

Differential Artery-Vein Analysis in OCT Angiography of Diabetic Retinopathy

BY

MATTIA CASTELNUOVO

Laurea Triennale, Politecnico di Milano, Milan, Italy, 2018

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Science in Bioengineering
in the Graduate College of the
University of Illinois at Chicago, 2021

Chicago, Illinois

Defense Committee:

Xincheng Yao, Chair and Advisor

James Lee, UIC

Enrico Gianluca Caiani, Politecnico di Milano

ACKNOWLEDGEMENTS

Firstly, I want to thank my UIC advisor Dr. Xincheng Yao for accepting me into his lab and helping me develop my research skills. I also want to thank all of my colleagues of the LIERI Lab for helping me and advising me during my permanence at UIC.

Secondly, I want to thank Prof. Enrico Caiani for his helpful advices and comments that helped me to craft a more complete study.

Thanks to Jenna Stephens, who helped all of us during every step of this journey.

Thanks to Judy and Daniel West, who always made me feel like at home.

Thanks to my family, who always supported me, and especially my parents, who made this journey possible in the first place.

Special thanks to Giulia, whose positiveness, warmth and sincere love helped me and still help me every day.

MC

TABLE OF CONTENTS

<u>CHAPTER</u>		<u>PAGE</u>
1	BACKGROUND	2
	1.1 Diabetic Retinopathy	2
	1.2 Blood Vessel-related Biomarkers for DR	3
	1.3 Fundus Imaging	4
	1.4 OCTA: the new standard?	6
	1.5 Artery-Vein Differentiation	7
2	MATERIALS AND METHODS	9
	2.1 Composition of the Dataset	9
	2.2 Classification of DR stages	10
	2.3 From OCTA image to binary map	10
	2.4 Artery-Vein Differentiation	13
	2.4.1 Bias Field Correction	15
	2.4.2 OCT Vessel Map Extraction	17
	2.4.3 Source Node Identification	20
	2.4.4 Profiles Extraction	21
	2.4.5 Feature Extraction and Classification of the Source Nodes	22
	2.4.6 Tracking the Vessel from Source to End Nodes ...	24
	2.4.7 Classifying the OCTA Binary Vessel Map	28
	2.5 Feature Extraction	29
	2.5.1 Separating Arteries from Veins	30
	2.5.2 Identifying the Branching Triangle	31
	2.5.3 Features Measurement	33
	2.6 Feature Sets	35
	2.7 Statistical Analysis	36
3	RESULTS	37
	3.1 Demographics	38
	3.2 ANOVAs	38
	3.2.1 Artery ANOVA	39
	3.2.2 Vein ANOVA	40
	3.2.3 A/V Ration ANOVA	41
	3.2.4 Weighted Average ANOVA	42
	3.3 Boxplots	43
	3.3.1 VBA Boxplot	43
	3.3.2 CBA1 Boxplot	44
	3.3.3 CBA2 Boxplot	45
	3.3.4 VBC Boxplot	46
	3.3.5 VWR1 Boxplot	47
	3.3.6 VWR2 Boxplot	48

TABLE OF CONTENTS (continued)

<u>CHAPTER</u>	<u>PAGE</u>
3.4 Student t-tests	49
3.4.1 Bonferroni Correction	49
3.4.2 Weighted Average CBA1 Individual Comparison t-tests	50
3.4.3 Weighted Average VBC Individual Comparison t-tests	50
3.4.4 Weighted Average VWR1 Individual Comparison t- tests	51
3.4.5 Artery VBC Individual Comparison t-tests	52
3.4.6 Artery VWR1 Individual Comparison t-tests	52
3.4.7 Vein VWR1 Individual Comparison t-tests	53
4 DISCUSSION	54
4.1 Comparison with the previous study	54
4.1.1 Differences between Angle- and Width-based Features	56
4.2 Sources of Differences	57
4.2.1 Differences in Vessel Map Generation	57
4.2.2 Correspondence of Analyzed Branchpoints	57
4.2.2.1 Example: a very common scenario	59
4.2.3 The Influence of Measurement Windows	61
4.3 Why the Bonferroni Correction	61
5 CONCLUSIVE REMARKS	63
APPENDIX	64
CITED LITERATURE	65
VITA	71

LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
I DEMOGRAPHICS OF CONTROL AND DR SUBJECTS	37
II STATISTICS OF ARTERY GEOMETRIC FEATURES	39
III STATISTICS OF VEIN GEOMETRIC FEATURES	40
IV STATISTICS OF A/V RATIO GEOMETRIC FEATURES	41
V STATISTICS OF WEIGHTED AVERAGE GEOMETRIC FEATURES	42
VI STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR WEIGHTED AVERAGE CBA1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST . . .	50
VII STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR WEIGHTED AVERAGE VBC, DUE THE SIGNIFICANCE OF THE ANOVA TEST	50
VIII STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR WEIGHTED AVERAGE VWR1, DUE THE SIGNIFICANCE OF THE ANOVA TEST	51
IX STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR ARTERY VBC, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST	52
X STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR ARTERY VWR1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST	52
XI STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR VEIN VWR1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST	53

LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1	OCTA images of a Healthy eye (left) and a Severe NPDR eye (right) from two different patients	3
2	Fundus Image with superimposed OCTA image covering an area of 6x6 mm	5
3	En-Face OCT image (left) and corresponding OCTA image (right) of a Healthy eye	7
4	Flowchart showing the broad steps of the Feature Extraction	11
5	Detailed description of the Geometric Feature Computation process	12
6	An overview of the Vessel Map generation	13
7	Flowchart showing the broad steps of the Artery-Vein Classification	14
8	Detailed description of the Source Node Classification process	15
9	Image Normalization steps for En-Face OCT	16
10	Image Normalization steps for Outer Layer OCT	17
11	En-Face OCT (left) and Outer Layer OCT (right) images of the same eye	18
12	OCT Vessel Map generation steps	19
13	Assessment of the Vessel Map extraction by superimposition with the Intensity Normalized En-Face OCT image	20
14	Source Node Classification steps	21
15	Normalized feature distribution in artery and vein vessels .	22
16	Overview of all the rotations of the quadrants relative to the center of the fovea	24
17	Core steps of the Blood Vessel Tracking algorithm	25
18	Main steps of the Blood Vessel Tracking algorithm	26
19	Artery-Vein classification in OCTA	29
20	Separation of the OCTA artery-vein map into the OCTA artery and vein maps	30
21	Overview of the Branching Triangle Identification and Features Measurement	32
22	Boxplot of the VBA distributions among all patient cohorts and all feature sets	43
23	Boxplot of the CBA1 distributions among all patient cohorts and all feature sets	44
24	Boxplot of the CBA2 distributions among all patient cohorts and all feature sets	45
25	Boxplot of the VBC distributions among all patient cohorts and all feature sets	46

LIST OF FIGURES (continued)

<u>FIGURE</u>		<u>PAGE</u>
26	Boxplot of the VWR1 distributions among all patient cohorts and all feature sets	47
27	Boxplot of the VWR2 distributions among all patient cohorts and all feature sets	48
28	False Branchpoint scenario	59
29	Non-detection of real branchpoints	60

ABBREVIATIONS

DR	Diabetic Retinopathy
DM	Diabetes Mellitus
NoDR	Patients with Diabetes but not Diabetic Retinopathy
NPDR	Non-Proliferative Diabetic Retinopathy
OCT	Optical Coherence Tomography
OCTA	Optical Coherence Tomography - Angiography
VBA	Vessel Branching Angle
CBA	Child Branching Angle
VBC	Vessel Branching Coefficient
VWR	Vessel Width Ratio

SUMMARY

Given the rising prevalence of Diabetes and thus Diabetic Retinopathy in the world, developing better techniques for early detection is paramount. This study aims to build upon previous studies that used geometric quantitative features in OCTA images to classify Non-Proliferative Diabetic Retinopathy (NPDR). Six geometric features, three angle-related (VBA, CBA1 and CBA2) and three width-related (VBC, VWR1 and VWR2), were automatically calculated in four distinct types of analyses made possible by the classification of OCTA vessel structures in arteries and veins using a clustering technique built around four features derived from the same vessels in the relative OCT images. Arteries' and veins' features were calculated separately, their ratios were computed and then their separate measurements were put together in a "Weighted Average" method. Comparative analysis of healthy patients, patients with diabetes but no retinopathy (NoDR) and patients with various stages of NPDR was conducted. This study found that sensitivity increased in width-related measurements and decreased in angle-related ones compared with the previous study. Some explanations are proposed as causes of this difference. Also, it was observed that the ratio measurements between artery values and vein values do not seem to contribute to the classification of patients and data seem to suggest that arteries and veins may be affected in circa the same way by NPDR with respect to the analyzed six features. VBC and VWR1 showed very good ability to separate between Control and NoDR patients from the NPDR patients even with the most conservative Bonferroni Correction.

AIMS

Blood vessel branchpoint analysis in Optical Coherence Tomography Angiography (OCTA) has been demonstrated for objective classification of Diabetic Retinopathy [4]. The aim of this work is to test if the branchpoint geometry features are dependent on artery and vein types, and thus to verify if differential artery and vein analysis can improve the sensitivity of OCTA detection of geometric feature distortion due to DR.

CHAPTER 1

BACKGROUND

1.1 Diabetic Retinopathy

Diabetic Retinopathy (DR) is a form of retinopathy found in Diabetes Mellitus (DM) patients and is the leading cause of blindness in people aged 20 to 64 in the US [1]. In 2004 it was estimated that around 40% of diabetic patients aged 40+ developed this retinopathy, bringing the number of affected people to more than 4 million Americans [2]. Worldwide, it was predicted that the people suffering from diabetes-related retinopathies will increase to close to 600 million by 2025 [3]. This will inevitably cause a strain of the healthcare systems around the developing countries especially. The early detection of this illness will save resources and will prevent partially or totally the related loss of vision [3].

The increased glucose concentration in the blood caused by DM can block the blood vessels in the retina and eventually break their walls. In the beginning, this condition is called NPDR (Non-Proliferative Diabetic Retinopathy), because the partially blocked vessels are not replaced by new ones. This leads to swelling of the nerve fibers and ever-increasing leaks into the retina. When the blood vessels begin to close due to the blockages, a more severe retinopathy stage is reached. Figure 1 shows the difference between a Healthy eye and an eye with Severe NPDR. New blood

vessels will be prompted to form but they will not be properly developed and the leakings will increase. Pressure will gradually build in the eye, leading to optic nerve damage, which will in turn lead to vision damage. Moreover, the retina can begin to detach from the back of the eye. Since the early detection of DR is paramount, the NPDR stage must be clearly diagnosable in order to reduce vision damage to the minimum.

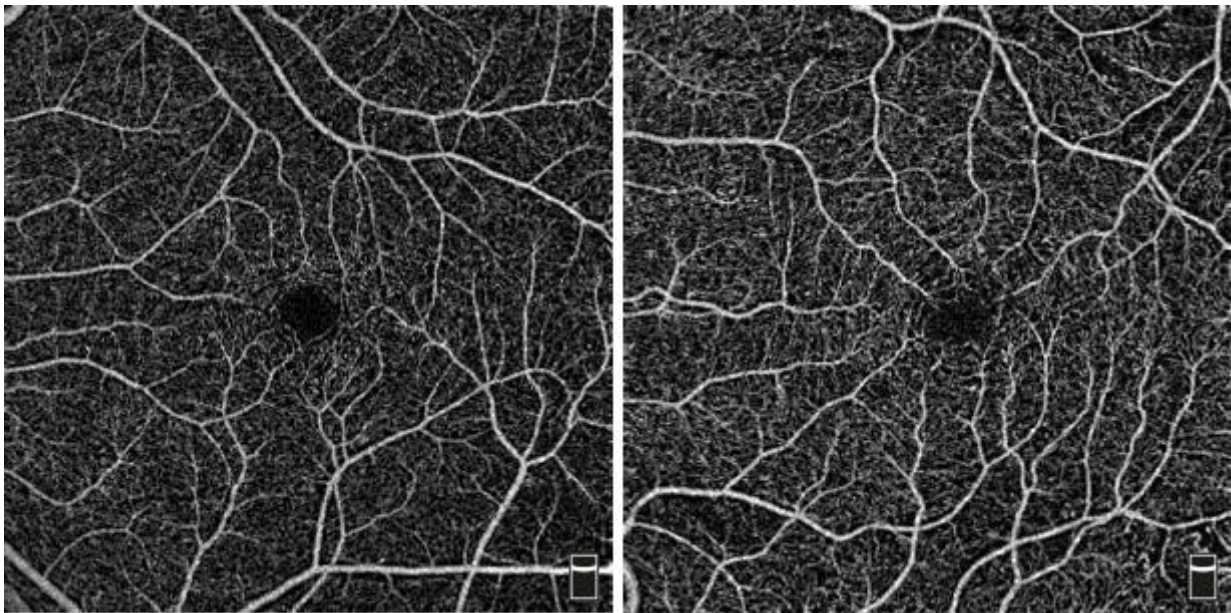


Figure 1: OCTA images of a Healthy eye (left) and a Severe NPDR eye (right) from two different patients

1.2 Blood Vessel-related Biomarkers for DR

In 1926, Murray analyzed the problem of oxygen transport and blood flow [5]. His reasoning relied on the hypothesis that blood flow should be as economic as possible in energetic terms. So, the blood vessels should have a physiological standard in terms of radius, number, branching, etc. Should these vessels deviate from this pattern, oxygen circulation will be impaired. In 1983, Ernest et Al. studied oxygen autoregulation in dogs with hyperglycemia and found that it impaired oxygen autoregulation [6]. In a retinal circulation study from 1996, DM patients were observed

to have a larger total volumetric blood flow rate (increasing with the illness duration) and “decreased regulatory responses to hyperoxia” with respect to control patients [7]. Advanced digital imaging has greatly helped this field of study. In 2010 a study performed on 1159 subjects in Sydney found correlations between DM (and DR) and abnormalities in the retinal vessel pattern [8]: “Longer diabetes duration was associated with larger arteriolar branching angle and increased arteriolar [tortuosity]” and higher blood glycemia “was associated with increased arteriolar tortuosity”. These changes occur before the retinopathy is present.

1.3 Fundus Imaging

“Fundus oculi” is Latin for “rear of the eye”. Fundus Imaging is therefore the imaging of the back of the eye, that is the retina. This technique has existed for over a century, and its major improvement came with the development of digital imaging. These images are taken with a special camera called a Fundus Camera, which photographs the retina directly using the pupil as both entrance and exit for its illuminating and imaging light rays [13]. In a Fundus Image, the macula, the optic disk and the blood vessels are clearly distinguishable. Moreover, the employment of RGB colored filters results in images in full-color and the use of particular fluorescent dyes increase the contrast of specific structures allowing for easier identification and segmentation. Figure 2 shows a Fundus Image with the respective 6x6 mm OCTA image superimposed, which is centered around the Macula. The optic disk is visible in the Fundus Image.

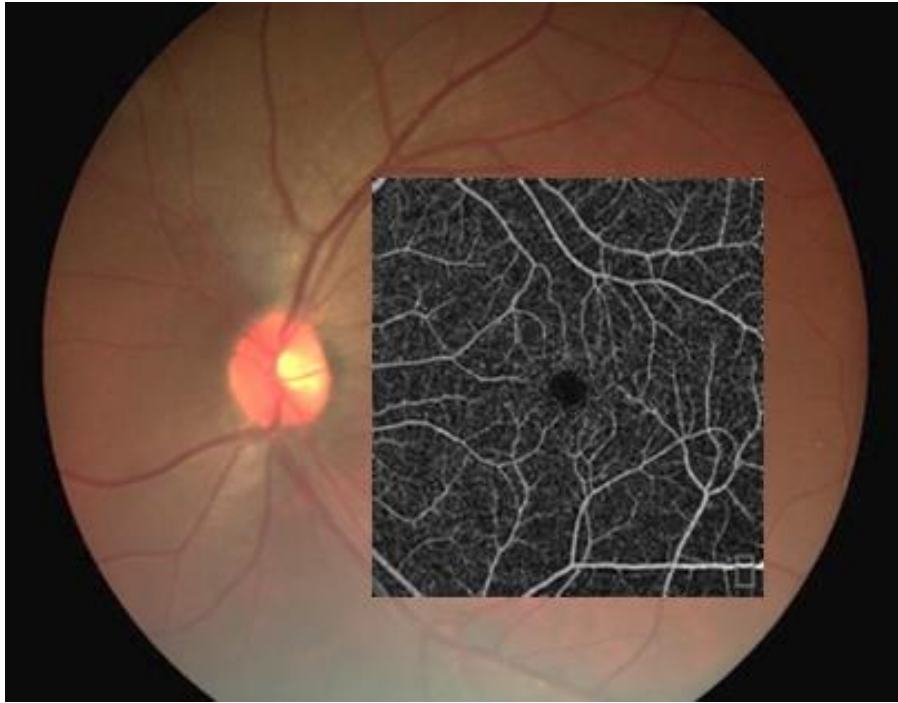


Figure 2: Fundus Image with superimposed OCTA image covering an area of 6x6 mm

Traditionally, Fundus Imaging has been used to “follow, diagnose and treat eye diseases.” [13]. On this basis, studies were performed to automatically identify DR based on blood vessel segmentation and lesion identification [9]. Algorithms have been developed to search for microaneurysms and small hemorrhages which can indicate the presence of abnormalities in blood circulation, reaching a 0.99 AUC (Area Under the Curve) with Deep Learning models. Other algorithms were created to segment the blood vessels and analyze them, but the presence of lesions and the not ideal quality of fundus images limited this procedure, which nonetheless managed to achieve over 97% accuracy with the use of Convolutional Neural Networks. The disadvantage of such procedures is that their operations are completely autonomous, and so their inner workings are not intelligible in case of errors.

In 2014, Habib et Al. [18] used fundus imaging to conduct a study on the geometric features of blood vessels of healthy patients as well as patients with NPDR and DR. The features for each bifurcation were measured manually and were not averaged for each image. This study found that

the diameters of the parent vessel as well as both the child vessels significantly varied between the cohorts. Moreover, two of the three bifurcation angles (corresponding to our VBA and CBA1 respectively) could significantly differentiate between the cohorts. These results will be referenced in the “Results” section of this study.

1.4 OCTA: the new Standard?

Optical Coherence Tomography Angiography (OCTA) is a noninvasive and noncontact imaging technique that is able to differentiate between blood vessels and the other tissues. It has seen numerous improvements in “sensitivity, acquisition speed and resolution” and “rapid and wide adoption in ophthalmology” in the last two decades [10]. Its noninvasive and tridimensional nature makes it preferable to invasive and bidimensional methods such as Fluorescein Angiography (FA). Moreover, there exist numerous techniques to improve blood vessel contrast in OCTA, in order to highlight “ocular structures with micrometer-scale depth resolution”. Figure 3 shows a En-Face OCT image and a OCTA image of the same eye, in which can be seen that the OCTA technique has a better representation of the smaller vessels. OCTA has become popular in identifying abnormalities in the vessel geometry and in the blood flow [36,37], and this makes it very important for the identification of retinopathies such as age-related macular degeneration [24-28], Diabetic Retinopathy [4,29-32,38-41], Sickle Cell Retinopathy [20-23], etc. In 2018, a study demonstrated that it was possible to distinguish between control and mild Diabetic Retinopathy patients using features extracted from segmented vessels and the foveal avascular zone (FAZ) [11,38]. The segmented vessels were extracted from OCTA images and the accuracy of this early detection was 94.3%. The introduction of non-dimensional markers minimizes “the effects of optical artifacts, refractive errors, and variation of image resolutions” [4]. Using geometric vessel-related markers is therefore a very promising option to identify NPDR and prevent any vision loss.

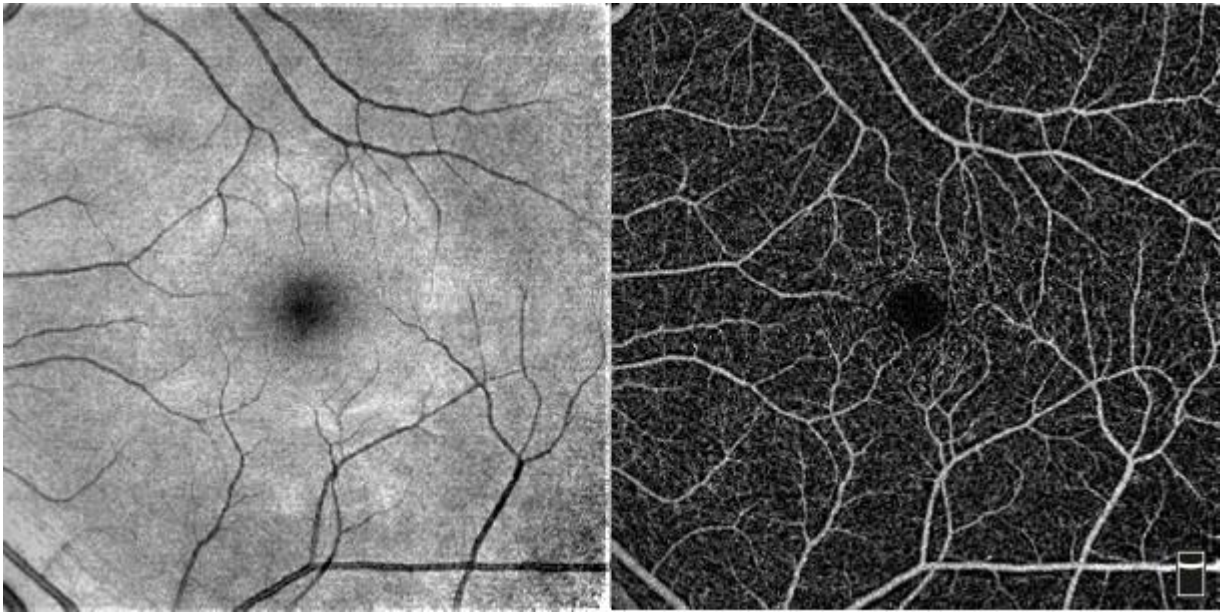


Figure 3: En-Face OCT image (left) and corresponding OCTA image (right) of a Healthy eye

Le and Al. [4], reported six markers, derived from vessel segmentation in OCTA images, and a procedure to distinguish between healthy patients and patients with NPDR, as well as between different stages of NPDR. The study found that four out of six of the proposed features had significantly different distributions between the various cohorts. This work will use the same markers but a different take on the procedure.

1.5 Artery-Vein Differentiation

Differentiating arteries from veins and analyzing them separately has been shown to increase sensitivity and shown if and how different illnesses affect them in a different way [21]. In this study, geometric features extracted from arteries and veins separately will be analyzed to verify if and for which features the sensitivity increases, decreases or does not change.

In the years, many methods have been proposed to classify arteries and veins in both Fundus Imaging [17,35] and OCTA, the latter ranging from Deep Learning [33] to Near Infrared Oximetry [34] to Features calculated in the corresponding OCT images [14]. This last method is the one that is used in this study to differentiate arteries from veins in the OCTA images.

CHAPTER 2

MATERIAL AND METHODS

The study was approved by the institutional review board of the University of Illinois at Chicago and follows the ethical standards stated in the Declaration of Helsinki.

2.1 Composition of the Dataset

The dataset used in this work is composed of 139 OCTA images of patients, both male and female, taken from the clinical database of the Lions of Illinois Eye Research Institute (LIERI), and partially overlaps with the dataset used in the work of Le et al. [4]. All the images were taken using an ANGIOVUE spectral domain OCTA system (Optovue, Fremont, CA), with a 70-kHz A-scan rate, an axial resolution of 5 μm , and a lateral resolution of 15 μm . All OCTA images were macular scans. The software used for exporting the images was ReVue (Optovue, Fremont, CA).

The images have different quality scores and were taken from patients divided in the following categories:

- 20 Control patients (without both Diabetes and Diabetic Retinopathy)
- 17 patients with Diabetes but not Diabetic Retinopathy (NoDR)
- 20 patients with Mild NPDR
- 20 patients with Moderate NPDR
- 20 patients with Severe NPDR

Each image was labeled by a specialized professional.

2.2 Classification of DR stages

According to the American Diabetes Association, DR stages are defined as such [12]:

- Mild NPDR: “Small areas of balloon-like swelling in the retina’s tiny blood vessels, called microaneurysms, occur at this earliest stage of the disease. These microaneurysms may leak fluid into the retina.”
- Moderate NPDR: “As the disease progresses, blood vessels that nourish the retina may swell and distort. They may also lose their ability to transport blood. Both conditions cause characteristic changes to the appearance of the retina and may contribute to DME.”
- Severe NPDR: “Many more blood vessels are blocked, depriving blood supply to areas of the retina. These areas secrete growth factors that signal the retina to grow new blood vessels.”

2.3 From OCTA image to binary map

For each OCTA scan, only the “Outer Retina” slice was extracted and processed. The processing was done with custom-made Matlab (Mathworks, Natick, MA) procedures.

In order to obtain a binary vessel map, the OCTA images are filtered using a Frangi 2D filter, that has the effect of enhancing the edges of the vessels, and then binarized using a single threshold. Moreover, the binary image is then cleaned of any spurious pixels or little segmented areas. These steps constitute the “preprocessing” part of the procedure. The complete pipeline is shown in Figure 4. Figure 5 contains a detailed view of the Geometric Feature Computation Process. Figure 6 shows the OCTA binary Vessel map that is obtained during pre-processing.

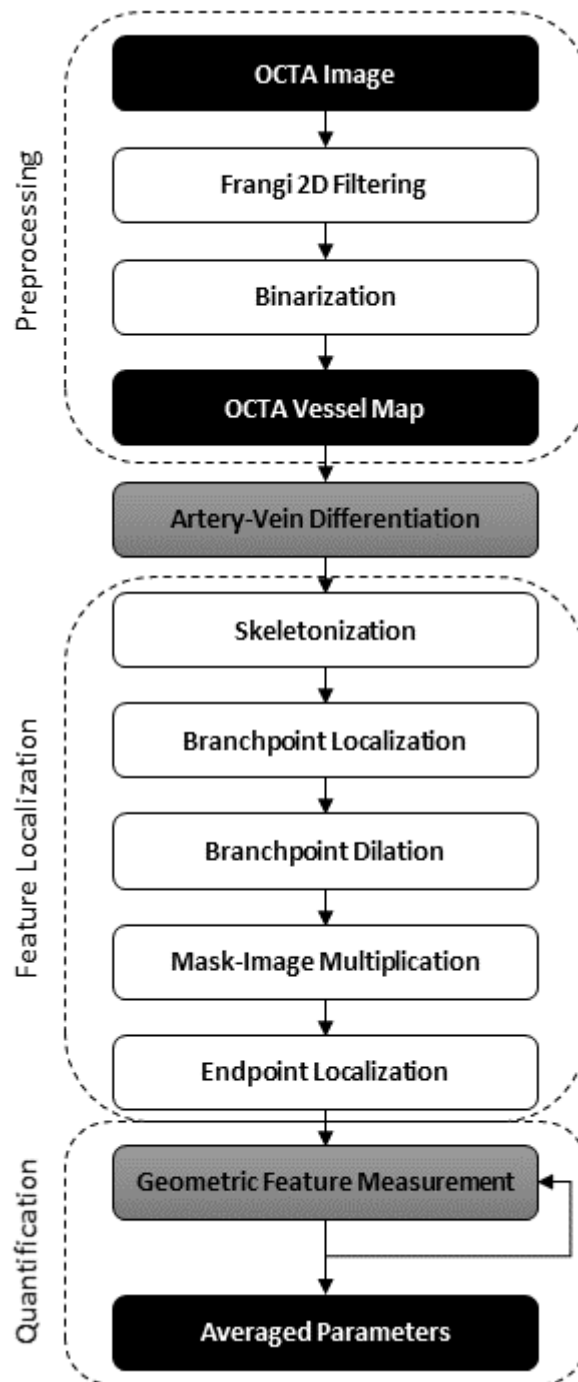


Figure 4: flowchart showing the broad steps of the Feature Extraction. The black boxes represent inputs and outputs while the grey ones stand for steps that will be detailed in the following pages. Adapted from *Fully automated geometric feature analysis in optical coherence tomography angiography for objective classification of diabetic retinopathy* [4], April 22 2019,

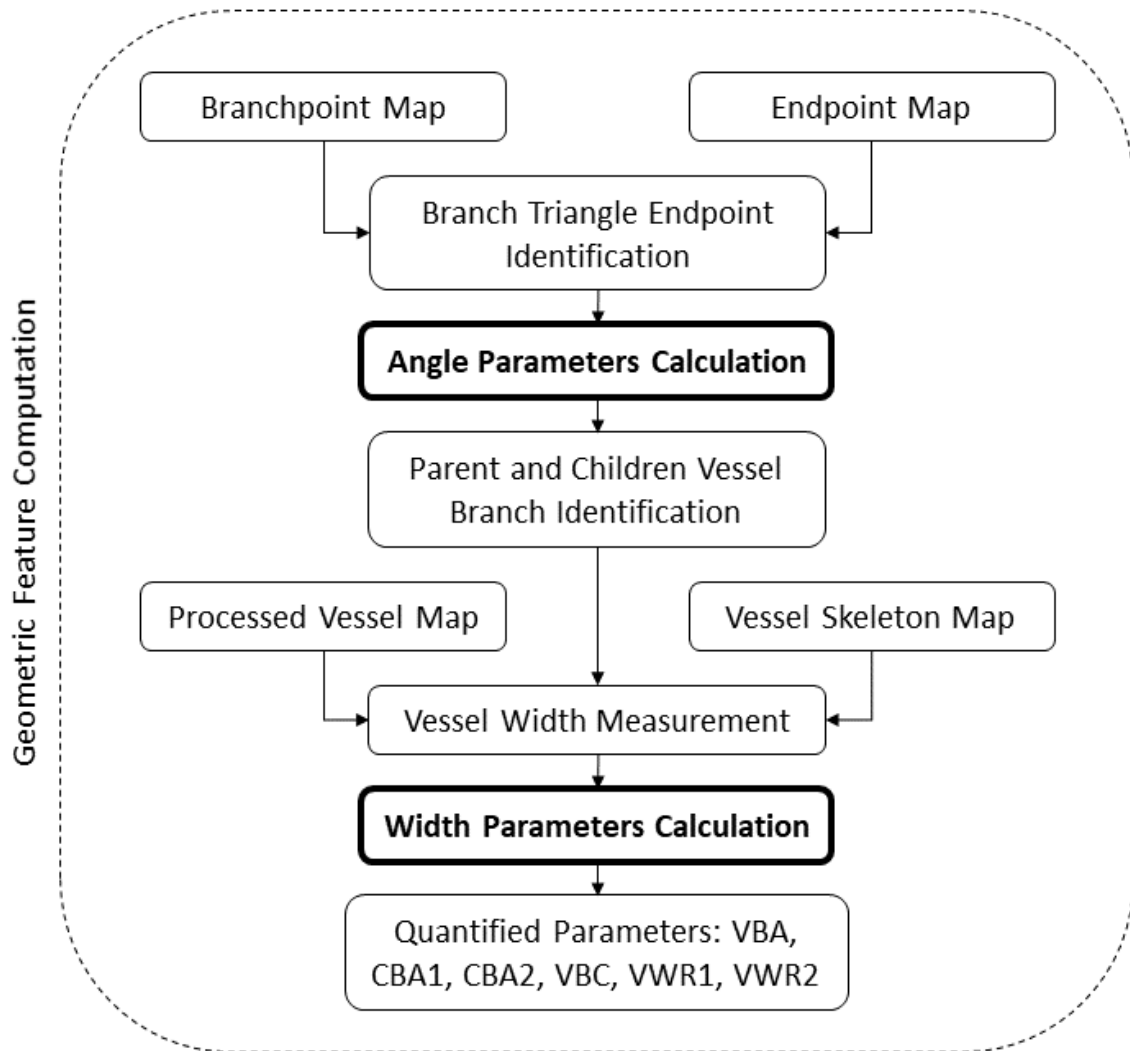


Figure 5: Detailed description of the Geometric Feature Computation process. Thick boxes show steps that output the values of the features. Adapted from *Fully automated geometric feature analysis in optical coherence tomography angiography for objective classification of diabetic retinopathy* [4], April 22 2019, Copyright 2019 by The Optical Society

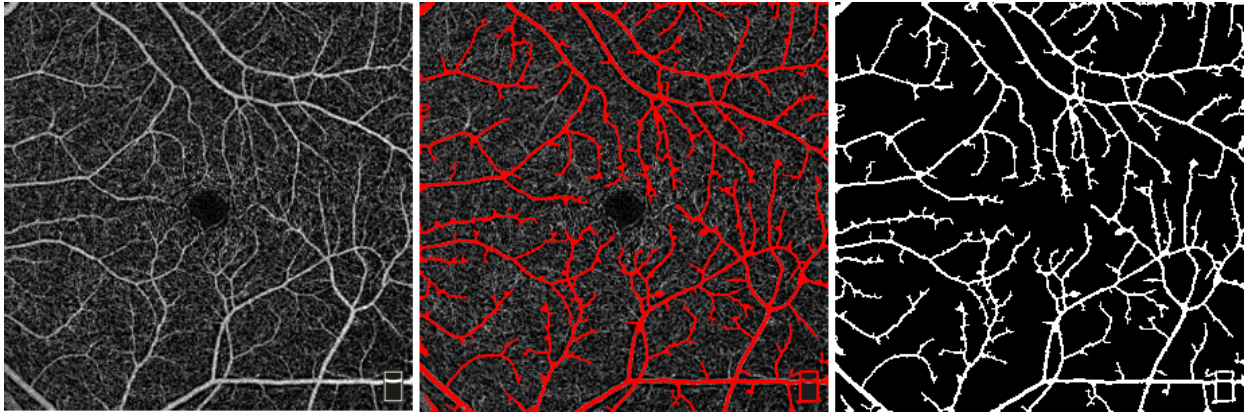


Figure 6: An overview of the Vessel Map generation. The OCTA Image (left) is processed to obtain the Vessel Map (right). In the middle, an overlay of the vessel map (in red) on the original OCTA image is shown.

2.4 Artery-Vein Differentiation

In 2019, Alam et Al. [14] developed a method to differentiate arteries and veins in OCT images employing four classifying features. In this study, a slight modification of the procedure is proposed. The only difference is in the OCT Vessel Map extraction phase and will be explained later.

In order to classify the blood vessels, en face OCT images, corresponding to the OCTA images in the dataset, were used. Since “the en face OCT and OCTA are naturally reconstructed from the same spectrogram data sets, the OCT artery-vein vessel map can be readily used to guide artery-vein differentiation in corresponding OCTA.” [14]. The complete pipeline of the Artery-Vein Differentiation is shown in Figure 7. Figure 8 contains a detailed view of the Source Node Classification process.

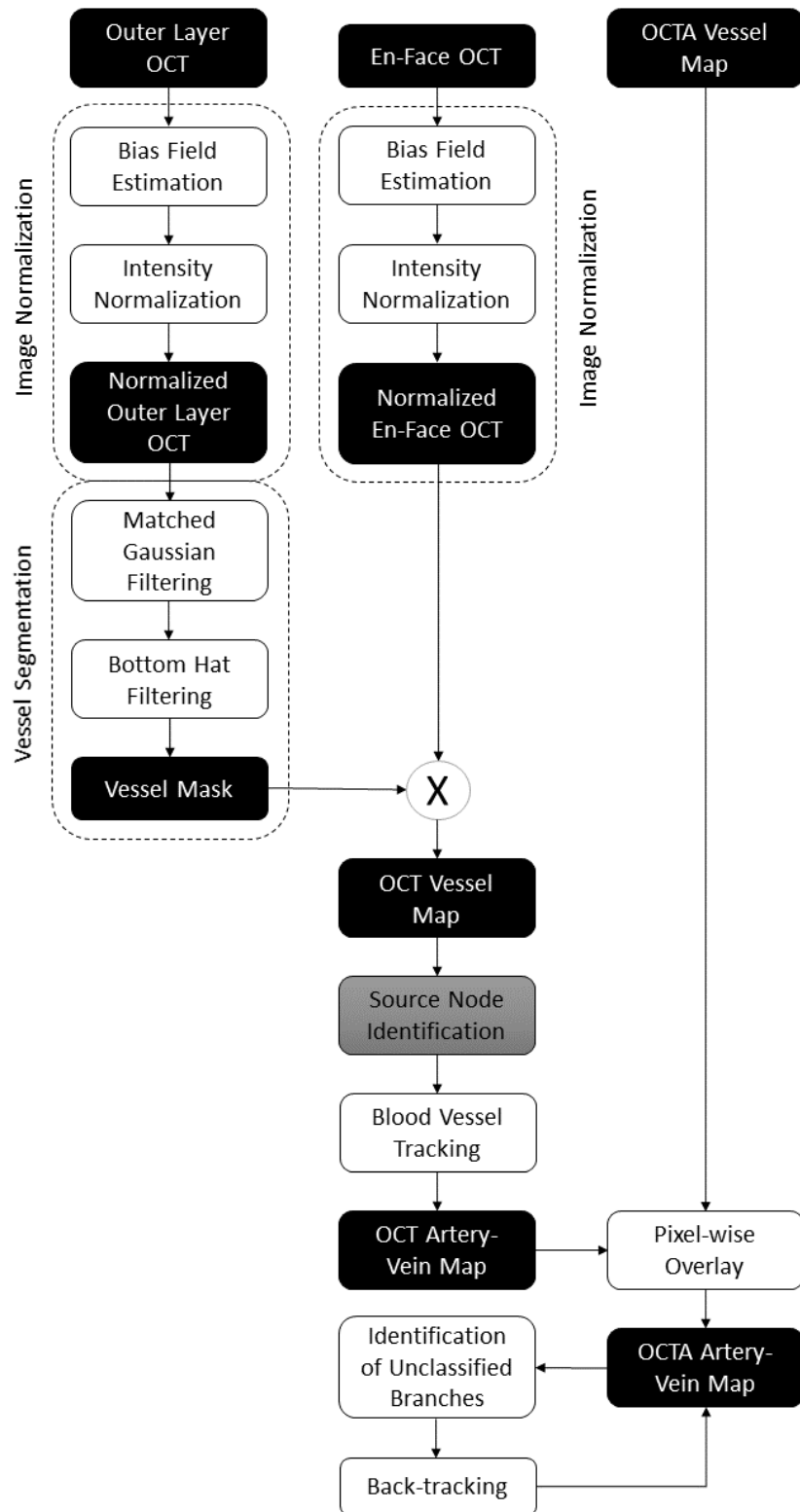


Figure 7: flowchart showing the broad steps of the Artery-Vein Classification. The black boxes represent inputs and outputs while the grey ones stand for steps that will be detailed in the following pages. Adapted from *OCT feature analysis guided artery-vein differentiation in*

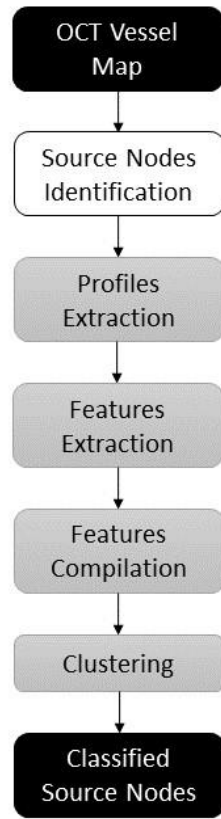


Figure 8: Detailed description of the Source Node Classification process. Black boxes show inputs and outputs of the process. Adapted from *OCT feature analysis guided artery-vein differentiation in OCTA* [14], March 29 2019, Copyright 2019 by The Optical Society

2.4.1 Bias Field Correction

According to Alam et Al. [14], since “the en face OCT images are often affected by intensity inhomogeneity”. Thus, “a bias correction technique is implemented to remove the intensity inhomogeneity before performing OCT feature analysis”. The aim of this step is to estimate the bias field of the image. Then, the image is divided by the estimated bias field in order to generate an image with normalized intensity.

A model for the en face OCT image is the following:

$$I(x, y) = b(x, y) I_{true}(x, y)$$

where $I(x, y)$ is the intensity profile of the en face OCT image, $b(x, y)$ is the bias field that is estimated, and $I_{true}(x, y)$ is the OCT image with normalized intensity. In turn, $b(x, y)$ can be defined as:

$$b(x, y) = \sum_{k=1}^M w_k \cdot G_k$$

where “ G is the set of smooth basis functions which ensures the smooth varying property of the bias field. 20 polynomials of the first degree are chose as the basis functions.” Given $M = 20$, the optimal coefficients $[w_1, w_2, \dots, w_M]$ have to be calculated. After this, the en face OCT image is divided by the bias field. This Bias-Field Correction is applied to two images (instead of one as in the previous paper) per patient: the En-Face OCT image and the Outer Layer OCT image. Figure 9 and 10 show this procedure applied to both images.

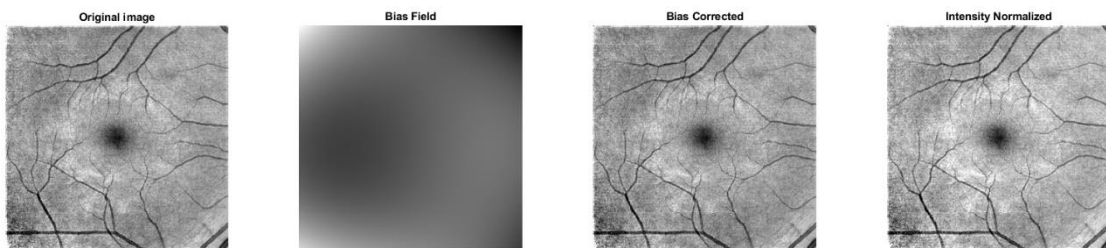


Figure 9: Image Normalization steps for En-Face OCT. From left to right: the original En-Face OCT image, the Bias Field extracted from the image, the image without the Bias Field and the image after Intensity Normalization. Adapted from *OCT feature analysis guided artery-vein differentiation in OCTA* [14], March 29 2019, Copyright 2019 by The Optical Society

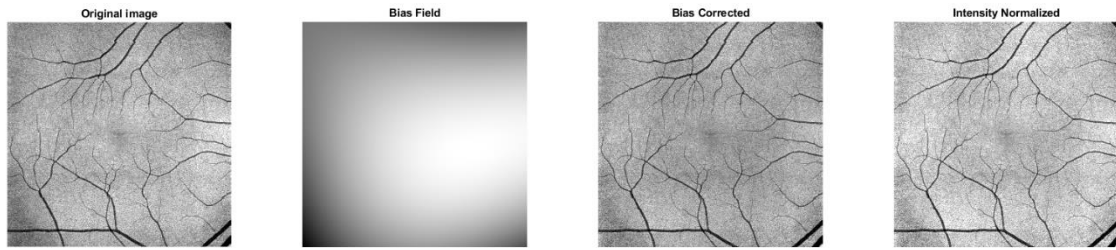


Figure 10: Image Normalization steps for Outer Layer OCT. From left to right: the original Outer Layer OCT image, the Bias Field extracted from the image, the image without the Bias Field and the image after Intensity Normalization. Adapted from *OCT feature analysis guided artery-vein differentiation in OCTA* [14], March 29 2019, Copyright 2019 by The Optical Society

2.4.2 OCT Vessel Map Extraction

After normalizing the images, the vessel map is extracted from the OCT Outer Layer image instead of the En-Face OCT one. This is because the former image contains the shadows of the vessels. These shadows cannot be differentiated in arteries and veins but have more contrast with respect to the background, thus making the vessel mask extraction easier. For every image, the correspondence between En-Face vessels and Outer Layer structures was checked. Figure 11 shows the difference between a En-Face OCT image and an Outer Layer OCT image of the same patient.

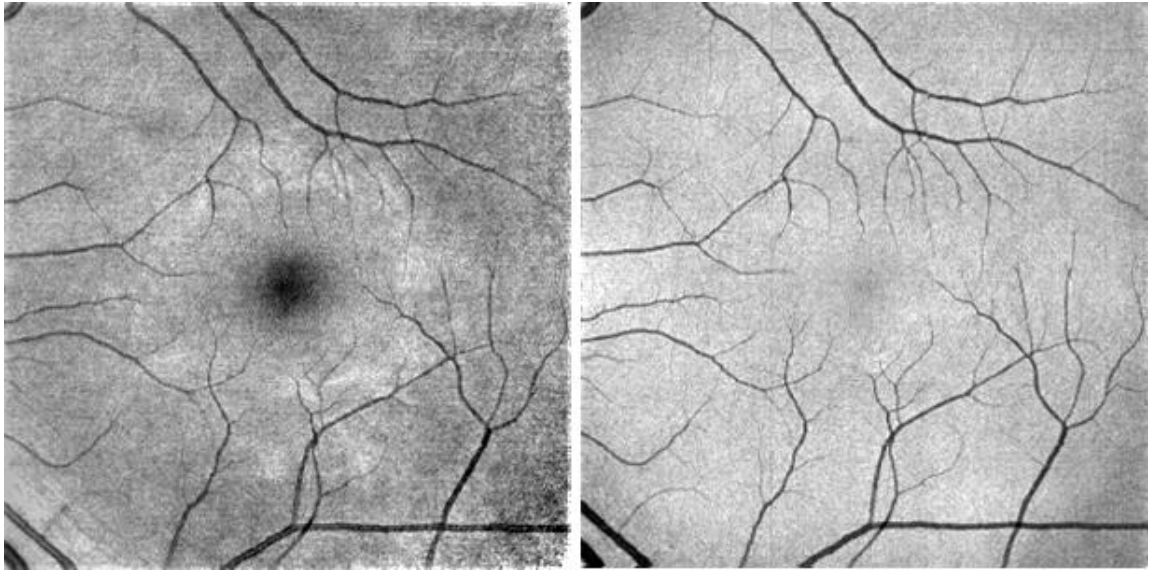


Figure 11: En-Face OCT (left) and Outer Layer OCT (right) images of the same eye. The latter image is cleaner and has more contrast between vessels and background.

Then, “a matched filtering method is combined with bottom hat filtering to enhance and extract the blood vessel map from the en face OCT image. For the match filtering, 2D Gaussian kernels of 12 different orientations and 10 different sizes are implemented to match blood vessels. These kernels cover all the blood vessel diameters and directions” [14] and are defined by Li et Al. [15] as:

$$I(x, y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y \cos \theta - x \sin \theta)^2}{2\sigma^2}}$$

where x, y are the coordinates of the pixels, θ is the rotation angle of each kernel (ranged 0 to π), and σ is the width of the kernel. Taking again from Alam et Al. [14]: “the en face OCT image is convolved with 120 kernels after subtracting the mean from each kernel. In the process of cross-correlation, the features matching the kernels (representing blood vessels of various widths) produce larger coefficient values. Larger coefficient values representing the vessel structures are selected to produce the segmented blood vessel map (Fig. 2(D)) from a maximum intensity projection of all the convolved images”.

In the original paper, the next step would be the application of a 20×20 pixels bottom-hat filter in order to further improve contrast between vessel structures and background. An improvement to this procedure was conceived. Instead of using only one bottom-hat filter, a selection of four bottom-hat filters with different structural elements and different sizes were applied to the image in parallel and their results were averaged. This was shown to better highlight the vessel structures, allowing for better masks to be extracted.

After these changes, the procedure of the paper was closely followed: “A global thresholding method is used to extract the binary blood vessel map (Fig. 2(E)). The binary map is multiplied with the intensity normalized en face OCT image to generate the segmented vessel map of en face OCT image (Fig. 2(F)). This en face OCT vessel map is used for further source node identification and artery-vein tracking.” [14]. Figure 12 shows these steps applied to images from an eye of a Control patient.

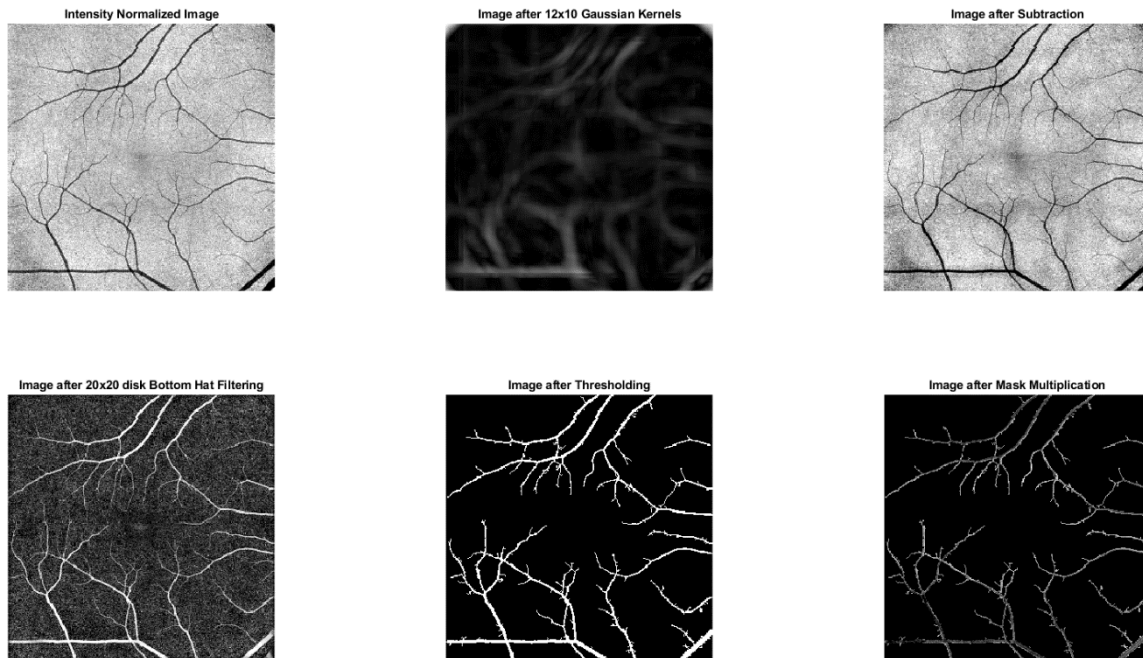


Figure 12: OCT Vessel Map generation steps. Top left: Intensity Normalized Outer Layer OCT image from the previous step. Top middle: mask generated by the Matched Gaussian filtering. Top right: image after subtraction of the mask. Bottom left: image after Bottom-Hat filtering. Bottom middle: image after single threshold binarization. Bottom right: result of the multiplication of the vessel map with the Intensity Normalized En-Face OCT image.

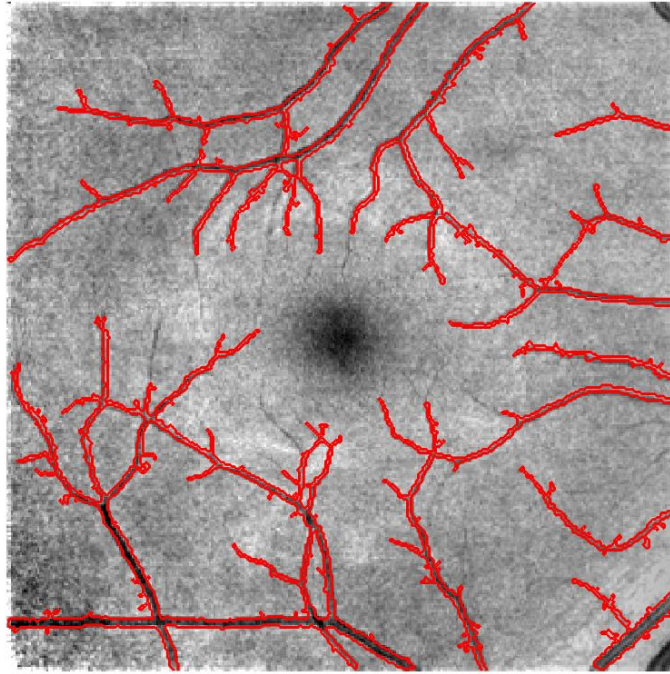


Figure 13: Assessment of the Vessel Map extraction by superimposition with the Intensity Normalized En-Face OCT image.

2.4.3 Source Node Identification

In order to classify the newly extracted vessels into arteries or veins, an ad hoc strategy is required. In Alam. Et Al. [14] a procedure of this kind is outlined. It involves several steps: the first one is to identify the source nodes, defined as the points in which different vessels cross the boundary of the image. The second step is to extract some features required to correctly identify each source node. The last step is to trace the vessels from the source nodes to their various endpoints in the image and thus reconstructing the entire vessel map with the vessels being marked as either veins or arteries.

Before the next step, the vessel map is skeletonized to create a skeleton map, in which each vessel is reduced to a line one pixel wide. To identify the source nodes a single line is drawn in the skeleton map following the borders of the images with a certain offset. Every white pixel crossed by this line is marked as a source node and thus is eligible for classification. Figure 14 (left) shows the result of this step.

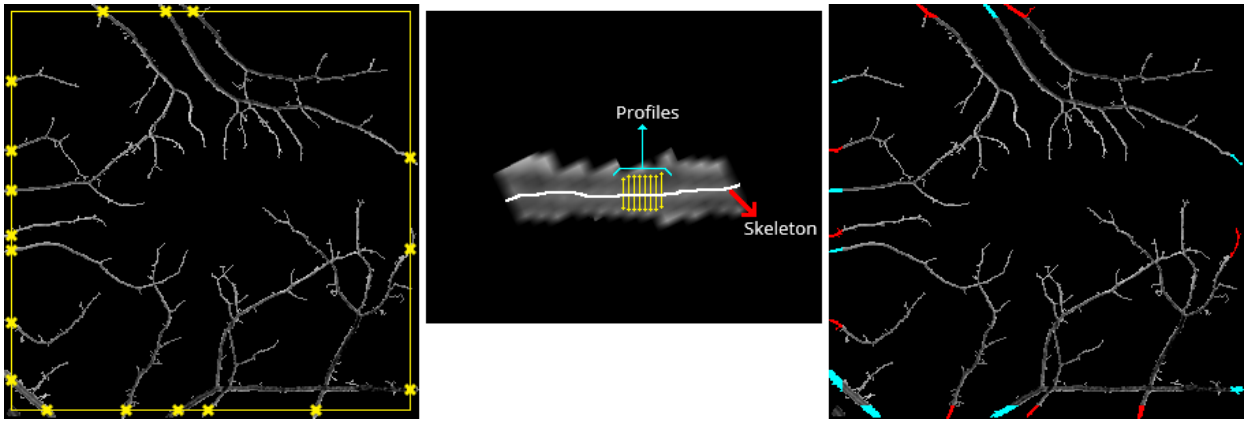


Figure 14: Source Node Classification steps. On the left, the Source Nodes are identified and are marked with yellow crosses. In the middle, an enlarged Source Node segment with respective skeleton (in white) and sample profiles (in yellow). On the right, the Source Nodes are identified as arteries (in red) or veins (in light blue). Adapted from *OCT feature analysis guided artery-vein differentiation in OCTA* [14], March 29 2019, Copyright 2019 by The Optical Society

2.4.4 Profiles Extraction

Using ‘profiles’ as defined in [16] is possible to introduce new classification features. Segments of the vessel in both the vessel map and skeleton map is taken near the source node. As explained in [14]: “The number of profiles depends on the size of the source node segments. The profiles are set to have a distance of two pixels among each other along the vessel skeleton. For example, if a vessel skeleton within the source node segment is 40 pixels long, the first profile will be extracted at the middle point, pixel 20. Then the profiles will be extracted at 2 pixel intervals in both directions. The maximum number of profiles is set at 20 and minimum is 5. In case of a segment smaller than 5 pixels long, profiles are taken from all of the pixels.”

A profile set, called ‘ PS_i ’, is thus created equal to:

$$PS_i = S \times N$$

where S is the number of segments for each source node, N is the number of profiles (ranging from a minimum of 5 to a maximum of 20), and i is the number of source nodes in the image. An

image profile set is defined as PS and comprises the PS_i for each source node in the image. For example, if an image has 16 source nodes, PS will be $[PS_1, PS_2, \dots, PS_{16}]$. Figure 14 (middle) shows some sample profiles in a Source Node segment.

2.4.5 Feature Extraction and Classification of the Source Nodes

Once all the profiles are extracted, four features are then measured for every single profile. These features are: “i) ratio of vessel width to central reflex, ii) average of maximum profile brightness, iii) average of median profile intensity, and iv) optical density of vessel boundary intensity compared to background intensity.” [14]. In Figure 15, the average normalized distributions for both arteries and vein features are shown.

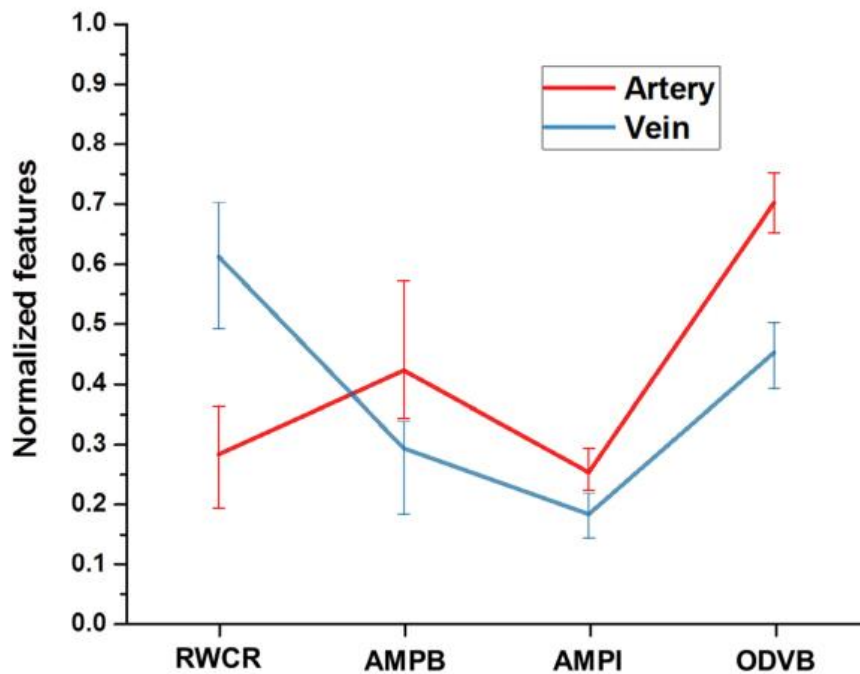


Figure 15: Normalized feature distribution in artery and vein vessels. RWCR: ratio of width to central reflex; AMPB: average of maximum profile brightness; AMPI: average of median profile intensity; ODVB: optical density of vessel boundary. Reprinted from *OCT feature analysis guided artery-vein differentiation in OCTA* [14], March 29 2019, Copyright 2019 by

After measuring these features, a K-means algorithm is employed to classify the source nodes. Each source node will be classified as being either artery or vein. The K-means method was chosen “due to its computational simplicity and efficiency”, and because its results were very reliable [14]. The two classes of the algorithm (Artery and Vein) have their centroids are set to the extremes (maximum and minimum) of the input set, in order for the cluster centers to be as far away as possible for each other. The metric used to calculate distance from the centroids was the Euclidean distance.

In order to maximize the K-means algorithm’s accuracy, three steps are performed:

1. the algorithm is applied (and the classification is performed) to every source node segment ($S * i$ datapoints)
2. four quadrants are identified in the image, starting from the fovea. Each quadrant has its K-means predictions calculated separately. Given the uneven intensity distribution inside quadrants of the images, this step minimizes its effect on the overall estimation.
3. the four quadrant are then rotated by 45 degrees eight times (to reach 360 degrees), as shown in Figure 16. The algorithm is applied to every rotation and every quadrant, for a total of eight estimations per source node. Moreover, rotating by 45 degrees implies that there is overlapping between the quadrants, in order to reduce even further the uneven intensity distribution. The final classification of a source node is therefore the combination of the eight predictions, in order to reduce the contribution of outliers.

Figure 14 (right) shows the result of the Classification step.

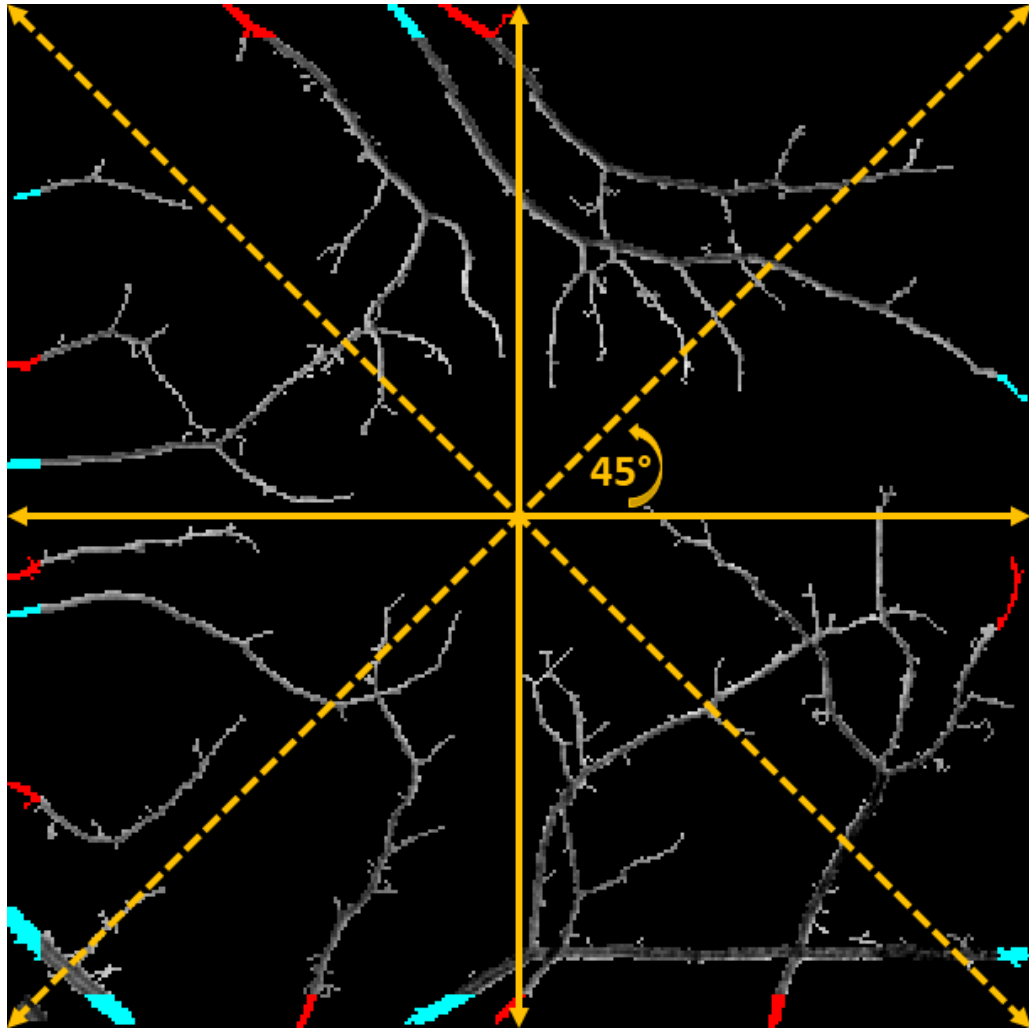


Figure 16: Overview of all the rotations of the quadrants relative to the center of the fovea.

Adapted from *OCT feature analysis guided artery-vein differentiation in OCTA* [14], March 29

2019, Copyright 2019 by The Optical Society

2.4.6 Tracking the Vessel from Source to End Nodes

After the classification of all the source nodes, the classification has to be extended to the entire vessels in the image. The algorithm to track the blood vessels from the source nodes to the various end nodes is described in [17]. Figure 17 shows the core steps of this algorithm.

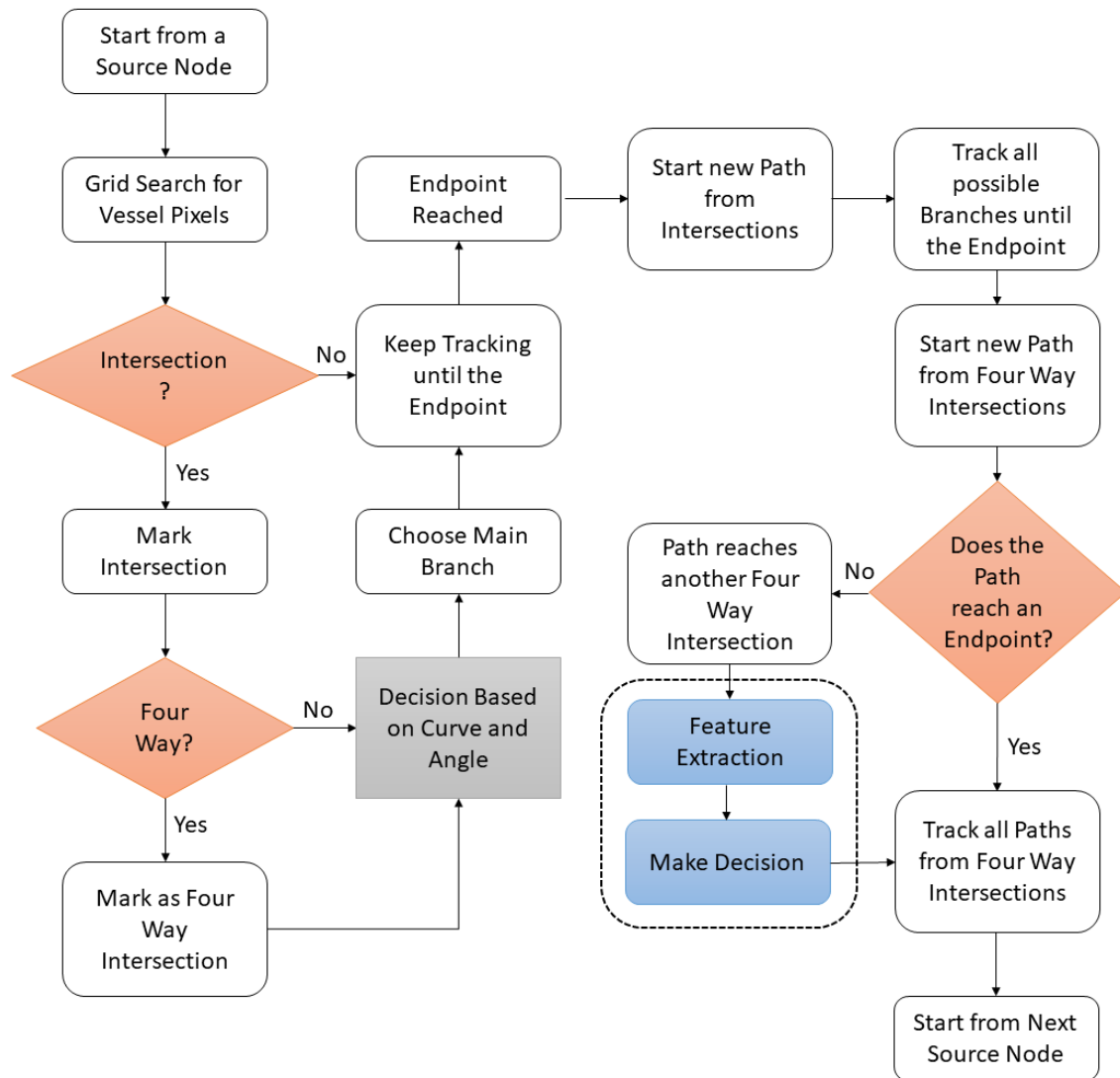


Figure 17: core steps of the Blood Vessel Tracking algorithm. Adapted from *Combining ODR and Blood Vessel Tracking for Artery- Vein Classification and Analysis in Color Fundus Images* [17], April 2018, Copyright 2018 by ARVO Journals

The tracking algorithm begins from a source node. It uses a system of concentric grids to identify vessel pixels from which to continue. The grid starts as 3×3 pixels and its size grows every time any vessel pixels are not found inside. When an intersection is found it is marked, and the algorithm follows the main branch, defined by curve and angle information.

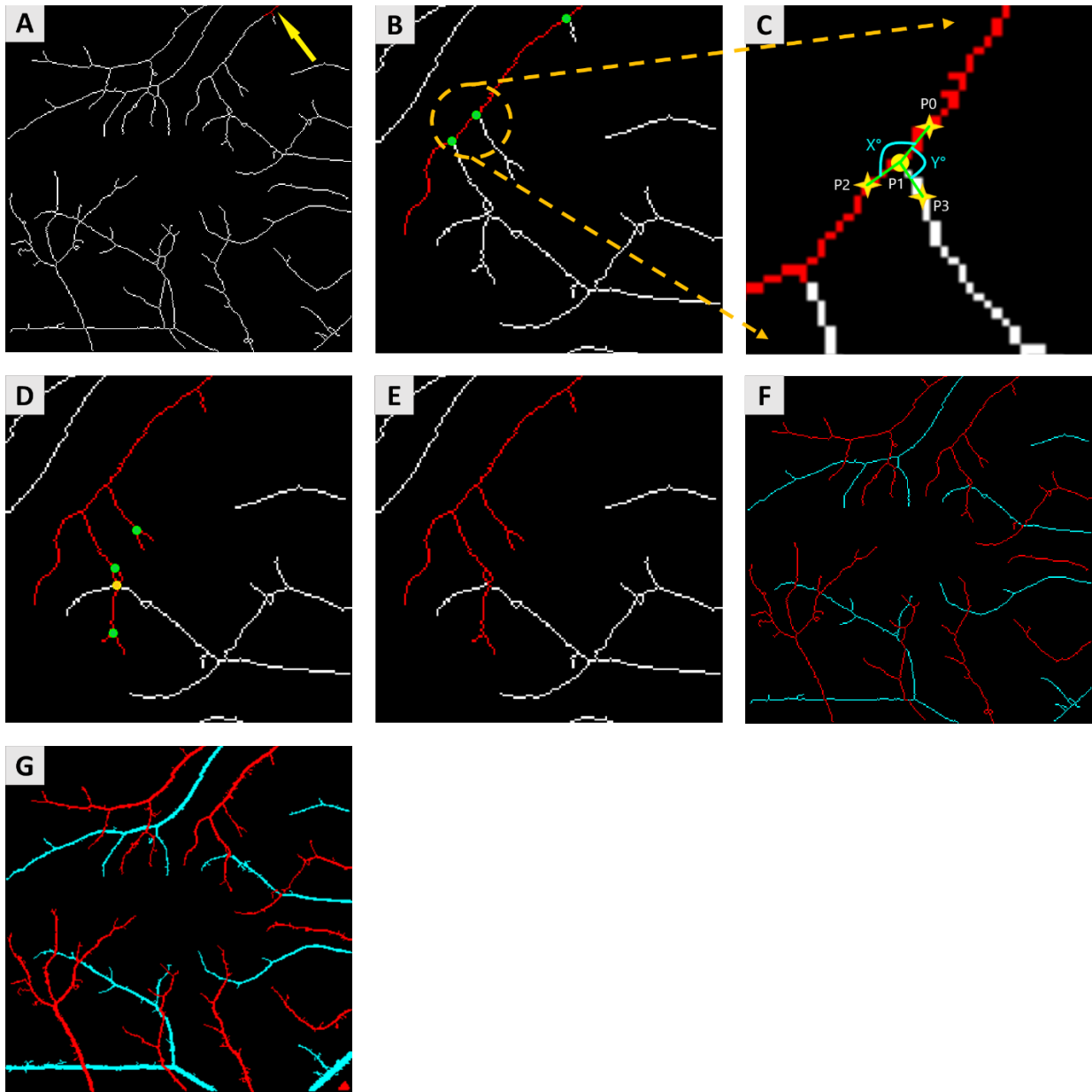


Figure 18: Main steps of the Blood Vessel Tracking algorithm. **A**: the algorithm starts from a source node. **B**: the main branch of the vessel is tracked; the *green dots* represent the encountered branch nodes. **C**: enlarged window showing the process of choosing the forward path in the vessel map. **D**: all remaining possible branches are identified; one four-way intersection is marked with a *yellow dot*. **E**: the algorithm takes a decision about the four-way intersection and the whole vessel is identified. **F**: these steps are repeated until the skeleton map is completely classified in arteries (*red*) and vein (*blue*). **G**: fully classified vessel map. Adapted from *Combining ODR and Blood Vessel Tracking for Artery- Vein Classification and Analysis in Color Fundus Images* [17], April 2018, Copyright 2018 by ARVO Journals

“P1 is the current point, and P2 and P3 are the candidate points for the forward path. A past-visited point P0 is chosen, and two separate curves are considered to compare. The first curve is P0-P1-P2, and second is P0-P1-P3. The curvature can be quantified using the distance metric between P0 (x1,y1) and P2 or P3 (x2,y2) which can be calculated as”: [17]

$$Curvature = \frac{1}{n} \sum_{i=1}^n \left(\frac{Geodesic\ Distance}{Euclidean\ Distance} \right)$$

The two distances are calculated between the two endpoints of the vessel branch. The Geodesic Distance is defined as:

$$Geodesic\ Distance = \int_{t_0}^{t_1} \sqrt{\left(\frac{dx(t)}{dt}\right)^2 + \left(\frac{dy(t)}{dt}\right)^2} dt$$

Where $x(t)$ is the curve P0-P1-P2, $y(t)$ is the curve P0-P1-P3 and $[t_0, t_1]$ is the interval that comprises both curves. The Euclidean Distance is defined as:

$$Euclidean\ Distance = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

The path with smaller curvature is therefore considered the main branch and it will be used by the algorithm to advance forward. The angle information is also taken into account: the branch with the greatest angle will be used to identify the main branch. After the first endpoint is reached, the tracking algorithm selects a marked intersection and continues from the other branch. If the marked intersection is a four-way intersection, then it remains marked and saved for later. Once all three-way intersections are tracked to their endpoints, the same procedure is repeated for another source node.

Regarding the four-way branchpoints, there are two possibilities: either it is a crossing between two different vessels or it is a multiple branching of the same vessel. To verify this, all the four branches are tracked backwards until they reach an intersection or an endpoint. If they do not reach any other vessel intersection they are classified as being part of the main branch. Otherwise,

the four branches must be classified, with one at least already so (the main branch). The algorithm uses various textural parameters to compare the two vessels and decide to which one the branches are assigned to. If these parameters do not provide enough information, then the remaining branches are classified as either arteries or veins by similarities with the textural parameters of all arteries and veins in the map. An overview of the results of the main steps of this algorithm is found in Figure 18.

2.4.7 Classifying the OCTA Binary Vessel Map

Once the skeleton map is completely classified, it is combined with the binary vessel map to generate a classified vessel map.

First, it is to be considered that “the en face OCT and OCTA images are generated from 3D projection of OCT B-scans and B-Scan speckle variance images, respectively. This means the structural coordinates of both images are same. Therefore, any process of image registration is not required for the overlaying process” [14]. To classify the greater vessels in the OCTA image, it is possible to transfer the classification directly from the same coordinates in the OCT classified map.

However, there are smaller vessels in the OCTA map that do not appear in the classified map. The procedure to classify them is very similar to the tracking algorithm explained above. It starts in the skeleton map: from the endpoint of an unclassified vessel and back-tracks until it reaches a classified vessel, and then begins again from another endpoint. Four-ways intersections are treated in the same way as before. After this passage, the OCTA vessel map is completely classified into arteries and veins. An overview of this step is found in Figure 19.

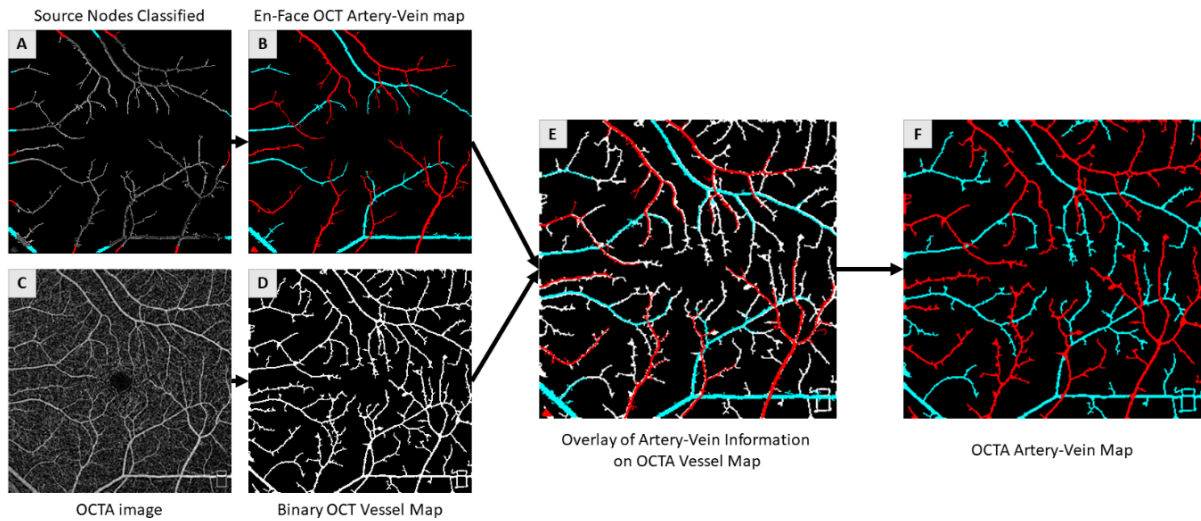


Figure 19: Artery-Vein classification in OCTA. **A:** classified source nodes in OCT vessel map.

B: OCT artery-vein map. **C:** original OCTA image. **D:** OCTA binary vessel map. **E:** OCT artery-vein map overlaid onto the OCTA binary vessel map. **F:** final OCTA artery-vein map.

Adapted from OCT feature analysis guided artery-vein differentiation in OCTA [14], March 29 2019, Copyright 2019 by The Optical Society

2.5 Feature Extraction

In this work, a total of six geometric features will be analyzed through a fully automatic custom-developed MATLAB procedure. Three of these features are based on branching angle information and the other three are based on vessel width information. Taking from Le et al. [4]: “Vessel angle-based features measure the degree of bifurcation and relative angle change among the parent and child vessel branches. The first feature is vessel branching angle (VBA: θ), which measures the overall degree of bifurcation between the child branches. The second and third features are the child branching angle 1 (CBA1: α_1) and child branching angle 2 (CBA2: α_2), which measures the direction of branching of each child vessel. The child angles can provide information relating to the deviation of direction of child vessels with respect to the original parent vessel direction.

Vessel width-based features measure the structural change of the vessels as a result of branching. The first vessel-width feature is vessel branching coefficient (VBC), which measures the effect of branching on vessel area between the parent, first (larger) child, and second (smaller) child. Similarly, the second and third vessel-width features are vessel width ratio 1 (VWR1) and width

ratio 2 (VWR2), which examines the change in width between individual child and parent vessels.”

2.5.1 Separating Arteries from Veins

Once the classified OCTA vessel map is ready, veins and arteries must be separated in two different maps. This procedure is very straightforward, since the previous algorithm saves the arteries with a red color and the veins with a blue color. A simple MATLAB procedure can easily extract the two different maps. The results of this procedure can be seen in Figure 20. However, there are gaps whenever a type of vessel (usually arteries) crosses over the other type (usually veins), leaving the latter vessel with a discontinuity. This is not a cause for concern since the geometric analysis focuses on branchpoints and the chance that two vessels cross over a branchpoint (five-way or six-way intersection) is low (but still possible).

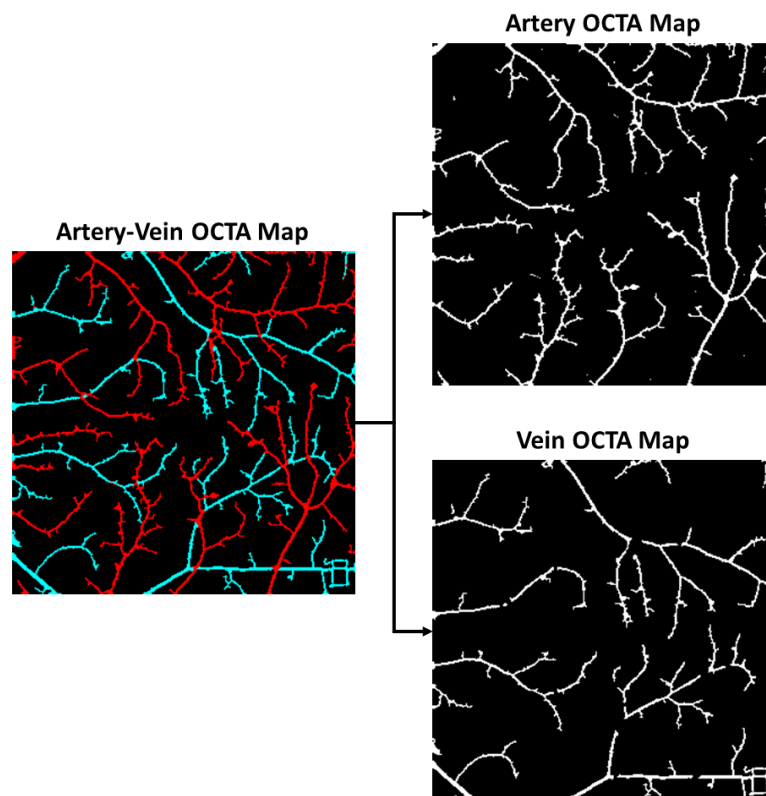


Figure 20: Separation of the OCTA artery-vein map into the OCTA artery and vein maps. The discontinuities in the OCTA vein map can be clearly seen.

2.5.2 Identifying the Branching Triangle

Once the arteries and veins are separated, the features are ready to be measured. The measurement occurs separately for arteries and veins.

In order to do that, every branchpoint in the skeleton map must be identified. MATLAB morphological functions are used, returning a binary map with only the vessels' branch points. These points are then dilated to form disks with a certain radius. This radius is arbitrarily defined as 7 pixels. The multiplication of this disk map with the skeleton map gives a map with only the branchpoints areas (the rest is eliminated). After this, another morphological function allows for the extractions of all endpoints in the resulting image. So, the endpoints of all the branches are identified, corresponding to the points in which the circumferences of the disks intersect with the skeleton map. "The preprocessing and branch triangle identification steps results in four images, i.e., the processed vessels, vessel skeleton, branchpoints, and endpoints, used for following feature measurement" [4].

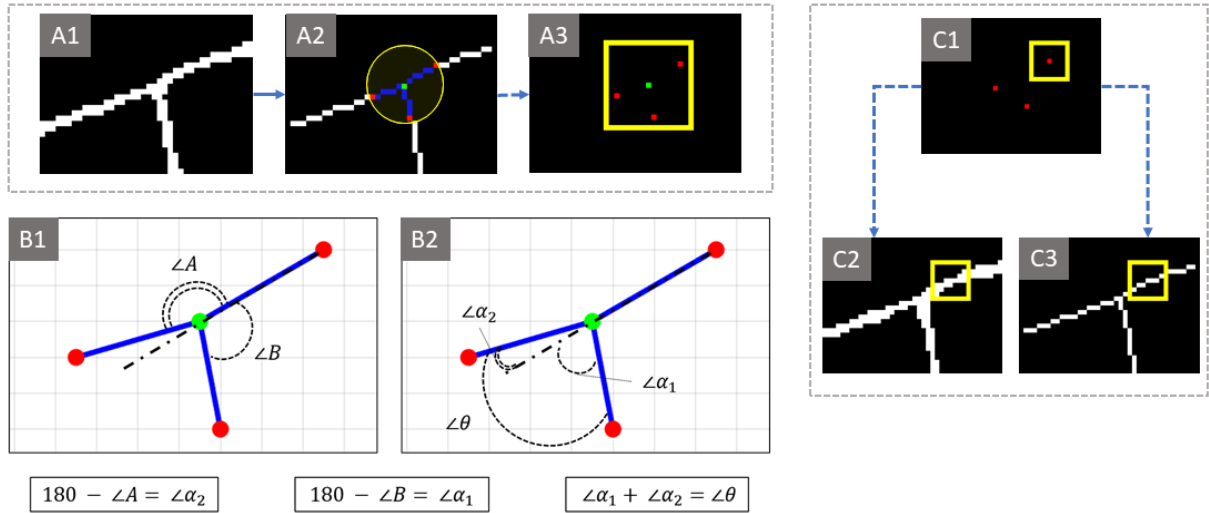


Figure 21: Overview of the Branching Triangle Identification and Features Measurement.

A: Feature Localization Step for an example branchpoint. **A1:** Processed vessel map.

A2: A composite image containing the vessel skeleton, where the *green pixel* represents the branchpoint, *red pixels* represent the end points, the *blue pixels* represent our vessels of interest, and the *yellow circle* represents the dilated area. **A3:** A composite image of the branchpoint (*green*) and endpoints (*red*), where the *yellow square* represents the window area. **B:** Illustration

of branch angle measurement. Angles A and B in **B1** are complementary angles used to calculate θ (VBA), α_1 (CBA1) and α_2 (CBA2) in **B2**. **C:** Illustration of window method used to calculate vessel width of one representative vessel branch. **C1:** A *yellow window* is centered around the endpoint (*red pixel*). The coordinates of the *yellow window* are used to measure the pixels in the processed vessel map (**C2**) and in the skeleton map (**C3**). Adapted from *Fully automated geometric feature analysis in optical coherence tomography angiography for objective classification of diabetic retinopathy* [4], April 22 2019, Copyright 2019 by The

2.5.3 Features Measurement

Once all branchpoints and endpoints are identified, the iterative feature measurement algorithm is run for each branchpoint with exactly 3 branchpoints. Four endpoints make sure that the respective branchpoint is always labeled as a crossing of two vessels, while branchpoints with a different number of endpoints are treated as errors and outright excluded from the analysis.

For each branchpoint, a square window 15 pixels wide and centered at the branchpoint itself is extracted. The window width is defined as $1 + 2 \times$ the dilation radius defined in 2.5.2 (arbitrarily set as 7). The window is used to save the coordinates of the endpoints that are inside of it in the endpoints image. Here is when the branchpoint can be excluded (if the number of endpoints is different than 3).

The coordinates of the recorded endpoints and the branchpoint in the center are then used to define segments (vectors), starting from the branchpoint and ending in the three endpoints and identifying three vessels: the parent and the two children. These three vectors are used to measure the three angles that are between them. These angles will be called θ , A and B. The angle θ is also the feature called VBA, and it is defined as the smallest of the three angles. This definition is based on the vessel branching's physiology. Angle A and B are the largest and second largest angle respectively, and together with θ they sum up to 360° . The two other angle features, CBA1 and CBA2, can be calculated with angles A and B:

$$CBA1 = 180^\circ - B$$

$$CBA2 = 180^\circ - A$$

where the first Child Branching Angle is always larger than the second one.

With the three angle features measured, the three endpoints can be classified to find the parent endpoint and the children endpoints. The two segments that bound the angle θ (VBA) have to be the children, so their endpoints are marked as such. Conversely, the remaining vector is marked

as the parent vessel. When all three vessels are identified, their widths have to be measured. In order to do this, we use another window method. This time, the window is centered around a single endpoint, and the width of the vessel is calculated as such:

$$width = \frac{\sum_{i=1, j=1}^n V(i, j)}{\sum_{i=1, j=1}^n S(i, j)}$$

where $V(i, j)$ is the number of pixels in the window area in the vessel map image, and $S(i, j)$ is the number of pixels in the window area in the skeleton map image. With this setup, the vessel map provides an area and the skeleton map provides a pseudo-length and the result is a representation of width of the vessel. This window method is repeated three time, once per vessel.

The final step is calculating the width-related features. The first one is the Vessel Branching Coefficient (VBC). It is obtained by the following equation:

$$VBC = \frac{(child_1 width)^2 + (child_2 width)^2}{(parent width)^2}$$

where $child_1 width$ is the width of the smaller child vessel, $child_2 width$ is the width of the larger child vessel and $parent width$ is the width of the parent vessel.

VWR1 and VWR2 are calculated as:

$$VWR1 = \frac{child_1 width}{parent width}$$

$$VWR2 = \frac{child_2 width}{parent width}$$

After this step, all six features are measured for one branchpoint. This measuring procedure is done for every non-excluded branchpoint in the image. The results of every branchpoint are thus averaged to produce one average value for each feature. The final six values are then recorded and saved in a spreadsheet with the results from the other images.

2.6 Feature Sets

In this study there are four feature sets, composed of six features each, making the number of extracted features per image up to a total of 24. The four sets are:

1. **ARTERY**: these are the six features that were extracted only from artery vessel maps.
2. **VEIN**: these are the six features that were extracted only from vein vessel maps.
3. **A/V RATIO**: these are the ratios between the six artery features and the six vein features:

$$A/V \text{ feature} = \frac{feature_A}{feature_V}$$

where $feature_A$ and $feature_V$ are the values of the same feature in the Artery and Vein case respectively.

The analyses on this dataset will provide us with information on if and how DR affects arteries and vein in a different way.

4. **WEIGHTED AVERAGE**: these six features were obtained as follows:

$$global \text{ feature} = \frac{feature_A \times n_A + feature_V \times n_V}{n_A + n_V}$$

where n_A and n_V are the number of non-excluded branchpoints in the artery and vein maps respectively.

These features are made to put together the information obtained from the Artery and Vein measurements by an average weighted by the number of analyzed branchpoints.

2.7 Statistical Analysis

There are two types of statistical analyses, both performed using Microsoft Excel (Microsoft Corporation, Redmont, WA). The first consists in a one-way, multi-label analysis of variance (ANOVA). It compares the statistical distributions of the features among the five patient groups (Control, NoDR, Mild NPDR, Moderate NPDR, Severe NPDR). This analysis was done four times, one for each feature set (Weighted Average, Artery, Vein and A/V Ratio). Whenever ANOVA found significantly different distributions, one versus one comparison of those features in the same cohorts followed. They were performed using a Student t-test and corrected using the Bonferroni Correction. Statistical significance was defined as having a P-value under 0.05.

CHAPTER 3

RESULTS

The database used for this study is the same as the ones used in [4]. Therefore, a direct comparison of the results can be made.

3.1 Demographics

The demographics of the dataset are detailed in the following table:

TABLE I: DEMOGRAPHICS OF CONTROL AND DR SUBJECTS

			Diabetic Retinopathy		
	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Number of Subjects	20	17	20	20	20
Sex (male)	12	6	11	12	11
Age (mean \pm SD), years	42 \pm 9.8	66.4 \pm 10.14	50.1 \pm 12.61	50.8 \pm 8.39	57.84 \pm 10.37
Age Range, years	25-71	49-86	24-74	32-68	41-73
Duration of Diabetes (mean \pm SD), years	-	-	19.64 \pm 13.27	16.13 \pm 10.58	23.40 \pm 11.95
Diabetes Type (% type II)	-	100	100	100	100
Insulin Dependent (Y/N)	-	14/3	7/13	12/8	15/5
HbA1C, %	-	5.9 \pm 0.7	6.5 \pm 0.6	7.3 \pm 0.9	7.8 \pm 1.3
HTN prevalence, %	10	17	45	80	80

Abbreviations: DR: diabetic retinopathy, SD: standard deviation, HbA1C: Glycated Hemoglobin, HTN: hypertension.

Reprinted from *Fully automated geometric feature analysis in optical coherence tomography angiography for objective classification of diabetic retinopathy* [4], April 22 2019, Copyright 2019 by The Optical Society.

There was not statistical significance in the distributions of age, sex and HTN between the various DR groups and the Control group (ANOVA p-value = 0.14; CHI-Square test p-values = 0.11 and 0.32 respectively).

3.2 ANOVAs

The analysis of variance in the four sets yielded results that were similar but at the same time different than the ones found in [4]. The details of this analyses are shown in the following four paragraphs.

To summarize, width-related features (VBC, VWR1, VWR2) seem to perform better than angle-related features (VBA, CBA1, CBA2) in distinguishing between the five cohorts, especially between Healthy and NPDR subjects.

Weighted Average values obtained by weighted average between arteries and veins seems to be the most sensitive of the four feature sets in separating Control and NoDR groups from the NPDR patients. It must be specified that Artery values are similar in sensitivity and that Weighted Average analysis contains nearly double the data points of the other three feature sets.

A curious finding is that the ratios between artery and vein values (A/V Ratios) do not appear to be significantly different between groups and (almost) all hover around the values of 1. More detailed results descriptions follow.

3.2.1 Artery ANOVA

Described in the following table are the average values, standard deviations and ANOVA p-values for the features extracted from the Artery Maps:

TABLE II: STATISTICS OF ARTERY GEOMETRIC FEATURES

Features	Control	NoDR	Mild	Moderate	Severe	p-values
VBA	83.71 ± 5.32	83.76 ± 4.57	84.06 ± 4.90	84.79 ± 6.02	81.98 ± 6.29	0.344
CBA1	58.89 ± 4.70	58.34 ± 4.07	58.92 ± 4.42	59.93 ± 4.89	57.44 ± 3.97	0.281
CBA2	24.81 ± 3.93	25.42 ± 4.08	25.14 ± 5.26	24.86 ± 3.34	24.54 ± 4.34	0.945
VBC	1.29 ± 0.15	1.33 ± 0.13	1.38 ± 0.11	1.39 ± 0.09	1.36 ± 0.10	0.019
VWR1	0.676 ± 0.08	0.709 ± 0.07	0.766 ± 0.06	0.751 ± 0.05	0.754 ± 0.04	6 E-08
VWR2	0.881 ± 0.06	0.870 ± 0.06	0.856 ± 0.05	0.882 ± 0.05	0.858 ± 0.05	0.246

Based on these p-values, we can assert that only two features, VBC and VWR1, may have the potential to differentiate between the groups. Looking at the very low p-value of VWR1 and the distributions of its values among the different cohorts, one could claim that the Control group and the NoDR group can clearly be separated from the NPDR ones and between themselves as well. Qualitatively, it was seen that VBA, CBA1, CBA2 and VWR2 follow the same trends found by Le et Al, while VBC and VWR1 roughly do so.

3.2.2 Vein ANOVA

Described in the following table are the average values, standard deviations and ANOVA p-values for the features extracted from the Vein Maps:

TABLE III: STATISTICS OF VEIN GEOMETRIC FEATURES

Features	Control	NoDR	Mild	Moderate	Severe	p-values
VBA	83.61 ± 3.75	81.43 ± 4.56	81.31 ± 6.93	85.84 ± 7.42	82.08 ± 6.78	0.062
CBA1	58.66 ± 3.64	56.44 ± 4.66	57.15 ± 6.62	59.46 ± 4.29	57.26 ± 5.95	0.248
CBA2	24.95 ± 3.36	25.00 ± 4.05	24.16 ± 5.27	26.37 ± 5.65	24.82 ± 4.93	0.618
VBC	1.31 ± 0.12	1.31 ± 0.15	1.37 ± 0.17	1.36 ± 0.11	1.29 ± 0.14	0.167
VWR1	0.695 ± 0.07	0.717 ± 0.07	0.757 ± 0.09	0.749 ± 0.06	0.742 ± 0.06	0.0065
VWR2	0.869 ± 0.07	0.855 ± 0.06	0.863 ± 0.06	0.863 ± 0.06	0.829 ± 0.07	0.065

Based on these p-values, we can assert that only one feature, namely VWR1, may have the potential to differentiate between the groups. While VBC and VWR1 p-values are clearly worse than in the Artery ANOVA, it must also be pointed out that VBA and VWR2 p-values are much lower in Veins compared to Arteries, albeit they do not pass the significance threshold of 0.05. VWR2 is nonetheless a little more significant than in the previous study (0.065 vs. 0.086). Qualitatively, it was seen that VBA, CBA1 and VWR2 follow the same trends found by Le et Al, while CBA2 and VWR1 roughly do so.

3.2.3 A/V Ratio ANOVA

Described in the following table are the average values, standard deviations and ANOVA p-values for the Ratio of the features extracted from Artery and Vein Maps:

TABLE IV: STATISTICS OF A/V RATIO GEOMETRIC FEATURES

Features	Control	NoDR	Mild	Moderate	Severe	p-values
VBA	1.00 ± 0.07	1.03 ± 0.07	1.04 ± 0.11	0.99 ± 0.11	1.00 ± 0.09	0.312
CBA1	1.01 ± 0.07	1.04 ± 0.12	1.05 ± 0.18	1.01 ± 0.10	1.01 ± 0.13	0.644
CBA2	1.01 ± 0.22	1.04 ± 0.25	1.03 ± 0.28	0.98 ± 0.25	1.03 ± 0.28	0.953
VBC	1.00 ± 0.16	1.02 ± 0.13	1.02 ± 0.14	1.03 ± 0.14	1.07 ± 0.14	0.389
VWR1	0.98 ± 0.15	0.99 ± 0.11	1.02 ± 0.12	1.01 ± 0.11	1.02 ± 0.10	0.619
VWR2	1.02 ± 0.11	1.02 ± 0.10	0.99 ± 0.08	1.03 ± 0.09	1.04 ± 0.12	0.495

Based on these p-values, we can see that no feature seems to possess the potential to differentiate between the groups. The close average values (hovering around a value of 1) may suggest that there is no significant difference between artery and vein values in every patient type, i.e., arteries and veins features may be influenced in the same way and/or in the same magnitude overall. The (relatively) high standard deviation also could suggest a high inter-individual variability of the A/V Ratio for some of these features, especially CBA2.

3.2.4 Weighted Average ANOVA

Described in the following table are the average values, standard deviations and ANOVA p-values for the features computed from the weighted average of the values extracted from Artery and Vein Maps:

TABLE V: STATISTICS OF WEIGHTED AVERAGE GEOMETRIC FEATURES

Features	Control	NoDR	Mild	Moderate	Severe	p-values
VBA	83.79 ± 3.52	82.97 ± 3.55	82.84 ± 3.91	85.10 ± 4.94	81.92 ± 5.32	0.090
CBA1	58.83 ± 3.68	57.78 ± 3.05	57.61 ± 2.23	59.79 ± 3.64	57.37 ± 3.15	0.039
CBA2	24.96 ± 2.51	25.18 ± 2.89	25.23 ± 3.12	25.31 ± 2.79	24.55 ± 3.45	0.85
VBC	1.29 ± 0.09	1.33 ± 0.10	1.39 ± 0.09	1.38 ± 0.05	1.35 ± 0.08	0.0013
VWR1	0.683 ± 0.06	0.712 ± 0.07	0.765 ± 0.05	0.749 ± 0.04	0.750 ± 0.04	4 E-09
VWR2	0.875 ± 0.04	0.868 ± 0.04	0.864 ± 0.04	0.874 ± 0.03	0.850 ± 0.04	0.069

Based on these p-values, we can assert that three features, CBA1, VBC and VWR1, may have the potential to differentiate between the groups. These features have lower p-values than either the Artery and Vein feature sets. This could be due to two separate influences: one suggesting that accounting for both types of vessels yields more information and a better capacity of separation between the cohorts, and the other considering that this type of analysis has nearly double the data points with respect to the other three feature sets. Looking at the very low p-value of VWR1 and the distributions of its values among the different cohorts, one could claim that the Control group and the NoDR group can clearly be separated from the NPDR ones and between themselves as well with ease. The Weighted Average Analysis obtained with the weighted average thus appears to be the most promising analysis in this study, although with the above specifications about the nearly doubled data points, which could lower the standard deviation which in turn lowers the p-values.

3.3 Boxplots

In order to better represent the distributions of the features among the four sets, a series of six boxplot (one per feature) will be shown below.

3.3.1 VBA Boxplot

The first boxplot shows the distributions of the VBA feature across all cohorts and all feature sets.

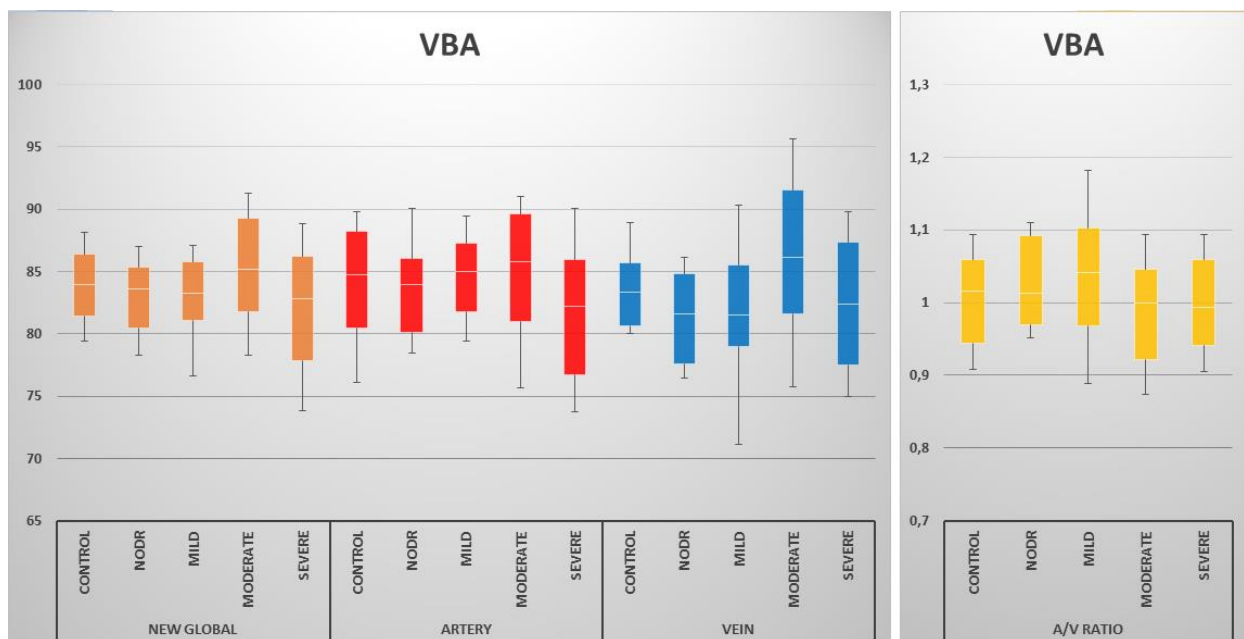


Figure 22: Boxplot of the VBA distributions among all patient cohorts and all feature sets

It can be seen that VBA distributions follow the same pattern in both Arteries and Veins. This makes the Weighted Average distributions have the same trend and the A/V Ratio ones to be almost horizontal and indistinguishable from one another.

3.3.2 CBA1 Boxplot

The second boxplot shows the distributions of the CBA1 feature across all cohorts and all feature sets.

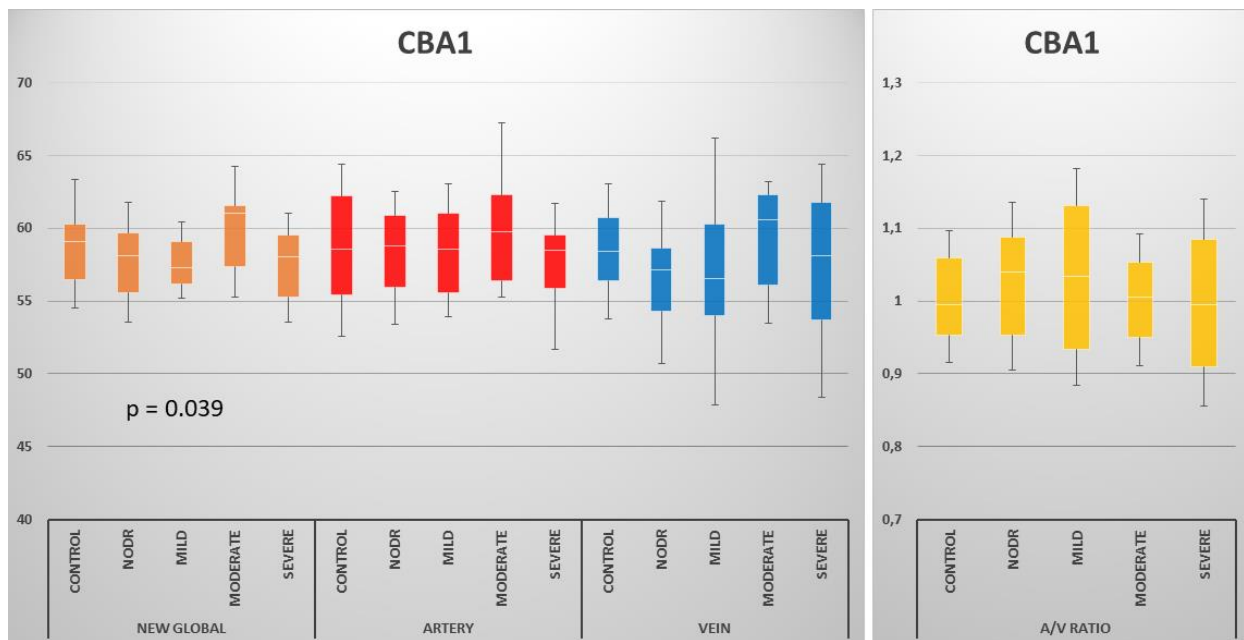


Figure 23: Boxplot of the CBA1 distributions among all patient cohorts and all feature sets

It can be seen that the Artery distributions cannot be differentiated between each other. The Weighted Average distributions have lower variance than both Artery and Vein ones while following a more contained version of the Vein distributions trends. This allows for a significant p-value, given the median value of the Moderate NPDR distribution. The A/V Ratio distributions are, as before, almost horizontal and indistinguishable from one another.

3.3.3 CBA2 Boxplot

The third boxplot shows the distributions of the CBA2 feature across all cohorts and all feature sets.

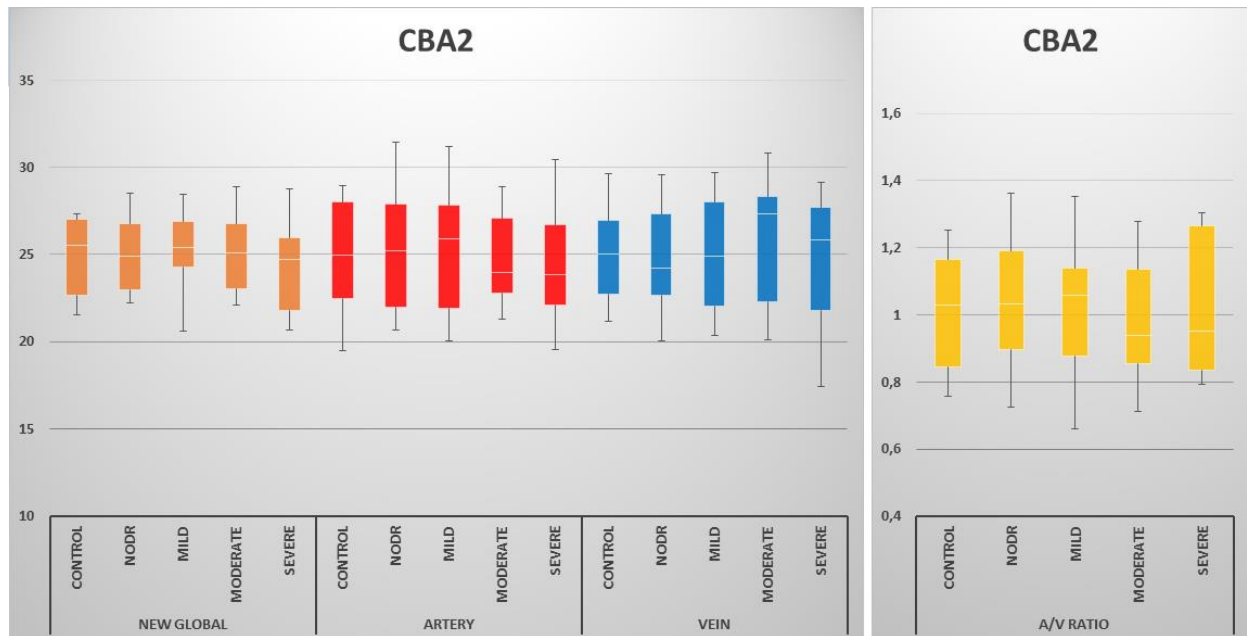


Figure 24: Boxplot of the CBA2 distributions among all patient cohorts and all feature sets

As was the case with the results found in [4], distributions are basically horizontal and indistinguishable from one another. Both Arteries and Vein smaller branch angles do not vary with the presence or staging of the illness in a significant way.

3.3.4 VBC Boxplot

The fourth boxplot shows the distributions of the VBC feature across all cohorts and all feature sets.

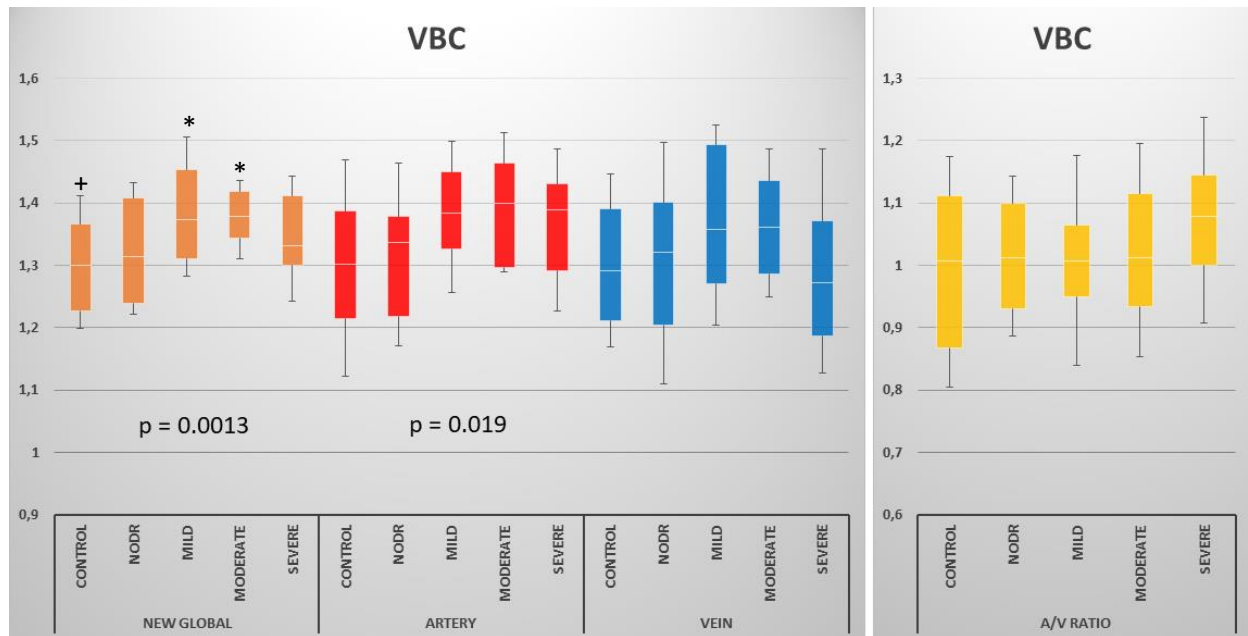


Figure 25: Boxplot of the VBC distributions among all patient cohorts and all feature sets

It can be seen that the Weighted Average distributions show a very clear ascending trend from healthy to NPDR patients. This trend also can be seen in the Artery distributions, while the Vein ones show a final descent in the Severe NPDR stage. This descent prompts an increase in the A/V Ratio but not enough to be clearly separated from the other distributions.

3.3.5 VWR1 Boxplot

The fifth boxplot shows the distributions of the VWR1 feature across all cohorts and all feature sets.

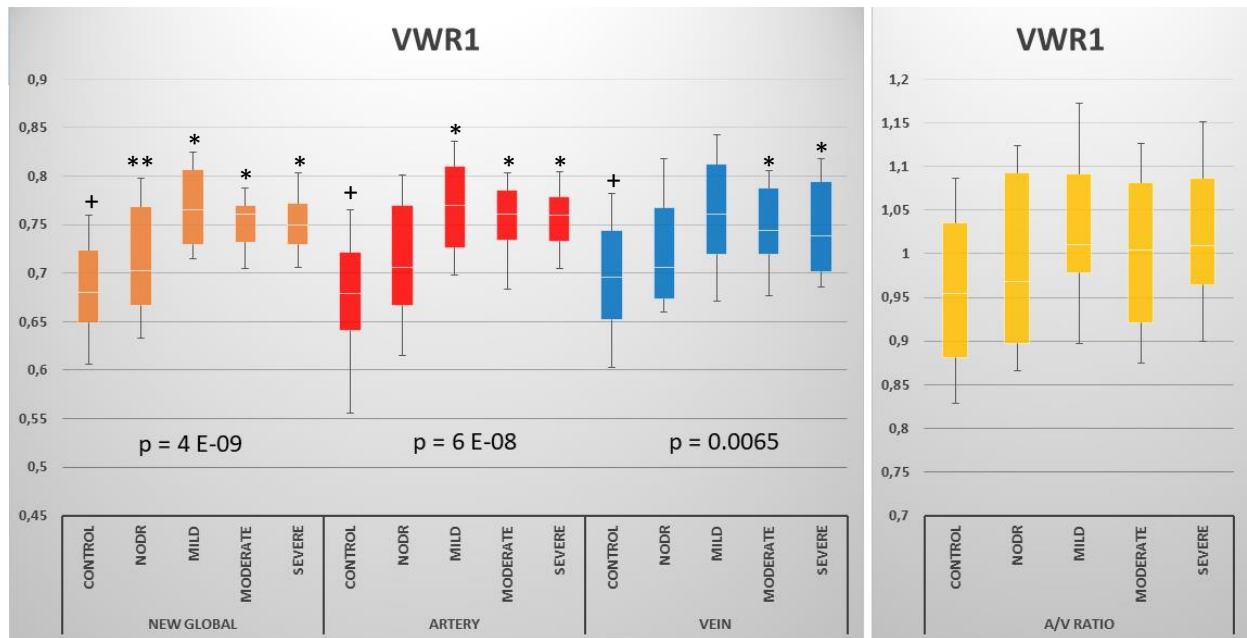


Figure 26: Boxplot of the VWR1 distributions among all patient cohorts and all feature sets

It can be seen that the Weighted Average, Artery and Vein distributions show a very clear ascending trend from healthy to NPDR patients. Since the trend is the same in arteries and veins, the A/V Ratio distributions are almost aligned horizontally. VWR1 shows the same trend as VBC, suggesting that the first child vessel, which is usually the smallest one, increase in diameter with respect to the parent vessel more than the second child vessel (usually the largest one). This is likely due to the blockage of the blood vessels cause by Diabetes, which threatens to totally block the smaller vessels. Thus, the rise in the diameter of the smaller vessels compared to the larger ones is to be expected, and can be reliably used to separate the Control, NoDR and NPDR cohorts.

3.3.6 VWR2 Boxplot

The sixth and final boxplot shows the distributions of the VWR2 feature across all cohorts and all feature sets.

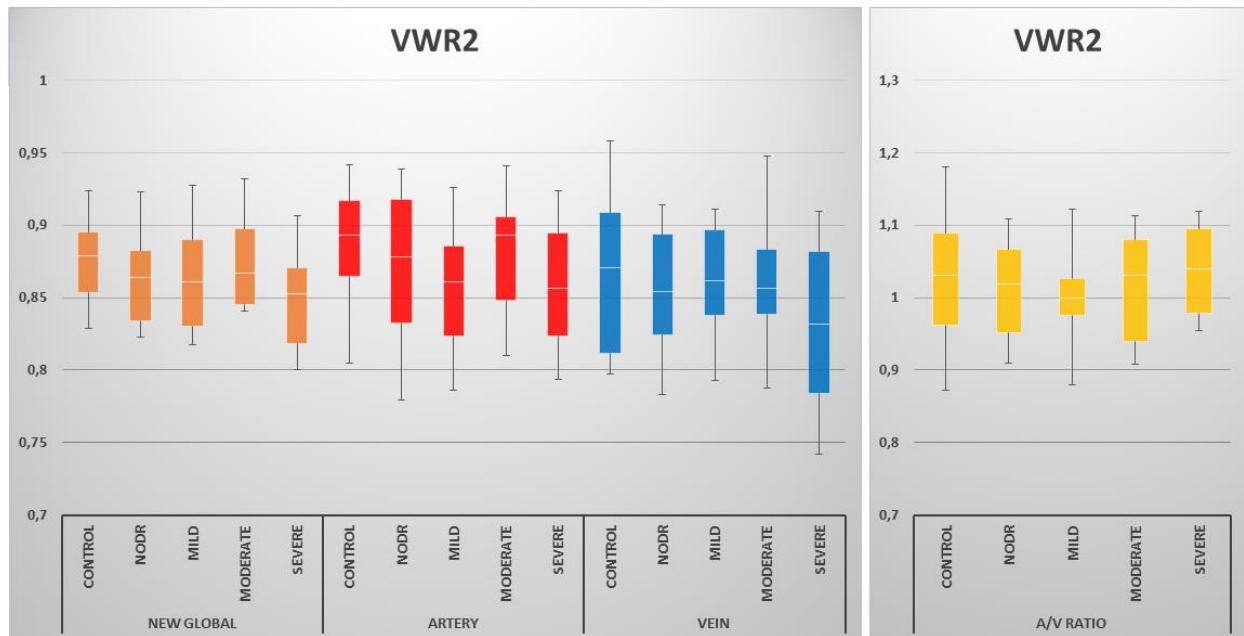


Figure 27: Boxplot of the VWR2 distributions among all patient cohorts and all feature sets

It can be seen that the Weighted Average and Artery distributions show a somewhat descending trend from healthy to NPDR patients. This trend is very much reduced in the Vein distributions. The A/V Ratio distributions are arranged in a very contained V shape.

3.4 Student t-tests

In order to better study the separation capabilities of the most promising features, a series of six Student t-tests was performed. The selected features were the ones which p-values were statistically significant, i.e. CBA1, VBC and VWR1 from the Weighted Average set, VBC and VWR1 from the Artery set and VWR1 from the Vein set. The Bonferroni Correction was used on the results of the t-tests to offset the number of tests performed per feature. This procedure is explained in the next paragraphs.

3.4.1 Bonferroni Correction

The Bonferroni Correction is based on the fact that the probability of observing rare events increases as the number of tested hypotheses increases [19]. Thus, the type I errors (false positives) are maximized. In this study, a total of ten hypotheses are tested for each feature. Dividing by ten the significance threshold will minimize the likelihood of false positives, but it will increase the likelihood of false negatives. Nevertheless, the p-values that will still be statistically significant are sure to be so. This, in turn, allows a better recognition of the differences between the various distributions. In the following tables, all shown p-values will have to be lower than 0.005 to achieve significance. In order to avoid false negative, the p-values that will be statistically significant without the correction but are not if the correction is applied will be marked with a * symbol. The significant p-values will be shown in **bold**.

3.4.2 Weighted Average CBA1 Individual Comparison t-tests

The individual t-test comparison for Weighted Average CBA1 feature is shown below:

TABLE VI: STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR WEIGHTED AVERAGE CBA1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST

	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Control					
NoDR	0.2498				
Mild NPDR	0.1304	0.8241			
Moderate NPDR	0.3549	0.0534	0.0233*		
Severe NPDR	0.0779	0.6033	0.7188	0.0136*	

It can be seen that there are no strictly significant p-values, even if, as seen in the CBA1 boxplot, the Moderate NPDR distribution could be differentiated from the Mild NPDR and Severe NPDR ones.

3.4.3 Weighted Average VBC Individual Comparison t-tests

The individual t-test comparison for Weighted Average VBC feature is shown below:

TABLE VII: STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR WEIGHTED AVERAGE VBC, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST

	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Control					
NoDR	0.3075				
Mild NPDR	0.000883	0.0369*			
Moderate NPDR	0.000171	0.0379*	0.6417		
Severe NPDR	0.0260*	0.4324	0.0773	0.0627	

It can be seen that the Control distribution can clearly be separated from the Mild and Moderate NPDR. Three p-values, namely Control vs. Severe, NoDR vs. Mild and NoDR vs. Moderate, are significant before the Bonferroni Correction but not significant after.

3.4.4 Weighted Average VWR1 Individual Comparison t-tests

The individual t-test comparison for Weighted Average VWR1 feature is shown below:

TABLE VIII: STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE COMPARISONS FOR WEIGHTED AVERAGE VWR1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST

	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Control					
NoDR	0.0904				
Mild NPDR	0.000001	0.0035			
Moderate NPDR	0.000004	0.0215*	0.2453		
Severe NPDR	0.0000002	0.0126*	0.2371	0.9111	

It can be seen that the Control group can be clearly separated from any NPDR group, but not from the NoDR group. This latter group can be distinguished from Mild NPDR patients but it is unclear if it's statistically different from Moderate and Severe NPDR distributions. This is the feature with the lowest p-value in the study, and the high separating power between healthy and NPDR patients could be easily exploited in a classifier.

3.4.5 Artery VBC Individual Comparison t-tests

The individual t-test comparison for Artery VBC feature is shown below:

TABLE IX: STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE
COMPARISONS FOR ARTERY VBC, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST

	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Control					
NoDR	0.4154				
Mild NPDR	0.0212*	0.1347			
Moderate NPDR	0.0061*	0.0554	0.7004		
Severe NPDR	0.0279*	0.2047	0.6311	0.3301	

It can be seen that, compared to the Weighted Average VBC individual comparison t-test table, these p-values are higher, as was the ANOVA p-value for this feature. However, Weighted Average analysis has the advantage of having nearly double the data points that lower the standard deviation and lower the p-values.

3.4.6 Artery VWR1 Individual Comparison t-tests

The individual t-test comparison for Artery VWR1 feature is shown below:

TABLE X: STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE
COMPARISONS FOR ARTERY VWR1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST

	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Control					
NoDR	0.1071				
Mild NPDR	0.000003	0.0054*			
Moderate NPDR	0.00009	0.0333*	0.3839		
Severe NPDR	0.000004	0.0109*	0.3867	0.8377	

It can be seen that, as was the case in the Weighted Average VWR1 analysis, the Control group can be clearly separated from any NPDR group, but not from the NoDR group.

3.4.7 Vein VWR1 Individual Comparison t-tests

The individual t-test comparison for Vein VWR1 feature is shown below:

TABLE XI: STUDENT T-TESTS RESULTS FOR INDIVIDUAL GROUP SIGNIFICANCE
COMPARISONS FOR VEIN VWR1, DUE TO THE SIGNIFICANCE OF THE ANOVA TEST

	Control	NoDR	Mild NPDR	Moderate NPDR	Severe NPDR
Control					
NoDR	0.2694				
Mild NPDR	0.0083*	0.1011			
Moderate NPDR	0.0047	0.1187	0.7162		
Severe NPDR	0.0040	0.1677	0.4650	0.6630	

It can be seen that the Vein VWR1 feature should be capable of separating the healthy patients from the ones with Moderate and Severe NPDR, but it is unclear if it can separate them effectively from the Mild NPDR cohort. Overall, it is clearly worse than the corresponding Artery p-values.

CHAPTER 4

DISCUSSION

To summarize, an automated procedure was used to differentiate the vessels of OCTA images into arteries and veins. Then, a fully automated method was used to measure and extract six geometric features (VBA, CBA1, CBA2, VBC, VWR1 and VWR2). Statistical analyses, ANOVA and Student t-tests were performed on the extracted parameters belonging to five different cohorts (Control, NoDR, Mild NPDR, Moderate NPDR and Severe NPDR subjects).

4.1 Comparison with the previous study

Looking at Table 3 in the paper from Le. et Al. [4], it can be seen that there are some differences in significances and trends of the various parameters. The differences will be examined first, and some causes for these differences will be presented next.

Looking at the p-values in the previously said table, it can be noted that the Angle-based features (VBA, CBA1 and CBA2) are, on average, less significant than in the non-differentiated study, while the Width-related features (VBC, VWR1 and VWR2) shown, on average, increased significance. In particular, VBA, which was the most promising feature in the previous study, is now above the significance threshold. Looking at the distributions, the mean values and standard deviations are higher, allowing for less separation capability. The trend of the VBA distribution is conserved in the Artery features, even if the averages are nearer to each other. In the Vein distributions, the average values are more far apart from each other (especially the Moderate NPDR values), but it is still not enough to reach significance. CBA1 and CBA2 also follow the same observations. Looking at the Student t-tests of the previous study [4], it can be noted that

VBA could not allow for a total Control-NoDR-NPDR separation, since Control and NoDR distributions could not be clearly separated from the Severe NPDR one.

Regarding the Width-based measurements, VBC and VWR1 have a very high significance and separating power. They also show a very clear trend, where the values get higher the more the NPDR progresses. The Student t-tests confirm that they can be used to differentiate between Control and NoDR subjects on one side and the NPDR subjects on the other. The trends of the previous study are not preserved and are even inverted. Looking at the Student t-tests of the previous study [4], it can be noted that both VBC and VWR1 could not separate between NoDR and any of the NPDR distributions, and also between Control and Moderate NPDR distributions. Another thing to note is that, similarly to this study, VBC had less separating power than VWR1.

It is worth noting that, although the trend of VBC appears inverted compared to the previous study, its distributions among the patient groups give us the same insights. From the work by Adam et al. [42] the estimated VBC for a perfectly symmetrical branching is circa 1.26, and any variation from that value implies an asymmetry in the branching. In the previous study the VBC values decreased with the onset and progression of NPDR, while in this study they increase, but in both cases the growing asymmetry is shown. That is true for the Artery and Weighted Average analysis: in the Vein distributions we find that the Severe NPDR values are closer to a situation of symmetry than the two other NPDR groups. This is in turn signaled by a little spike in the A/V Ratio, but not enough to be detected as significantly different from the other distributions.

4.1.1 Differences between Angle- and Width-based Features

The different performance of the Angle and Width-based parameters, especially with regard to the same parameters in the old study, is worth investigating since it may offer insights on the physiology of the illness or on the method itself. As said before, Angle-based features have decreased significance and sensitivity while Width-based features have increased significance and sensitivity.

Since $CBA1 + CBA2 = VBA$, the three features suffer from the fact that a measuring error can trickle down and influence the others. This influences the distributions in the five patient cohorts, and in turn influences the significance of the features themselves. Compared to the previous paper, VBA averages are closer together, which means that also CBA1 and CBA2 averages are closer together. Also, a higher standard deviation is inherited from VBA by the other two angle-related features. The lower significance may be explained in this way.

When measuring quantities, taking ratios of two or more measurements can help in reducing or enhancing the sensitivity of any single measurement. The width-related features are constructed in such a way. It is natural that, if any differences exist between the parent and vessel diameters distributions, the three features distributions have enhanced sensitivity. This is exactly what has been found out in this study: the p-values of VWR1 in particular are very low, the trends are clear and are conserved between Artery, Vein and Weighted Average analyses. Moreover, since the ratio is used to quantify the relationship between (in this case) the parent vessel width and the children vessel width, it cannot detect changes in the vessels' widths that still result in the same relationship between them.

4.2 Sources of Differences

In this section, the most prominent sources of the differences shown in the previous paragraph will be discussed.

The sources that will be presented below are three:

1. the differences in the vessel map generation
2. the analyzed branchpoints and their correspondence
3. the measurement window cleanness

4.2.1 Differences in Vessel Map Generation

The artery-vein classification algorithm used in this study to generate the AV vessel maps is an extra step which may have an influence on the measurements of the features. The previous results [4] were indeed measured on unclassified vessel maps. In order to maintain the automatic characteristics of the said classification algorithm, its resulting vessel maps were not touched or cured. The two similar but different processes in these two studies could result in differences in the measurements of the features. It is necessary to remind that, used on unclassified maps, the algorithm used for this study outputs the same results as the ones used by Le et al.

4.2.2 Correspondence of Analyzed Branchpoints

A complete study of the branchpoints analyzed by the algorithm was conducted. In this sub-study, every branchpoint of artery and vein maps was manually inspected with the objective to see how many branchpoints were analyzed in the classified and unclassified maps and how many of those branchpoint were analyzed in both cases. The latter metric was defined as Branchpoint Correspondence. Three main findings are described below:

The first finding is that the combined Artery and Vein maps have, on average, 1 to 2 analyzed branchpoint more than the unclassified maps. The only notable exception is the NoDR cohort, in which there are, on average, 5 branchpoints more than in the unclassified maps.

The second finding is that the Branchpoint Correspondence metric is, on average, around 31.6 %. This means than little less than one third of the branchpoints are analyzed in both the Artery and Vein maps and the unclassified maps. This percentage varies and is highest in the Moderate NPDR patients and lowest in the NoDR patients. This last figure also derives from the fact that the NoDR cohort has more analyzed branchpoints.

The third finding regards the number of analyzed branchpoints in Artery maps compared to Vein maps. On average, Artery maps have 3 to 4 more analyzed branchpoints that Vein maps. This number is lower in Healthy patients (1 to 2 more) and is higher in Severe NPDR patients (5 to 6 more). This could be explained by looking back at Figure 20: since arteries tend to cross over veins and not vice versa, the Vein map suffer from discontinuities of its vessels. If these crossings happen near a vein branchpoint, they will disrupt the identification and the analysis of its features.

In the retina, arteries tend to cross over veins. In the unclassified maps, this is not an issue, but in the classified maps it can become so. That is because, some veins are split by arteries crossing over them. In turn, this results in more branchpoint exclusions in veins since a vein branching can be cut off by an artery. This explains the observed difference in numbers of analyzed branchpoints in vein maps compared to artery maps that was described as the third major finding above. It must be noted that, in order to affect any measurement, the crossing must occur close or right above a vein branching.

This sub-study offers many insights on the differences between the showed results and the ones from Le at Al. [4]. Below, an example of a very common situation encountered in the sub-study is shown to better understand the high impact of these metrics on the results.

4.2.2.1 Example: two common scenarios

Manually analyzing the Branchpoint Correspondence between classified and unclassified images, two common scenarios were identified that could account for a significant part of the differences between the two types of images.

The first one is about branchpoints that appear in unclassified maps but not in the classified ones. Since the images are classified into Artery and Vein maps, “false” branchpoints (not representing a real branching) like the one shown below are separated and thus not taken into account by the algorithm. Classifying the images cleans the final values of the features of any noise caused by these occurrences, of which an example is shown in Figure X. An arteriole terminates its course close to a vein: without separating arteries from veins, a “false” branchpoint is identified where, in fact, no such branchpoint is really present.

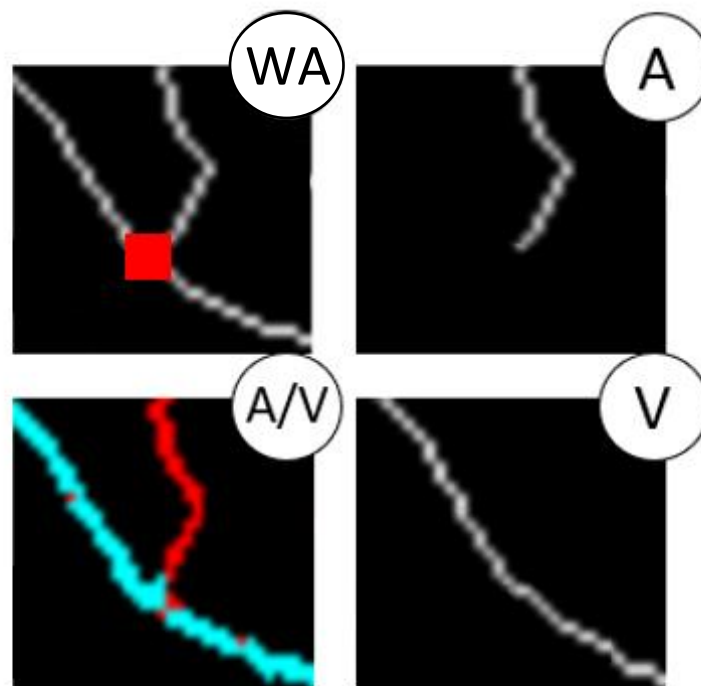


Figure 28: False branchpoint scenario. **WA**: Weighted Average unclassified vessel map. The *red square* represents an identified branchpoint. **A/V**: Artery-Vein differentiated vessel map.

A and V: Artery and Vein map.

The second scenario is shown below in Figure X. In this example, in the unclassified Weighted Average skeleton map (marked as G) an artery branching and a vein branching intersect in a way that the two branchpoints are too close and one of them (the vein one) is misidentified as a four-way intersection. Red squares were used to signal branchpoints that were not excluded. After classifying the vessels and separating the Artery and Vein maps, the two branchpoints are correctly identified and analyzed by the algorithm. Identifying the right branchpoints and excluding the bad ones leads to better endpoints identification and different values being measured and counted.

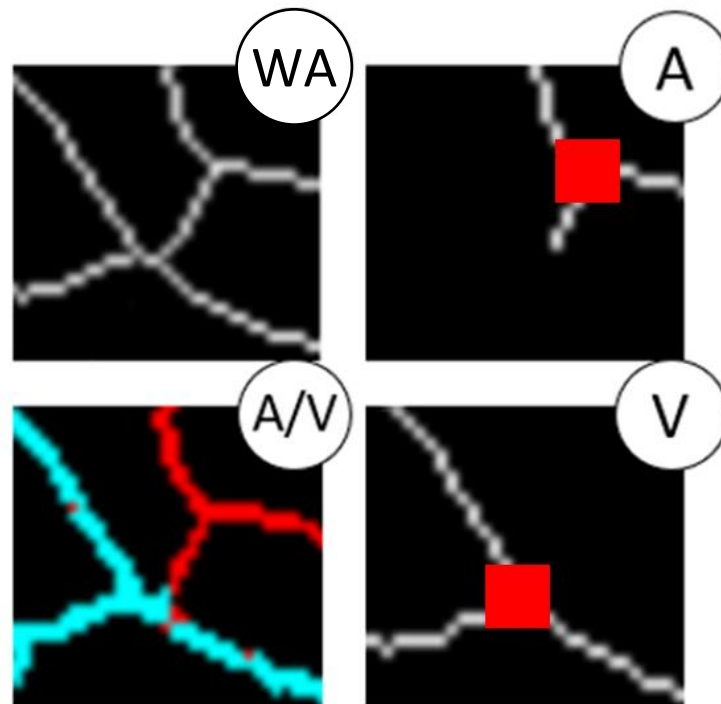


Figure 29: Non-detection of real branchpoints. **WA**: Weighted Average unclassified vessel map. **A/V**: Artery-Vein differentiated vessel map. **A and V**: Artery and Vein map. The *red squares* represent identified branchpoints.

4.2.3 The Influence of Measurement Windows

This reasoning starts from the fact that, in order to measure vessel width, a window centered on the endpoint is identified and its contents are extracted for the computations. In the classified and separated Artery and Vein maps, these windows will obviously contain less “noise”. This “noise” is composed of many possible components: mainly endings and/or parts of other vessels and/or the same types of vessels, etc. Single vessel’s width measurements are surely improved by a cleaner and emptier measurement window. If we couple the resulting more accurate values with the use of ratios, we should obtain better features with better sensitivity. This is reflected by the fact that width-based features have greater significance and separating power between the various groups. Width-related features have clearer trends, lower standard deviations and more separated distributions, and this paragraph may be the major reason why.

4.3 Why the Bonferroni Correction

In this study, the Bonferroni Correction was employed to correct the p-values of the Student t-tests. It is the most conservative interpretation of results of multiple statistical tests, giving the priority to minimizing the false positives (type I statistical errors). On the other hand, it also has the effect of increasing type II statistical errors, i.e., false negatives. Having this information, the resulting p-values should be interpreted in this way:

- **p-values < 0.005:** the two distributions can be regarded as significantly different and thus separable.
- **p-values > 0.05:** the two distributions cannot be regarded as significantly different and thus cannot be separated.
- **0.005 < p-values < 0.05:** the two distributions fall in a sort of limbo. The Bonferroni correction makes them not significant in order to minimize false positives but may incur in a false negative case. In the t-tests performed on the six significant features found in the

ANOVA analyses, 14 p-values fall in this category. It cannot be excluded that false negatives could be present in these 14 p-values, as cannot be excluded that the Bonferroni Correction correctly avoided some cases of false positives.

To summarize, p-values that retain significance after the correction are sure to be true positives, while the ones that do not achieve it without the correction are sure to be true negatives. The others in between are either false negatives (after the correction) or false positives (before the correction).

CHAPTER 5

CONCLUSIVE REMARKS

To summarize, differentiating arteries and vein in OCTA images to detect Diabetic Retinopathy through the extraction of the six geometric features has given some insights about the process and the illness itself.


The first insight is that a Weighted Average analysis of the feature seems to possess the best separating power compared to the other types of analysis explored in this study. However, it must be noted that it had nearly double the data points of the other three feature sets. In this light, the Artery distributions contain a similar amount of separating power. The differences between using an unclassified map and computing the weighted average of the values from the classified maps has been explored in detail.

The second insight is that, with respect to the six features analyzed in this study, there seems to be no statistical difference in how NPDR affects arteries and veins. Almost all of the A/V Ratios hover around a value of 1. There are some exceptions, but they cannot be clearly separated from the other distributions. This may suggest that the first stages of Diabetic Retinopathy affect arteries' and vein' branching features with the same magnitude. In future studies, other features may show different behaviors.

The third insight is that the width-related features seem to be more capable of discriminating between the patients' groups than the angle-related ones. Changes in the blood vessels' width seem to be much more informative than the ones that occur in the relative angles in bifurcations.

APPENDIX

Material Permission: since all the reprinted or adapted figures and tables were published on Open Access journals, the Creative Commons License applies.



Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)


This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).


You are free to:


Share — copy and redistribute the material in any medium or format

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

**Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**NonCommercial** — You may not use the material for [commercial purposes](#).

**NoDerivatives** — If you [remix, transform, or build upon](#) the material, you may not distribute the modified material.

No additional restrictions — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

CITED LITERATURE

1. Michael M. Engelgau, Linda S. Geiss, Jinan B. Saaddine, et al. The Evolving Diabetes Burden in the United States. *Ann Intern Med.* 2004;140:945-950. [Epub ahead of print 1 June 2004]. doi:10.7326/0003-4819-140-11-200406010-00035
2. Kempen, J. H., O'Colmain, B. J., Leske, M. C., Haffner, S. M., Klein, R., Moss, S. E., Taylor, H. R., Hamman, R. F., West, S. K., Wang, J. J., Congdon, N. G., & Friedman, D. S. (2004). The Prevalence of Diabetic Retinopathy among Adults in the United States. *Archives of ophthalmology*, 122(4), 552-563. <https://doi.org/10.1001/archopht.122.4.552>
3. Pandey SK, Sharma V. World diabetes day 2018: Battling the Emerging Epidemic of Diabetic Retinopathy. *Indian J Ophthalmol.* 2018;66(11):1652-1653. doi:10.4103/ijo.IJO_1681_18
4. Le D, Alam M, Miao BA, Lim JI, Yao X. Fully automated geometric feature analysis in optical coherence tomography angiography for objective classification of diabetic retinopathy. *Biomed Opt Express.* 2019;10(5):2493-2503. Published 2019 Apr 22. doi:10.1364/BOE.10.002493
5. Murray CD. The Physiological Principle of Minimum Work: I. The Vascular System and the Cost of Blood Volume. *Proc Natl Acad Sci U S A.* 1926;12(3):207-214. doi:10.1073/pnas.12.3.207
6. Ernest JT, Goldstick TK, Engerman RL. Hyperglycemia impairs retinal oxygen autoregulation in normal and diabetic dogs. *Invest Ophthalmol Vis Sci.* 1983;24(7):985-989.
7. Grunwald JE, DuPont J, Riva CE. Retinal haemodynamics in patients with early diabetes mellitus. *Br J Ophthalmol.* 1996;80(4):327-331. doi:10.1136/bjo.80.4.3

CITED LITERATURE (continued)

8. Sasongko MB, Wang JJ, Donaghue KC, et al. Alterations in retinal microvascular geometry in young type 1 diabetes. *Diabetes Care*. 2010;33(6):1331-1336. doi:10.2337/dc10-0055
9. Gupta A., Chhikara R., Diabetic Retinopathy: Present and Past, *Procedia Computer Science*, 2018, Vol. 132: 1432-1440
10. Simon S. Gao, Yali Jia, Miao Zhang, Johnny P. Su, Gangjun Liu, Thomas S. Hwang, Steven T. Bailey, David Huang; Optical Coherence Tomography Angiography. *Invest. Ophthalmol. Vis. Sci*. 2016;57(9):OCT27-OCT36
11. Eladawi, N., Elmogy, M., Khalifa, F., Ghazal, M., Ghazi, N., Aboelfetouh, A., Riad, A., Sandhu, H., Schaal, S. and El-Baz, A. (2018), Early diabetic retinopathy diagnosis based on local retinal blood vessel analysis in optical coherence tomography angiography (OCTA) images. *Med. Phys.*, 45: 4582-4599
12. Solomon SD, Chew E, Duh EJ, et al. Diabetic Retinopathy: A Position Statement by the American Diabetes Association [published correction appears in *Diabetes Care*. 2017 Jun;40(6):809] [published correction appears in *Diabetes Care*. 2017 Jul 13;:]. *Diabetes Care*. 2017;40(3):412-418. doi:10.2337/dc16-2641
13. *Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging*, 2nd Edition, Patrick J. Saine and Marshall E. Tyler, Butterworth-Heinemann Medical; ISBN: 0750673729
14. Alam M, Toslak D, Lim JI, Yao X. OCT feature analysis guided artery-vein differentiation in OCTA. *Biomed Opt Express*. 2019;10(4):2055-2066. Published 2019 Mar 26. doi:10.1364/BOE.10.002055

CITED LITERATURE (continued)

15. Li C., Xu C., Anderson A.W., Gore J.C. (2009) MRI Tissue Classification and Bias Field Estimation Based on Coherent Local Intensity Clustering: A Unified Energy Minimization Framework. In: Prince J.L., Pham D.L., Myers K.J. (eds) Information Processing in Medical Imaging. IPMI 2009. Lecture Notes in Computer Science, vol 5636. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02498-6_24
16. Vázquez S.G., Barreira N., Penedo M.G., Saez M., Pose-Reino A. (2010) Using Retinex Image Enhancement to Improve the Artery/Vein Classification in Retinal Images. In: Campilho A., Kamel M. (eds) Image Analysis and Recognition. ICIAR 2010. Lecture Notes in Computer Science, vol 6112. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-13775-4_6
17. M. Alam, T. Son, D. Toslak, J. I. Lim, and X. Yao, “Combining ODR and Blood Vessel Tracking for Artery- Vein Classification and Analysis in Color Fundus Images,” Transl. Vis. Sci. Technol. 7(2), 23 (2018).
18. Habib, M.S., Al-Diri, B., Hunter, A. et al. The association between retinal vascular geometry changes and diabetic retinopathy and their role in prediction of progression – an exploratory study. BMC Ophthalmol 14, 89 (2014). <https://doi.org/10.1186/1471-2415-14-89>
19. Bonferroni, C. E., Teoria statistica delle classi e calcolo delle probabilità, Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze (1936)
20. Alam M, Thapa D, Lim JI, Cao D, Yao X. Quantitative characteristics of sickle cell retinopathy in optical coherence tomography angiography. Biomed Opt Express. 2017;8(3):1741-1753. Published 2017 Feb 23. doi:10.1364/BOE.8.001741

CITED LITERATURE (continued)

21. Minhaj Alam, Jennifer I. Lim, Devrim Toslak, Xincheng Yao; Differential Artery–Vein Analysis Improves the Performance of OCTA Staging of Sickle Cell Retinopathy. *Trans. Vis. Sci. Tech.* 2019;8(2):3. doi: <https://doi.org/10.1167/tvst.8.2.3>.
22. Sambhav K, Grover S, Chalam KV. Temporal Thinning In Sickle Cell Retinopathy Is Associated With Diminished Perfusion On Octa And Dense Scotoma On Microperimetry. *Retin Cases Brief Rep.* 2019;13(4):308-313. doi:10.1097/ICB.0000000000000610
23. Grego, L., Pignatto, S., Alfier, F. et al. Optical coherence tomography (OCT) and OCT angiography allow early identification of sickle cell maculopathy in children and correlate it with systemic risk factors. *Graefes Arch Clin Exp Ophthalmol* 258, 2551–2561 (2020). <https://doi.org/10.1007/s00417-020-04764-y>
24. Arya M, Sabrosa AS, Duker JS, Waheed NK. Choriocapillaris changes in dry age-related macular degeneration and geographic atrophy: a review. *Eye Vis (Lond)*. 2018;5:22. Published 2018 Sep 15. doi:10.1186/s40662-018-0118-x
25. Perrott-Reynolds R, Cann R, Cronbach N, et al. The diagnostic accuracy of OCT angiography in naive and treated neovascular age-related macular degeneration: a review. *Eye (Lond)*. 2019;33(2):274-282. doi:10.1038/s41433-018-0229-6
26. Moreira-Neto CA, Moulton EM, Fujimoto JG, Waheed NK, Ferrara D. Choriocapillaris Loss in Advanced Age-Related Macular Degeneration. *J Ophthalmol*. 2018;2018:8125267. Published 2018 Jan 30. doi:10.1155/2018/8125267
27. Usman M, Iqbal K, Ali MH, Nafees K. Features and Diagnostic Accuracy of Optical Coherence Tomography Angiography in Neovascular Age-related Macular Degeneration. *Cureus*. 2019;11(12):e6485. Published 2019 Dec 28. doi:10.7759/cureus.6485

CITED LITERATURE (continued)

28. Arrigo A, Aragona E, Bordato A, et al. Quantitative OCTA parameter variations after treatment of macular neovascularization secondary to age-related macular degeneration [published online ahead of print, 2020 Dec 7]. *Retina*. 2020; 10.1097/IAE.0000000000003065. doi:10.1097/IAE.0000000000003065
29. Tey KY, Teo K, Tan ACS, et al. Optical coherence tomography angiography in diabetic retinopathy: a review of current applications. *Eye Vis (Lond)*. 2019;6:37. Published 2019 Nov 18. doi:10.1186/s40662-019-0160-3
30. Khadamy J, Abri Aghdam K, Falavarjani KG. An Update on Optical Coherence Tomography Angiography in Diabetic Retinopathy. *J Ophthalmic Vis Res*. 2018;13(4):487-497. doi:10.4103/jovr.jovr_57_18
31. Alonso-Plasencia M, Abreu-González R, Gómez-Culebras MA. Structure-Function Correlation Using OCT Angiography And Microperimetry In Diabetic Retinopathy. *Clin Ophthalmol*. 2019;13:2181-2188. Published 2019 Nov 11. doi:10.2147/OPTH.S220877
32. Cano J, O'Neill WD, Penn RD, et al. Classification of advanced and early stages of diabetic retinopathy from non-diabetic subjects by an ordinary least squares modeling method applied to OCTA images. *Biomed Opt Express*. 2020;11(8):4666-4678. Published 2020 Jul 27.
33. M. Alam, D. Le, T. Son, J. I. Lim, and X. Yao, "AV-Net: deep learning for fully automated artery-vein classification in optical coherence tomography angiography," *Biomedical optics express* 11, 5249-5257 (2020).
34. 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 (Maywood)* 244, 813-818 (2019).

CITED LITERATURE (continued)

35. M. Alam, D. Toslak, J. I. Lim, and X. Yao, "Color Fundus Image Guided Artery-Vein Differentiation in Optical Coherence Tomography Angiography," *Invest Ophthalmol Vis Sci* 59, 4953-4962 (2018).
36. X. Yao, M. N. Alam, D. Le, and D. Toslak, "Quantitative optical coherence tomography angiography: A review," *Exp Biol Med* (Maywood) 245, 301-312 (2020).
37. M. Alam, D. Le, J. I. Lim, R. V. P. Chan, and X. Yao, "Supervised Machine Learning Based Multi-Task Artificial Intelligence Classification of Retinopathies," *J Clin Med* 8(2019).
38. M. Alam, Y. Zhang, J. I. Lim, R. V. P. Chan, M. Yang, and X. Yao, "Quantitative Optical Coherence Tomography Angiography Features for Objective Classification and Staging of Diabetic Retinopathy," *Retina* 40, 322-332 (2020).
39. Y. T. Hsieh, M. N. Alam, D. Le, C. C. Hsiao, C. H. Yang, D. L. Chao, and X. Yao, "OCT Angiography Biomarkers for Predicting Visual Outcomes after Ranibizumab Treatment for Diabetic Macular Edema," *Ophthalmol Retina* (2019).
40. D. Le, M. Alam, C. K. Yao, J. I. Lim, Y. T. Hsieh, R. V. P. Chan, D. Toslak, and X. Yao, "Transfer Learning for Automated OCTA Detection of Diabetic Retinopathy," *Transl Vis Sci Technol* 9, 35 (2020).
41. M. Alam, D. Le, J. I. Lim, and X. Yao, "Vascular Complexity Analysis in Optical Coherence Tomography Angiography of Diabetic Retinopathy," *Retina* (2020).
42. J. A. Adam, "Blood vessel branching: beyond the standard calculus problem," *Math. Mag.* 84(3), 196–207 (2011)

VITA

NAME: Mattia Castelnovo

EDUCATION: Master of Science in Bioengineering, University of Illinois at
Chicago, May 2021, USA

Bachelor's Degree in Biomedical Engineering, Jul 2018,
Politecnico di Milano, Italy

LANGUAGE SKILLS: Italian – Native speaker

English – Professional knowledge and proficiency:

- IELTS Certification obtained in Apr 2018
(Score 8/10)
- A.Y. 2018-2021: Lessons and Exams attended in
English
- A.Y. 2019-2020: One year of study abroad in
Chicago, Illinois