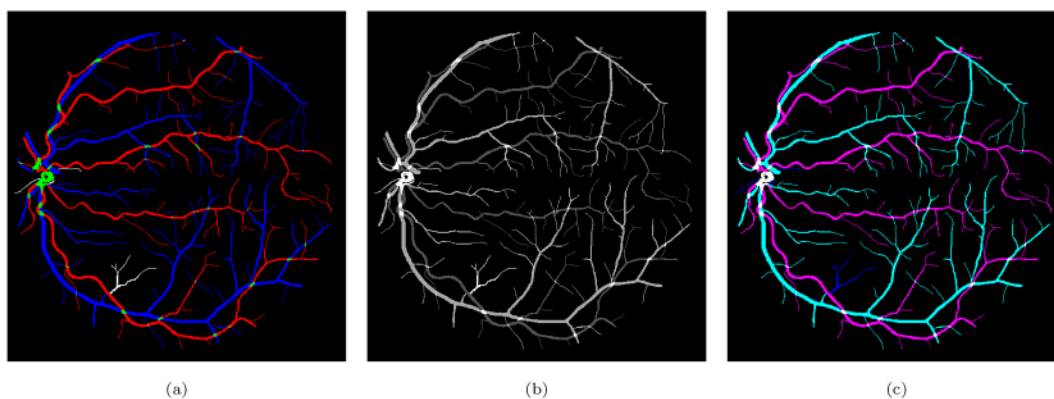

(a) 　　　　　　　　　　　　　(b) 　　　　　　　　　　　　　(c)

Fig. 7. Example of RITE retinography ground truth and its adaptations for the MS and traditional approaches. (a) Original RITE ground truth. (b) Adapted ground truth for the traditional approach. (c) Adapted ground truth for the MS approach.

In all the cases, the parameters of the networks are initialized using the He et al. [65] initialization method with uniform distribution, as in [61,64].

As the training set contains only 15 images, we use online data augmentation to artificially increase the training samples. Thus, on each epoch, random transformations are applied to the input images. These transformations consist of combinations of random color and intensity variations, slight affine transformations (rotation, scaling and shearing) along vertical and horizontal flipping. These transformations are the same as those used in [60,62].

The proposed methodology is implemented in Python 3 using the open source framework PyTorch. Training and development was performed on a single NVIDIA Tesla P40 GPU with a memory size of 24 GB. The CPU was an Intel Xeon Gold 6146 CPU @ 3.20GHz, and the memory size, 16 GB. In such a system, the complete training of a U-Net network following the proposed methodology takes about 1.5 h. In the test phase, it takes less than 0.01 s to segment one image using the GPU, and about 10 s using the CPU alone.

2.7. Quantitative evaluation

The quantitative evaluation of the presented approaches is performed by comparing the predicted vascular segmentation, as well as the identification of arteries and veins, against the manual ground truth annotations. As we are interested in both the segmentation and the classification, we divide the evaluation into three parts, focused on each one of these aspects and on the joint performance. Also, only for the proposed approach, we include an extra evaluation focused on assessing the performance of the approach on the complex vessel crossings localization task.

The first part is focused on assessing the segmentation performance on the different structures of interest: arteries, veins and full vascular tree. For each structure, we compute the Receiver Operating Characteristic (ROC) and Precision-Recall curves, by considering each segmentation target as a binary classification problem against the background. For the arteries and veins, we only take into account the pixels inside the ROI, excluding uncertain vessels and vessel crossings. The latter are excluded for performing a fair comparison with the traditional approach, which does not include vessel crossings in both arteries and veins probability maps. In each case, the positive class is the structure of interest: arteries, veins and vessels, and the negative class, everything else.

The ROC and PR curves are built applying a variable threshold to the predicted probability maps, so that it is possible to perform the evaluation without selecting an specific decision threshold. Also, to summarize these curves, we computed the Area Under the Curve (AUC) values in each case.

It is worth noting that we include the ROC and PR curves due to different reasons. The former is included due to its wide use in the literature as a default segmentation metric, while the latter is included due to its greater sensitiveness when the target classes are unbalanced [66]. In our case, for the segmentation of the different structures, the number of samples of the positive class, i.e. the vessels of each kind, is significantly lower than the number of samples from the negative class, which cover the whole retinal background and non-target structures. In this scenario, PR analysis is more convenient, as it presents a higher

sensitiveness to changes in the number of false positives.

The second part of the evaluation is focused on assessing the performance of the approaches in artery/vein classification. This scenario represents the standard evaluation setting adopted by the state-of-the-art works for A/V classification. Thus, we follow the exact same evaluation procedure that is used in those works [37–41]. Specifically, we build the confusion matrix for the artery and vein classes for the pixels that were annotated as "vessel" in the ground truth, excluding the uncertain vessels and the vessel crossings. To that end, we consider the class with the highest probability at the network output. We report Sensitivity, Specificity and Accuracy computed from this matrix, considering arteries as the positive class. In this case, as with many prior works, we are reporting the results for the system at the operating point that results from the network training. Nevertheless, as this operating point highly depends on the training set and the used loss, the metrics derived from this confusion matrix are not suitable for comparing with systems at other operating points. Performing a threshold optimization to select other suitable operating point for this classification, as for example maximizing a balanced accuracy on the validation set [39] also depends on the used validation set and metric. This does not improve the comparability as long as there is not a gold standard metric used as reference to evaluate the task. Thus, in addition to reporting the performance of the selected system, we provide the ROC curves along with the AUC-ROC metric, considering a varying threshold for the artery-vein classification of the ground truth classified vessels.

Finally, in order to evaluate the joint artery-vein classification and the vascular segmentation, we also follow the common approach in some previous works. In this way, we provide the Sensitivity, Sensibility and Accuracy in the vessel-background classification using the trained system, along with the artery-vein classification performance metrics. However, this approach, although being adopted by the state-of-the-art works in A/V classification, is not providing a joint evaluation of artery-vein classification of the segmented vasculature. This evaluation includes all the pixels that were labelled as "vessel" in the ground truth, assuming that the predicted vessels masks are perfect. Nonetheless, this assumption is not true, as the models can only detect part of the total number of vessel pixels. In a real scenario, only these pixels would be classified into artery and vein classes, not all the pixels that actually belongs to vessels, as the standard evaluation assumes. Therefore, the standard evaluation does not truly reflects the overall performance of the models.

To overcome this issue, we include a different evaluation method, which constitutes the third part of the quantitative evaluation. This method is intended to determine, simultaneously, the performance of the approaches regarding the A/V classification and the vessels segmentation. For this end, we plot the vessel/background classification sensitivity against the A/V classification accuracy for multiple thresholds. In addition, to compute the latter for the different thresholds, only the detected vessel pixels labelled as such in the ground truth are taken into account (i.e. the intersection of detected and annotated vessels). The curve resulting from applying this evaluation clearly depicts the overall performance of the networks regarding the two involved tasks.

As commented above, we include an extra evaluation for the proposed approach. This evaluation is focused on assessing the performance of the approach in the vessels crossings localization task. For this purpose, it is necessary, first of all, to define how the output of the network and the ground truth are translated to a set of coordinates representing the vessel crossings. In this process, the first step is to obtain the ground truth and the predicted segmentation masks of the crossings. This is achieved by performing an element-wise product of the artery and vein channels, as the vessel crossings belong to both classes. Then, for the prediction, we use an intensity threshold followed by a small dilation to merge very close regions. Once the prediction is converted to binary and dilated, we perform a connected-component analysis (CCA) and compute the centroids of the detected regions for both the prediction and the ground truth. The coordinates of the detected centroids are the

ones used in the evaluation. As gold standard, as it is done in [64], we consider a predicted crossing a True Positive when it is within a certain distance d of a crossing from the ground truth. Otherwise, it is considered a False Positive. Also, a crossing can be predicted only once. Thus, in each case, only the closest prediction within the distance d is considered as a True Positive. The crossings that are not located within this distance of any ground truth crossing are considered False Negatives. This analysis is performed for a representative distance of 10 pixels ($d = 10$). Finally, in order to build the PR curve, we use these measurements to compute the Precision and Recall metrics for a moving threshold over the entire test set.

2.8. Experimental details

To quantitatively evaluate the performance of the MS and the traditional approaches in the SSCAV task, we trained the same models using both approaches. Also, to evaluate the effect of the image preprocessing, we trained the networks using both the preprocessed images and the original ones. Thus, we compare a total of 4 alternatives. In addition, to take into account the stochasticity of the networks training, we performed 5 training repetitions with random initialization for each considered alternative.

Then, for the different alternatives, we performed the evaluation described in Section 2.7. In each case, we built the mean ROC and PR curves (calculating its corresponding mean AUC values) and we computed all the selected metrics for the A/V classification (Sensitivity, Specificity and Accuracy) and its corresponding mean values. Furthermore, we performed the joint evaluation of the segmentation and the classification tasks. In each case, the standard deviation value is also included.

Ultimately, in order to determine if the results of the proposed approach are significantly better (statistically) than those of the traditional approach, we have performed a one-sided Student's *t*-test of the means of the AUC-ROC, AUC-PR and Accuracy metrics.

3. Results and discussion

In Fig. 8, the mean ROC and PR curves in RITE-test for the networks that were trained using the BCE3 and CE4 losses with and without image preprocessing ("enhanced" and "original", respectively) are depicted. As commented in the previous sections, we include three ROC and PR curves, one for each structure of interest: arteries (A), veins (V), and Vascular Tree (VT)—that is, vessels—. Fig. 9 depicts the ROC and PR curves for the different networks in the DRIVE-test set. Additionally, all AUC values of the curves are summarized in Table 2. Complementing the segmentation results, Fig. 10 depicts the ROC and PR curves of the different approaches for the A/V classification in the RITE-test set. These curves are built considering only the pixels labelled as "vessel" in the ground truth, being "artery" the positive class. On the other hand, Table 3 reports the classification results in RITE-test for the different trained models. Fig. 12 depicts the vessels classification sensitivity against the artery/vein classification accuracy of the different models and multiple thresholds in RITE-test dataset. Fig. 11 depicts the PR curve for vessel crossing localization in RITE-test set for the networks trained using the MS approach. Lastly, Fig. 13 shows representative examples of the predicted probability maps that were obtained by the trained models using both losses and the enhanced retinographies as input. In each case, the figure depicts the RGB composition of the predicted maps for a better clarity.

As it can be observed in the vasculature segmentation results of Figs. 8 and 9 as well as Table 2, the best performance is achieved by the proposed MS approach using the BCE3 loss, either using the enhanced or the original retinography as input, and regardless of the target class. For each structure of interest, i.e. arteries, veins and vessels, both the AUC-ROC and AUC-PR values are higher with the MS approach than with the traditional CE4 one. Furthermore, AUC-PR and AUC-ROC for vessels are
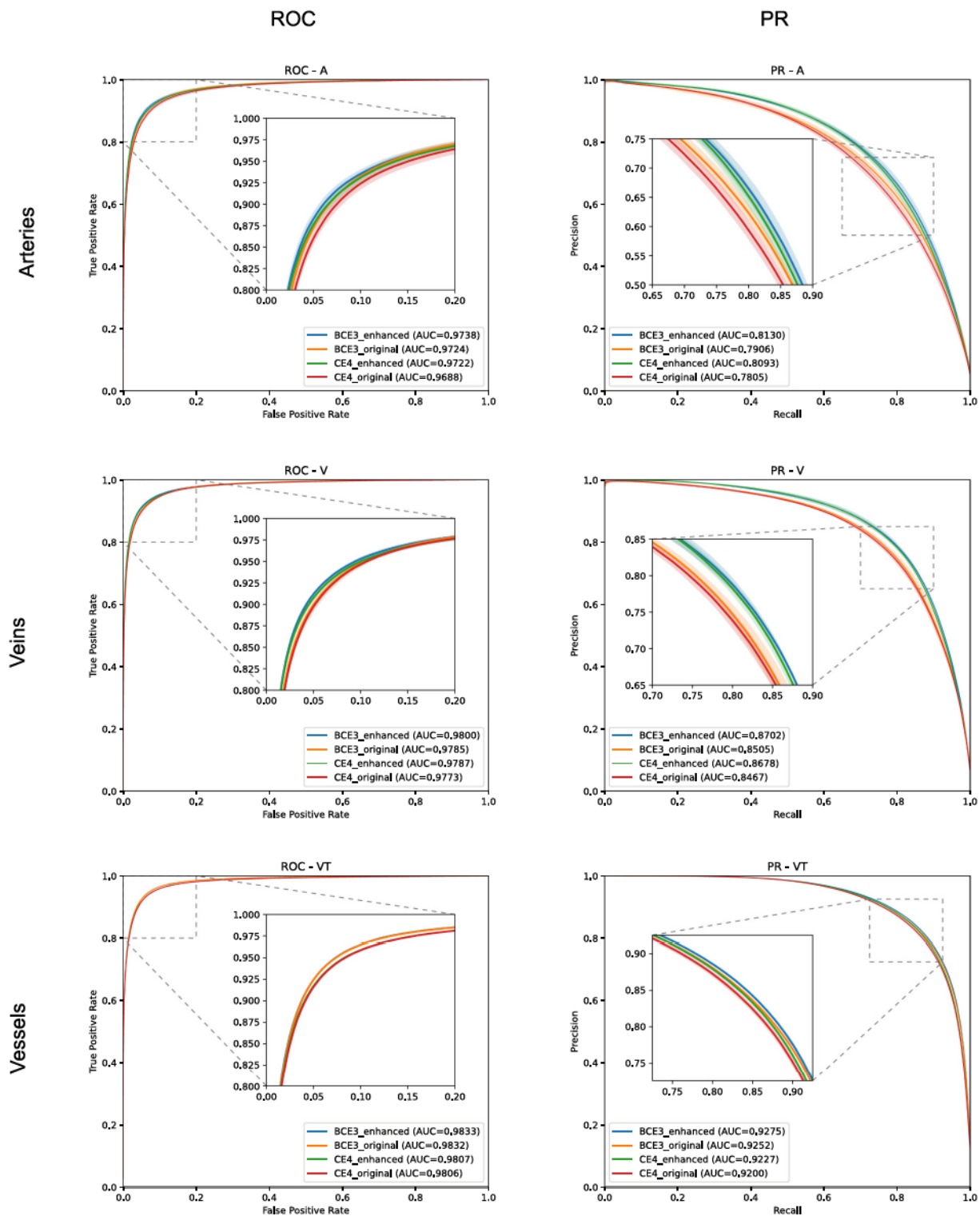
Fig. 8. ROC and PR curves for the different structures in RITE-test for the networks trained using the MS and the traditional approaches and the enhanced and the original images.

significantly higher for the MS approach than for CE4 ($p < 0.001$), as well as other statistics like AUC-ROC for arteries ($p < 0.05$) and AUC-ROC por veins ($p < 0.05$).

In addition to the higher segmentation performance using BCE3, crossings are handled in a more intuitive way. In this case, crossing pixels are simply assigned to both artery and vein classes at a time and the network is able to detect them (see PR curve depicted in Fig. 11)

while allowing to achieve a continuous segmentation of both the arterial and venular trees. For CE4, differently, the crossings are mostly treated as a separate class (along with the uncertain vessels), either to detect them [39], or to let the network detect the artery or the vein above the other [36,37,40,41]. There is also some approach in which the crossings belong to the same class as that of the upper vessel [38]. Nevertheless, whichever of these alternatives gives raise to incomplete segmentation
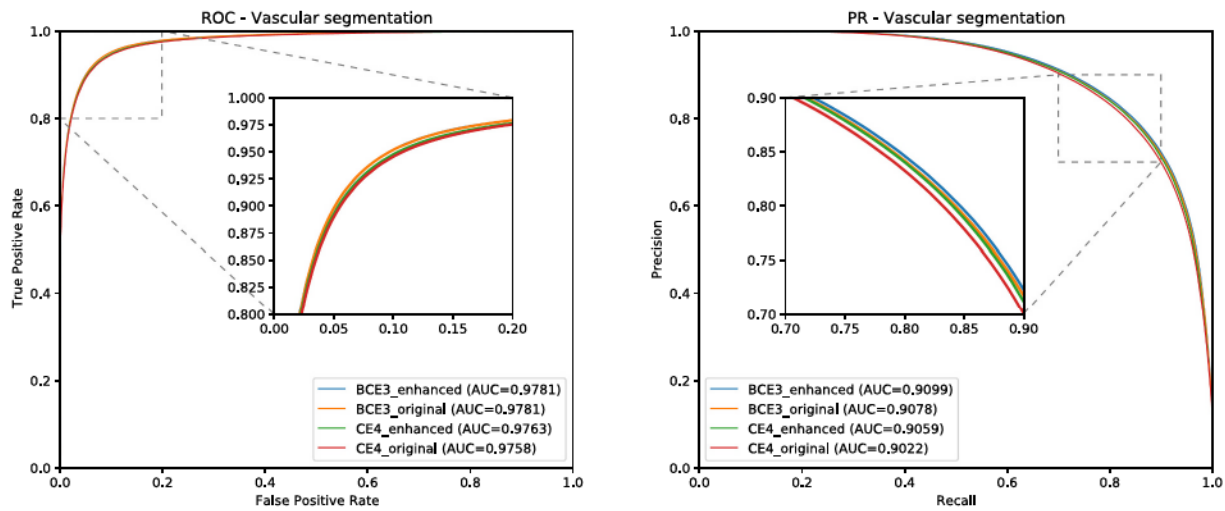
Fig. 9. Vascular segmentation ROC and PR curves in DRIVE-test for the networks trained using the MS and the traditional approaches and the enhanced and the original images.

Table 2
Segmentation results.

| Model	RITE-test						DRIVE-test	
	Arteries		Veins		Vessels		Vessels	
	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
BCE3 original	97.24 ± 0.10	79.06 ± 0.77	97.85 ± 0.12	85.05 ± 0.44	98.32 ± 0.03	92.52 ± 0.08	97.81 ± 0.03	90.78 ± 0.09
CE4 original	96.88 ± 0.32	78.05 ± 0.82	97.73 ± 0.12	84.67 ± 0.50	98.06 ± 0.06	92.00 ± 0.21	97.58 ± 0.04	90.22 ± 0.12
BCE3 enhanced	97.38 ± 0.21	81.30 ± 0.73	98.00 ± 0.05	87.02 ± 0.38	98.33 ± 0.04	92.75 ± 0.12	97.81 ± 0.02	90.99 ± 0.07
CE4 enhanced	97.22 ± 0.22	80.93 ± 0.89	97.87 ± 0.12	86.78 ± 0.54	98.07 ± 0.04	92.27 ± 0.08	97.63 ± 0.03	90.59 ± 0.07
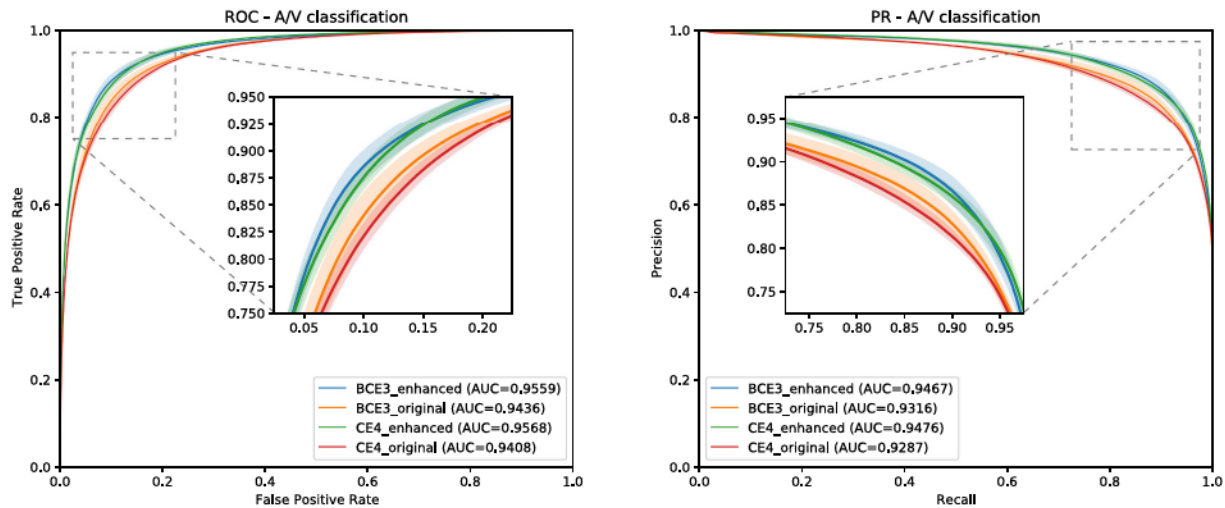
Fig. 10. A/V classification ROC and PR curves in RITE-test for the networks trained using the MS and the traditional approaches and the enhanced and the original images.

maps for both arteries and veins. Multiple examples of this effect can be found in Fig. 14. Taking all this into account, it can be stated that the MS strategy favors a better segmentation of the different structures, and handles the different cases in a much simpler way.

Regarding the artery/vein discrimination, although the MS alternatives achieve highly positive results (see Table 3), it cannot be affirmed that they perform better than the traditional alternatives. Considering the variability, the A/V classification results of both approaches are not significantly different. However, for the vessel/background discrimination, the situation is different. In this case, consistently with the higher AUC values for vessels segmentation, the MS approach achieves significantly better results ($p<0.01$) for both Accuracy and AUC-ROC. As can be seen in Table 3, when training with the original images, the sensitivity values of the MS approach are, on average, more than 2.5% higher than those of the traditional approach. Similarly, when training with the enhanced images, the improvement is about 1%. We focus on

Table 3
Classification results.

| Model	Artery/Vein				Vessel/background			
	Sens. (%)	Spec. (%)	Acc. (%)	AUC-ROC (%)	Sens. (%)	Spec. (%)	Acc. (%)	AUC-ROC (%)
BCE3 original	86.27 ± 1.53	88.64 ± 1.63	87.47 ± 0.85	94.36 ± 0.49	80.26 ± 0.52	98.49 ± 0.07	96.23 ± 0.01	98.32 ± 0.03
CE4 original	84.52 ± 1.53	88.46 ± 1.01	86.55 ± 0.50	94.08 ± 0.38	77.52 ± 0.82	98.70 ± 0.12	96.08 ± 0.04	98.06 ± 0.06
BCE3 enhanced	87.47 ± 2.09	90.89 ± 0.68	89.24 ± 0.73	95.59 ± 0.40	79.12 ± 1.22	98.65 ± 0.13	96.16 ± 0.05	98.33 ± 0.04
CE4 enhanced	87.24 ± 1.24	90.26 ± 0.69	88.78 ± 0.53	95.68 ± 0.37	78.07 ± 1.64	98.67 ± 0.19	96.05 ± 0.05	98.07 ± 0.04
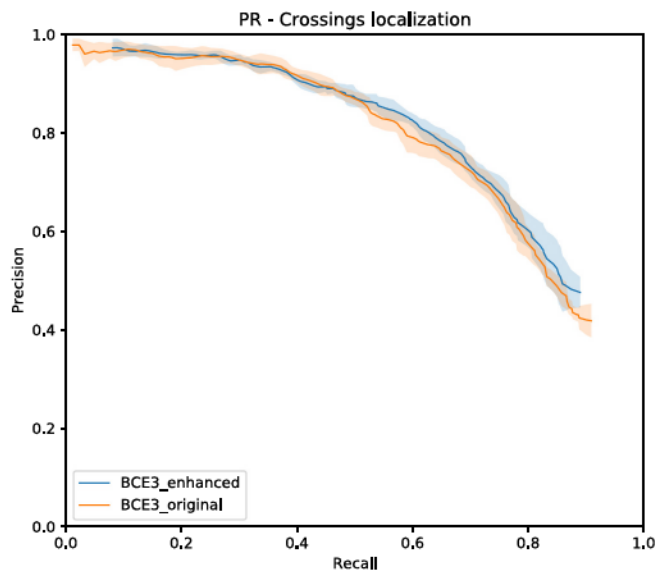
Fig. 11. Vessels crossings localization PR curve in the RITE-test set for multiple thresholds and the different evaluated models trained using the MS approach.

sensitivity, at a similar level of specificity, due to the positive class that is "vessel", and the sensitivity measures the proportion of positives that were correctly identified. Specificity and accuracy, although they are also relevant, are much less sensitive in this scenario, since they take into account the background pixels, much more numerous and easier to classify.

The results depicted in Fig. 12 are also in this line. Both methods, the
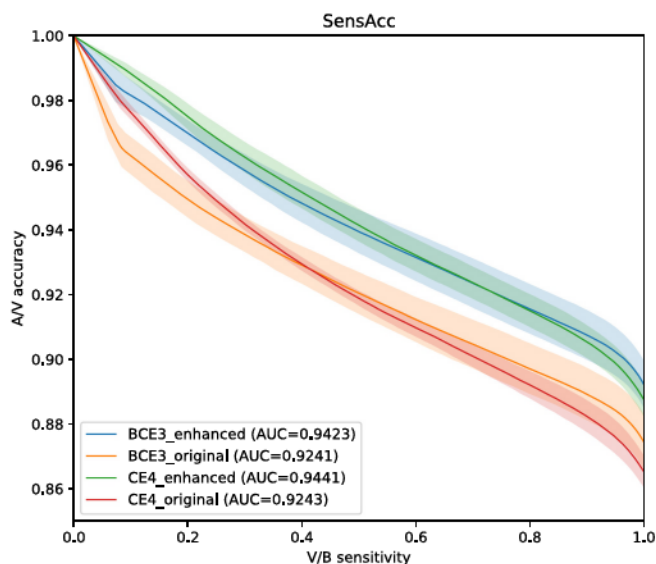

Fig. 12. Vessels classification sensitivity against artery/vein classification accuracy in RITE-test for multiple thresholds and the different evaluated models.

traditional CE4 loss and the proposed MS with BCE3 loss, perform similarly when evaluated in simultaneous vascular segmentation and artery/vein classification. It is remarkable, however, that for a high vascular segmentation sensitivity the proposed approach achieves a slightly higher artery/vein classification accuracy.

In light of all these results together, it can be stated that the MS approach represents the more convenient alternative, since it is able to detect more vessels with similar A/V classification accuracy.

Another relevant point that can be observed in the different figures is the beneficial effect of the preprocessing for the A/V classification. As can be seen in Table 3 and Fig. 12, the performance of the networks is, in general, better when they are trained using the enhanced images than when they are trained using the original ones. In fact, this improvement can be clearly appreciated in Fig. 13, which depicts the probability maps that were predicted by the networks for a RITE-test retinography. When comparing the predictions of the "enhanced" and the "original" alternatives, it can be observed that some manifest classification errors disappear, an thus the arteries and veins continuity is slightly better preserved. Regarding the vessels segmentation, in light of the results, the preprocessing does not lead to such a significant improvement in the networks performance.

3.1. Comparison with the state of the art

Table 4 shows the comparison with the state-of-the-art works for the joint vessel segmentation and the A/V classification tasks. In this comparison, we include the results that were reported by relevant methods that have been evaluated on the publicly available RITE-test set. However, some of the works do not evaluate the vessel segmentation performance of their methods [38,41], although the proposed methods aim at performing the vascular segmentation. As previously stated, the evaluation of the artery/vein classification task in the state-of-the-art works, and herein adopted for comparison, considers the ground truth vessels. However, this evaluation does not take into account if the networks are not able to detect all the vessels, which is relevant as in a real case scenario only the detected vessel pixels would be classified into artery and vein. In this way, if a model presents a high performance on the A/V classification, but is not able to detect most of the vessels, its overall performance is low. Consequently, an appropriate evaluation of the SSCAV should include not only an assessment of how well the model classifies the vessels, but also an assessment of how well the model detects those vessels.

Moreover, the methods in the state of the art report the results for a fixed operating point system, with a varying preference for false positives or false negatives among the works. As this complicates the direct comparison between the state-of-the-art results, we complement the results in Table 4 with the ROC curves in Fig. 15, for the artery/vein classification and the vascular segmentation tasks. In these graphs, we represent the ROC curves for our proposed system (BCE3 loss and enhanced retinographies as input) along with the point representations of the systems in Table 4. Finally, in order to provide a context for the performance in the vasculature segmentation of the proposed system, we provide a comparison with reference works aiming at the vasculature segmentation in Table 5. Also, we complement the results shown in this table with the ROC curve depicted in Fig. 16, as it was done for the RITE-test results. The reported results use the DRIVE-test set, which is
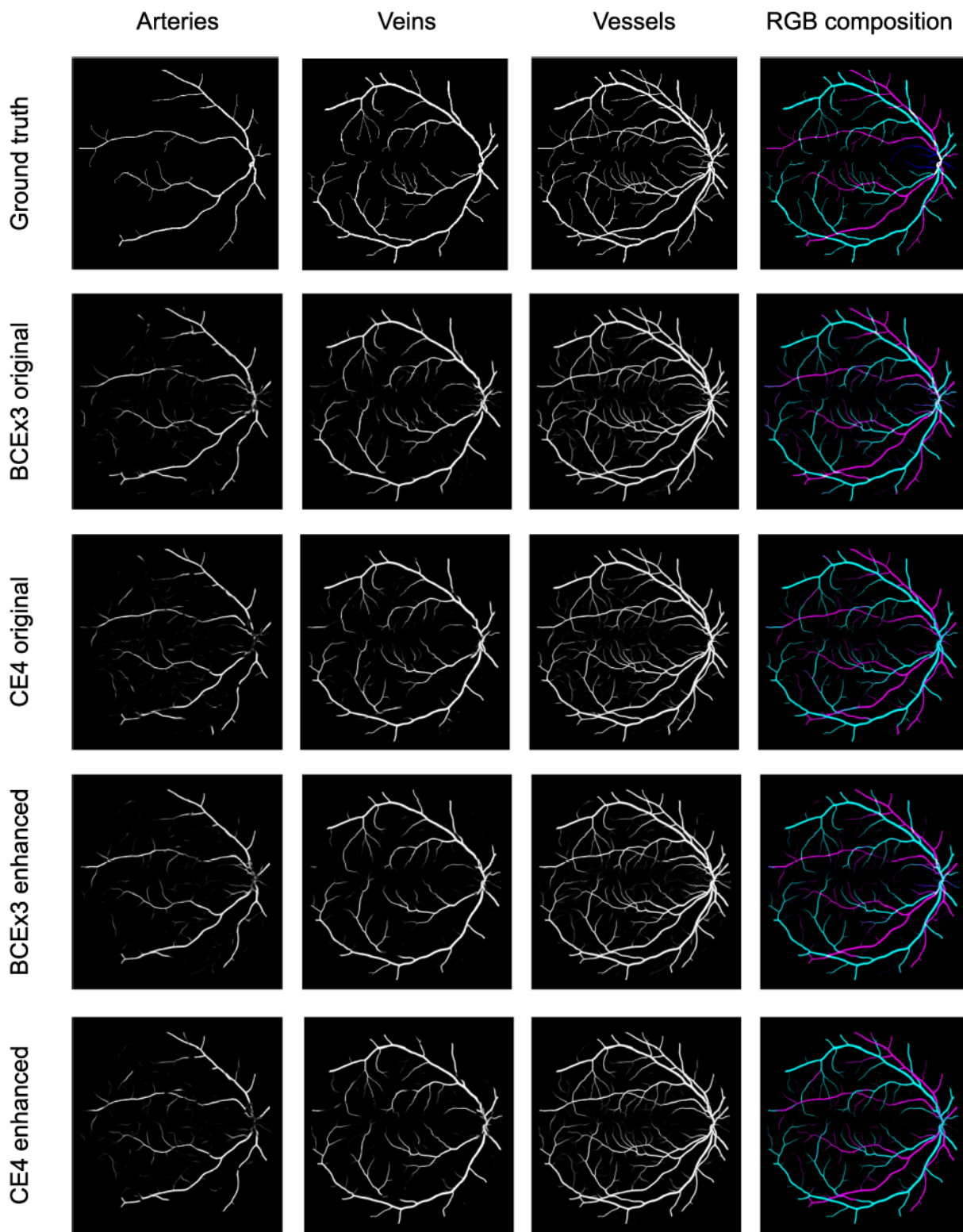
Fig. 13. Examples of generated probability maps of the different classes and its RGB composition.

composed of the same images as in the RITE-test set. However, in this case, the ground truth segmentations considered as gold standard are those provided by the first expert, and thus the obtained results are slightly different from the results reported for RITE.

As it can be observed in the comparison results, the proposed approach achieves highly competitive results in both the artery/vein classification and the vascular segmentation tasks. In this regard, it is especially relevant that our approach achieves the best performance in the vascular segmentation task among the works aiming at solving both SSCAV tasks, and even a competitive performance with the state-of-the-art specialized methods. This is particularly relevant since, in a real A/V classification scenario, only the detected vessels would be classified into arteries and veins. Regarding the artery/vein classification of the ground truth vessels, however, the achieved results are slightly below the most
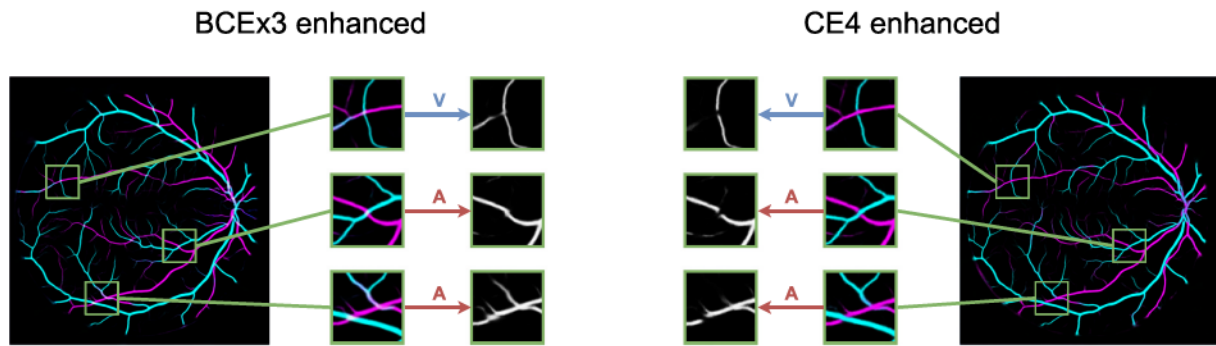
Fig. 14. Examples of arteries, veins and vessels probability maps (RGB compositions) generated by the models trained using the MS and the traditional approaches with preprocessed images.
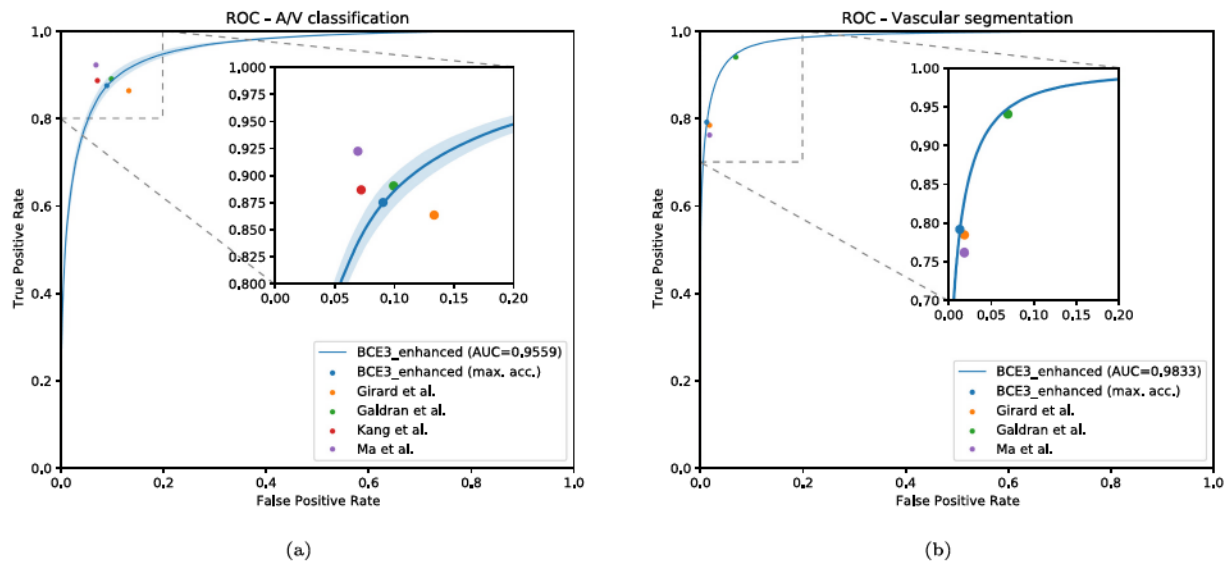

(a)

(b)

Fig. 15. ROC curves in the RITE-test set for our proposed system along with the point representations of the state-of-the-art approaches for the (a) artery/vein classification and (b) vascular segmentation tasks.
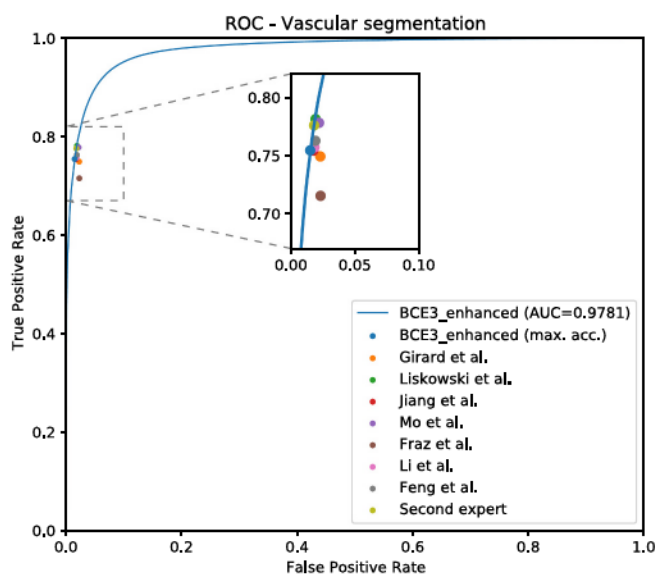

Fig. 16. ROC curve in the DRIVE-test set for our proposed system along with the point representations of the state-of-the-art approaches for the vascular segmentation task.

recent works [40,41].

In order to provide insight on the comparison results, it should be noticed that our method proposes a straightforward application of a plain U-Net architecture [49], to full resolution images, by simply adjusting the used loss. Thus, the research herein described aims at providing methodological advances by simplifying the used methods, gaining insight on why they work, taking advantages of more appropriate ways of posing the end-to-end training of the target tasks, and avoiding the use of any bells and whistles. In this regard, some of the prior works trained and applied the networks to a patch-wise [37,40], instead of using the full-size images as our case. We think that it is more convenient to take full advantage of the FCNN architectures to pose end-to-end training approaches that use full-size images as inputs, even if the patch-wise balancing and the augmentation approaches can take minor advantages on the results, depending on the used datasets. Several methods, including the best performing ones, used ad hoc weights to adjust the importance of each class on the loss [38,40,41], or even heuristic weights to balance the importance of each vessel within each image [40,41]. While we could adjust the class weights, as well as the importance of the difficult target structures, to improve the overall provided results of our approach, we think that such approach requires appropriate validation methodologies that are not usually practical giving the large training times of deep neural networks. There are also interesting ideas which can boost the performance of any approach, and could be potentially applied to our approach in future work, such as the

Table 4

Comparison with the state of the art for the joint vessels segmentation and Artery/Vein classification task in the RITE dataset. The best results are highlighted in bold.

Method	Year	Artery/Vein classification			Vessels segmentation			
		Sens. (%)	Spec. (%)	Acc. (%)	Sens. (%)	Spec. (%)	Acc. (%)	AUC-ROC (%)
Girard et al. [37]	2019	86.3	86.6	86.5	78.4	98.1	95.7	97.2
Galdran et al. [39]	2019	89	90	89	**94**	93	93	95
Kang et al. [41]	2020	88.63	92.72	90.81	–	–	–	–
Ma et al. [40]	2019	**92.2**	**93.0**	92.6	76.16	98.11	95.70	98.10
Hemelings et al. [38]	2019	–	–	**94.87**	–	–	–	–
Proposed	2020	87.47 ± 2.09	90.89 ± 0.68	89.24 ± 0.73	79.12 ± 1.22	**98.65 ± 0.13**	**96.16 ± 0.05**	**98.33 ± 0.04**

Table 5

Comparison with the state of the art exclusively for the vessels segmentation task in DRIVE dataset. Here, the first expert is considered as gold standard. The best results are highlighted in bold.

Method	Sens. (%)	Spec. (%)	Acc. (%)	AUC-ROC (%)
Girard et al. [37]	74.9	97.7	94.8	96.4
Liskowski et al. [24]	**78.11**	98.07	95.35	97.90
Jiang et al. [28]	75.40 ± 0.05	98.25 ± 0.01	96.24 ± 0.01	**98.10**
Mo et al. [67]	77.79	97.80	95.21	97.82
Fraz et al. [68]	71.52	97.68	94.30 ± 0.72	–
Li et al. [69]	75.69	98.16	95.27	97.38
Feng et al. [70]	76.25	98.09	95.28	96.78
Hervella et al. [61]	–	–	–	97.82
Second expert	77.57	98.19	**96.37**	–
Proposed (enhanced)	75.42 ± 1.23	**98.49 ± 0.13**	95.45 ± 0.05	97.81 ± 0.02
Proposed (original)	76.46 ± 0.52	98.36 ± 0.07	95.54 ± 0.02	97.81 ± 0.03

use of pretrained networks [40,41] (ours are trained from scratch), enforcing the vascular tree classification coherence using an ad hoc postprocessing approach [37,41], or the use of deep supervision to enhance the backpropagation of gradients [40,41]. Notwithstanding we did not include any of these mechanisms, the achieved results are highly competitive in artery/vein classification, and certainly better in the vascular segmentation, which emphasizes the potential of the proposed approach.

The only additional mechanism from the state of the art that we included in our approach is the ad hoc preprocessing. Initially, however, our intention was not to include it, since we consider that any neural network with an adequate capacity should be able to learn such simple preprocessing. Notwithstanding, its wide use in the literature and its significant impact in the classification results made us to finally incorporate one of these methods. In this sense, it is singular the contrast between the important improvement in classification and the almost null impact in segmentation. This observation could motivate future research.

Finally, it should be noticed that the proposed method presents an additional advantage over the rest of the state-of-the-art approaches: it allows to detect vessel crossings with a satisfactory performance. This aspect is distinctive of our work, since there is no other in the state of the art that allows to perform this task. Additionally, our approach gives raise to fully-connected arterial and venular trees. This factor facilitates the structural analysis of both types of vessels, and it is useful for the future application of methods for addressing the vascular coherence issue. Furthermore, it remarkable that these advantages are achieved with a simple change in the loss; a change that also simplifies the approach of the problem and significantly improves the performance on the vascular segmentation task.

4. Conclusions

In this work, we have proposed a novel approach for the simultaneous segmentation and classification of the retinal arteries and veins (SSCAV) using FCNNs. This approach decomposes this joint task into three segmentation problems: the segmentation of arteries, veins and the whole vascular tree. To train the networks following this approach, we have proposed a novel loss named BCE3, that combines the independent segmentation losses of the three classes of interest.

To evaluate the potential and advantages of this approach, we performed a comprehensive comparative study with the approach that currently represents the state of the art in the simultaneous segmentation and classification task. Contrary to ours, this approach uses Cross-Entropy as loss. Furthermore, we also compared the results achieved by our method with the current state-of-the-art results for both the SSCAV and exclusively the vascular segmentation task in the RITE and DRIVE datasets.

The results provided by the different experiments demonstrate that our proposal achieves a competitive performance in the classification of arteries and veins, as well as significantly improving the retinal vasculature segmentation. Complementarily, it allows to detect the complex vessel crossings with a satisfactory performance.

On the other hand, the comparison with the state of the art in the SSCAV and the vasculature segmentation tasks shows that our method achieves highly competitive results. In the SSCAV, it is remarkable that it surpasses the segmentation results of all the state-of-the-art approaches whereas preserving a high classification performance in comparison with the most recent methods. This aspect is valuable, as the current state of the art evaluate the vasculature classification on all the vessel pixels of the ground truth, but, in a real scenario, only the detected vessel pixels would be classified into arteries and veins. In this sense, the greater vessel detection effectiveness of our method represents a potential advantage. Furthermore, it is worth noting that these results are achieved with a direct adaptation and design of the used loss, and avoiding the use of any bells and whistles. In this regard, some of the mechanisms included in the most recent works (e.g. transfer learning, deep supervision or ad hoc postprocessing) are compatible with our approach, and could probably improve its results.

In addition to the above, our method presents some inherent advantages. First, it is simple, since it consists of the straightforward application of a FCNN with a custom loss. Second, it provides continuous arterial and venular trees, also directly providing vasculature segmentation masks. Third, it allows to detect vessel crossings with an adequate effectiveness. These advantages are distinctive of our approach regarding to the current state of the art.

Notwithstanding, our approach also presents some potential aspects for further improvement. One of them, which also affects to the rest of the works that address the SSCAV, is the vessels classification incoherence. That is, the presence of manifestly misclassified segments of vessels in the output images. Although some works have proposed methods to mitigate these errors (see Section 1.1), they do not provide a fully learning-based approach capable of solving this problem for the most part of examples. With the current state-of-the-art methodologies, the classification between arteries and veins using FCNNs still relies on

relatively low-level features. Thus, the problem of getting the networks to learn the high-level structure of the arteries and veins is still unsolved. Along with this issue, there is also room for improvement in the state of the art regarding another aspect: the image preprocessing. Although this technique demonstrated a positive impact in the A/V classification results, it does not improve the vascular segmentation. Also, it constitutes an ad hoc stage that does not allow to train the networks in a truly end-to-end setting (from the original data to the target decisions). These two issues represent interesting fields for further research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] Kanski JJ, Bowling B. Clinical ophthalmology: a systematic approach. 7th ed. Elsevier Health Sciences; 2011.

[2] Abràmoff MD, Garvin MK, Sonka M. Retinal imaging and image analysis. IEEE Rev Biomed Eng 2010;3:169–208. https://doi.org/10.1109/RBME.2010.2084567.

[3] Rosenfeld PJ, Brown DM, Heier JS, Boyer DS, Kaiser PK, Chung CY, et al. Ranibizumab for neovascular age-related macular degeneration. N Engl J Med 2006;355(14):1419–31. https://doi.org/10.1056/NEJMoa054481 (pMID: 17021318).

[4] da Silva AVB, Gouvea SA, da Silva APB, Bortolon S, Rodrigues AN, Abreu GR, et al. Changes in retinal microvascular diameter in patients with diabetes. Int J Gen Med 2015;8:267–73. https://doi.org/10.2147/IJGM.S83749.

[5] Vandewalle E, Pinto L Abegão, Olafsdottir OB, De Clerck E, Stalmans P, Van Calster J, et al. Oximetry in glaucoma: correlation of metabolic change with structural and functional damage. Acta Ophthalmol 2014;92(2):105–10. https://doi.org/10.1111/aos.12011.

[6] Ramm L, Jentsch S, Peters S, Augsten R, Hammer M. Investigation of blood flow regulation and oxygen saturation of the retinal vessels in primary open-angle glaucoma. Graefes Arch Clin Exp Ophthalmol 2014;252(11):1803–10. https://doi.org/10.1007/s00417-014-2766-4.

[7] Wang S, Xu L, Wang Y, Wang Y, Jonas JB. Retinal vessel diameter in normal and glaucomatous eyes: the Beijing eye study. Clin Exp Ophthalmol 2007;35(9):800–7. https://doi.org/10.1111/j.1442-9071.2007.01627.x.

[8] Mitchell P, Leung H, Wang J, Rochtchina E, Lee A, Wong T, et al. Retinal vessel diameter and open-angle glaucoma: the blue mountains eye study. Ophthalmology 2005;112(2):245–50. https://doi.org/10.1016/j.ophtha.2004.08.015 (cited By 162).

[9] Mathers C, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006;3(11):2011–30. https://doi.org/10.1371/journal.pmed.0030442 (cited By 5959).

[10] Ross R. Atherosclerosis - an inflammatory disease. N Engl J Med 1999;340(2):115–26. https://doi.org/10.1056/NEJM199901143400207 (cited By 17782).

[11] Prokofyeva E, Zrenner E. Epidemiology of major eye diseases leading to blindness in europe: a literature review. Ophthalmic Res 2012;47(4):171–88. https://doi.org/10.1159/000329603.

[12] Staal J, Abràmoff MD, Niemeijer M, Viergever MA, Ginneken B van. Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging 2004;23(4):501–9. https://doi.org/10.1109/TMI.2004.825627.

[13] Patton N, Aslam T, Macgillivray T, Pattie A, Deary IJ, Dhillon B. Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures. J Anat 2005;206(4):319–48. https://doi.org/10.1111/j.1469-7580.2005.00395.x.

[14] Fraz M, Remagnino P, Hoppe A, Uyyanonvara B, Rudnicka A, Owen C, et al. Blood vessel segmentation methodologies in retinal images – a survey. Comput Methods Prog Biomed 2012;108(1):407–33. https://doi.org/10.1016/j.cmpb.2012.03.009.

[15] Kawasaki R, Cheung N, Wang JJ, Klein R, Klein BE, Cotch MF, et al. Retinal vessel diameters and risk of hypertension: the multiethnic study of atherosclerosis. J Hypertens 2009;27(12):2386–93. https://doi.org/10.1097/HJH.0b013e3283310f7e.

[16] Ikram MK, de Jong FJ, Vingerling JR, Witteman JCM, Hofman A, Breteler MMB, et al. Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? The Rotterdam study. Invest Ophthalmol Vis Sci 2004;45(7):2129–34. https://doi.org/10.1167/iovs.03-1390.

[17] Sun C, Wang JJ, Mackey DA, Wong TY. Retinal vascular caliber: systemic, environmental, and genetic associations. Surv Ophthalmol 2009;54(1):74–95. https://doi.org/10.1016/j.survophthal.2008.10.003.

[18] Y. Hatanaka, T. Nakagawa, A. Aoyama, X. Zhou, T. Hara, H. Fujita, M. Kakogawa, Y. Hayashi, Y. Mizukusa, A. Fujita, Automated detection algorithm for arteriolar narrowing on fundus images, in: 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, 2005, pp. 286–289. doi:https://doi.org/10.1109/IEMBS.2005.1616400.

[19] Jiang X, Mojon D. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. IEEE Trans Pattern Anal Mach Intell 2003;25(1):131–7. https://doi.org/10.1109/TPAMI.2003.1159954.

[20] Nain D, Yezzi A, Turk G. Vessel segmentation using a shape driven flow. In: Medical image computing and computer-assisted intervention – MICCAI. vol. 3216. LNCS; 2004. p. 51–9. https://doi.org/10.1007/978-3-540-30135-6_7.

[21] Tolias YA, Panas SM. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. IEEE Trans Med Imaging 1998;17(2):263–73. https://doi.org/10.1109/42.700738.

[22] Sinthanayothin C, Boyce JF, Cook HL, Williamson TH. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. Br J Ophthalmol 1999;83(8):902–10. https://doi.org/10.1136/bjo.83.8.902.

[23] Marín D, Aquino A, Gegundez-Arias ME, Bravo JM. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Trans Med Imaging 2011;30(1):146–58. https://doi.org/10.1109/TMI.2010.2064333.

[24] Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks. IEEE Trans Med Imaging 2016;35(11):2369–80. https://doi.org/10.1109/TMI.2016.2546227.

[25] Fu H, Xu Y, Lin S, Wong DW Kee, Liu J. DeepVessel: retinal vessel segmentation via deep learning and conditional random field. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. Medical image computing and computer-assisted intervention – MICCAI 2016. Cham: Springer International Publishing; 2016. p. 132–9. https://doi.org/10.1007/978-3-319-46723-8_16.

[26] Fu H, Xu Y, Wong DWK, Liu J. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI); 2016. p. 698–701. https://doi.org/10.1109/ISBI.2016.7493362.

[27] Dasgupta A, Singh S. A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI); 2017. p. 248–51. https://doi.org/10.1109/ISBI.2017.7950512.

[28] Jiang Z, Zhang H, Wang Y, Ko S-B. Retinal blood vessel segmentation using fully convolutional network with transfer learning. Comput Med Imaging Graph 2018;68:1–15. https://doi.org/10.1016/j.compmedimag.2018.04.005.

[29] Oliveira A, Pereira S, Silva CA. Retinal vessel segmentation based on fully convolutional neural networks. Expert Syst Appl 2018;112:229–42. https://doi.org/10.1016/j.eswa.2018.06.034.

[30] Jin Q, Meng Z, Pham TD, Chen Q, Wei L, Su R. DUNet: A deformable network for retinal vessel segmentation. Knowl-Based Syst 2019;178:149–62. https://doi.org/10.1016/j.knosys.2019.04.025.

[31] Relan D, MacGillivray T, Ballerini L, Trucco E. Retinal vessel classification: sorting arteries and veins. In: 2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2013. p. 7396–9. https://doi.org/10.1109/EMBC.2013.6611267.

[32] 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; 2014. p. 142–5. https://doi.org/10.1109/EMBC.2014.6943549.

[33] Dashtbozorg B, Mendonça AM, Campilho A. An automatic graph-based approach for artery/vein classification in retinal images. IEEE Trans Image Process 2014;23(3):1073–83. https://doi.org/10.1109/TIP.2013.2263809.

[34] Estrada R, Allingham MJ, Mettu PS, Cousins SW, Tomasi C, Farsiu S. Retinal artery-vein classification via topology estimation. IEEE Trans Med Imaging 2015;34(12):2518–34. https://doi.org/10.1109/TMI.2015.2443117.

[35] Welikala R, Foster P, Whincup P, Rudnicka A, Owen C, Strachan D, et al. Automated arteriole and venule classification using deep learning for retinal images from the uk biobank cohort. Comput Biol Med 2017;90:23–32. https://doi.org/10.1016/j.compbiomed.2017.09.005.

[36] Xu X, Wang R, Lv P, Gao B, Li C, Tian Z, et al. Simultaneous arteriole and venule segmentation with domain-specific loss function on a new public database. Biomed Opt Express 2018;9(7):3153–66. https://doi.org/10.1364/BOE.9.003153.

[37] Girard F, Kavalec C, Cheriet F. Joint segmentation and classification of retinal arteries/veins from fundus images. Artif Intell Med 2019;94:96–109. https://doi.org/10.1016/j.artmed.2019.02.004.

[38] Hemelings R, Elen B, Stalmans I, Keer K Van, Boever P De, Blaschko MB. Artery–vein segmentation in fundus images using a fully convolutional network. Comput Med Imaging Graph 2019;76:101636. https://doi.org/10.1016/j. compmedimag.2019.05.004.

[39] Galdran A, Meyer M, Costa P, MendonÇa A Campilho. Uncertainty-aware artery/ vein classification for retinal images. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019); 2019. p. 556–60. https://doi.org/10.1109/ ISBI.2019.8759380.

[40] Ma W, Yu S, Ma K, Wang J, Ding X, Zheng Y. Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, et al., editors. Medical image computing and computer assisted intervention – MICCAI 2019. Cham: Springer International Publishing; 2019. p. 769–78. https://doi.org/10.1007/978-3-030-32239-7_85.

[41] Kang H, Gao Y, Guo S, Xu X, Li T, Wang K. AVNet: A retinal artery/vein classification network with category-attention weighted fusion. Comput Methods Prog Biomed 2020;195:105629. https://doi.org/10.1016/j.cmpb.2020.105629.

[42] Hu Q, Abràmoff MD, Garvin MK. Automated separation of binary overlapping trees in low-contrast color retinal images. In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N, editors. Medical image computing and computer-assisted intervention – MICCAI 2013. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 436–43. https://doi.org/10.1007/978-3-642-40763-5_54.

[43] Qureshi TA, Habib M, Hunter A, Al-Diri B. A manually-labeled, artery/vein classified benchmark for the drive dataset. In: Proceedings of the 26th IEEE international symposium on computer-based medical systems; 2013. p. 485–8. https://doi.org/10.1109/CBMS.2013.6627847.

[44] Orlando JI, Breda J Barbosa, Keer K van, Blaschko MB, Blanco PJ, Bulant CA. Towards a glaucoma risk index based on simulated hemodynamics from fundus images. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Cham: Springer International Publishing; 2018. p. 65–73. https://doi.org/ 10.1007/978-3-030-00934-2_8.

[45] Rothaus K, Jiang X, Rhiem P. Separation of the retinal vascular graph in arteries and veins based upon structural knowledge. Image Vis Comput 2009;27(7): 864–75. https://doi.org/10.1016/j.imavis.2008.02.013 (7th IAPR-TC15 Workshop on Graph-based Representations (GbR 2007)).

[46] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998:2278–324. https://doi.org/10.1109/ 5.726791.

[47] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th international conference on neural information processing systems - volume 1, NIPS'12. Red Hook, NY, USA: Curran Associates Inc; 2012. p. 1097–105. https://doi.org/ 10.1145/3065386.

[48] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: 2015 IEEE conference on Computer Vision and Pattern Recognition (CVPR); 2015. p. 3431–40. https://doi.org/10.1109/ CVPR.2015.7298965.

[49] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer International Publishing; 2015. p. 234–41. https://doi.org/ 10.1007/978-3-319-24574-4_28.

[50] Zamperini A, Giachetti A, Trucco E, Chin KS. Effective features for artery-vein classification in digital fundus images. In: 2012 25th IEEE international symposium on Computer-Based Medical Systems (CBMS); 2012. p. 1–6. https://doi.org/ 10.1109/CBMS.2012.6266336.

[51] Saez M, González-Vázquez S, González-Penedo M, Barceló MA, Pena-Seijo M, de Tuero G Coll, et al. Development of an automated system to classify retinal vessels into arteries and veins. Comput Methods Prog Biomed 2012;108(1):367–76. https://doi.org/10.1016/j.cmpb.2012.02.008.

[52] Hao S, Zhou Y, Guo Y. A brief survey on semantic segmentation with deep learning. Neurocomputing 2020;406:302–21. https://doi.org/10.1016/j. neucom.2019.11.118.

[53] Szegedy C, Liu Wei, Yangqing Jia P Sermanet, Reed S, Anguelov D, Erhan D, et al. Going deeper with convolutions. In: 2015 IEEE conference on Computer Vision and Pattern Recognition (CVPR); 2015. p. 1–9. https://doi.org/10.1109/ CVPR.2015.7298594.

[54] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 770–8. https://doi.org/10.1109/CVPR.2016.90.

[55] Zhu J-Y, Zhang R, Pathak D, Darrell T, Efros AA, Wang O, et al. Toward multimodal image-to-image translation. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. Advances in neural information processing systems 30. Curran Associates, Inc; 2017. p. 465–76.

[56] Isola P, Zhu J, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: 2017 IEEE conference on Computer Vision and Pattern Recognition (CVPR); 2017. p. 5967–76. https://doi.org/10.1109/CVPR.2017.632.

[57] Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, et al. U-net: deep learning for cell counting, detection, and morphometry. Nat Methods 2019;16(1): 67–70. https://doi.org/10.1038/s41592-018-0261-2.

[58] Fu H, Cheng J, Xu Y, Wong DWK, Liu J, Cao X. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Trans Med Imaging 2018;37(7):1597–605. https://doi.org/10.1109/ TMI.2018.2791488.

[59] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88. https://doi.org/10.1016/j.media.2017.07.005.

[60] Morano J, Hervella ÁS, Barreira N, Novo J, Rouco J. Multimodal transfer learning-based approaches for retinal vascular segmentation. In: European Conference on Artificial Intelligence (ECAI), vol. 325 of frontiers in artificial intelligence and applications; 2020. p. 1866–73. https://doi.org/10.3233/FAIA200303.

[61] Hervella Álvaro S, Rouco J, Novo J, Ortega M. Learning the retinal anatomy from scarce annotated data using self-supervised multimodal reconstruction. Appl Soft Comput 2020;91:106210. https://doi.org/10.1016/j.asoc.2020.106210.

[62] Hervella Álvaro S, Rouco J, Novo J, Ortega M. Self-supervised multimodal reconstruction of retinal images over paired datasets. Expert Syst Appl 2020;161: 113674. https://doi.org/10.1016/j.eswa.2020.113674.

[63] Kingma DP, Ba J. Adam: a method for stochastic optimization. In: 3rd international conference on learning representations, ICLR, San Diego, CA, USA, May 7–9, 2015, conference track proceedings; 2015.

[64] Hervella Álvaro S, Rouco J, Novo J, Penedo MG, Ortega M. Deep multi-instance heatmap regression for the detection of retinal vessel crossings and bifurcations in eye fundus images. Comput Methods Prog Biomed 2020;186:105201. https://doi. org/10.1016/j.cmpb.2019.105201.

[65] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), ICCV, Washington, DC, USA, 2015, pp. 1026–1034. doi:https://doi.org/10.1109/ICCV.2015.123.

[66] Davis J, Goadrich M. The relationship between precision-recall and roc curves. In: Proceedings of the 23rd international conference on machine learning, ACM, ICML '06, Association for Computing Machinery, New York, NY, USA; 2006. p. 233–40. https://doi.org/10.1145/1143844.1143874.

[67] Mo J, Zhang L. Multi-level deep supervised networks for retinal vessel segmentation. Int J Comput Assist Radiol Surg 2017;12(12):2181–93. https://doi. org/10.1007/s11548-017-1619-0.

[68] Fraz M, Barman S, Remagnino P, Hoppe A, Basit A, Uyyanonvara B, et al. An approach to localize the retinal blood vessels using bit planes and centerline detection. Comput Methods Prog Biomed 2012;108(2):600–16. https://doi.org/ 10.1016/j.cmpb.2011.08.009.

[69] Li Q, Feng B, Xie L, Liang P, Zhang H, Wang T. A cross-modality learning approach for vessel segmentation in retinal images. IEEE Trans Med Imaging 2016;35(1): 109–18. https://doi.org/10.1109/TMI.2015.2457891.

[70] Feng S, Zhuo Z, Pan D, Tian Q. Ccnet: a cross-connected convolutional network for segmenting retinal vessels using multi-scale features. Neurocomputing 2020;392: 268–76. https://doi.org/10.1016/j.neucom.2018.10.098.