

Early Hypertensive Retinopathy detection by  
leveraging salient regions using deep learning

2022

Mónica Zamarrón Pérez



# Universidad Autónoma de Querétaro

## Facultad de Ingeniería

Early Hypertensive Retinopathy detection by  
leveraging salient regions using deep learning

### Tesis

Que como parte de los requisitos para  
obtener el grado de

### **Maestro en Ciencias en Inteligencia Artificial**

Presenta:

Mónica Zamarrón Pérez

Dirigido por:

Dr. Saúl Tovar Arriaga

Co-Director:

Dra. Renata García Franco

Querétaro, Qro. a Agosto 2022



Dirección General de Bibliotecas y Servicios Digitales  
de Información



Early Hypertensive Retinopathy detection by  
leveraging salient regions using deep learning

**por**

Mónica Zamarrón Pérez

se distribuye bajo una [Licencia Creative Commons  
Atribución-NoComercial-SinDerivadas 4.0  
Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/).

**Clave RI:** IGMAC-300598



Universidad Autónoma de Querétaro

Facultad de Ingeniería

Maestría en Ciencias en Inteligencia Artificial

**Early Hypertensive Retinopathy detection by leveraging salient regions using deep learning**

TESIS

Que como parte de los requisitos para obtener el grado de  
Maestro en Ciencias en Inteligencia Artificial

Presenta:

**Ing. Mónica Zamarrón Pérez**

Dirigido por:

**Dr. Saúl Tovar Arriaga**

Co-Dirigido por:

**Dra. Renata García Franco**

SINODALES

Dr. Saúl Tovar Arriaga

Presidente

Dra. Renata García Franco

Secretario

Dr. Jesús Carlos Pedraza Ortega

Vocal

Dr. Marco Aceves Fernández

Suplente

Dr. Juan Manuel Ramos Arreguín

Suplente

Centro Universitario, Querétaro, Qro.

Agosto 2022

México



*This thesis is dedicated to my husband for his never-ending support, and my parents who gave me the love for medical research.*



# Acknowledgments

This master's thesis was successfully accomplished thanks to CONACYT (Consejo Nacional de Ciencia y Tecnología) funding, and the support of Postgraduate Faculty of Engineering at Universidad Autónoma de Querétaro.

I would like to thank my advisor, Dr. Saúl Tovar Arriaga, as well as Dr. Jesús Carlos Pedraza Ortega, for giving me their assistance and guidance when needed. Likewise, I am deeply grateful to the INDEREB (Instituto de la Retina del Bajío) team, especially Dra. Renata Garcia Franco, for providing numerous fundus images that made possible this research along with her medical knowledge and expertise. Finally, I would like to thank my husband for his endless support, this achievement is ours.





# Abstract

Hypertension is a disease with high incidence and prevalence characterized by its central role in retinal vascular damage, known as hypertensive retinopathy. Retinal vascular changes seen in fundus images can be employed as biomarkers to measure generalized vessels constrictions through the Arteriolar-to-Venular Ratio (AVR). Advances in Artificial Intelligence and straightforward medical image acquisition have allowed the development of innovative support systems for the automatic diagnosis of eye diseases.

This work presents an automatic system that uses a U-net-based neural network to classify arteries, veins, and the optic disc from fundus images through semantic segmentation to calculate the AVR and determine whether a patient has abnormal results and thus the disease. The system was tested on three datasets. Two are publicly available: DRIVE and INSPIRE-AVR, each consisting of 40 fundus images; the third database was created with the INDEREB aid. It has two versions with 105 photos in total and will be accessible to the public. The A/V classification results achieved with the proposed method are 95.74% accuracy, 97.37% specificity, and 93.85% sensitivity on DRIVE, 95.15% accuracy, 97.88% specificity, and 92.27% sensitivity on INSPIRE-AVR. Additionally, AVR results from the proposed system were compared against experts with no significant statistical differences.

**Keywords:** *hypertensive retinopathy, Artificial Intelligence, AVR, automatic diagnosis, U-net, A/V classification, semantic segmentation.*



# Resumen

La hipertensión es una enfermedad de alta incidencia y prevalencia caracterizada por su papel central en el daño vascular retiniano, conocido como retinopatía hipertensiva. Los cambios vasculares de la retina que se observan en las imágenes de fondo de ojo pueden emplearse como biomarcadores para medir las constricciones generalizadas de los vasos a través de la Relación Arteriola-Venula (AVR). Los avances en la Inteligencia Artificial y la fácil adquisición de imágenes médicas han permitido el desarrollo de sistemas innovadores de apoyo para el diagnóstico automático de enfermedades oculares.

Este trabajo presenta un sistema automático que utiliza una red neuronal basada en la U-net para clasificar arterias, venas y el disco óptico a partir de imágenes de fondo de ojo con segmentación semántica para calcular el AVR y determinar si un paciente tiene resultados anormales y, por tanto, la enfermedad. El sistema se probó con tres conjuntos de datos. Dos de ellos están disponibles públicamente: DRIVE e INSPIRE-AVR, cada uno de los cuales consta de 40 imágenes de fondo de ojo; la tercera base de datos se creó con la ayuda del INDEREB. Esta tiene dos versiones con 105 fotos en total y será accesible al público. Los resultados de la clasificación A/V obtenidos con el método propuesto son: 95,74% de precisión, 97,37% de especificidad y 93,85% de sensibilidad en DRIVE, 95,15% de precisión, 97,88% de especificidad y 92,27% de sensibilidad en INSPIRE-AVR. Además, los resultados de AVR del sistema propuesto se compararon con los de los expertos, sin que hubiera diferencias estadísticas significativas.

**Palabras clave:** *retinopatía hipertensiva, Inteligencia Artificial, AVR, diagnóstico automático, U-net, clasificación A/V, segmentación semántica.*



# Symbols and Abbreviations

$\gamma$	Focal loss coefficient
$\Sigma$	Summing junction
$\varphi$	Activation function
ADAM	Adaptative Moment Estimation
AI	Artificial Intelligence
ANN	Artificial Neural Network
AVR	Arteriolar-to-Venular Ratio
A/V	Artery/Vein
CDF	Cumulative Distribution Function
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
CRAE	Central Retinal Artery Equivalent
CRVE	Central Retinal Vein Equivalent
CWF	Category-attention Weighted Fusion
DD	Disc Diameter
DL	Deep Learning
DRIVE	Digital Retinal Images for Vessel Extraction
FCN	Fully Convolutional Network
FN	False Negative
FOV	Field of View
FP	False Positive
GPU	Graphics Processing Unit

H	Height
HR	Hypertensive Retinopathy
INDEREB	Instituto de la Retina del Bajío
INSPIRE-AVR	Iowa Normative Set for Processing Images of the Retina for Arteriolar-to-Venular Ratio
ML	Machine Learning
OD	Optic Disk
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest
RGB	Red Green Blue
ROI	Region of Interest
SH	Systemic Hypertension
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
VSR	Vascular Structure Reconstruction
W	Width
WHO	World Health Organization

# Contents

<b>Acknowledgments</b>	
<b>Abstract</b>	<b>i</b>
<b>Resumen</b>	<b>iii</b>
<b>Symbols and Abbreviations</b>	<b>v</b>
<b>Contents</b>	<b>vi</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Artificial Intelligence in Medicine . . . . .	1
1.1.1 Fundus Images . . . . .	1
1.2 Hypertension . . . . .	2
1.2.1 Hypertensive Retinopathy . . . . .	3
1.2.2 Stages of the disease . . . . .	3
1.2.3 AVR Calculation . . . . .	4
1.3 Hypothesis . . . . .	6
1.4 Objectives . . . . .	6
1.4.1 General Objective . . . . .	6
1.4.2 Specific Objectives . . . . .	6
<b>2 Justification</b>	<b>7</b>
<b>3 Problem Description</b>	<b>8</b>
<b>4 Theoretical Foundation</b>	<b>9</b>
4.1 Machine Learning . . . . .	10
4.2 Deep Learning . . . . .	10
4.2.1 Computer Vision . . . . .	12
4.2.1.1 Classification . . . . .	13
4.2.1.2 Classification + localization . . . . .	13

4.2.1.3	Object detection	13
4.2.1.4	Semantic segmentation	13
4.2.2	Neural Networks	14
4.2.2.1	U-net	15
4.3	Image Preprocessing	16
4.3.1	Color spaces	16
4.3.1.1	RGB	17
4.3.1.2	Lab	17
4.3.1.3	HSV	17
4.3.2	CLAHE	17
4.3.3	Histogram Equalization	18
4.3.4	Gamma Correction	18
4.4	State of the Art	19
<b>5</b>	<b>Methodology</b>	<b>22</b>
5.1	Image acquisition	23
5.2	Pre-processing	27
5.2.1	Color and Size adjustment	27
5.2.2	Data augmentation	29
5.3	Processing	30
5.3.1	Network architecture	30
5.3.1.1	Hyperparameters	30
5.3.1.2	Training	32
5.3.2	Semantic segmentation	33
5.4	Post-processing	35
5.4.1	AVR computation	35
5.4.1.1	Convex Hull	36
5.4.1.2	Software	39
<b>6</b>	<b>Results</b>	<b>42</b>
6.1	Evaluation metrics	42
6.2	A/V classification	43
6.3	AVR against experts	46
6.4	Early Hypertensive detection through AVR	49
<b>7</b>	<b>Conclusion</b>	<b>52</b>
<b>Appendix A</b>	<b>Python Libraries</b>	<b>53</b>
<b>Appendix B</b>	<b>Complete results on INSPIRE-AVR dataset</b>	<b>54</b>
<b>Appendix C</b>	<b>Complete results on INDEREBv1 dataset</b>	<b>55</b>
<b>Appendix D</b>	<b>Complete results on INDEREBv2 dataset</b>	<b>56</b>



<b>Appendix E</b>	<b>INDEREB Database</b>	<b>57</b>
E.1	Version 1	57
E.2	Version 2	64

# List of Figures

1.1	Digital Fundus images. (a) Healthy fundus, (b) Fundus with HR symptoms.	2
1.2	Retinal fundus images of hypertensive retinopathy grades. (a) Grade 1, (b) Grade 2, (c) Grade 3, (d) Grade 4 [1].	4
1.3	The proposed approach for AVR computation. (a) Original image, (b) Image after its pre-processing and semantic segmentation, where the CRAE and CRVE were measured.	5
1.4	Example of retinal image used for AVR	6
4.1	Artificial Intelligence and its subdivisions.	10
4.2	Representation of a (a) Biological Neuron and an (b) Artificial Neural Network.	11
4.3	Difference between (a) Machine Learning and (b) Deep Learning.	12
4.4	Computer vision. (a) Original image, (b) Classification + localization, (c) Object detection and (d) Semantic segmentation.	13
4.5	Architecture of a Convolutional Neural Network [2].	14
4.6	Convolution operation on a matrix of size 5×5 with a kernel of size 3×3.	15
4.7	Architecture of traditional U-net neural network [3].	15
4.8	Color spaces: (a) RGB, (b) Lab and (c) HSV.	16
4.9	Preprocessing on (a) Original image with (b) CLAHE.	17
4.10	Preprocessing on (a) Original image with (b) Histogram Equalization.	18
4.11	Preprocessing on (a) Original image with (b) Gamma Correction.	19
5.1	Proposed approach of a fully unsupervised system for hypertension retinopathy detection.	22
5.2	DRIVE Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) Arteries and Veins.	23
5.3	INSPIRE-AVR Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) AVR.	23
5.4	INDEREB Version 1 Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) Arteries and Veins.	24
5.5	INDEREB Version 2 Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) Systemic Hypertension presence and AVR.	24
5.6	INDEREB Version 2. Distribution on (a) Gender (b) Systemic Hypertension.	25
5.7	INDEREB Version 2. Age distribution against (a) Gender (b) Systemic Hypertension.	25
5.8	CLAHE. (a) Original, (b) Lab, (c) HSV, (d) RGB.	27
5.9	Histogram Equalization. (a) Original, (b) Lab, (c) HSV, (d) RGB.	28
5.10	Gamma Correction. (a) Original, (b) Lab, (c) HSV, (d) RGB.	29

5.11 Data augmentation via random transformations. (a) Original, (b) Crop, (c) Flip, (d) Stretch. . . . .	29
5.12 Architecture of a U-net neural network . . . . .	30
5.13 ReLU activation function . . . . .	31
5.14 Semantic Segmentation on the DRIVE database. (a) Original, A/V: (b) Ground truth, and (c) Model prediction. . . . .	33
5.15 Semantic Segmentation on the INSPIRE-AVR database. (a) Original, A/V: (b) Ground truth, (c) Model prediction, OD: (d) Ground truth, (e) Model prediction. . . . .	33
5.16 Semantic Segmentation on the INDEREB v1 database. (a) Original, A/V: (b) Ground truth, (c) Model prediction, OD: (d) Ground truth, (e) Model prediction. . . . .	34
5.17 Semantic Segmentation on the INDEREB v2 database. (a) Original, A/V: (b) Ground truth, (c) Model prediction, OD: (d) Ground truth, (e) Model prediction. . . . .	34
5.18 Proposed method on INSPIRE-AVR dataset. (a) Input image, (b) A/V classification, (c) OD location, (d) ROI and (e) AVR measure in 0.1 steps. . . . .	35
5.19 (a) Set of points and its (b) Convex Hull. . . . .	36
5.20 Diameter estimation through measuring the Euclidian distance of each pixel against the others. (d) Longest distance ( <i>diameter</i> ) obtained from random examples: (a), (b) and (c). . . . .	37
5.21 Blood vessel selection. (a) A/V prediction (b) over original image, (c) A/V pixels overlap- ping the circumferences (d) in original image. . . . .	38
5.22 Convex Hull for width estimation. (a) Image conformed by pixels that intersected with the disks, (b) Close-up at Euclidian distance calculation for each pixel. . . . .	38
5.23 AVR calculation webapp. . . . .	39
5.24 Sidebar options. (a) How to use the tool, (b) Display uploaded images, (c) About. . . . .	40
5.25 AVR calculation preparation. . . . .	40
5.26 AVR Results. (a) Displaying and (b) Without displaying uploaded images. . . . .	41
6.1 A/V Classification results on DRIVE. (a) Ground Truth, (b) Prediction, (c) Misclassification in prediction. . . . .	44
6.2 A/V Classification results on INSPIRE-AVR. (a) Ground Truth, (b) Prediction, (c) Misclas- sification in prediction. . . . .	44
6.3 A/V Classification results on INDEREBv1. (a) Ground Truth, (b) Prediction, (c) Misclassi- fication in prediction. . . . .	45
6.4 A/V Classification results on INDEREBv2. (a) Ground Truth, (b) Prediction, (c) Misclassi- fication in prediction. . . . .	46
6.5 Bland-Altman plots on the INSPIRE-AVR dataset for proposed results against (a) Observer 1, (b) Observer 2, and (c) Niemeijer's method. . . . .	47
6.6 Bland-Altman plot on the INDEREBv2 dataset for results with proposed system against Expert. . . . .	48
6.7 AVR distribution with HR diagnosis for (a) Expert results and (b) Proposed system results. . . . .	49
6.8 AVR distribution regarding Gender with HR diagnosis for (a) Expert results and (b) Pro- posed system results. . . . .	50
6.9 AVR distribution regarding Age and Gender with HR diagnosis for (a) Expert results and (b) Proposed system results. . . . .	51

# List of Tables

1.1	Classification of hypertension based on blood pressure in adults.	2
1.2	Keith-Wagener-Barker Classification of Hypertensive Retinopathy	3
4.1	Main applications of AI in Ophthalmology.	9
4.2	Brief summary of the current state-of-the-art.	21
5.1	Database INDEREB version 2.	26
5.2	Hyperparameters of proposed network.	31
6.1	Comparison with the state of the art for A/V Classification in the Drive dataset.	43
6.2	Comparison with the state of the art for A/V Classification in the INSPIRE-AVR dataset.	44
6.3	A/V Classification results obtained on the private dataset INDEREBv1.	45
6.4	A/V Classification results obtained on the private dataset INDEREBv2.	45
6.5	Comparison with the stare of the art for AVR calculation on the INSPIRE-AVR dataset.	46
6.6	Summarized results of AVR calculation on the INDEREBv1 dataset.	48
6.7	Comparison of AVR results on the INDEREBv2 dataset against expert.	48
6.8	Comparison of ML models metrics for Early Hypertensive Retinopathy detection.	51

# Introduction

## 1.1 Artificial Intelligence in Medicine

Artificial intelligence (AI) is a broad term referring to computers able to simulate intelligent behavior with a minimum of human involvement [4]. John McCarthy first mentioned the word AI during a conference in 1956. However, Alan Turing is the one who defined the term as the ability of machines to mimic human performance in cognitive tasks. Since then, researchers have investigated the possible applications of AI in every medical field [5].

Knowledge and experience are the most important factors when delivering a successful diagnosis. In humans, this usually comes with time, which simultaneously is a limitation along with the amount of information the human mind can retain. Modern medicine requires gathering and analyzing a vast volume of data to solve challenging clinical problems [6]. Hence, the development of AI in medicine has aimed to aid health care workers in formulating an accurate diagnosis, making decisions, and predicting outcomes.

AI has significantly impacted the ophthalmology field since constant improvements in diagnostic tools have offered an exceptional insight into the retina condition and ocular diseases of patients. Digital fundus images allow the AI to evaluate rapidly and non-invasively millions of morphological features in diverse datasets. Methodologies based on machine learning (ML) and deep learning (DL) can detect and quantify pathological characteristics in almost every retinal disease [7].

Nowadays, computer capabilities allow the processing and evaluating of large amounts of data in real-time. State-of-the-art systems include ANNs (Artificial Neural Networks), fuzzy logic, genetic algorithms, and hybrids.

### 1.1.1 Fundus Images

Retinal fundus images have allowed specialists to diagnose medical conditions due to the use of vascular changes as biomarkers [8] since the vessel tortuosity and its calibre can be associated with specific diseases, one of those is the hypertension retinopathy [9, 10].

In the past few years, fundus digital images have gained popularity to study microvascular changes in retinas. This is due to its non-invasive approach, low cost and straightforwardness to perform. Furthermore,

most eye lesions are visually seen in this type of image [11]. Figure 1.1 shows a normal fundus image and another one with advanced hypertensive retinopathy (HR).

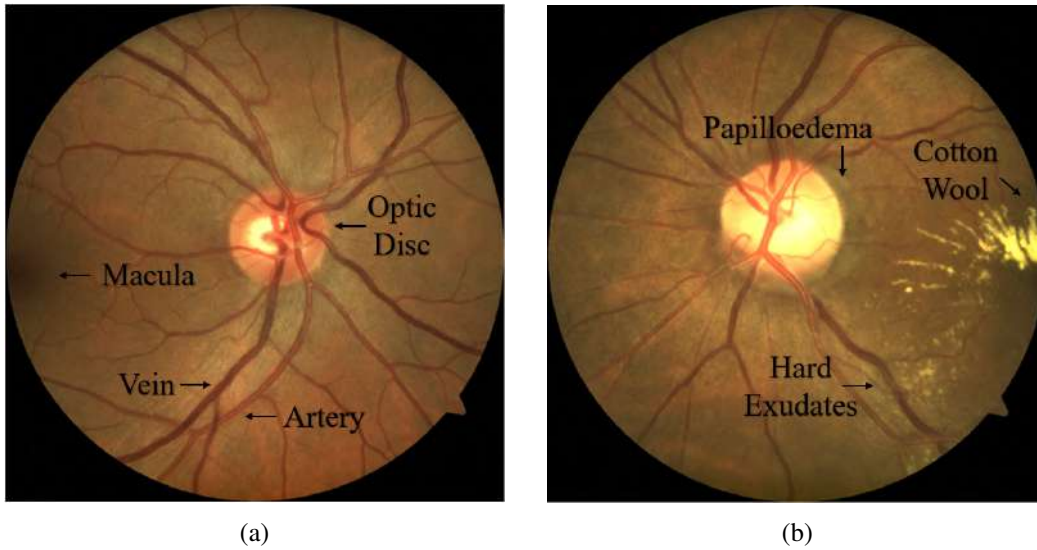


Figure 1.1: Digital Fundus images. (a) Healthy fundus, (b) Fundus with HR symptoms.

## 1.2 Hypertension

Systemic Arterial Hypertension is often asymptomatic, thus considered as a silent killer since it rarely causes symptoms in the early stages, without treatment it could advance until reaching severe conditions [12]. It causes persistent high blood pressure, leading to multi-systemic effects such as retinopathy and death in extreme cases [13, 14].

In a healthy person, each heartbeat pumps blood from the heart into the blood vessels through all the body. Blood pressure is created by the blood being pushed against the artery walls as it is pumped by the heart. Hence, the higher the pressure, the harder the heart has to work to pump blood and if it is left uncontrolled, hypertension can lead to heart attacks, strokes, kidney failure and blindness. Most people are diagnosed until they may not have access to treatment or are not able to control their illness over the long term [15]. The grades of hypertension according to blood pressure ranges in adults can be found in Table 1.1.

<i>Grade</i>	<i>Systolic</i>	<i>Diastolic</i>
Normal	120 - 129 mmHg	80 - 84 mmHg
1	140 - 159 mmHg	90 - 99 mmHg
2	160 - 179 mmHg	100 - 109 mmHg
3	>180 mmHg	>110 mmHg

Table 1.1: Classification of hypertension based on blood pressure in adults.

### 1.2.1 Hypertensive Retinopathy

Hypertension has a main role in retinal vascular damage, which is known as hypertensive retinopathy, this condition causes a series of retinal microvascular changes in several parts of the eye in response to the high blood pressure [16].

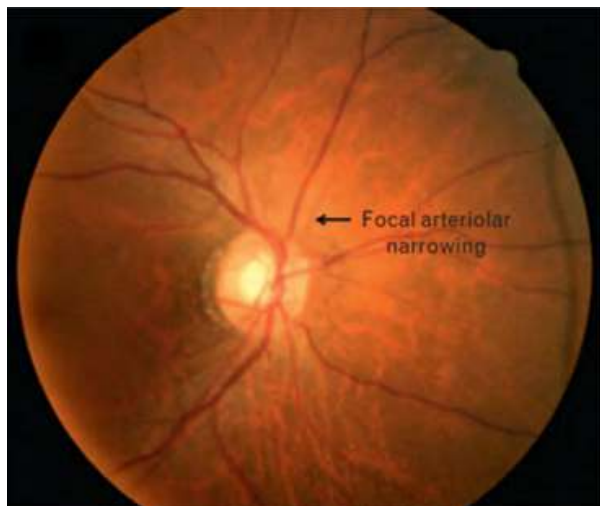
This disease usually manifests in sequential phases of damage and although there exist several classifications, the most widely used is the hypertensive retinopathy grading system which was created in 1939 by Keith-Wagener-Barker [17], shown in Table 1.2 along with its systemic associations.

### 1.2.2 Stages of the disease

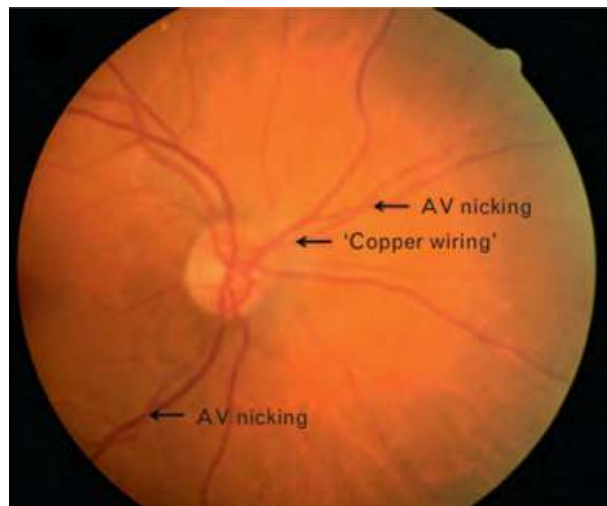
Grade	Signs	Systemic Association
1	Mild generalized retinal arteriolar narrowing or sclerosis	None
2	Definite focal narrowing and arteriovenous crossings Moderate to marked sclerosis of the retinal arterioles Exaggerated arterial light reflex	Modest association with risk of clinical stroke, subclinical stroke, coronary heart disease.
3	Definite focal narrowing and arteriovenous crossings Exudates Cotton wool spots Sclerosis and spastic lesions of retinal arterioles	Strong association with risk of clinical stroke, subclinical stroke, cognitive decline, death from cardiovascular causes.
4	Grade III signs and papilledema	Strong association with death.

Table 1.2: Keith-Wagener-Barker Classification of Hypertensive Retinopathy

Hypertensive retinopathy starts with evident retinal arteries narrowing due to vasospasm and vasoconstriction, as shown in Figure 1.2a. Eventually, the high blood pressure leads to arteries thickening over veins where both cross between them as shown in Figure 1.2b.



(a)



(b)

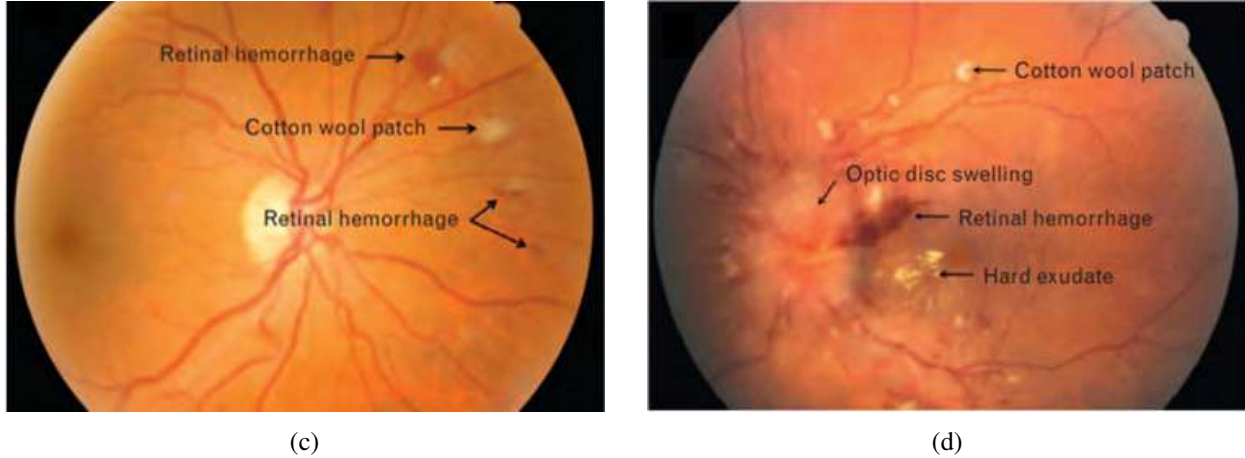


Figure 1.2: Retinal fundus images of hypertensive retinopathy grades. (a) Grade 1, (b) Grade 2, (c) Grade 3, (d) Grade 4 [1].

The third phase occurs if the hypertension continues uncontrolled, here the blood-retinal barrier is disrupted causing exudates, cotton wool spots, vessel endothelial changes (microaneurysms) and hemorrhages into the retina, as shown in Figure 1.2c. The last stage occurs with chronic hypertension, where the optic nerve and macula are damaged, as shown in Figure 1.2d. [18]

The study between retinal micro- and macro-circulation are key for diagnosis of certain diseases that cause early retinal vascular changes, such as vessel and arteriolar widths alterations. Arteriolar narrowing is related to initial damage caused by hypertension and venular widening is associated with hyperglycemia and obesity [19].

### 1.2.3 AVR Calculation

Retinal vascular alterations can be monitored to determine the stage and severity of the disease. Several metrics have been developed for that purpose, being the Arteriolar-to-Venular Ratio (AVR) the most commonly used. It measures generalized arteriolar constriction, which is the first early sign of hypertensive retinopathy [20]. However, relevant vascular alterations are troublesome to be manually assessed in practice as the method is arduous and time-consuming [21]. To accurately compute it is necessary to first select the six largest arterioles and six largest venules passing through a previously defined region, as seen in Figure 1.3.

Since the AVR is commonly used as a biomarker, a formula was created to measure the vessels width. It is computed as shown in equation 1.1, where CRAE corresponds to central retinal artery equivalent and CRVE to central retinal vein equivalent [22].

$$AVR = \frac{CRAE}{CRVE} \quad (1.1)$$

However, the AVR measurement is not a straightforward task. First, vessels are classified into arteries and veins to then be measured in an area concentric to the optic disc. However, experts do not use all retinal vessels and have to decide which ones are appropriate for calculus [23].



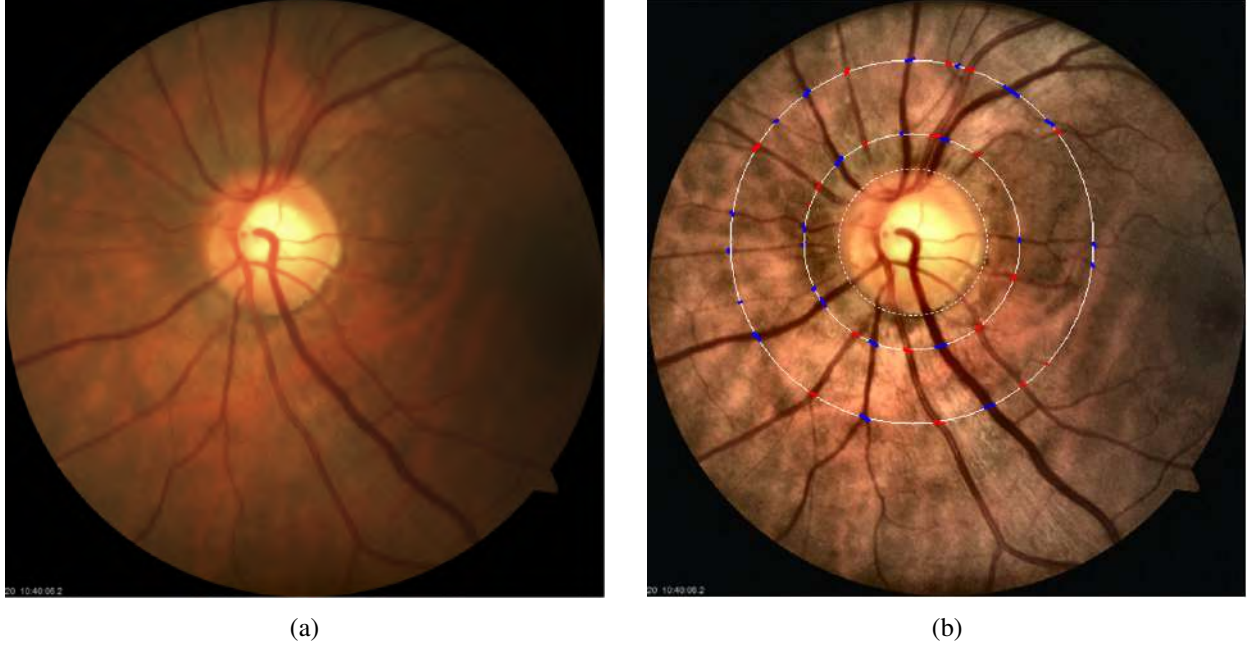


Figure 1.3: The proposed approach for AVR computation. (a) Original image, (b) Image after its preprocessing and semantic segmentation, where the CRAE and CRVE were measured.

The most used method to calculate both CRAE and CRVE was proposed by Knudtson et al., which provides two equations [24].

$$W_{artery} = 0.88\sqrt{W_1^2 + W_2^2} \quad (1.2)$$

$$W_{vein} = 0.95\sqrt{W_1^2 + W_2^2} \quad (1.3)$$

The Knudtson method consists of selecting the 6 largest arterioles and venules to then apply the corresponding equation [1.2] or [1.3] to the first pair of vessels, where  $W_1$  represents the narrowest artery and  $W_2$  the widest, the process is repeated until a remaining vessel is carried over to the final iteration to derive the CRAE or CRVE and finally the equation [1.1] is applied.

Current computer software tools make it possible to leverage digital fundus images by measuring the retinal vessels calibre to evaluate objective-wise the degree of arteriolar narrowing. Yet, the vast majority of software available is semi-automated, therefore needing human interaction to estimate accurately calibres [25, 26, 27].

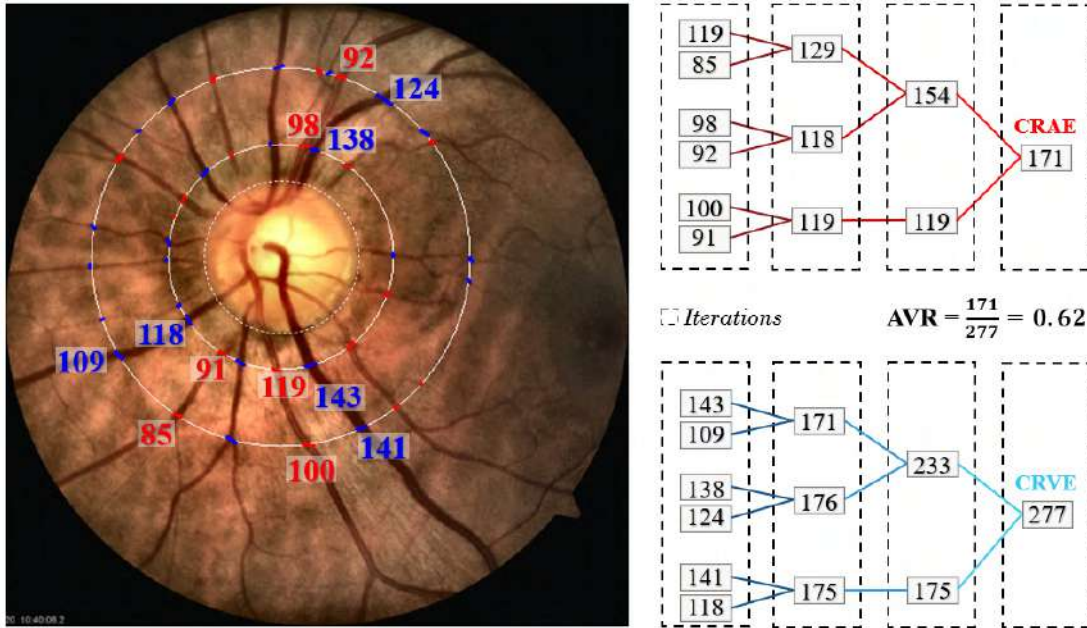


Figure 1.4: Example of retinal image used for AVR

### 1.3 Hypothesis

A deep learning based system, that uses leveraging salient regions to segment and classify veins and arteries, can be used to accurately and reliably compute the arteriovenous ratio (AVR) from fundus images to aid retinologist determine whether or not a patient has grade 1 hypertensive retinopathy.

### 1.4 Objectives

#### 1.4.1 General Objective

To develop a deep learning based system through leveraging salient regions to segment and classify veins and arteries to compute the AVR biomarker.

#### 1.4.2 Specific Objectives

- To select two public databases and fundus images from INDEREB database with detailed annotations for retinal blood vascular patterns.
- To label and pre-process fundus images with algorithms for enhancement and luminosity normalization to be able to segment retinal blood vessels and arteries to measure AVR.
- To develop a neural network system able to automatically locate the optic disc and to segment and classify blood vessels to compute AVR.
- To train and test the developed system with both public and private datasets to evaluate and report results.

---

# Justification

Hypertension complications are responsible for 9.4 million deaths worldwide every year. There is a significant gain in global health and economy in the early detection and adequate control of hypertension. However, its prevention and supervision are quite complex, especially in low and middle income countries [15]. In Mexico, the disease has a high incidence as well as prevalence, about 30.7% of adult population have it. However, only 1 out of 3 is aware of it and half of them receive treatment [28].

Eyes can be affected by hypertension in several ways, the most damaging is the development of retinopathy. Additionally, they are the only organ in the human body that can display in vivo the vascular changes due to systemic hypertension. A person's overall health and mortality have a short and long-term correlation with hypertensive retinopathy symptoms and its severity [29]. Hypertensive retinopathy grading is shown in Table 1.2 and fundus images of the four grades of HR can be seen in Figure 1.2.

In order to get people diagnosed with hypertensive retinopathy it is necessary to have a specialized inspection from an experienced ophthalmologist. However, with the number of eye diseases increasing, most health facilities cannot keep up with rapid diagnosis. Thus, delaying a timely detection of hypertensive retinopathy which may generate complications. This problematic is majorly seen in third world countries.

Non-invasive visualization of retinal structures, such as fundus images, give an unprecedented solution in modern medicine as the most powerful real-time diagnostic tool. Currently, imaging diagnostic is the most efficient application of AI, since it can use algorithms to detect and learn features from large volumes of imaging data, recognize disease-specific patterns and correlate features to gain innovative scientific insight [7].

Nowadays, there are several AI methods based on machine learning (ML) and deep learning (DL) that can detect diabetic retinopathy, glaucoma, age-related macular degeneration and other retinal diseases. However, the field of hypertensive retinopathy detection, being not as popular as the other diseases, still has much potential to investigate and develop new innovative methods.

One of the major advantages of artificial intelligence in retina is the possibility of working with the AVR in a more precise and faster way than any retinologist could perform manually. Hence, AVR measurement offers experts aid when doing an early detection of hypertensive retinopathy, allowing patients to get a timely treatment and avoid complications.

---

## Problem Description

Most countries suffer from low doctor to patient ratio, which overburdens the patient care system, causing diagnosis and proper treatment become error prone and time intensive. Additionally, a great amount of data is generated every day in clinics and hospitals, but it is rarely made public and even less employed for computer aided diagnostics [30].

In recent years, with the availability of fundus digital cameras and the consequent interest of retinal images classification for medical diagnosis, several state-of-the-arts methods have appeared in the literature. However, the challenge of recognizing microvascular lesions in retinal eye images, which are caused by high blood pressure, still remains because of the fundal variability of each patient [31].

Current AI methods for retinal analyses differ greatly in their applicability, interpretability and reliability in different datasets and diseases but its overall potential prevails, from automated detection, diagnostic grading to therapy guidance. The potential benefit of AI in retina is further expanded by enabling large scale management prediction and prognostic, thus empowering ophthalmologists to provide rapid and high-quality diagnosis, and successfully dealing with the 21st century complexity of ophthalmology [7].

Present methodologies are divided in several steps, the first one consists of detecting the region of interest, then a general vessel segmentation is made to classify the prediction's results into veins and arteries to finally measure the AVR. This makes the overall algorithm quite extensive.

The AVR is crucial for HR diagnosis, since the ratio between the arteries and veins estimates the arteriolar narrowing objectively. Therefore, fundus image segmentation is necessary for AI methods to be able to detect those biomarkers. Unfortunately, large datasets are not available which affects greatly the deep learning process to recognize the pattern of the data in real life application. To be able to improve the current state-of-the-art models this investigation worked closely with the INDEREB to create a larger dataset and support our results with their medical expertise.

# Theoretical Foundation

McCulloch and Pitts made the first approach to AI in 1943 by trying to simulate the function of a single neuron. They proposed a learning model comparable to the process in neurons in the human brain based on True/False statements that used a threshold to output a value [32].

As models continued to develop, Alan Turing created a methodology to evaluate their intelligence in 1950 [33]. The Turing Test questions whether an interrogator could discern between human and machine responses [34]. However, AI was formally introduced as a term and science by Marvin Minsky and John McCarthy in 1956 at a Dartmouth College conference [35].

Due to the exponential growth in the last decade, AI has advanced to create complex algorithms capable of performing tasks similar to a human brain and will continue to impact health care, including vision and eye care, significantly. However, AI diagnosis-based results need to be fully comprehended by retina specialists. To achieve it, experts must become familiar with the ML and DL decision methods used to make clinical diagnostics.

Thanks to advances in AI and growing interest in its applications, ophthalmology has been able to leverage fundus images, due to the clarity of morphological areas such as the retina, by combining them with computer vision. Three main applications have been developing in recent years, as seen in the following Table.

<i>Technique</i>	<i>Examples</i>	<i>Application</i>
Classification	Disease Type	Automated diagnosis
	Disease Stage	
Segmentation	Detect lesion	Creation of biomarkers
	Measure features	
	Calculate shape or volume	
Prediction	Value of another measure	Disease prognosis
	Blood pressure	
	Visual acuity	

Table 4.1: Main applications of AI in Ophthalmology.

ML is a subset of AI that represent the set of algorithms that are trained on data, learn from it and improve with experience, while DL is a subset of ML, it attempts to solve problems as a human would do, this is achieved by using multi-layered structures of algorithms called neural networks [7]. This can be better understood as shown in Figure 4.1.

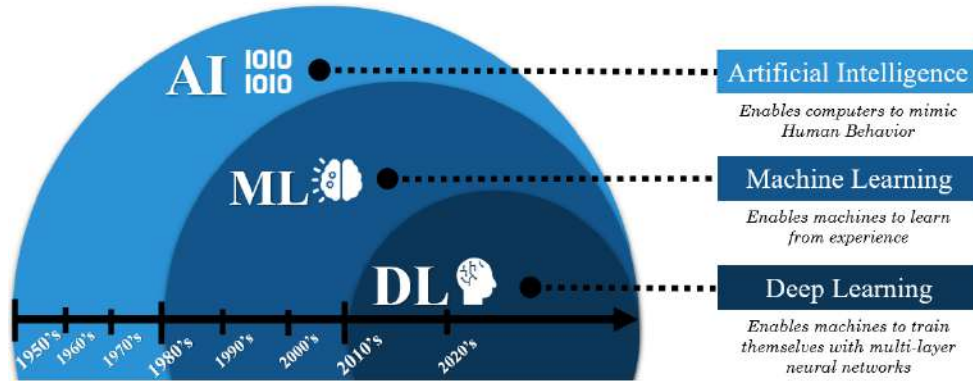


Figure 4.1: Artificial Intelligence and its subdivisions.

## 4.1 Machine Learning

It is an application of artificial intelligence where a machine learns from experience and makes future predictions. The techniques are used to automatically find patterns within complex data and use the knowledge about a specific problem to predict future events.

The algorithms are categorized into supervised learning, where human intervention is required, unsupervised learning, where no intervention is needed and reinforcement learning.

In supervised learning, the machine learns from labeled data. The labels help the algorithm to correlate the features and modify the model accordingly. In unsupervised learning, the algorithm tries to discover patterns with unclassified and unlabeled data. In reinforcement learning, the algorithm looks for the most significant reward through trial and error. It has three components: learner, environment, and actions [36].

## 4.2 Deep Learning

It is inspired by the structure of the human brain. It uses artificial neural networks, which are a multi-layered structure of several activation functions that mimic human decision making process.

The layers in neural networks are able to learn an implicit representation of the raw data on their own since the feature extraction is already part of the process [37]. For instance, an image can be represented in a variety of ways, either as a vector of intensity values per pixel or as a set of edges and regions of a specific shape.

An ANN may consist of two sets of neurons, one receives an input and the other sends an output. The algorithm have an input layer that modifies the input and passes it to the next layer and can use several processing layers to create multiple linear and non-linear transformations [38]. Figure 4.2 shows a biological neuron and an ANN.

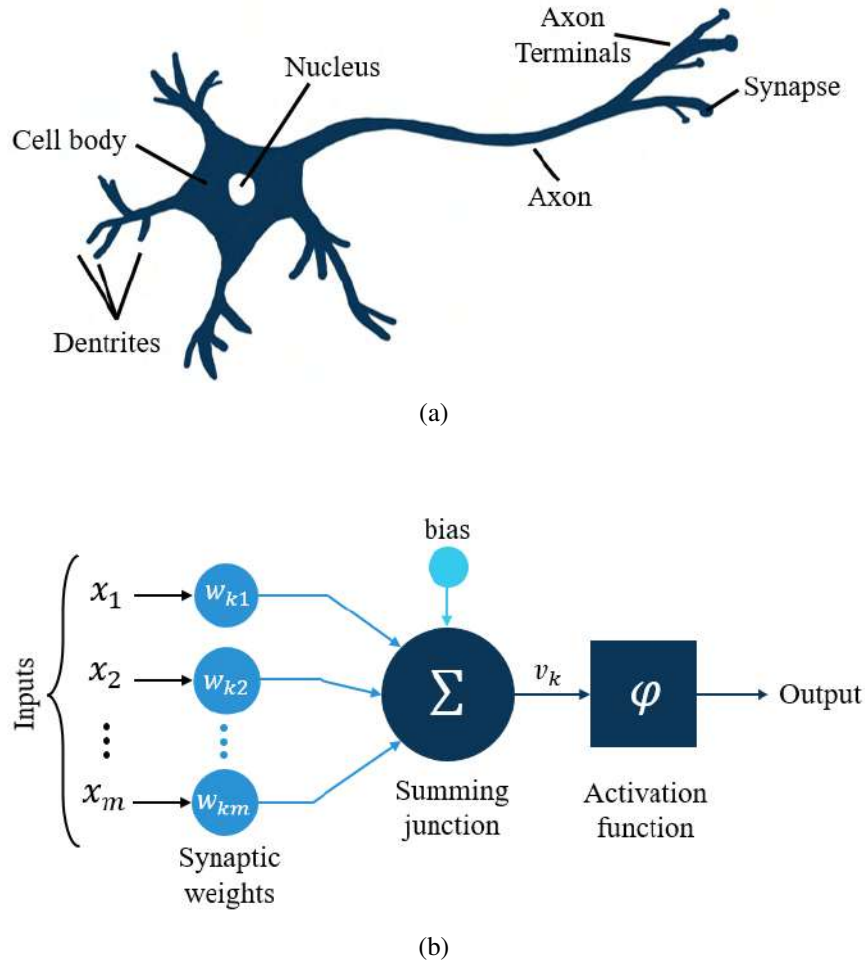


Figure 4.2: Representation of a (a) Biological Neuron and an (b) Artificial Neural Network.

The main advantage of DL is the ability to replace handmade features with efficient unsupervised or semi-supervised feature learning techniques. In the field of computer vision, neural networks have shown that they can outperform ML algorithms on tasks like object recognition. The main difference between Machine Learning and Deep Learning can be seen in Figure 4.3.

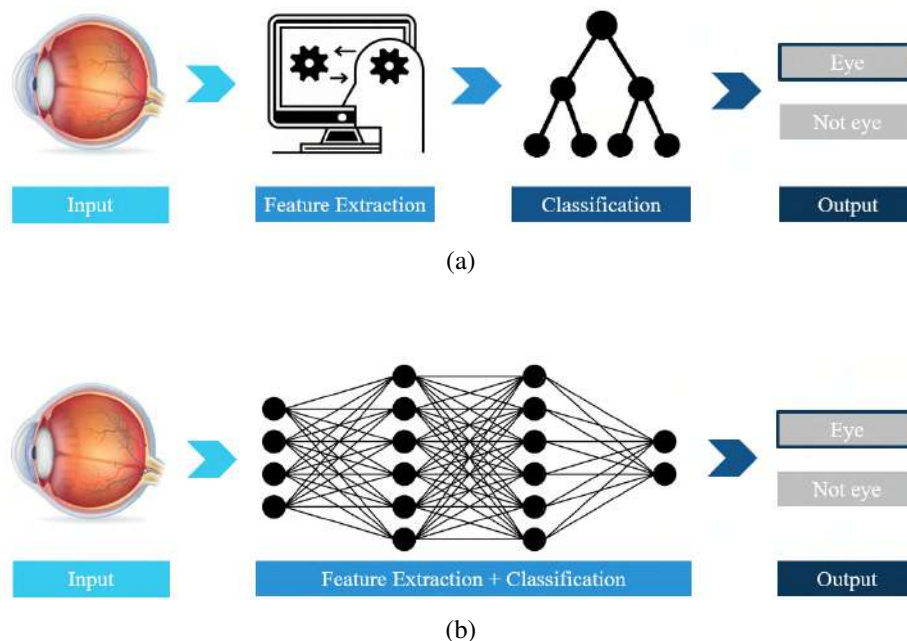
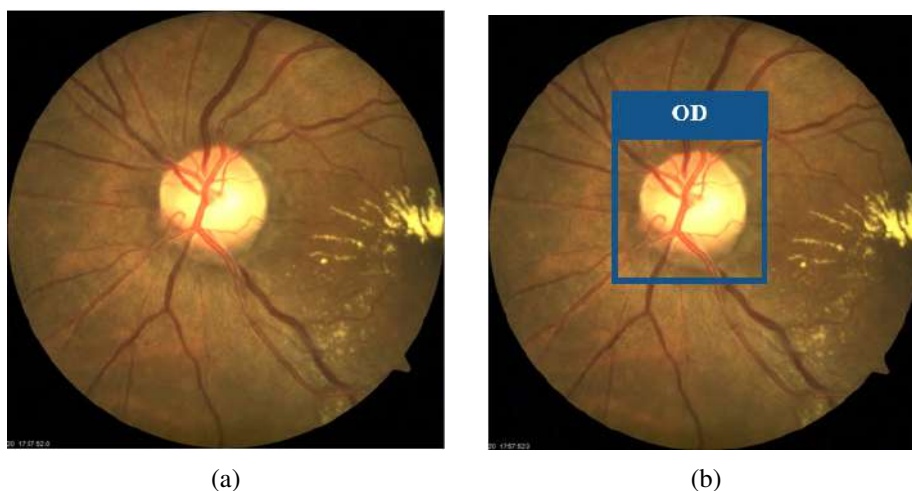


Figure 4.3: Difference between (a) Machine Learning and (b) Deep Learning.

### 4.2.1 Computer Vision

Deep learning has enabled computer vision to advance rapidly in recent years. It is an interdisciplinary field of AI that allows computers and systems to derive meaningful information from digital images, videos, and other visual inputs.

It works much like human vision, with the difference that humans have the advantage of a lifetime of context to train how to distinguish objects, if they move and if there is something wrong with an image. There are several levels of understanding, and for each of these, there is a defined problem, Figure [4.4](#).





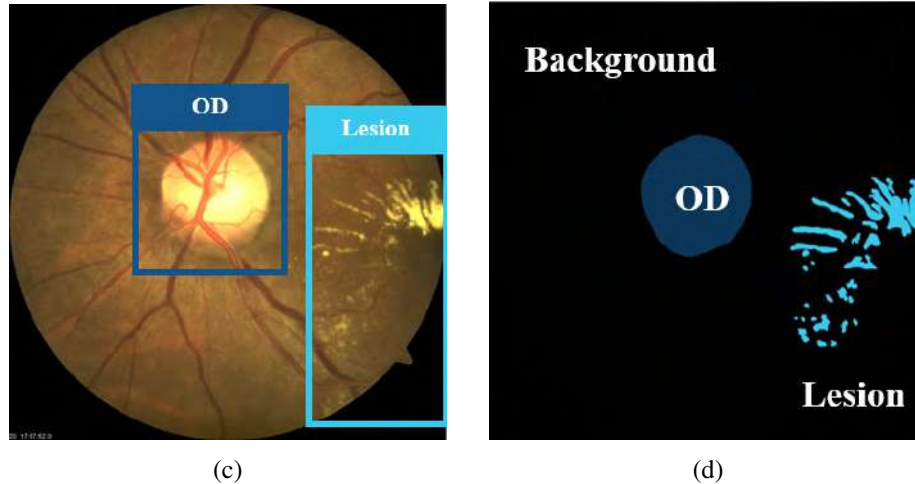


Figure 4.4: Computer vision. (a) Original image, (b) Classification + localization, (c) Object detection and (d) Semantic segmentation.

#### 4.2.1.1 Classification

Given an image, the computer is expected to predict a discrete label of the main object. In image classification, it is assumed that there is only one object, as shown in Figure 4.3.

#### 4.2.1.2 Classification + localization

The algorithm defines the discrete label and locates where exactly the object is in the image. It is implemented using a bounding box. It is assumed that there is only one object per image, as shown in Figure 4.4b.

#### 4.2.1.3 Object detection

The detection is not limited to a single object but may contain multiple objects. The task is to classify and locate all objects in the image, as shown in Figure 4.4c.

#### 4.2.1.4 Semantic segmentation

It is the process of dividing an image into multiple segments of specific regions. The objective is to label each pixel of an image with the class corresponding to what is being represented. Each pixel of the image is predicted, commonly known as dense prediction.

The objective is to take an image of size  $W \times H \times 3$  and generate a  $W \times H$  matrix containing the class IDs corresponding to all pixels [39]. The semantic segmentation is different from object detection in that it does not predict any bounding boxes around the objects. Furthermore, it does not distinguish between different instances of the same object, as shown in Figure 4.4d.

## 4.2.2 Neural Networks

In 1980, a model to simulate the human visual system was proposed by Fukushima, a simpler version of this technique is considered a Convolutional Neural Network (CNN). LeCun et al. developed a CNN called LeNet in 1989, which has five layers: one input, three hidden and one output. The input layer consists of a 16x16x9 pixel image, the three hidden layers involve two convolutional layers, a pooling layer and one fully connected layer [2]. A general architecture of a CNN can be seen in Figure 4.5.

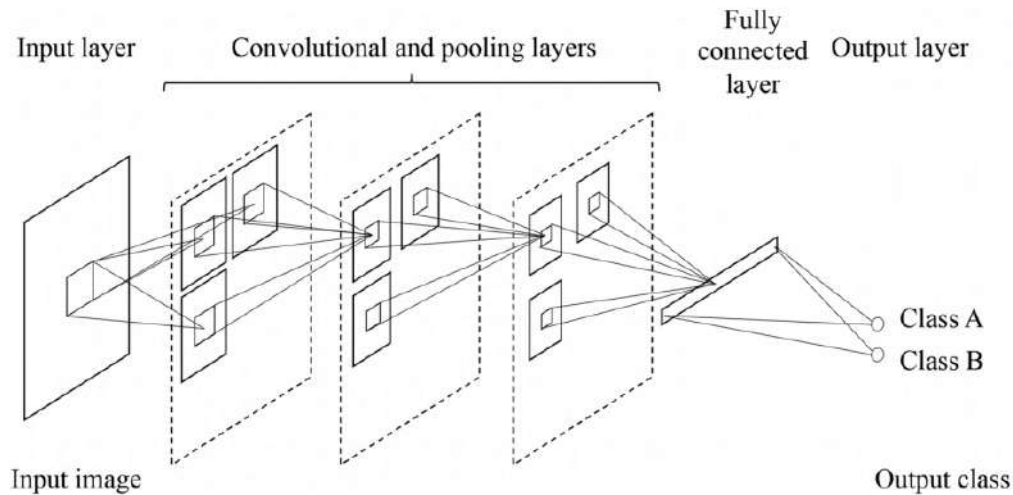


Figure 4.5: Architecture of a Convolutional Neural Network [2].

The Convolutional Neural Network (CNN) is one of the most popular artificial neural network (ANN) architectures, mostly used to solve difficult image-driven pattern recognition. The neurons within CNNs are organized in three dimensions, height and width, which are the dimensionality of the input and the depth. However, depth does not denote the number of layers within the neural network. CNNs contain three types of stacked layers, convolutional, pooling and fully-connected.

CNN architecture can be further explained as follows: the input layer consists of the image's pixel values. The convolutional layer uses learnable kernels, which are small in dimension and extends along the input's depth. Hence, when data encounters a convolutional layer, the last one convolves each filter over the spatial dimensionality of the input, creating a 2D activation map.

Figure 4.6 explains how the convolutional layer works. The pooling layer down samples the spatial dimensionality of the input, therefore reducing the number of parameters within the activation. The fully-connected layers produce scores to be used for classification. The rectified linear unit (ReLU) can be used between these layers to improve performance [40].

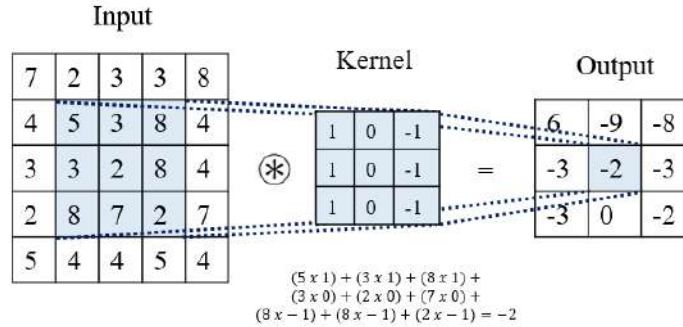


Figure 4.6: Convolution operation on a matrix of size 5x5 with a kernel of size 3x3.

#### 4.2.2.1 U-net

Deep learning applications in medical imaging gained interest in recent years. However, researchers were primarily interested in automatically recognizing and segmenting lesions in images. A breakthrough solution was proposed by Ronnberger et al. [3] in 2015 at the MICCAI conference, a Fully Convolutional Network named U-Net. The network consists of three main sections: an encoder, a bottleneck and a decoder. Due to its U-shape, it is capable of training at high speed with only a small amount of data needed.

In segmentation tasks, the image not only needs to be converted into a vector but also to be reconstructed from it, which is precisely the task of the U-Net. The Figure 4.7 shows the network's structure.

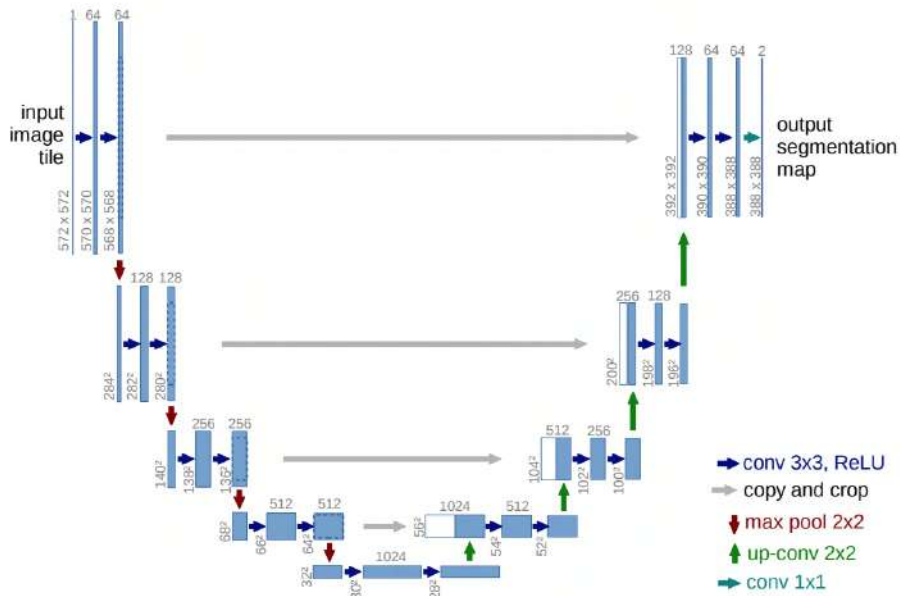


Figure 4.7: Architecture of traditional U-net neural network [3]

Traditionally, it is two-dimensional which allows working with fewer training images than other networks while still obtaining a more accurate segmentation. The central idea is to use sequential layers while the pooling operation is linked to the downsampling operator. As a result, those layers increase the output resolution. For localization, the contracted path's high resolution is mixed with the upsampled output so sequential convolutional layers can focus on fine features and produce a more accurate segmentation [41, 42].

The encoder is constructed with four down-sampling blocks of 3x3 convolutions, each with a ReLU as activation function and a 2x2 max pooling. Hence, the resolution is reduced by half during each down-sampling. The decoder consists of four up-sampling blocks of 3x3 convolutions using ReLU, and a 2x2 transposed convolution. The result is then concatenated with its corresponding feature map resolution through a skip connection. In consequence, features from the encoder are sent to the decoder at different resolution levels, producing a more thorough outcome.

### 4.3 Image Preprocessing

Nowadays fundus images are commonly used in the diagnosis of retinal diseases. However, the images often have problems with either low contrast or noise. The aim of image pre-processing is to have an image that is more suitable than the original for further analysis such as segmentation or classification of retinal blood vessels.

Usually, images do not have uniform illumination due to factors such as the pupil not being fully dilated, differences in retinal pigmentation, and camera specifications and limitations. In addition, the digitization process can cause low-intensity white noise. Thus, preprocessing is an essential step in medical image processing, as it helps remove the noise present in the image and improve some image characteristics.

#### 4.3.1 Color spaces

A color space is a specific organization of colors, usually in three dimensions, that allows us a reproducible representation of colors, both digitally and analogically. In general, color can be measured by its brightness, hue, lightness, luminance, chroma, and saturation. There are many color spaces, some are better for certain applications. In this research, three-color spaces were used: RGB, Lab, and HSV.

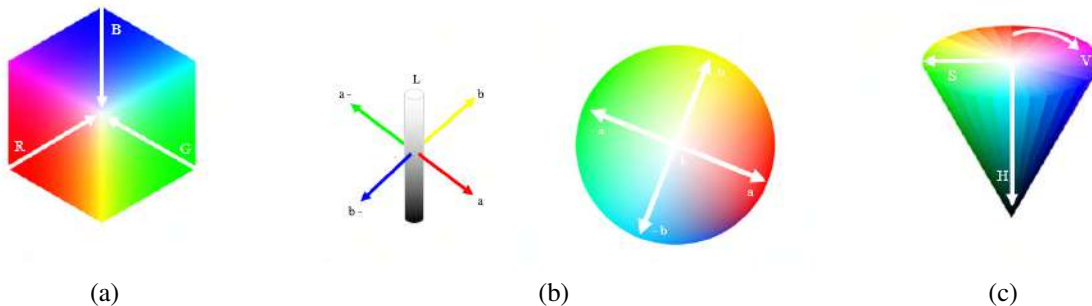


Figure 4.8: Color spaces: (a) RGB, (b) Lab and (c) HSV.

### 4.3.1.1 RGB

RGB is an acronym for *Red*, *Green*, and *Blue*, its representation can be seen in Figure 4.8a. It describes a chromatic model consisting of a mix of only those primary colors. However, the RGB model is not able to deliver an accurate description of the three colors representing its essence. It is because several shades of red, green, and blue can be used in RGB models, hence the additive synthesis may vary.

### 4.3.1.2 Lab

In this color space, there is one channel for *L* (Luminosity) and two other color channels *a* and *b*, known as chromaticity layers. The *a* layer indicates where the color falls along the red-green axis, while the *b* layer indicates where the color falls along the blue-yellow axis, this can be appreciated in Figure 4.8c.

### 4.3.1.3 HSV

This space is intended to be easily interpretable and readable by a human since it uses more familiar terms when talking about color, it is shown in Figure 4.8b. *H* stands for color intensity (hue), it is an angle in the range  $[0, 2\pi]$  with red at angle 0, green at angle  $\frac{2\pi}{3}$ , blue at angle  $\frac{4\pi}{3}$ , and red again at angle  $2\pi$ . *S* stands for saturation, which describes the purity of the hue according to a white reference. *V* stands for value, a percentage ranging from 0 to 100.

## 4.3.2 CLAHE

Contrast-Limited Adaptive Histogram Equalization is a method that improves the problem of low contrast in digital images, especially in medical images. Figure 4.9 shows an original image and the CLAHE result.

It basically operates by limiting the contrast by clipping the histogram at a predefined value before calculating the cumulative distribution function (CDF). This limits the slope of the CDF and hence of the transformation function. The value to which the histogram is clipped depends on the normalization of the histogram and thus on the size of the neighborhood region [43].



Figure 4.9: Preprocessing on (a) Original image with (b) CLAHE.

### 4.3.3 Histogram Equalization

This technique is a non-linear process that aims to enhance image brightness in a way that is especially suitable for human visual analysis. It seeks to modify an image in such a way as to produce one with a flatter histogram, in which all levels are equiprobable [44].

Its performance can be compelling, as it matches well the properties of human vision. However, noise in the image acquisition process will affect the shape of the original histogram and thus the equalized version, which is its weakness. In Figure 4.10 the Histogram Equalization can be appreciated along with the original image.

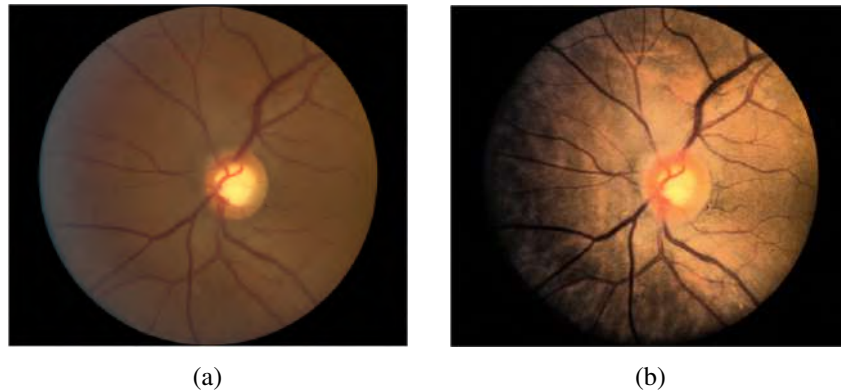


Figure 4.10: Preprocessing on (a) Original image with (b) Histogram Equalization.

### 4.3.4 Gamma Correction

The gamma correction is simply a power-law transformation, except for low luminances, where it is linear to avoid having an infinite derivative at zero luminance. An example of this technique can be seen in Figure 4.11.

The image pixel intensities must be scaled from [0, 255] to [0, 1]. From there, the following equation is applied to obtain a gamma-corrected image:

$$O = \frac{I^{\frac{1}{G}}}{255} \cdot 255 \quad (4.1)$$

Where  $I$  is the input image and  $G$  is the gamma value. The output image  $O$  is reduced to the range [0; 255]. Gamma values  $< 1$  will shift the image to the darker end of the spectrum, while gamma values  $> 1$  will make the image appear lighter [45].

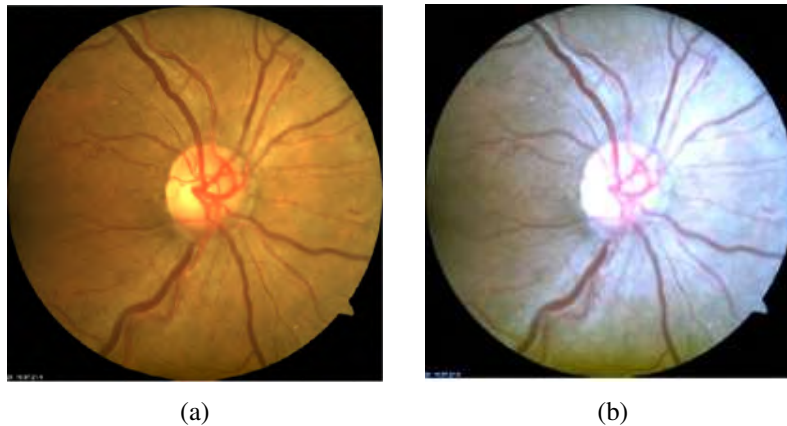


Figure 4.11: Preprocessing on (a) Original image with (b) Gamma Correction.

#### 4.4 State of the Art

Several automatic systems to detect retinal abnormalities have been previously proposed in the past. However, the early detection of hypertensive retinopathy has not been widely explored.

An automated retinopathy detection through digital image analysis presents significant benefits since it requires less time to process numerous images and offers less subjectivity than existing observer-based systems [46]. Hence, to develop a reliable system based on AVR measurements, it is necessary to have an accurate A/V classification, and the location of the optic disc (OD) in each image to measure the vessels width inside the determined diameters from the OD.

However, several works have focused only on retinal blood vessels segmentation, while the classification of those vessels into arteries and veins is not as explored in current literature. Usually, the A/V classification is considered a two steps job, where a general vessel segmentation is made first. Then, another technique is used to determine which vessel is either artery or vein from the previous result. The downside of that method is that if the obtained segmentation were poor, not all veins and arteries would be classified. The automated classification remains as an unexploited area in the field of retinal image processing [46], opening the possibility to be further improved. Nevertheless, with the advances in Deep Learning, a new challenging approach has surfaced, attempting to segment and classify arteries and veins simultaneously.

Traditionally, the A/V classification is only made after the blood vessels have been segmented. Hence, as a first step, it is necessary to have available segmentation masks, manually annotated vessels pixels or an own automatic vasculature segmentation of each fundus image [47].

Most A/V classification methods use machine learning techniques to learn local features. The structural information from the vascular tree is frequently used to improve the classification [14, 48]. Another approach for tracing branches in the vascular tree has been graph-based analysis with global topology criteria [8].

Akbar et al. [14] extracted features from a complete vascular map as input to a support vector machine (SVM) with a radial basis function (RBF) classifier. Srinidhi et al. [48] disentangled the vascular tree into

many subtrees, illustrating the vascular systems' topological and spatial connection and labeled them using hand-crafted features trained with a random forest (RF) classifier. Huang et al. [8] used a feature selection technique based on a genetic search to find the best feature subset for a supervised classifier.

Despite significant advancements in graph-based approaches, retinal A/V separation remains a challenge. To optimally leverage the underlying structural properties of a vascular network, most techniques rely on the geometrical study of vessel key points. Nevertheless, it usually requires a large computing cost and a long run time.

In recent years, deep-learning-based semantic segmentation techniques, essentially pixel-level classification, have been used for fundus image analysis obtaining impressive results in lesions and vascular segmentation, as well as A/V classification. Some works have approached the task as a three-class semantic segmentation [49, 21, 50, 51] (1) artery, (2) vein and (3) background, while others [52] have added a fourth class denominated as (4) uncertain.

Xu et al. [49] normalized images with histogram matching and improved a fully convolutional network (FCN) to simultaneously segment retinal arterioles and venules. Girard et al. [21] implemented a convolutional neural network (CNN) to identify vessels into arteries and veins by segmenting and classifying them together, followed by propagation of the output into a graph representation of the retinal vasculature. Ma et al. [50] designed a multi-task neural network with a spatial activation to perform a simultaneous vessel segmentation and A/V classification to enhance results. Kang et al. [50] integrated a category-attention weighted fusion (CWF) module into a segmentation network and employed a graph-based vascular structure reconstruction (VSR) to decrease noisy inconsistencies.

A novel method has been proposed to decompose the classification into three individual semantic segmentation of an artery, vein, and vascular tree. This multi-segmentation (MS) approach developed by Morano et al. [47] uses an FCN that delivers accurate segmentation masks for the three target vascular trees.

A summary of the previously mentioned contributions is shown in Table 4.2. It presents the authors along with their used techniques and the results achieved in each of the public databases this thesis is working with.

In this thesis, the main contribution is a system that computes a reliable AVR for hypertension retinopathy detection, which involves an automatic A/V and OD semantic segmentation through a U-net that uses colored retinal fundus images. Furthermore, we implemented a novel loss function that allows the model to concentrate more on learning difficult instances.



<i>Autor</i>	<i>Year</i>	<i>Technique</i>	<i>Database</i>	<i>Metrics</i>		
				<i>Sensitivity</i>	<i>Specificity</i>	<i>Accuracy</i>
Akbar et al. [14]	2018	SVM + RBF	INSPIRE-AVR	94.25 %	95.47 %	95.10%
Huang et al. [8]	2018	Feature selection	DRIVE	70.90 %	73.80 %	72 %
		with genetic search	INSPIRE-AVR	89.60 %	91.30 %	92 %
Xu et al. [49]	2018	FCN	DRIVE	87 %	<b>98 %</b>	83.20 %
Srinidhi et al. [48]	2019	Vascular tree + RF	DRIVE	<b>96.60 %</b>	92.90 %	<b>94.70 %</b>
			INSPIRE-AVR	<b>96.90 %</b>	<b>96.60 %</b>	<b>96.80 %</b>
Galdran et al. [52]	2019	CNN with semantic segmentation	DRIVE	89 %	90 %	89 %
			INSPIRE-AVR	82 %	78 %	80 %
Girard et al. [21]	2019	CNN	DRIVE	86.30 %	86.60 %	86.50 %
Ma et al. [50]	2019	Multi-task ANN	DRIVE	93.40 %	95.50 %	94.50 %
		with spatial activation	INSPIRE-AVR	92.40 %	91.30 %	91.60 %
Kang et al. [51]	2020	CWF + VSR	DRIVE	88.30 %	92.71 %	90.62 %
Morano et al. [47]	2021	Multisegmentation	DRIVE	87.47 %	90.89 %	89.24 %

Table 4.2: Brief summary of the current state-of-the-art.

# Methodology

The proposed method consists of 5 main stages, Figure 5.1 shows the proposed system's flow diagram, as well as its submodules.

The first step is the *Image acquisition*, where two public databases were selected, and another one was created with the INDEREB aid. It is comprised of two versions, the first one consists of 46 images and the second one of 59 images, both can be found in the Appendix. Next, *Preprocessing* entails enhancing retinal fundus images as well as data augmentation. It is followed by the critical stage: *Processing*, which is where the U-net uses semantic segmentation for the identification of arteries, veins, and optic disc. Then the AVR is computed during the *Post-Processing* to finally compare results against experts evaluations in the last stage.

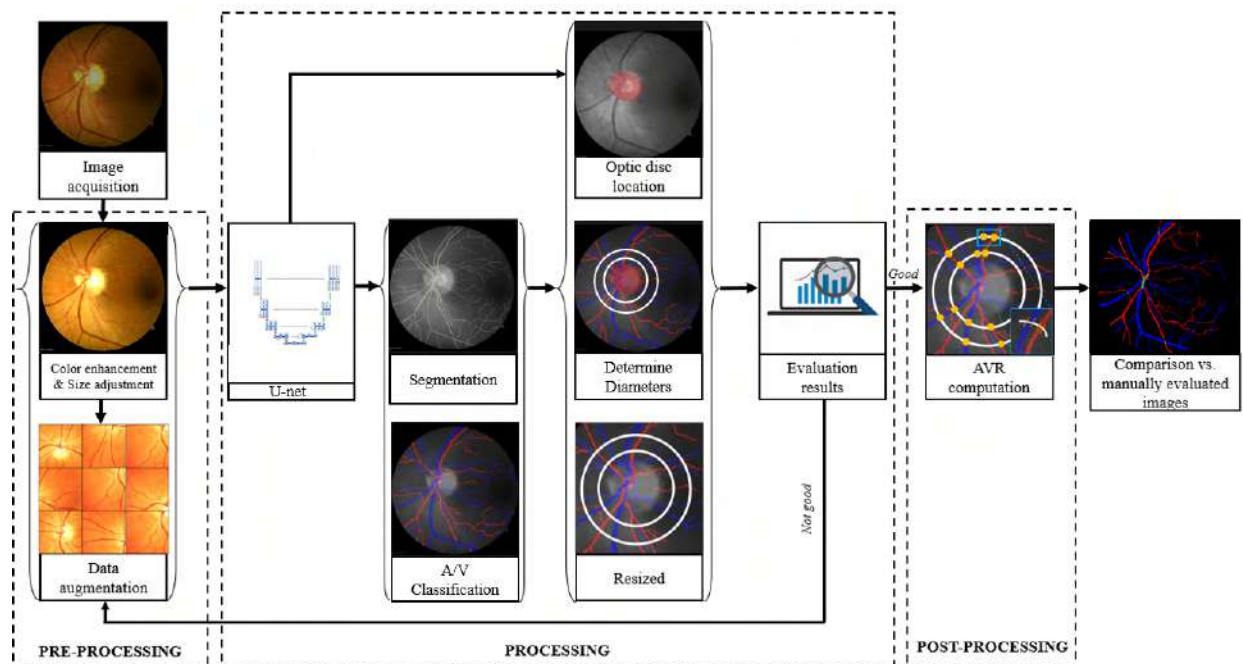


Figure 5.1: Proposed approach of a fully unsupervised system for hypertension retinopathy detection.

## 5.1 Image acquisition

This research works with fundus images of human eyes. There are several public databases available online that enable comparative studies on segmentation of blood vessels in retinal images. However, the selected public databases are: DRIVE [53] and INSPIRE-AVR [54].

DRIVE [53] (*Digital Retinal Images for Vessel Extraction*) is a dataset for retinal vessel segmentation, widely used to enable comparative studies. It comprises a total of 40 color fundus photos, with a resolution of 584 x 565 pixels. The images were collected using a Canon CR5 non-mydratic 3CCD camera with FOV equivalent to 45 degrees. The ground truth for the A/V classification of the dataset has been provided [55], which is based on the original DRIVE's 2nd manual vessel reference. Besides, a manual annotation of the OD was created for training purposes of the proposed network. An example of the DRIVE dataset can be seen in Figure 5.2.

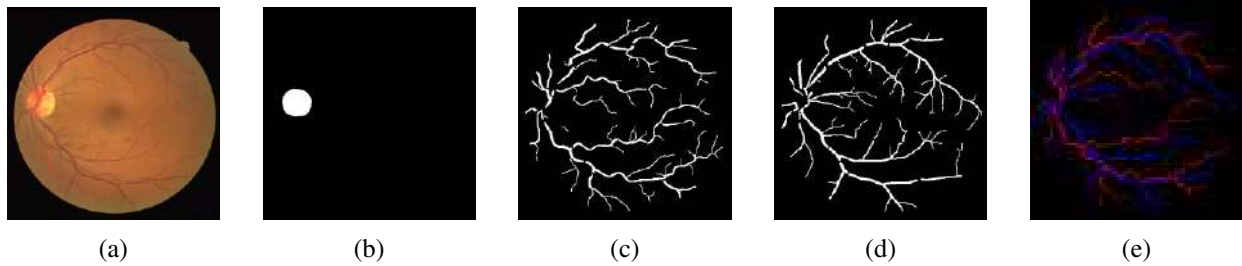


Figure 5.2: DRIVE Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) Arteries and Veins.

INSPIRE-AVR [54] (*Iowa Normative Set for Processing Images of the REtina*) consists of 40 color photographs centered at the optic disc. It also provides a reference standard for the AVR, based on the average of two experts' evaluations. Each image has a resolution of 2392x2048 pixels. The ground truth for the A/V classification of the dataset has also been provided [46]. An example of the INSPIRE-AVR dataset can be seen in Figure 5.3.

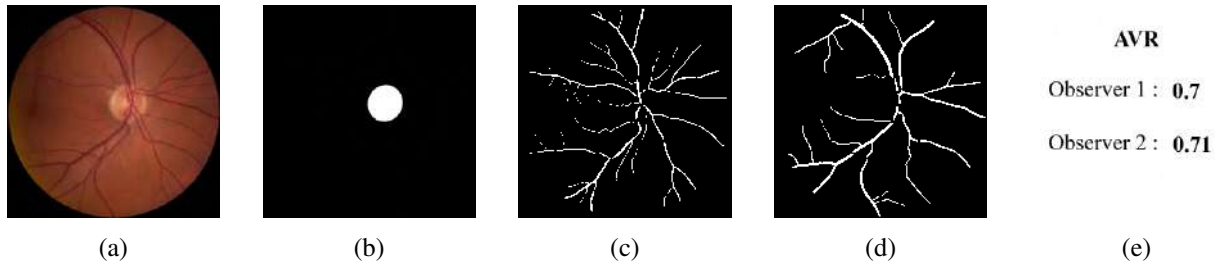


Figure 5.3: INSPIRE-AVR Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) AVR.

Besides the public datasets, a private one was created with a total of 105 images centered at the optic disc that were provided by the INDEREB. This new database is divided in two versions.

INDEREB version 1 consists of 46 images centered at the optic disc with a resolution of 1620 x 1444 pixels. The images were obtained using a camera Carl Zeiss, visucam nm/fa mydriatic with 45 and 30 degrees and a FOV of 30 degrees. The ground truth for A/V classification were provided by experts and manually annotated using Photoshop. Besides, a manual annotation of the OD was created for training purposes of the proposed network. The complete database can be found in the Appendix [E.1](#) and an example of the dataset is shown in [5.4](#).

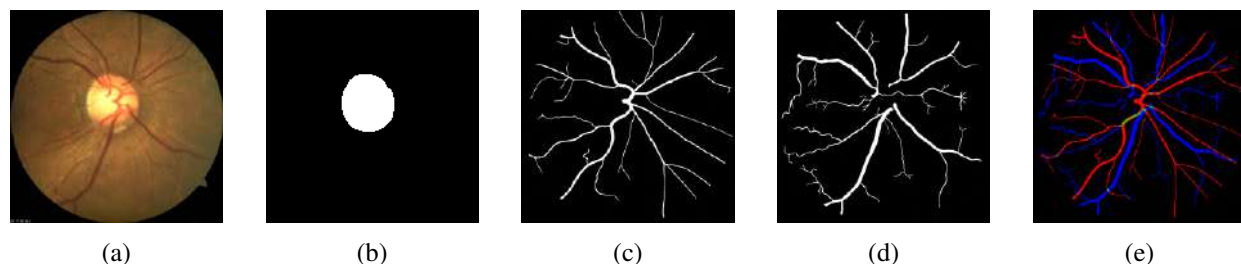


Figure 5.4: INDEREB Version 1 Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) Arteries and Veins.

Fundus images in INDEREB version 1 were selected with the following exclusion criteria:

- Smokers
- Diabetics
- Cataracts
- $> +2.00$  Diopters

INDEREB version 2 consists of 59 images centered at the optic disc with a resolution of 3912 x 3910 pixels. The camera used to capture the fundus images is a ZEISS CLARUS 700, which has an incorporated tool named Caliper that allows experts to take an accurate measure of blood vessels widths. The A/V ground truth was provided by experts and manually labeled using Krita. Likewise, a manual annotation of the OD was created for training purposes of the proposed network. The complete database can be seen in Appendix [E.2](#) and an example of the dataset is shown in [5.5](#).

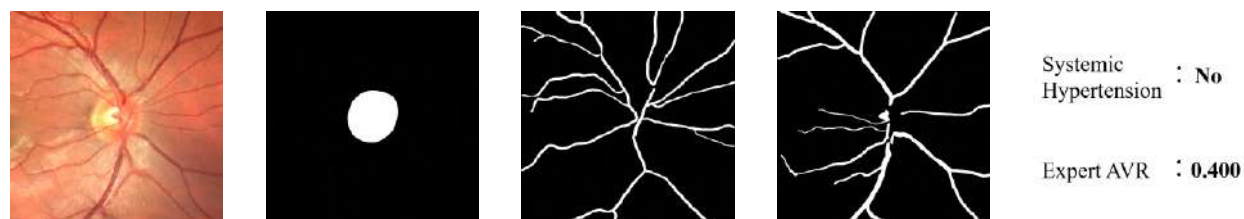


Figure 5.5: INDEREB Version 2 Database. (a) Fundus image, Ground Truths: (b) OD, (c) Arteries, (d) Veins, (e) Systemic Hypertension presence and AVR.

Fundus images in INDEREB version 2 were selected with the following exclusion criteria:

- Smokers
- Diabetics
- Cataracts
- Glaucoma
- Myopia
- > +6.00 Diopters
- Systemic disease
- Previous eye surgery
- Vasoconstrictor medication

This database was created with the intention of having balanced data, hence almost half of the images belong to patients who have *Systemic Hypertension*, as seen on Figure 1.2. Unfortunately, it is not an easy task to find enough *Males* that pass the exclusion criteria, which is why there are more *Females* fundus images in the database, Figure 5.6b.

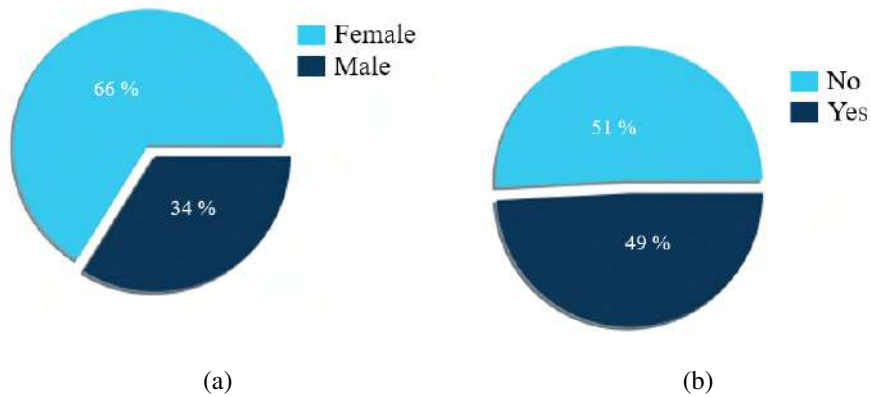


Figure 5.6: INDEREB Version 2. Distribution on (a) Gender (b) Systemic Hypertension.

Additionally, the patient's *Age* was considered an important criteria in the creation of this database. It was determined to establish the range from 17 to 75 years old. The *Age* density against *Gender* can be appreciated in Figure 5.7a and against the disease diagnostic in Figure 5.7b.

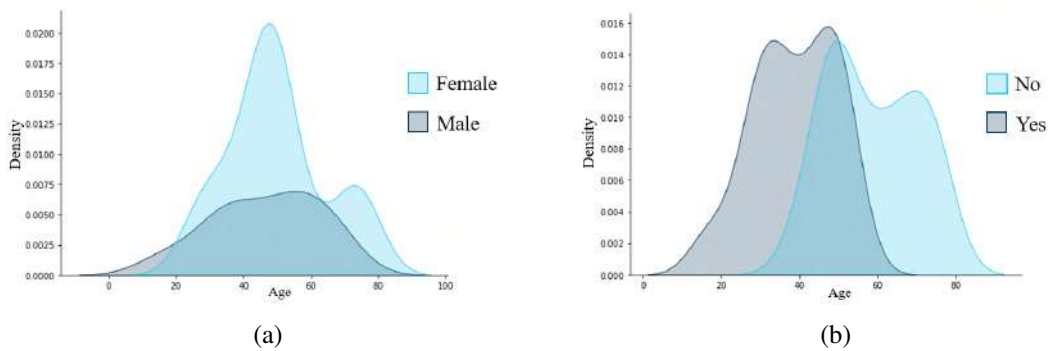


Figure 5.7: INDEREB Version 2. Age distribution against (a) Gender (b) Systemic Hypertension.

The detailed information of the database INDEREB Version 2 can be found in Table 5.1. It includes the file name, as well as its corresponding Gender, Age, whether the fundus image belongs to a patient with Systemic Hypertension or not, and the AVR gold standard provided by experts.

File	Gender	Age	Systemic Hypertension	Expert AVR
Image001	Male	32	No	0.400
Image002	Male	51	No	0.531
Image003	Male	51	No	0.518
Image004	Male	64	Yes	0.719
Image005	Female	50	Yes	0.602
Image006	Female	50	Yes	0.554
Image007	Male	63	Yes	0.500
Image008	Male	63	Yes	0.648
Image009	Female	75	Yes	0.722
Image010	Female	75	Yes	0.655
Image011	Female	65	Yes	0.632
Image012	Female	48	Yes	0.610
Image013	Female	48	Yes	0.590
Image014	Female	47	Yes	0.660
Image015	Female	47	Yes	0.612
Image016	Female	49	No	0.578
Image017	Female	49	No	0.631
Image018	Female	31	No	0.497
Image019	Female	31	No	0.471
Image020	Male	33	No	0.584
Image021	Male	33	No	0.647
Image022	Male	35	No	0.528
Image023	Male	35	No	0.486
Image024	Female	28	No	0.443
Image025	Female	28	No	0.490
Image026	Male	17	No	0.557
Image027	Male	17	No	0.606
Image028	Male	41	No	0.534
Image029	Male	41	No	0.622
⋮	⋮	⋮		

File	Gender	Age	Systemic Hypertension	Expert AVR
⋮	⋮	⋮		
Image030	Female	54	No	0.638
Image031	Female	40	No	0.694
Image032	Female	40	No	0.694
Image033	Female	46	No	0.596
Image034	Male	51	No	0.497
Image035	Male	51	No	0.475
Image036	Female	45	No	0.530
Image037	Female	45	No	0.720
Image038	Female	75	Yes	0.654
Image039	Female	50	Yes	0.591
Image040	Female	50	Yes	0.565
Image041	Female	74	Yes	0.561
Image042	Female	74	Yes	0.509
Image043	Female	49	Yes	0.618
Image044	Female	49	Yes	0.509
Image045	Female	49	No	0.746
Image046	Female	49	No	0.628
Image047	Female	27	No	0.540
Image048	Female	27	No	0.595
Image049	Female	72	Yes	0.673
Image050	Female	72	Yes	0.627
Image051	Female	59	Yes	0.607
Image052	Female	41	Yes	0.532
Image053	Female	41	Yes	0.554
Image054	Male	68	Yes	0.547
Image055	Male	68	Yes	0.542
Image056	Female	35	No	0.585
Image057	Female	51	Yes	0.508
Image058	Male	57	Yes	0.528
Image059	Male	57	Yes	0.496

Table 5.1: Database INDEREB version 2.

## 5.2 Pre-processing

### 5.2.1 Color and Size adjustment

Since this research works with several different databases, it is necessary to enhancing the color of all fundus images and adjusting the size at 512 x 512 pixels. This stage is brief but critical.

The selected techniques for this experimentation seek to highlight the inherent characteristics of retinal blood vessels, such as color, since arteries tend to present a lighter tone than veins. Additionally, the aim is to mitigate illumination variations to obtain a more robust segmentation. It was determined to prove the techniques: CLAHE, Histogram Equalization and Gamma Correction in three color spaces: Lab, HSV, and RGB.

First, CLAHE is applied in an fundus image with medium brightness. In LAB, Figure 5.8b, it clearly highlights the blood vessels, while in HSV, Figure 5.8c, it creates more contrast in the blood vessels, yet the smaller ones are not very detailed. Finally, the Figure 5.8d shows the image preprocessed in each of its channels in RGB. Here, the difference between veins and arteries is appreciated, but it also adds noise as green color which may results in a poor classification.

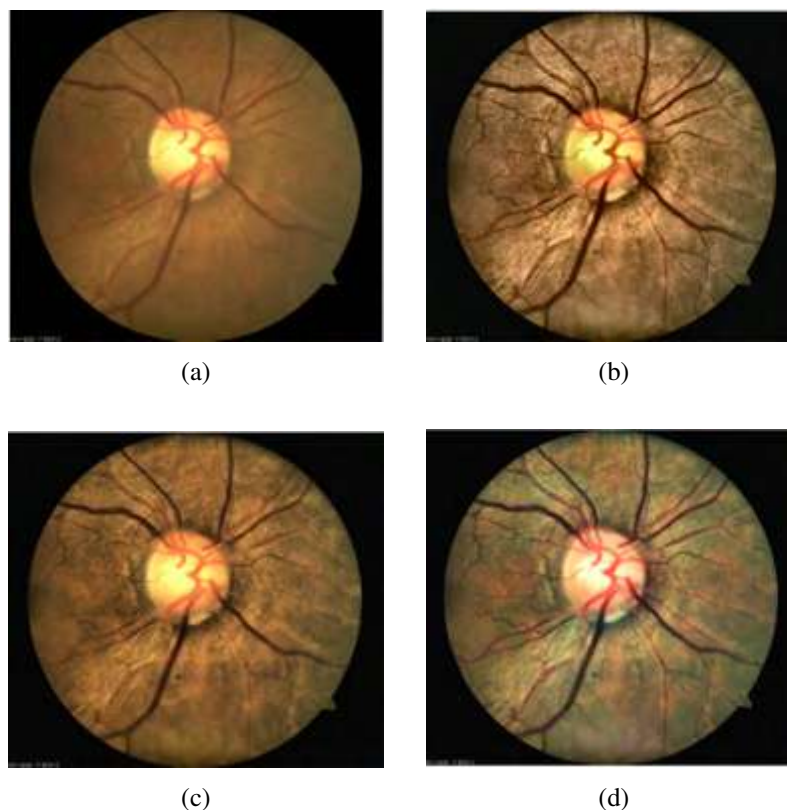


Figure 5.8: CLAHE. (a) Original, (b) Lab, (c) HSV, (d) RGB.

The next technique to evaluate is Histogram Equalization in a low brightness image. In Lab, Figure 5.9b, it permits an acceptable appreciation of the retinal vessels, even smaller ones. While in HSV would seem to

have a superior performance, its disadvantage is that the edges of the fundus image end up being very dark and thus, harder to classify. Finally, applying preprocessing to each channel is not a good approach for this method, as shown in the Figure [5.9d](#)

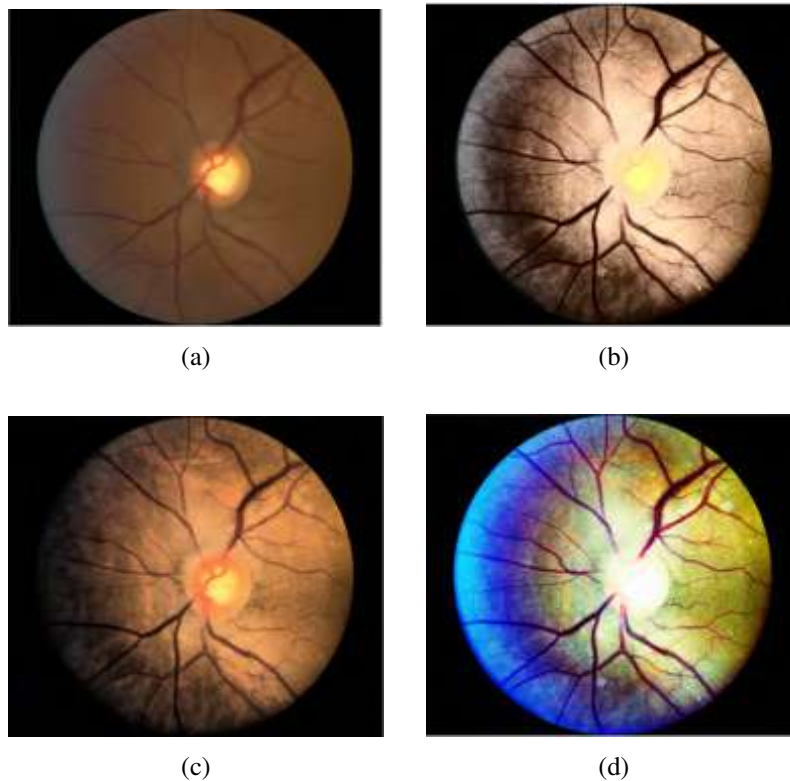


Figure 5.9: Histogram Equalization. (a) Original, (b) Lab, (c) HSV, (d) RGB.

Finally, Gamma Correction is applied for an image with high brightness. In Lab the image brightness increases, which makes the vessels at the darker ends noticeable, yet causing less color difference between arteries and veins, Figure [5.10b](#). It can be seen that in HSV the technique has no significant influence, Figure [5.10c](#). On the other hand, applying the preprocessing to each channel in RGB, Figure [5.10d](#), does not benefit us in terms of image interpretability.

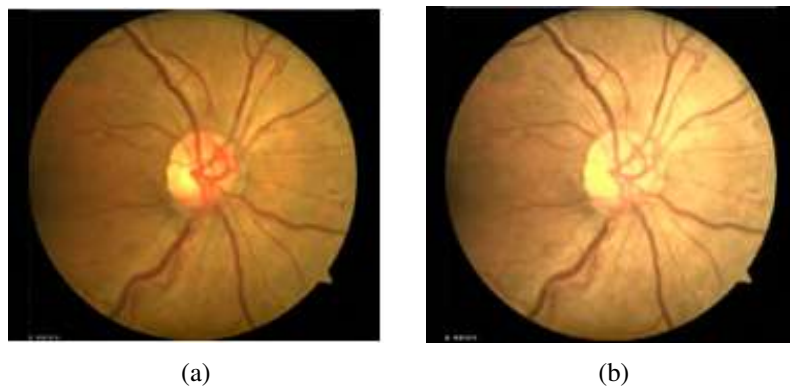






Figure 5.10: Gamma Correction. (a) Original, (b) Lab, (c) HSV, (d) RGB.

Image preprocessing can improve the quality, as well as the accuracy of classification models or contribute nothing when an incorrect technique or color space is used. In this experimentation three methods were shown, however, it is observed that CLAHE brings a great potential in the Lab color space since is able to highlight the retinal vessels without adding too much luminosity. Hence, this research uses CLAHE in the Lab color space to preprocess all databases.

It has been demonstrated the importance of having an adequate fundus image preprocessing for retinal vessel segmentation and classification using deep learning techniques. This is because, by highlighting the intrinsic characteristics of the images, in this case arteries and veins, the algorithm used for retinal vessel prediction will be able to improve its results in terms of sensitivity, specificity and accuracy.

### 5.2.2 Data augmentation

It is used when there is limited data or small amounts of it. Since the three databases used in this research have less than 100 images each, it can be considered as a problem of insufficient information, which may difficult the DL model task to correctly learn without over-fitting.

The premise of data augmentation is to create new data via random realistic alterations to the existing dataset, such as cropping, flipping, stretching, among others. An example can be seen in Figure 5.11. Luckily, Tensorflow offers several libraries that allows us to increase the diversity of each dataset by applying transformations into images, which are the ones used in this research.

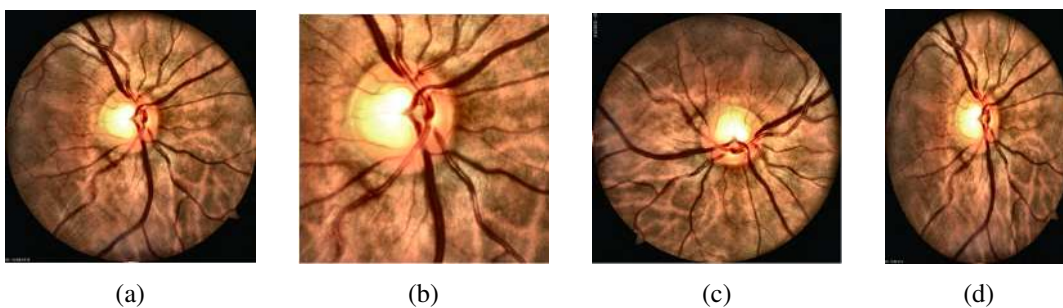


Figure 5.11: Data augmentation via random transformations. (a) Original, (b) Crop, (c) Flip, (d) Stretch.

## 5.3 Processing

### 5.3.1 Network architecture

This research uses an improved U-net model [3] that allows simultaneously segmentation in three-channel images of arterioles and venules, and in a separate instance, optic discs. An illustration of the proposed model can be seen in Figure 5.12. The network is constructed with two main paths denominated as contracting (*encoder*) and expanding (*decoder*), including skip connections through concatenation.

The encoder is constructed with four down-sampling blocks of 3x3 convolutions, each with a ReLU as an activation function and a 2x2 max pooling. Hence, the resolution is reduced by half during each down-sampling. The decoder consists of four up-sampling blocks of 3x3 convolutions using ReLU, and a 2x2 transposed convolution. The result is then concatenated with its corresponding feature map resolution through a skip connection. In consequence, features from the encoder are sent to the decoder at different resolution levels, producing a more thorough outcome.

The input is configured to receive a three-channel RGB image of size 3x512x512. The output segmentation map is adjusted at the expanding path, where the softmax activation is used to allow predictions for three classes: 'artery', 'vein', and 'background'. In contrast to the U-Net model, which crops the images after each convolution, our model produces a segmented image of the same size as the input.

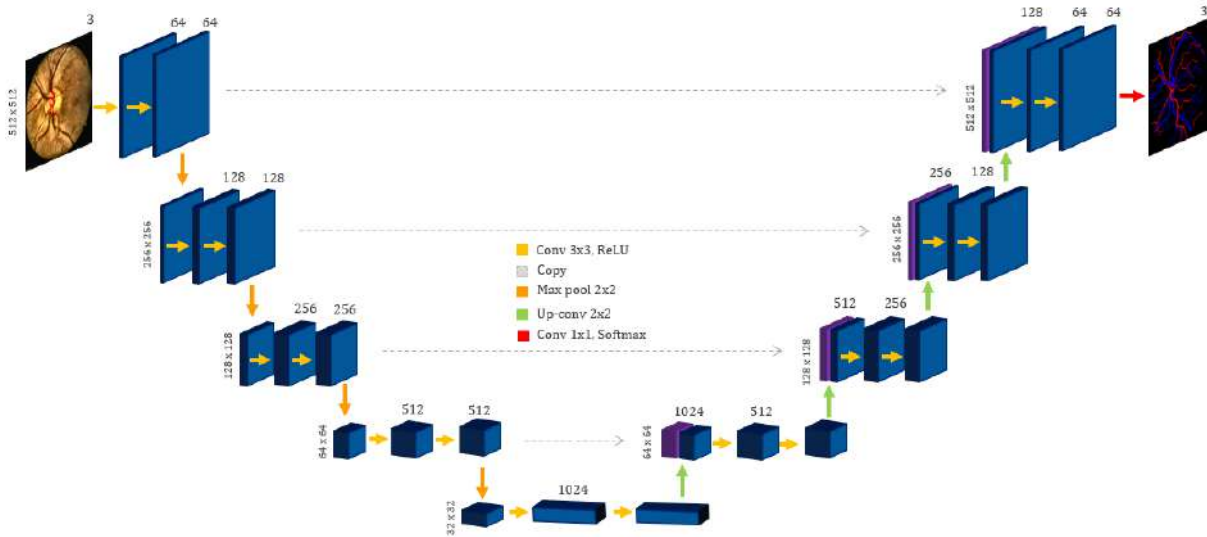


Figure 5.12: Architecture of a U-net neural network

#### 5.3.1.1 Hyperparameters

In ML, the way an artificial neural network is trained to learn is controlled by its hyperparameters, such as learning rate, optimizer, batch size, loss function and so on. A brief summary of the hyperparameters used in the proposed network can be seen in Table 5.2.

<i>Hyperparameter</i>	<i>Value</i>
U-Net depth	4
Activation function	ReLU
Learning rate	$1 \times 10^{-3}$
Batch size	32
Optimizer	ADAM
Loss function	Focal Tversky
No. training epochs	200

Table 5.2: Hyperparameters of proposed network.

This work was implemented in Python 3 (v3.7.0) using Tensorflow (v2.8.0). Training and testing were performed on a 2.30 GHz Intel® Xeon with an NVIDIA Tesla P100-PCIE-16GB. Primary libraries required for the development of the algorithm can be found in Appendix [A](#).

### Activation function

The activation functions purpose is to transform an ANN output signal into an input for the next layer in the architecture, which inherently affects the accuracy of the prediction. Its primary use is to introduce non-linearity in models [\[56, 57\]](#). Therefore, it is crucial to select the optimal type for the specific situation in Deep Learning models.

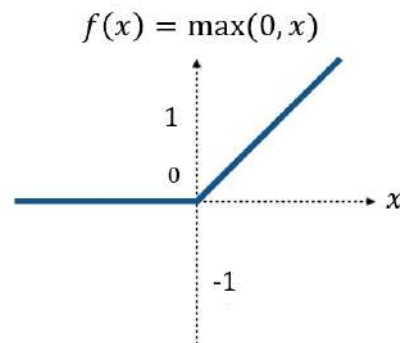


Figure 5.13: ReLU activation function

There are several activation functions being used in the literature, and this work will use Relu, Figure [5.13](#) shows the line plot. It is unique and perfect for this research because not all neurons are activated simultaneously, unlike its counterparts.

### Optimizer

Almost all challenges in ML and DL face a minimization problem in the weights of an ANN. Since most objective functions are stochastic, Kingma et al. [\[58\]](#) develop an efficient method to optimize. Its innovation resides in only needing first-order gradients along with small memory requirements. As its name states, ADAM (Adaptive moment estimation) adapts individual learning rates from estimations of gradients' first and second moments.

## Loss function

The proposed method uses a non-traditional loss function to address the inherent problem of unbalanced classes in fundus images. This is due to most of the pixels belonging to the 'background' and only the minority of them representing the vasculature of the retina. Selecting the most suitable loss is critical as it initiates the neural network's learning process.

The Focal Tversky Loss weights down easy examples to focus on hard ones using a  $\gamma$  coefficient ranging from [1,3] as shown in equation [5.1], where TI indicates the Tversky index [59].

$$FTL = \sum_c (1 - TI_c)^\gamma \quad (5.1)$$

The TI, shown in equation [5.2], can be considered as a generalization of the Dice score that improves the flexibility in balancing False Positives (FP) and False Negatives (FN).

$$TI_c = \frac{\sum_{i=1}^N p_{ic}g_{ic} + \epsilon}{\sum_{i=1}^N p_{ic}g_{ic} + \alpha \sum_{i=1}^N p_{ic}g_{i\bar{c}} + \beta \sum_{i=1}^N p_{i\bar{c}}g_{i\bar{c}} + \epsilon} \quad (5.2)$$

where,  $p_{ic}$  represents the predicted label and  $g_{ic}$  the ground truth. Furthermore,  $p_{i\bar{c}}$  corresponds to the pixel  $i$  probability of being class  $\bar{c}$ . The hyperparameters  $\alpha$  and  $\beta$  can be tuned to improve the sensitivity rate in unbalanced classes.

The proposed approach in [60] was to use a  $\gamma > 1$  so the loss function can focus on missclassifications. The best performance used  $\gamma = \frac{4}{3}$  and a higher  $\alpha$  value to improve the model convergence. Thus, this paper uses the same  $\gamma$  value with an  $\alpha = 0.7$  and  $\beta = 0.3$ .

### 5.3.1.2 Training

The neural network was initialized with a learning rate at  $1 \times 10^{-4}$ , the selected optimizer is Adam with both decay rates as-is from the original work in [58] ( $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ). Moreover, early stopping is introduced when the loss function does not change after 25 interactions.

The model was trained for 200 epochs. However, since each dataset contains less than 50 images, patches were augmented from training images during each sampling. The data augmentation techniques employed were elastic deformations such as horizontal and vertical flips, width and height shifts, random rotation and zoom.

### 5.3.2 Semantic segmentation

After the proposed improved U-Net has been trained, it is used to create semantic segmentation for the retinal A/V vasculature (artery and vein) in each of the four available databases. However, the OD segmentation was not implemented on the DRIVE dataset since it is not centered on the OD, thus it can not be successfully segmented in all cases.

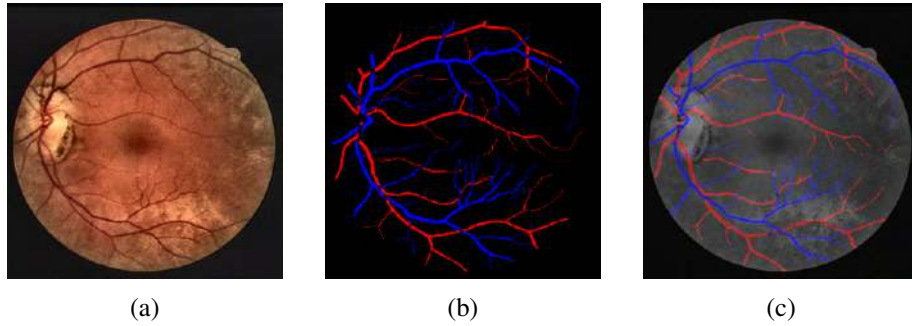


Figure 5.14: Semantic Segmentation on the DRIVE database. (a) Original, A/V: (b) Ground truth, and (c) Model prediction.

The semantic segmentation predictions done with the proposed network for the DRIVE database can be seen in Figure 5.14, while the results on the INSPIRE-AVR are shown in Figure 5.15. The predictions made for the INDEREB are visible in Figure 5.16 for version 1, and Figure 5.17 for version 2.

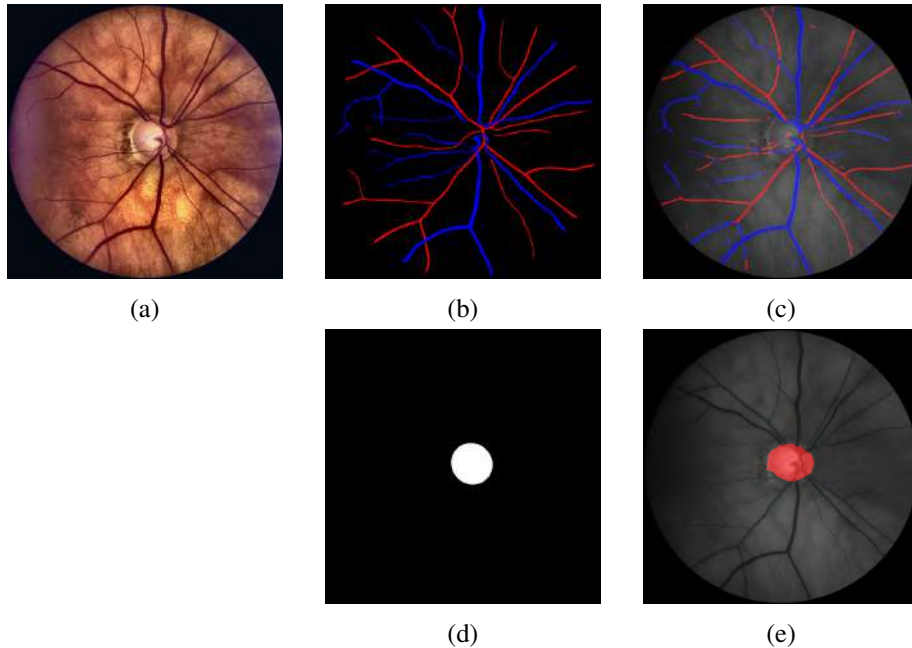


Figure 5.15: Semantic Segmentation on the INSPIRE-AVR database. (a) Original, A/V: (b) Ground truth, (c) Model prediction, OD: (d) Ground truth, (e) Model prediction.

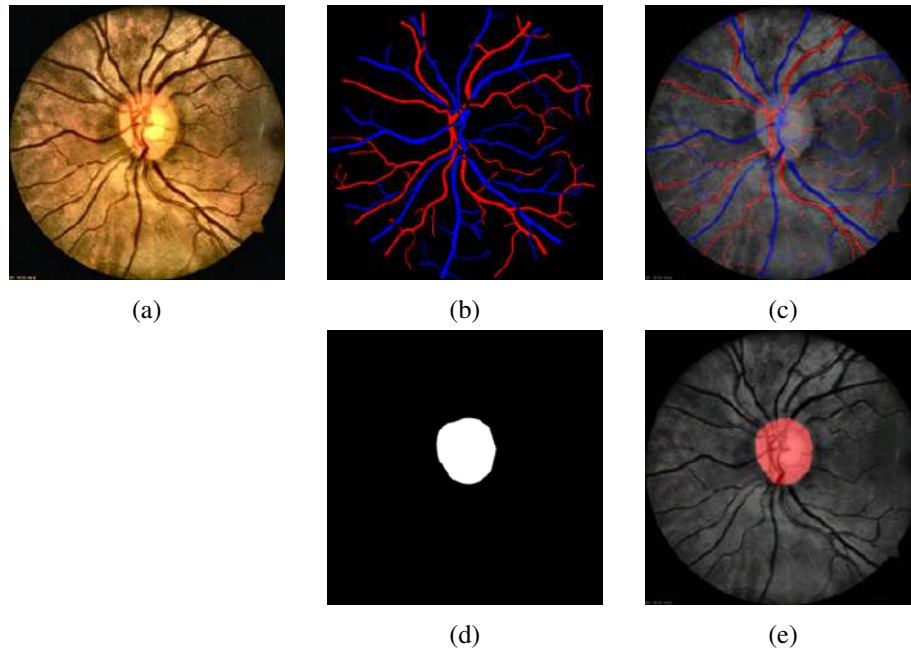


Figure 5.16: Semantic Segmentation on the INDEREB v1 database. (a) Original, A/V: (b) Ground truth, (c) Model prediction, OD: (d) Ground truth, (e) Model prediction.

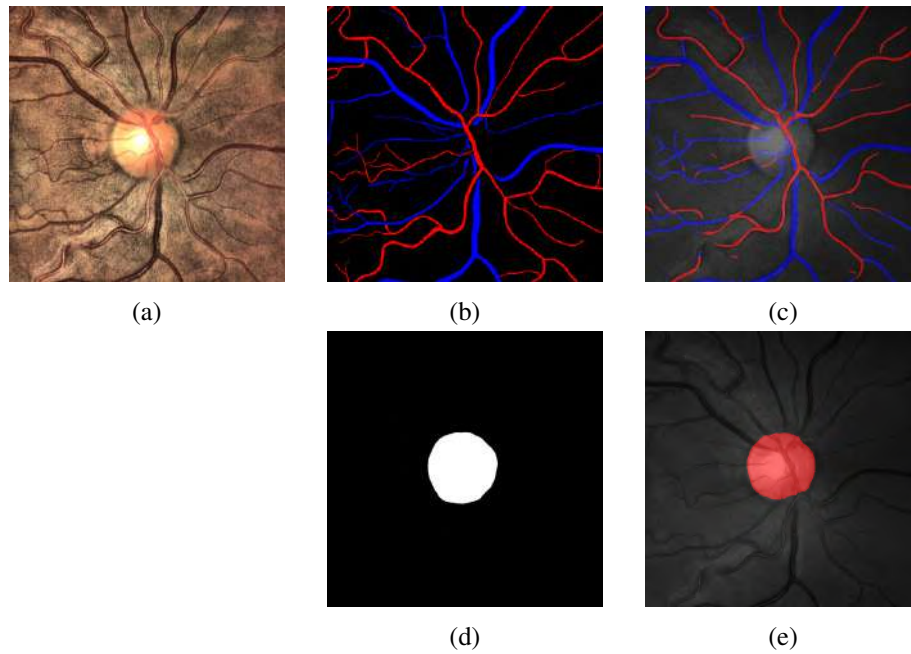


Figure 5.17: Semantic Segmentation on the INDEREB v2 database. (a) Original, A/V: (b) Ground truth, (c) Model prediction, OD: (d) Ground truth, (e) Model prediction.

## 5.4 Post-processing

In this stage, the model's prediction of A/V is used to estimate arteries and veins width using a technique named convex hull at  $[1 \text{ DD} - 1.5 \text{ DD}]$  from the OD prediction. Then equations [1.2](#), [1.2](#) are used with the width results to then employ equation [1.1](#) to obtain the AVR measurement.

### 5.4.1 AVR computation

As stated before, an automatic calculation requires a system capable of As stated before, an automatic calculation requires a system capable of delivering an accurate A/V classification along with OD localization to determine where the diameters will be located, also known as the region of interest (ROI), to then proceed to measure vessels' widths and finally compute the AVR. Figure [5.18](#) shows a brief summary of the workflow on the INSPIRE-AVR dataset.

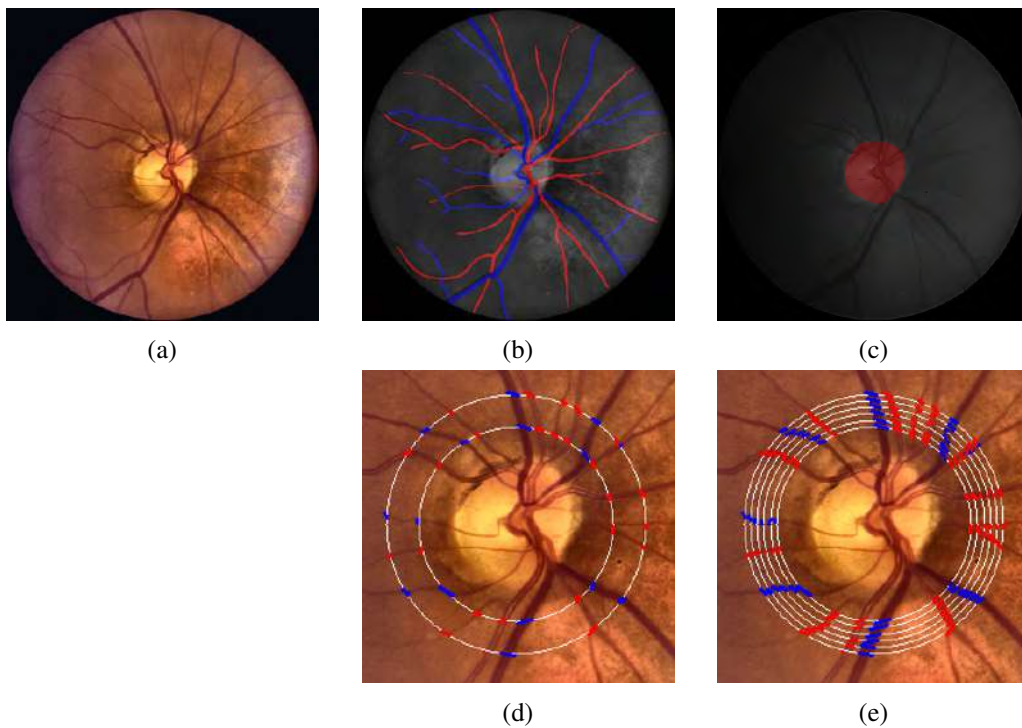


Figure 5.18: Proposed method on INSPIRE-AVR dataset. (a) Input image, (b) A/V classification, (c) OD location, (d) ROI and (e) AVR measure in 0.1 steps.

The ROI is the area that delimits where the vessels' width will be measured. It comprises two defined diameters centered at the OD, as shown in Figure [5.18d](#). The proposed method uses the same diameters as Niemeijer et al. [\[54\]](#), from 1 DD (Disk Diameter) to 1.5 DD, to maintain a golden standard and to have comparable results. The width of arteries and veins crossing at the established circumferences are measured in steps of 0.1 from the lower diameter to the upper one, as can be seen in Figure [5.18e](#). It is important to note that since all vessels have been classified, the system measures them all with a technique named Convex Hull.

During each step, the system compares all resulting widths and selects the 6 widest arteries for CRAE calculation using equation 1.2 and then selects the 6 widest veins for CRVE calculation using equation 1.3. After both CRAE and CRVE have been calculated, an AVR measure is computed using equation 1.1. Once all steps have been completed, each corresponding AVR result is averaged to deliver a final value.

### 5.4.1.1 Convex Hull

In computational geometry, a convex hull represents the smallest possible convex polygon, without inward bents, that contains a set of points in low dimensional Euclidean spaces, as seen in Figure 5.19.

There have been proposed plenty of algorithms that can construct the convex hull on a collection of points. This thesis works with the function implemented in OpenCV, which is based on Sklansky's [61] algorithm.

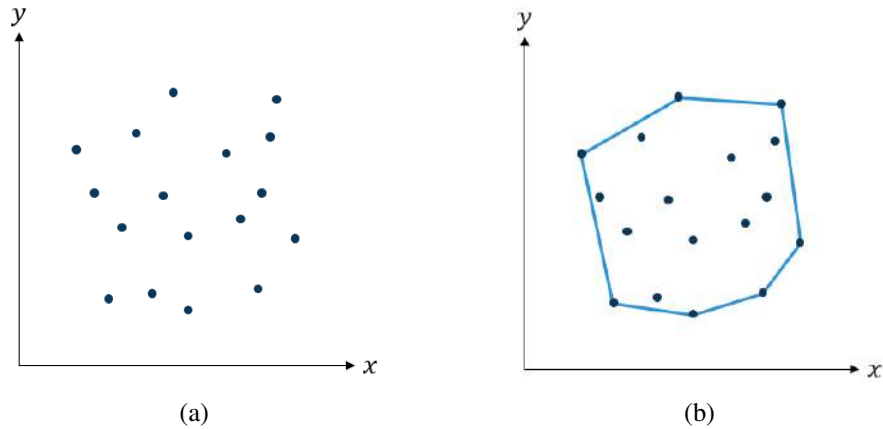


Figure 5.19: (a) Set of points and its (b) Convex Hull.

## ROI

The bounding circles of the measurement area are created with OpenCV, however, to create them it is necessary to know which coordinates correspond to the center of the OD, as well as its diameter. First, the predicted image is binarized, then it is determined where the highest concentration of white exists, and from these, the center of mass is estimated to obtain the centroid  $(\bar{x}, \bar{y})$  of the figure through its moments using equations 5.3, 5.4, and 5.5.

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (5.3)$$

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad (5.4)$$

$$\bar{y} = \frac{M_{01}}{M_{00}} \quad (5.5)$$



Then, the convex hull is applied to the ovoid figure, this technique returns a collection of pixels that form the perimeter. The Euclidean distance of each pixel is calculated against all the others to create a matrix of distances. The maximum distance recorded is taken, which is the one that corresponds to the diameter of the OD. These values will allow the circles to be drawn at the correct distance and size. An example can be seen in Figure 5.20.

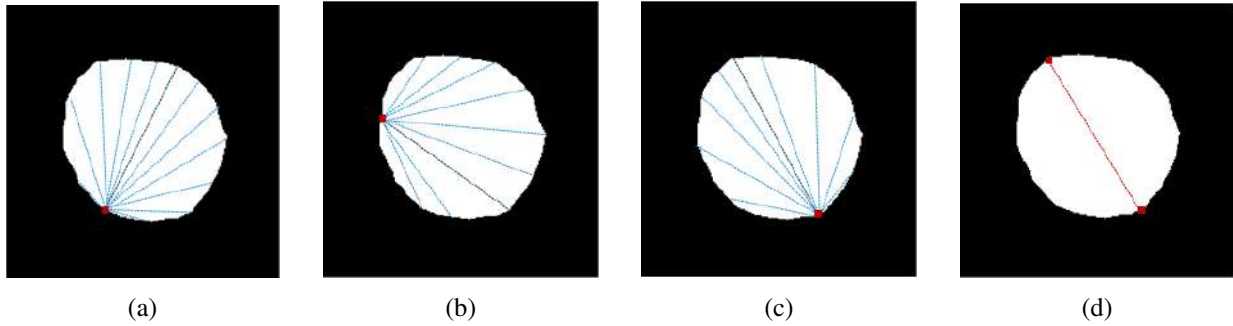
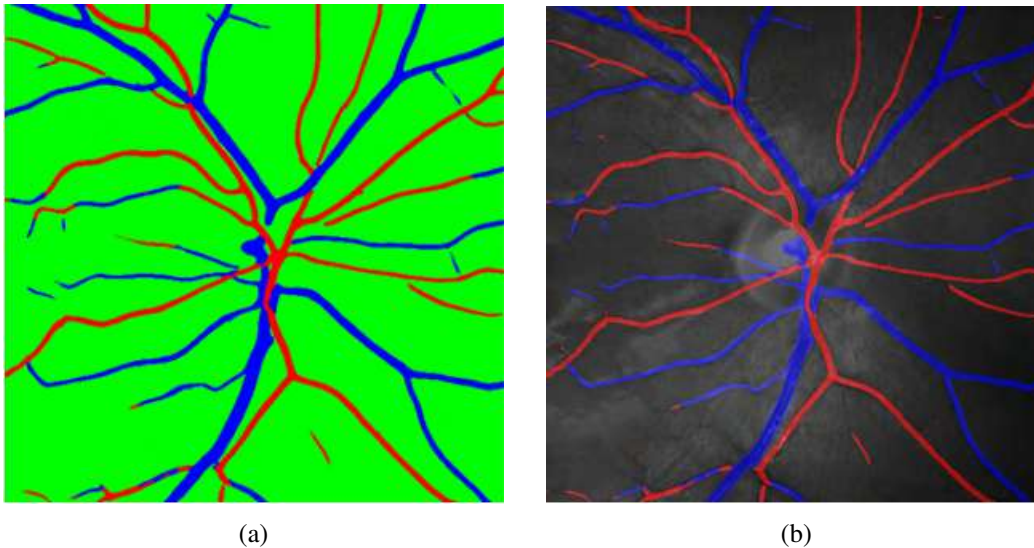


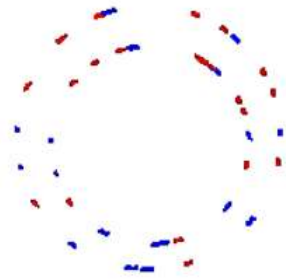
Figure 5.20: Diameter estimation through measuring the Euclidean distance of each pixel against the others. (d) Longest distance (*diameter*) obtained from random examples: (a), (b) and (c).

### Vessel width

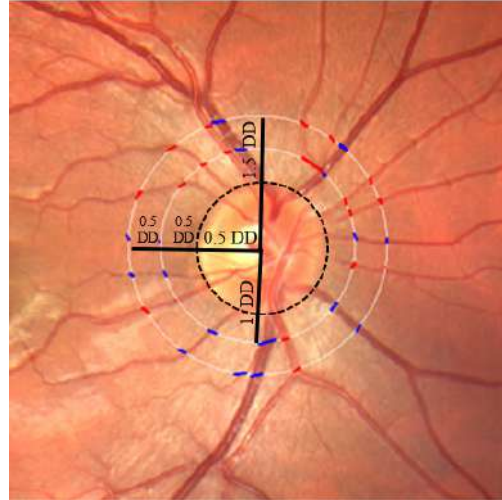
Thanks to the semantic segmentation, the prediction for the A/V classification is made up of three colors, as seen in Figure 5.21a, where R represents the arteries, G the background, and B the veins. A self-explanatory flow of blood vessel selection to be measured can be seen in Figure 5.21.

From the predicted image, two disks are drawn, one at 1 DD and the other at 1.5 DD from the OD, an example over the original image is shown in figure 5.21c. Then, a new image is created with the pixels in the R and B matrix of the original image that overlaps with the circumferences, it can be seen in Figure 5.21d.





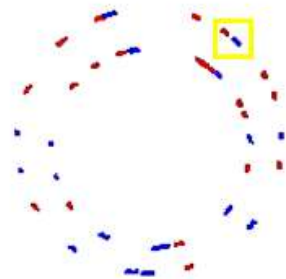
(c)



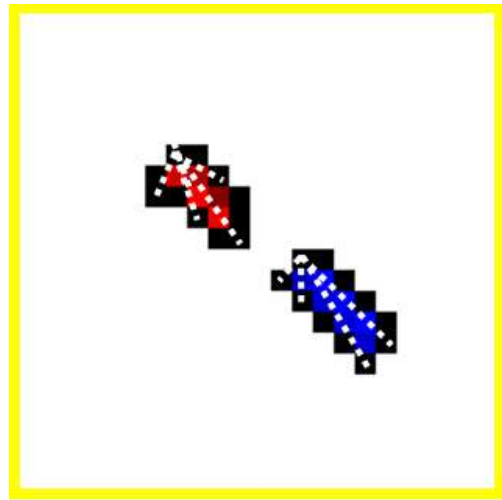
(d)

Figure 5.21: Blood vessel selection. (a) A/V prediction (b) over original image, (c) A/V pixels overlapping the circumferences (d) in original image.

The resulting image is conformed by small sections of pixels that intersected with the disks, as shown in Figure 5.22a. The convex hull is calculated for each of those, and again, the Euclidean distance of each pixel is calculated against the others. A close-up can be seen in Figure 5.22b, to obtain the maximum distance from the edge of the retinal blood vessel to the edge of itself, which corresponds to the width of the vein or artery.



(a)



(b)

Figure 5.22: Convex Hull for width estimation. (a) Image conformed by pixels that intersected with the disks, (b) Close-up at Euclidian distance calculation for each pixel.

### 5.4.1.2 Software

An application to calculate the AVR measurement at 1-1.5 DD in a fundus image was created using Streamlit [62]. It is a Python open-source framework commonly used to create and easily share custom web apps for machine learning and data science.

This software runs local and can be seen in your browser at `http://localhost:8501`.

The main app is shown in Figure 5.23. It has a sidebar with instructions on how to use the tool, Figure 5.24a shows it expanded, it also has a display option for the user, whether or not to show the images used to calculate the AVR, which can be seen in Figure 5.24b. Likewise, it has an About section that mentions that the tool was created for this thesis, as seen in Figure 5.24c.

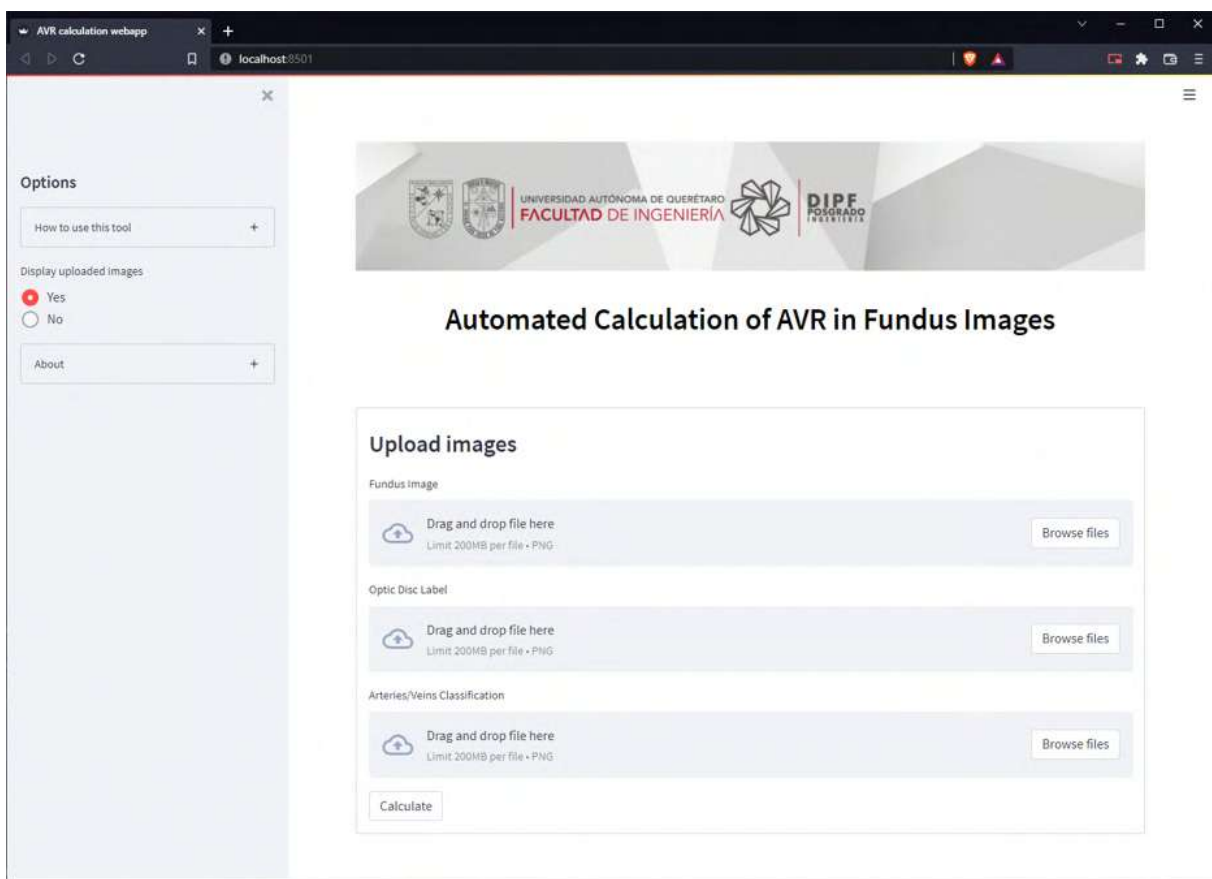


Figure 5.23: AVR calculation webapp.

This tool needs the original fundus image, the OD label and the A/V classification that the model previously predicted. After all images have been uploaded and the user has decided if all images are to be displayed, the user only need to click on *Calculate* to obtain the results. An example of this can be appreciated in Figure 5.25.

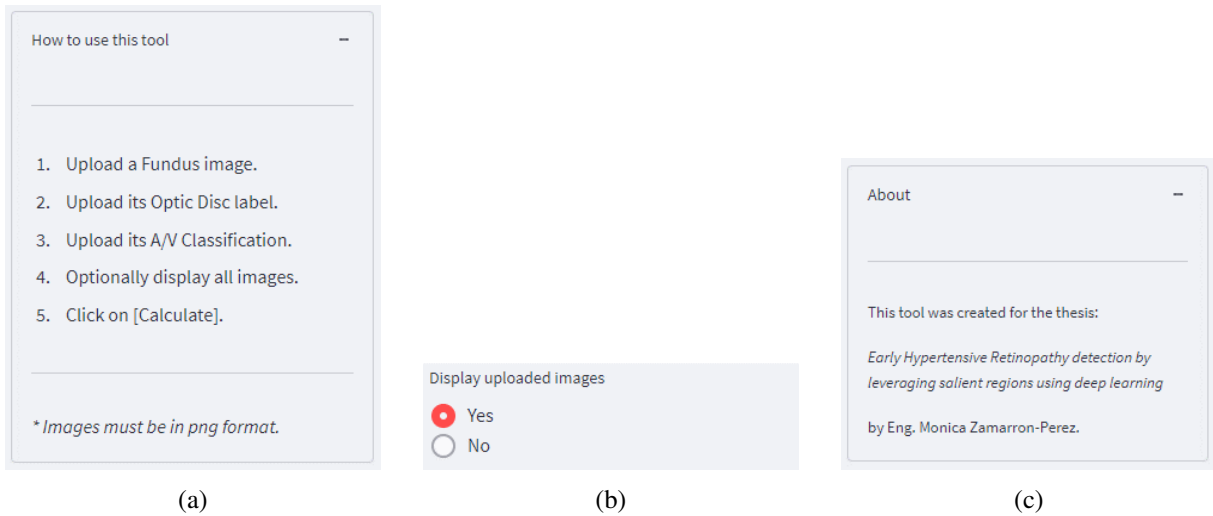


Figure 5.24: Sidebar options. (a) How to use the tool, (b) Display uploaded images, (c) About.

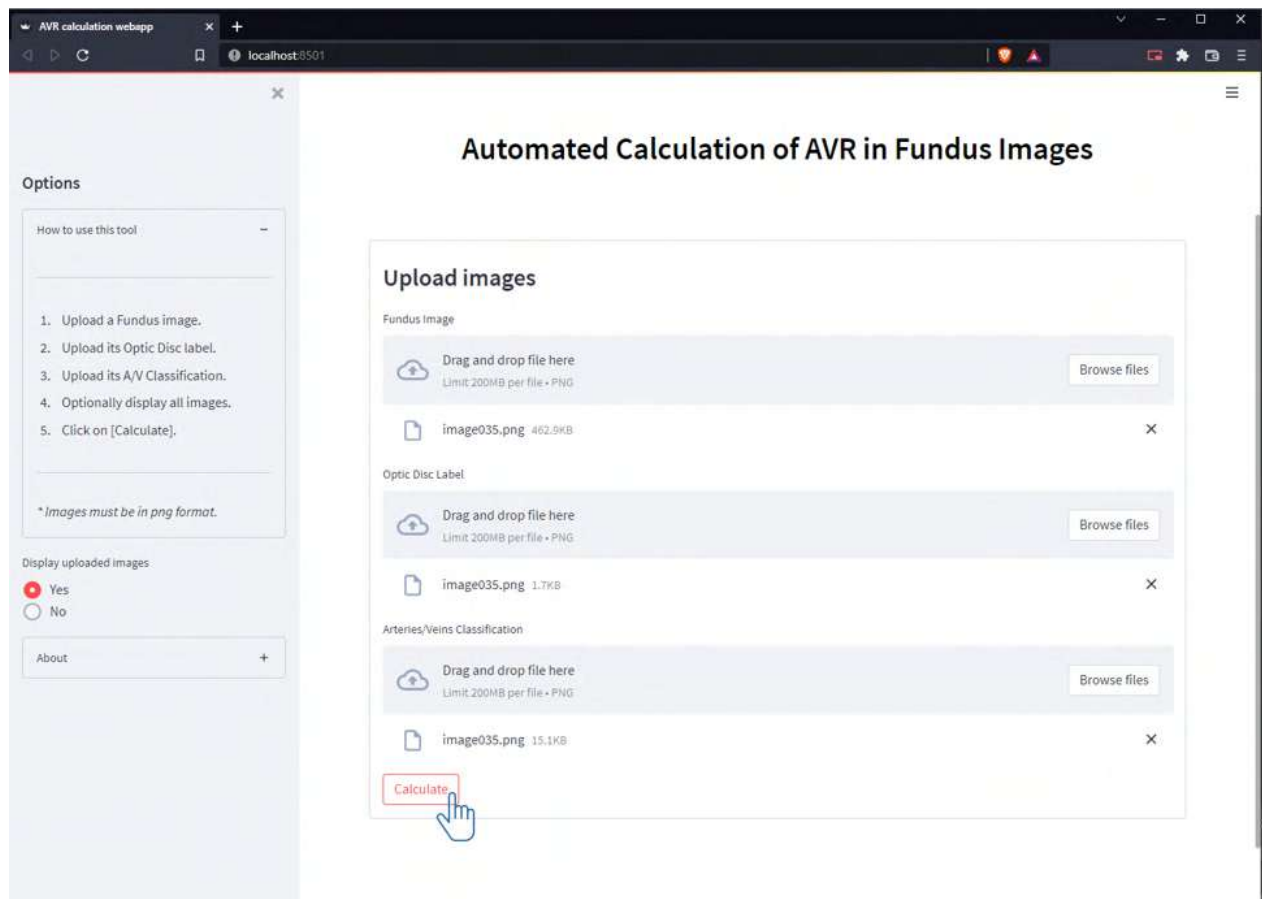
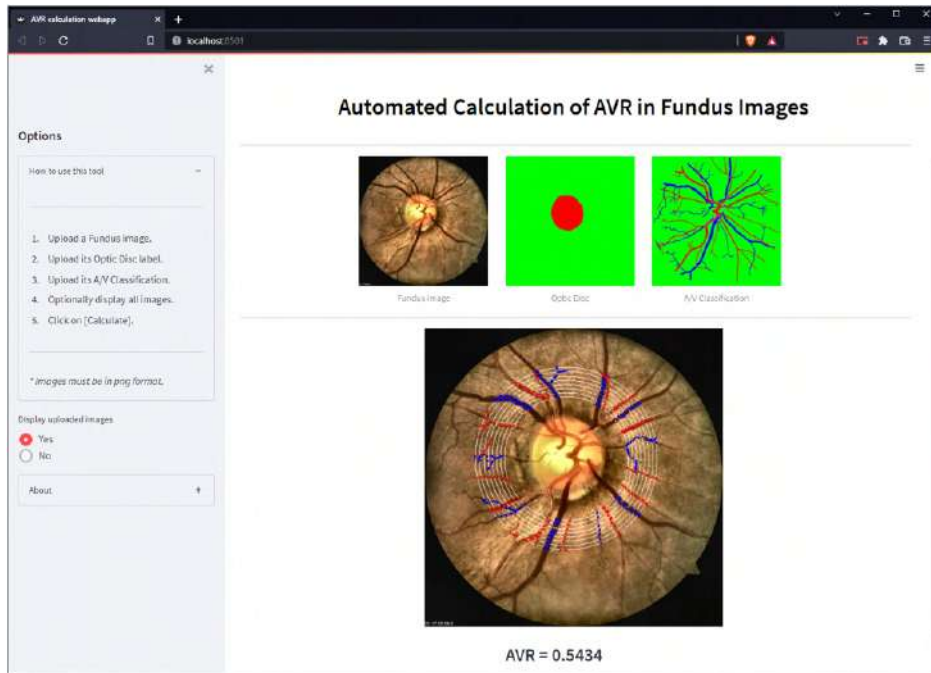
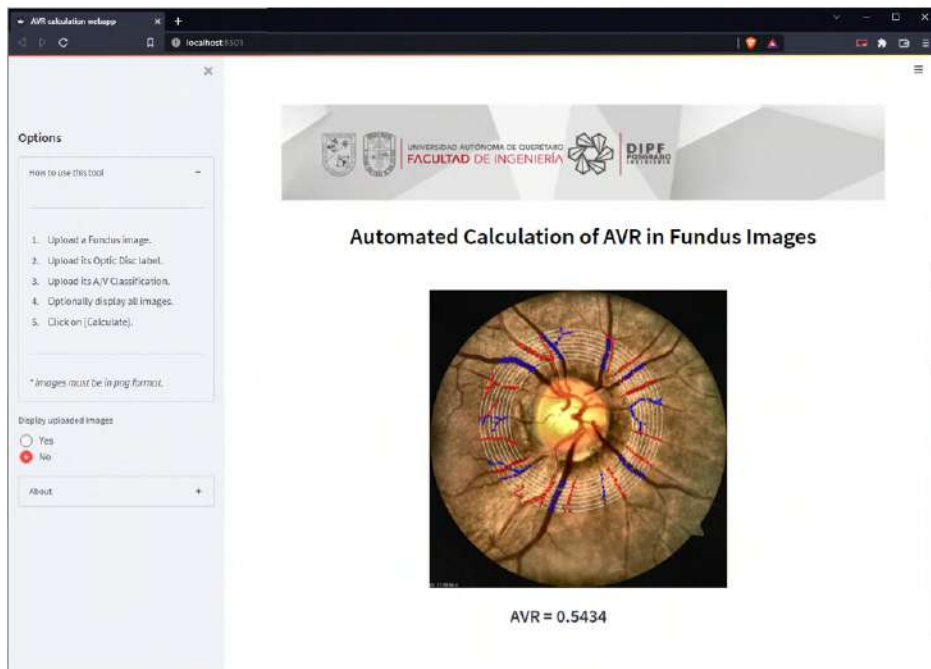


Figure 5.25: AVR calculation preparation.

Figure 5.26a shows the AVR results on a random image that belongs to the INDEREB dataset with the display uploaded images activated, while Figure 5.26b shows the results without the uploaded images.



(a)



(b)

Figure 5.26: AVR Results. (a) Displaying and (b) Without displaying uploaded images.

# Results

This chapter covers obtained results for blood vessel classification using the proposed U-net on three databases: DRIVE, INSPIRE-AVR, and INDEREB. Additionally, it compares the obtained metrics against the state-of-the-art works. Finally, the AVR results on INSPIRE-AVR and INDEREB are compared against those delivered by experts to analyze the proposed system.

## 6.1 Evaluation metrics

It is crucial to select the correct metrics to evaluate the model's performance against the current state-of-the-art after building it. Most works have employed sensitivity, specificity and accuracy, which calculation is based on four possible outcomes of the data.

**TP** -Positive classified correctly.

**TN** - Negative classified correctly.

**FP** - Negative classified as positive.

**FN** - Positive classified as negative.

Sensitivity, also known as Recall, expresses how well the model can detect the class, meaning from all the positive classes, how many were predicted correctly. It is calculated as in equation [6.1](#).

$$Sensitivity = \frac{TP}{TP + FN} \quad (6.1)$$

Specificity, also known as Precision, defines how reliable a model is in determining whether an element belongs to that class. In easier terms, from all classes predicted as positive, how many are actually positive. It is calculated as equation [6.2](#).

$$Specificity = \frac{TP}{TP + FP} \quad (6.2)$$

Accuracy determines how many positive and negative observations were successfully classified, calculated as in the equation [6.3](#).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6.3)$$

## 6.2 A/V classification

The retinal blood vessel classification was completed using a proposed U-net previously described in Chapter 5. This chapter compares this research’s results against recent works, from 2018 up to 2022, in two public databases and reports the results obtained in the dataset created with INDEREB aid.

The A/V classification was first implemented on DRIVE since that database has been used a comparative standard from several years now. Table [6.1](#) shows a comparison between state-of-the-art works and its results against the proposed method.

It can be appreciated that the *sensibility* was successfully improved up to 93.85% due to the focal Tversky loss implementation while maintaining satisfactory results regarding its *specificity* of 97.37%, and a higher *accuracy* than other works have achieved at 95.74%. This is due to having predictions with both sensibility and specificity similarly high, it improved the overall *accuracy*.

Method	Year	Sens.	Spec.	Acc.
Xu et al. <a href="#">[49]</a>	2018	87 %	<b>98 %</b>	83.2 %
Huang et al. <a href="#">[8]</a>	2018	70.90 %	73.80 %	72 %
Girard et al. <a href="#">[21]</a>	2019	86.3 %	86.6 %	86.5 %
Galdran et al. <a href="#">[52]</a>	2019	89 %	90 %	89 %
Ma et al. <a href="#">[50]</a>	2019	93.4 %	95.5 %	94.5 %
Srinidhi et al. <a href="#">[48]</a>	2019	<b>96.60 %</b>	92.90 %	94.70 %
Kang et al. <a href="#">[51]</a>	2020	88.30 %	92.71 %	90.62 %
Morano et al. <a href="#">[47]</a>	2021	87.47 %	90.89 %	89.24 %
Proposed	2022	93.85 %	97.37 %	<b>95.74%</b>

Table 6.1: Comparison with the state of the art for A/V Classification in the Drive dataset.

An example of an obtained prediction on the DRIVE dataset is shown in Figure [6.1](#). It can be appreciated that in smaller retinal blood vessels there can be a misclassification between artery and vein. However, in most cases, it consists of only a minor portion of pixels around the endpoints. Hence, why the metrics are not significantly affected. Besides, all state-of-the-art works face this type of problem.

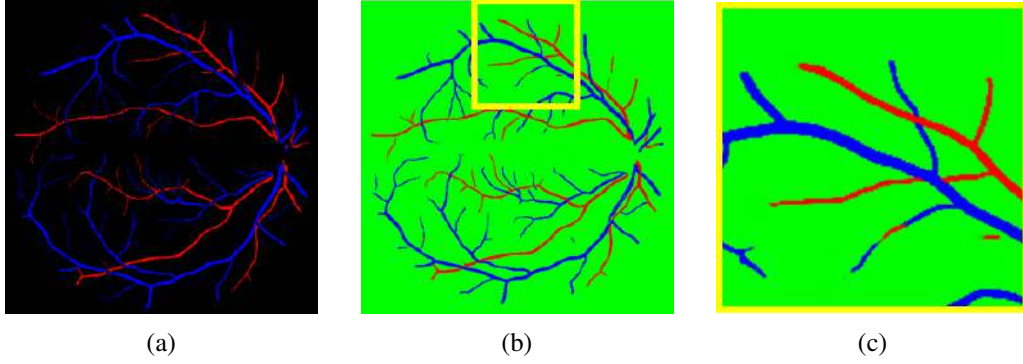


Figure 6.1: A/V Classification results on DRIVE. (a) Ground Truth, (b) Prediction, (c) Misclassification in prediction.

Similar behavior can be observed in Table 6.2, where works on the INSPIRE-AVR dataset for A/V classification are compared. The proposed method achieves a state-of-the-art value for the *specificity* at 97.88%, while providing great results for its *sensitivity* at 92.27%. The *accuracy* is among the top three at 95.15%.

Method	Year	Sens.	Spec.	Acc.
Akbar et al. [14]	2018	94.25 %	95.47 %	95.10 %
Huang et al. [8]	2018	89.60 %	91.30 %	92 %
Srinidhi et al. [48]	2019	<b>96.90 %</b>	96.60 %	<b>96.80 %</b>
Galdran et al. [52]	2019	82 %	78 %	80 %
Ma et. al [50]	2019	92.4 %	91.3 %	91.6 %
Proposed	2022	92.27 %	<b>97.88 %</b>	95.15 %

Table 6.2: Comparison with the state of the art for A/V Classification in the INSPIRE-AVR dataset.

Figure 6.2 enables a visual assessment of the INSPIRE-AVR dataset's resulting prediction. As in the previous database, smaller blood vessels are difficult to predict correctly for the neural network, as seen in Figure 6.2c. However, it is essential to note that since this thesis is focused on calculating the AVR near the OD, those edge cases do not impact the methodology and expected results.

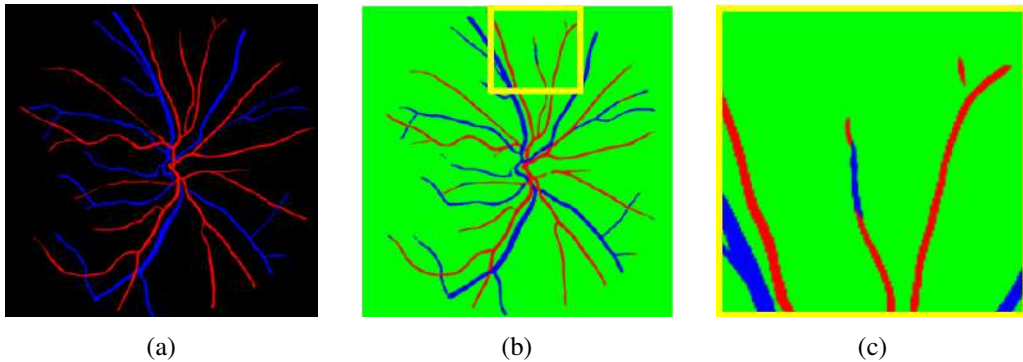


Figure 6.2: A/V Classification results on INSPIRE-AVR. (a) Ground Truth, (b) Prediction, (c) Misclassification in prediction.



Most importantly, obtained metrics on DRIVE and INSPIRE-AVR demonstrate that the proposed method has high reliability across different datasets, even though their differences are apparent, one centered at the macula and the other at the Optic Disc.

Finally, the neural network was tested on both versions of the created INDEREB dataset, which will be publicly available along with the ground truth provided by an expert. Table 6.3 shows the results obtained on the INDEREB version 1, while Table 6.4 displays the metrics on the version 2.

Method	Year	Sens.	Spec.	Acc.	FPR
Proposed	2022	90.55 %	96.93 %	93.97%	3.07%

Table 6.3: A/V Classification results obtained on the private dataset INDEREBv1.

Method	Year	Sens.	Spec.	Acc.	FPR
Proposed	2022	90.15 %	95.82 %	92.54%	4.18%

Table 6.4: A/V Classification results obtained on the private dataset INDEREBv2.

It can be seen that both versions have similar metrics regarding its results. Figure 6.3 shows the main problem in the predictions, the end of the vascular trees are harder to predict and classify due to their small size. Something similar happens in Figure 6.4, corresponding to version 2. However, in this dataset, the misclassification is more obvious where the vessels are crossing between classes.

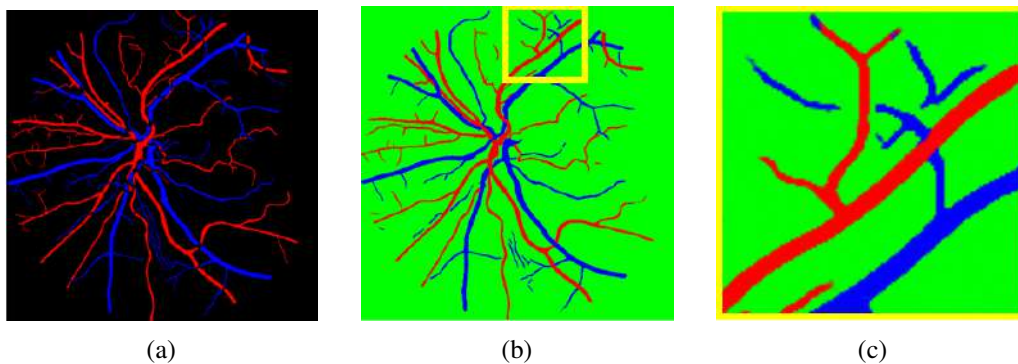


Figure 6.3: A/V Classification results on INDEREBv1. (a) Ground Truth, (b) Prediction, (c) Misclassification in prediction.

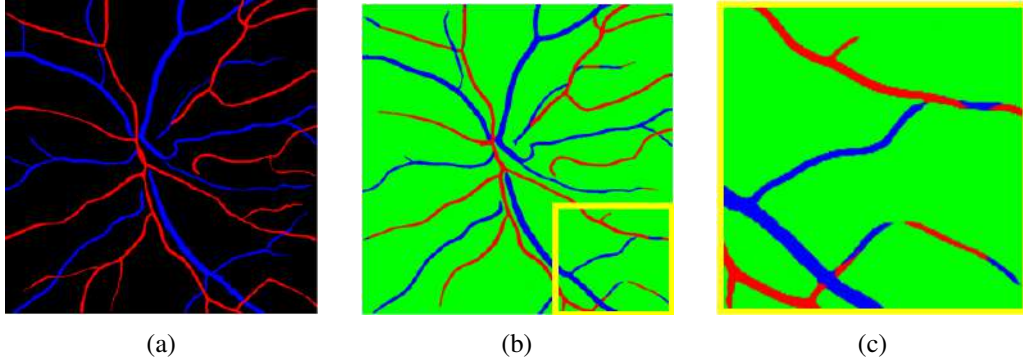


Figure 6.4: A/V Classification results on INDEREBv2. (a) Ground Truth, (b) Prediction, (c) Misclassification in prediction.

### 6.3 AVR against experts

The proposed algorithm for vessels width estimation was previously described on Chapter 5. It is capable of calculating each artery and vein width from the neural network predictions, to then select the 6 widest vessels to provide a final AVR measure.

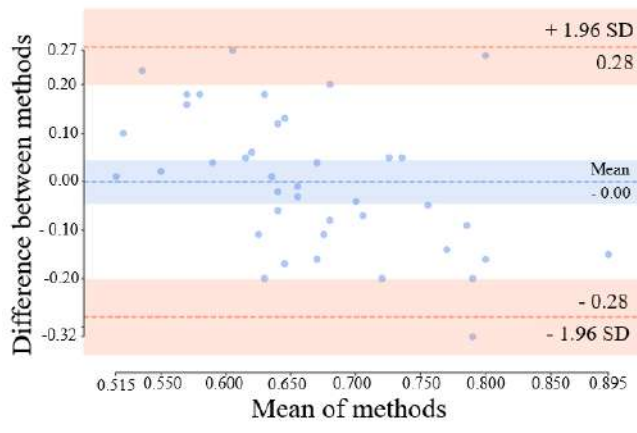
Table 6.4 shows a summary of results obtained for the INSPIRE-AVR dataset. The ground truth was first presented by Niemeijer et. al [54] and it is conformed with results from their proposed approach as well as two experts. For comparison, Observer 1 was determined as gold standard and used as reference to calculate errors. The complete comparison of AVR results for each image can be found on Appendix B.

	<b>Obs. 1</b>	Obs. 2		Niemeijer et al.		Proposed	
	<b>AVR</b>	AVR	Error	AVR	Error	AVR	Error
<b>Mean</b>	0.67	0.66	0.05	0.67	0.06	0.67	0.07
<b>STD</b>	0.08	0.08	0.05	0.07	0.04	0.09	0.04
<b>Min</b>	0.52	0.45	0	0.55	0.01	0.50	0
<b>Max</b>	0.93	0.85	0.29	0.81	0.15	0.87	0.15

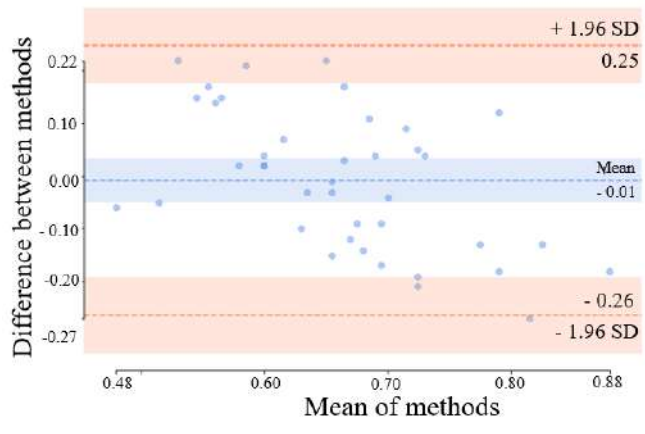
Table 6.5: Comparison with the state of the art for AVR calculation on the INSPIRE-AVR dataset.

To evaluate the proposed system, a Mann-Whitney U test was implemented under the null hypothesis  $H_0$  that both distributions of populations, Observer 1 and proposed, are equal (p-value = 0.488). The same applies using Observer 2 (p-value = 0.473) and with Niemeijer's method (p-value = 0.338).

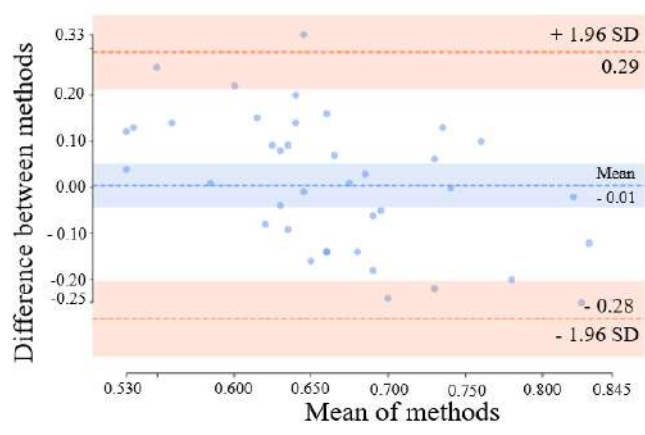
Furthermore, the results obtained were compared against Observer 1, Observer 2, and Niemeijer's method through Bland-Altman plots [63, 64], as shown in Fig. 6.5. Those graphs, unlike correlation, measure the agreement between data and allow visual inspection of the difference between methods against its mean.



(a)



(b)



(c)

Figure 6.5: Bland-Altman plots on the INSPIRE-AVR dataset for proposed results against (a) Observer 1, (b) Observer 2, and (c) Niemeijer's method.

The main limitation in the dataset INDEREBv1 is that experts only gave A/V labels to be considered as ground truth, there is no AVR reference to compare with. However, both previous Mann-Whitney U test and Bland-Altman plots have demonstrated that the proposed system is reliable across different datasets.

The AVR results on INDEREBv1 are summarized in Table 6.6, however, the complete results for each fundus image can be found on Appendix C.

	Mean	STD	Min	Max
AVR	0.66	0.13	0.45	0.91

Table 6.6: Summarized results of AVR calculation on the INDEREBv1 dataset.

The newest version of the INDEREB dataset does have an AVR gold standard, which was measured using Caliper, an incorporated tool in the ZEISS CLARUS 700 camera used to capture the fundus images. The results obtained with the proposed system are compared against them, as can be seen in Table 6.7. The complete results for each image can be seen on Appendix D.

	AVR		
	Expert	Proposed	Error
Mean	0.58	0.59	0.1
Std dev	0.08	0.12	0.09
Min	0.40	0.39	0
Max	0.75	0.86	0.33

Table 6.7: Comparison of AVR results on the INDEREBv2 dataset against expert.

Additionally, the results obtained were compared through a Bland-Altman plot, to visually assess the data agreement against its mean. This can be seen in Figure 6.6.

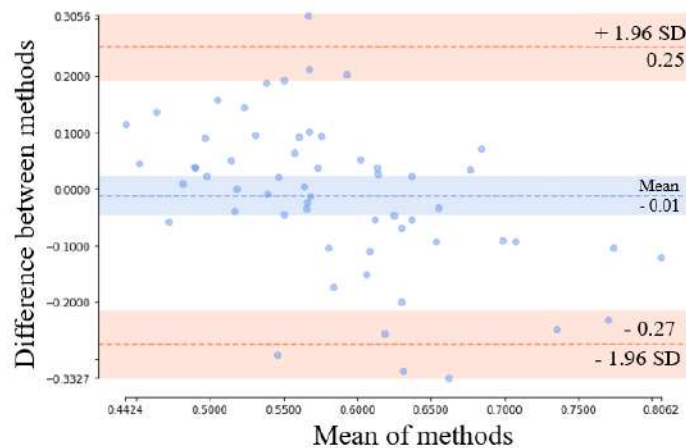


Figure 6.6: Bland-Altman plot on the INDEREBv2 dataset for results with proposed system against Expert.

## 6.4 Early Hypertensive detection through AVR

Several works determined an AVR of 0.66 up to 0.7 corresponds to a healthy fundus, while anything below that metric is a sign of Hypertensive Retinopathy. This assumption is based on the first sign of the disease, where retinal blood vessels narrow. However, there is not a scientific study that supports that asseveration.

Since public datasets specially made for AVR purposes usually only have one or two experts delivering an AVR measurement for each image, there is no information regarding which one corresponds to a patient that suffers from Hypertensive Retinopathy. Hence, there is not enough empirical evidence that supports establishing a 0.66 AVR value as a rule of thumb.

This thesis had its main purpose of furthering the investigation by creating the dataset INDEREBv2, where a diagnostic made by retina specialists is available along with its corresponding AVR. The first step was to demonstrate it is not possible to establish a strong division between healthy and unhealthy eyes by only a metric without considering more factors such as *Gender* and *Age*.

Figure 6.7 shows how the AVR is distributed along with the diagnosis of the disease. It is obvious that 0.66 does not separate both classes, having the disease and not, since there are fundi with a lower AVR and yet are considered healthy. In contrast, fundi with a higher AVR value can also have the disease.

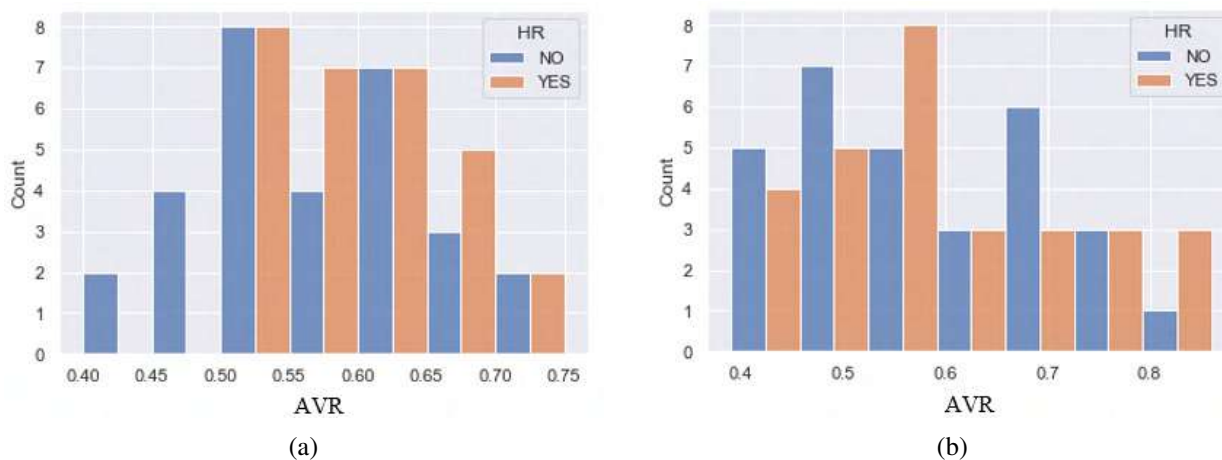


Figure 6.7: AVR distribution with HR diagnosis for (a) Expert results and (b) Proposed system results.

The next step was to analyze the AVR values regarding *Gender*, as can be seen in Figure 6.8. However, the problem remains since there is no clear separation between classes, neither for women nor men. Nevertheless, it is important to note that both distributions are notably different. Thus, using only one factor instead of several is not enough to decide how the AVR measurements are related to early hypertensive detection.

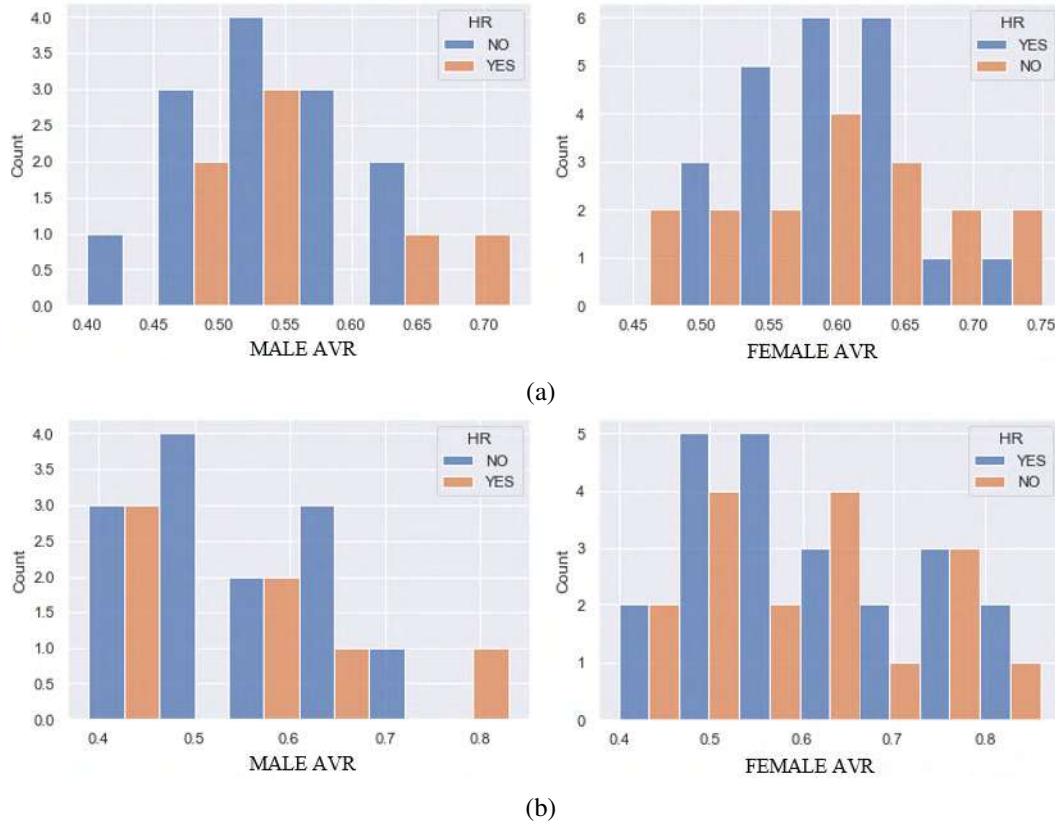


Figure 6.8: AVR distribution regarding Gender with HR diagnosis for (a) Expert results and (b) Proposed system results.

Then, a 3D scatter plot was created using *Age*, *Gender*, *AVR*, and the HR diagnosis to evaluate all available factors, as shown in Figure 6.9. It seems that older males are more easily differentiated than those with the disease. While something similar happens with older females, there is still a problematic area between those in their 30-40s.

Finally, to leverage AI and its applications, several Machine Learning techniques were implemented to the INDEREBv2 data to find the best classifier to predict early Hypertensive Retinopathy with available attributes: *AVR*, *Age*, and *Gender*. Since the diagnosis label is known, metrics of the results can be compared.

The library used for this experimentation is PyCaret [65], an open-source that automates ML workflows. It has a module specialized in Classifications for binary or multiclass problems with over 18 ready-to-use algorithms. It has a useful function named Compare Models, which trains and evaluates the performance of all methods available using cross-validation. As shown in Table 6.8, only the best five algorithms will be reported for this investigation.

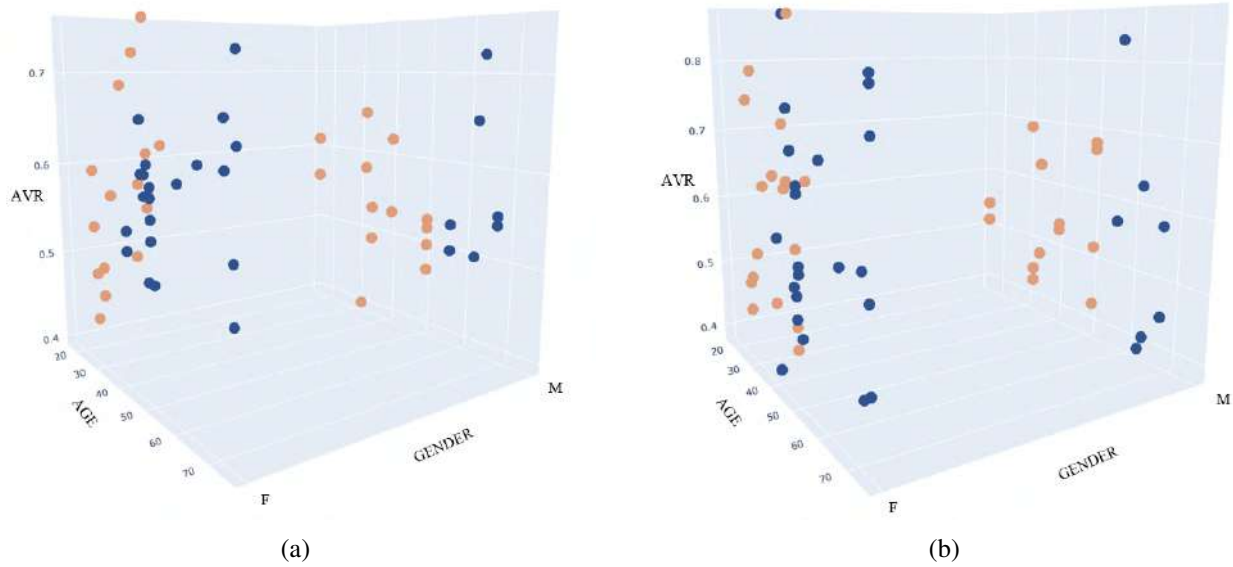


Figure 6.9: AVR distribution regarding Age and Gender with HR diagnosis for (a) Expert results and (b) Proposed system results.

<i>Model</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Accuracy</i>
Random Forest Classifier	91.67 %	92.50 %	88 %
Ridge Classifier	90 %	89.33 %	86 %
Gradient Boosting Classifier	90 %	85.83 %	85.50 %
Extra Trees Classifier	91.67 %	89.17 %	85.50%
Logistic Regression	90 %	84.33 %	83.50 %

Table 6.8: Comparison of ML models metrics for Early Hypertensive Retinopathy detection.

It has been proven that using only AVR measurements is not enough to determine whether a patient has or not early hypertensive retinopathy. However, a more accurate classifier can be developed when using multiple factors. Random Forest is the best method with this data, delivering good results but still having less than 90% accuracy. Maybe by increasing the number of features in the model, such as blood pressure and genetic predisposition, the prediction's accuracy may improve.

---

## Conclusion

This investigation proposed a comprehensive methodology that successfully retrieves veins and arteries from fundus images to compute a reliable and accurate AVR, since it produces results no statistically different from the experts. Furthermore, it demonstrated that the AVR by itself is not sufficient to grant an automatic early hypertensive retinopathy detection. However, by adding the *Age* and *Gender* variables, a Random Forest Classifier detected the early stages of the disease with a 88% accuracy.

This work created a new dataset with 59 fundus images centered at the OD and its corresponding A/V classification, AVR measurement, and a diagnosis, all made by experts. Likewise, a web application was developed to facilitate users' non-familiar with the whole process of obtaining an AVR measure from a fundus image.

It is essential to point out that a standard method for AVR computation needs to be established for reliability and repeatability prognosis in different datasets. In this thesis, the work proposed by Neijmar et al. was used as a golden standard. However, several works still determine where the vessels are measured almost arbitrarily, thus complicating future analysis and comparative improvements under the same methodology.

The proposed approach first enhances the fundus images with CLAHE. It then leverages deep learning capabilities by combining A/V classification, through semantic segmentation, with OD segmentation results using a U-Net. Finally, it automatically determines the ROI and measures the ratio between the arterioles and venules from 1-1.5 DD.

The most critical step while correctly calculating the AVR is having an accurate A/V classification as an initial point since the width of predicted vessels needs to be measured. The analysis made in this investigation has shown that the semantic segmentation can be significantly improved by using the Focal Tversky loss.

In conclusion, it has been proven that the proposed method could aid ophthalmologists in delivering consistent early hypertensive retinopathy diagnoses through a compilation of variables such as Age, Gender, and AVR measurements, but further investigation is needed.



# Python Libraries

```
keras==2.8.0
Keras-Preprocessing==1.1.2
keras-vis==0.4.1
matplotlib==3.2.2
numpy==1.21.6
nvidia-ml-py3==7.352.0
opencv-contrib-python==4.1.2.30
opencv-python==4.1.2.30
pandas==1.3.5
Pillow==7.1.2
plotly==5.5.0
PyCaret==2.3.6
scikit-image==0.18.3
scikit-learn==1.0.2
scipy==1.4.1
seaborn==0.11.2
sklearn-pandas==1.8.0
statsmodels==0.10.2
streamlit==1.9.0
tensorboard==2.8.0
tensorflow==2.8.0
```

---

## Complete results on INSPIRE-AVR dataset

Image	Exp. 1	Exp. 2		Niemeijer et al.		Proposed	
	AVR	AVR	Error	AVR	Error	AVR	Error
image001	0.7	0.71	0.01	0.62	0.08	0.8	0.1
image002	0.76	0.75	0.01	0.81	0.05	0.71	0.05
image003	0.66	0.69	0.03	0.81	0.15	0.58	0.08
image004	0.75	0.75	0	0.76	0.01	0.70	0.05
image005	0.53	0.61	0.08	0.59	0.06	0.63	0.1
image006	0.93	0.76	0.17	0.8	0.13	0.87	0.06
image007	0.63	0.68	0.05	0.7	0.07	0.75	0.12
image008	0.7	0.65	0.05	0.59	0.11	0.6	0.1
image009	0.65	0.64	0.01	0.68	0.03	0.5	0.15
image010	0.78	0.75	0.03	0.74	0.04	0.68	0.1
image011	0.65	0.65	0	0.58	0.07	0.66	0.01
image012	0.67	0.65	0.02	0.74	0.07	0.74	0.07
image013	0.64	0.71	0.07	0.68	0.04	0.67	0.03
image014	0.69	0.76	0.07	0.77	0.08	0.79	0.1
image015	0.56	0.85	0.29	0.57	0.01	0.53	0.03
image016	0.64	0.74	0.1	0.7	0.06	0.63	0.01
image017	0.57	0.62	0.05	0.6	0.03	0.57	0
image018	0.62	0.58	0.04	0.59	0.03	0.73	0.11
image019	0.64	0.61	0.03	0.67	0.03	0.59	0.05
image020	0.68	0.68	0	0.67	0.01	0.72	0.04
image021	0.52	0.45	0.07	0.55	0.03	0.52	0
image022	0.62	0.63	0.01	0.58	0.04	0.72	0.1
image023	0.67	0.63	0.04	0.71	0.04	0.59	0.08
image024	0.71	0.62	0.09	0.67	0.04	0.6	0.11
image025	0.57	0.58	0.01	0.59	0.02	0.68	0.11
image026	0.72	0.76	0.04	0.74	0.02	0.64	0.08
image027	0.65	0.64	0.01	0.63	0.02	0.59	0.06
image028	0.56	0.49	0.07	0.69	0.13	0.54	0.02
image029	0.73	0.61	0.12	0.6	0.13	0.78	0.05
image030	0.64	0.63	0.01	0.66	0.02	0.72	0.08
image031	0.63	0.68	0.05	0.61	0.02	0.65	0.02
image032	0.72	0.7	0.02	0.68	0.04	0.8	0.08
image033	0.59	0.61	0.02	0.61	0.02	0.7	0.11
image034	0.61	0.59	0.02	0.71	0.1	0.57	0.04
image035	0.65	0.61	0.04	0.68	0.03	0.59	0.06
image036	0.74	0.64	0.1	0.59	0.15	0.7	0.04
image037	0.69	0.62	0.07	0.64	0.05	0.65	0.04
image038	0.82	0.79	0.03	0.72	0.1	0.87	0.05
image039	0.61	0.64	0.03	0.7	0.09	0.67	0.06
image040	0.74	0.62	0.12	0.81	0.07	0.83	0.09

---

## Complete results on INDEREBv1 dataset

Image	AVR System	Image	AVR System
image001	0.68	⋮	⋮
image002	0.69	image024	0.82
image003	0.71	image025	0.55
image004	0.54	image026	0.52
image005	0.89	image027	0.73
image006	0.85	image028	0.68
image007	0.90	image029	0.55
image008	0.47	image030	0.57
image009	0.52	image031	0.75
image010	0.91	image032	0.48
image011	0.90	image033	0.47
image012	0.70	image034	0.53
image013	0.77	image035	0.50
image014	0.69	image036	0.60
image015	0.55	image037	0.56
image016	0.60	image038	0.55
image017	0.75	image039	0.64
image018	0.79	image040	0.70
image019	0.48	image041	0.68
image020	0.79	image042	0.81
image021	0.58	image043	0.68
image022	0.85	image044	0.45
image023	0.68	image045	0.55
⋮	⋮	image046	0.63

## Complete results on INDEREBv2 dataset

Image	Gender	Age	SH	AVR		
				Expert	Proposed	Error
image001	M	32	NO	0.40	0.69	0.29
image002	M	51	NO	0.53	0.68	0.15
image003	M	51	NO	0.52	0.52	0.00
image004	M	64	YES	0.72	0.41	0.31
image005	F	50	YES	0.60	0.65	0.05
image006	F	50	YES	0.55	0.66	0.11
image007	M	63	YES	0.50	0.39	0.11
image008	M	63	YES	0.65	0.63	0.02
image009	F	75	YES	0.72	0.80	0.08
image010	F	75	YES	0.65	0.79	0.14
image011	F	65	YES	0.63	0.59	0.04
image012	F	48	YES	0.61	0.70	0.09
image013	F	48	YES	0.59	0.53	0.06
image014	F	47	YES	0.66	0.75	0.09
image015	F	47	YES	0.61	0.86	0.25
image016	F	49	NO	0.58	0.48	0.10
image017	F	49	NO	0.63	0.45	0.19
image018	F	31	NO	0.50	0.54	0.04
image019	F	31	NO	0.47	0.79	0.32
image020	M	33	NO	0.58	0.43	0.16
image021	M	33	NO	0.65	0.45	0.19
image022	M	35	NO	0.53	0.63	0.10
image023	M	35	NO	0.49	0.48	0.01
image024	F	28	NO	0.44	0.50	0.06
image025	F	28	NO	0.49	0.75	0.26
image026	M	17	NO	0.56	0.54	0.02
image027	M	17	NO	0.61	0.51	0.09
image028	M	41	NO	0.53	0.54	0.01
image029	M	41	NO	0.62	0.53	0.09
⋮	⋮	⋮				

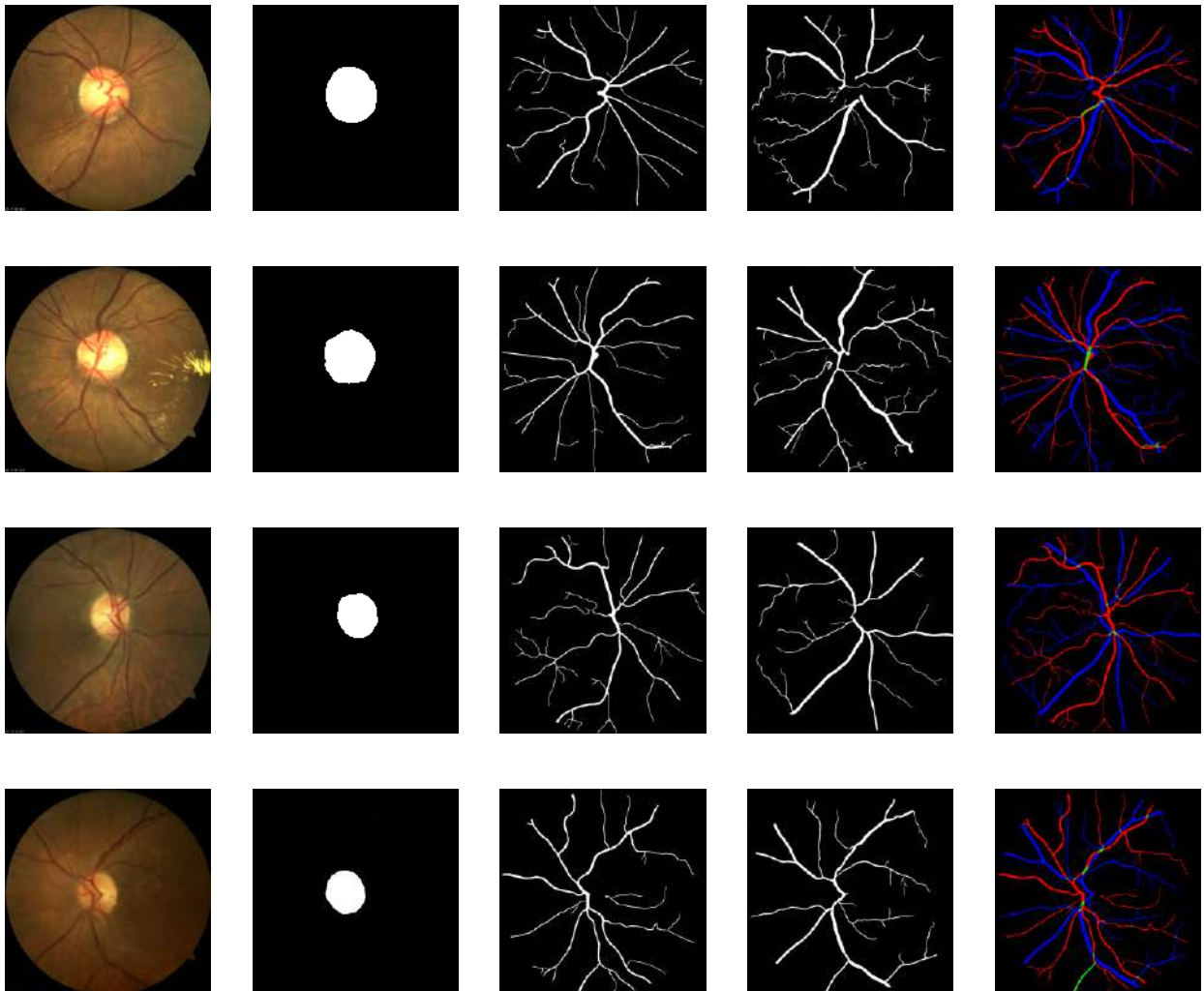
Image	Gender	Age	SH	AVR		
				Expert	Proposed	Error
⋮	⋮	⋮				
image030	F	54	NO	0.64	0.67	0.03
image031	F	40	NO	0.69	0.49	0.20
image032	F	40	NO	0.69	0.66	0.03
image033	F	46	NO	0.60	0.66	0.07
image034	M	51	NO	0.50	0.67	0.17
image035	M	51	NO	0.47	0.43	0.05
image036	F	45	NO	0.53	0.73	0.20
image037	F	45	NO	0.72	0.65	0.07
image038	F	75	YES	0.65	0.74	0.09
image039	F	50	YES	0.59	0.55	0.04
image040	F	50	YES	0.57	0.56	0.00
image041	F	74	YES	0.56	0.57	0.01
image042	F	74	YES	0.51	0.47	0.04
image043	F	49	YES	0.62	0.52	0.10
image044	F	49	YES	0.51	0.49	0.02
image045	F	49	NO	0.75	0.86	0.11
image046	F	49	NO	0.63	0.58	0.05
image047	F	27	NO	0.54	0.49	0.05
image048	F	27	NO	0.60	0.45	0.14
image049	F	72	YES	0.67	0.46	0.21
image050	F	72	YES	0.63	0.60	0.03
image051	F	59	YES	0.61	0.70	0.09
image052	F	41	YES	0.53	0.40	0.14
image053	F	41	YES	0.55	0.58	0.02
image054	M	68	YES	0.55	0.58	0.04
image055	M	68	YES	0.54	0.45	0.09
image056	F	35	NO	0.58	0.64	0.05
image057	F	51	YES	0.51	0.47	0.04
image058	M	57	YES	0.53	0.57	0.04
image059	M	57	YES	0.50	0.83	0.33

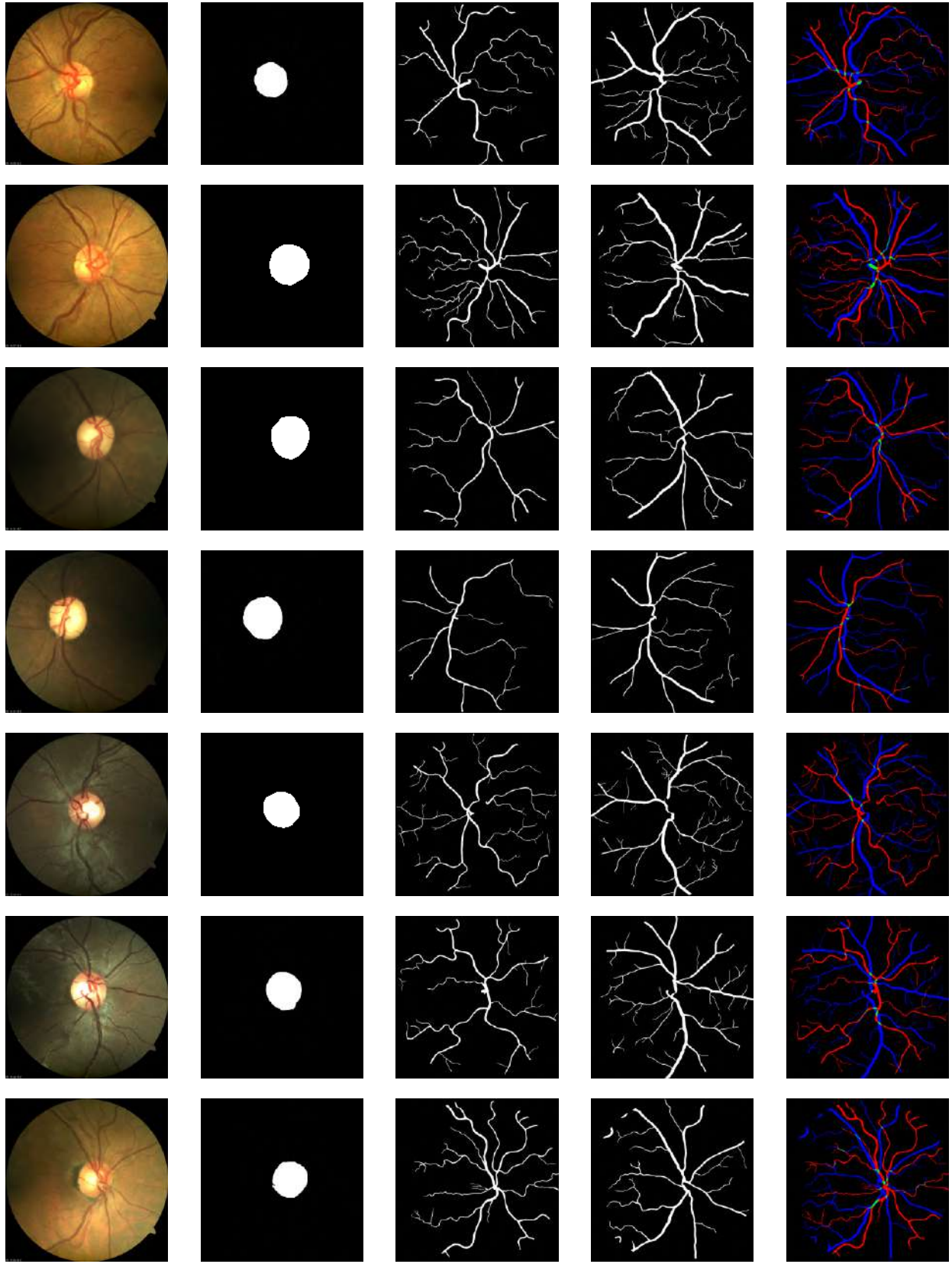
---

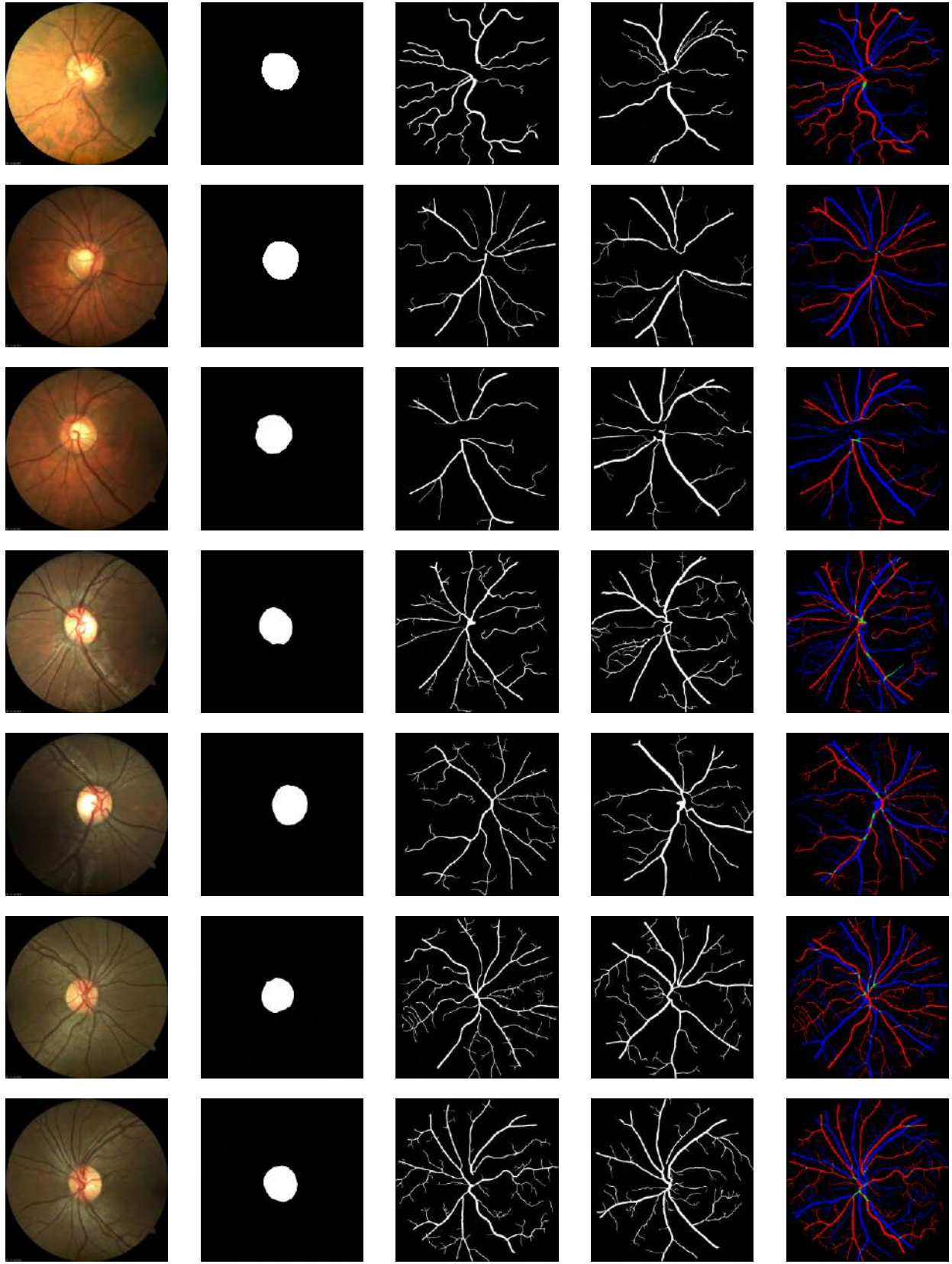
# INDEREB Database

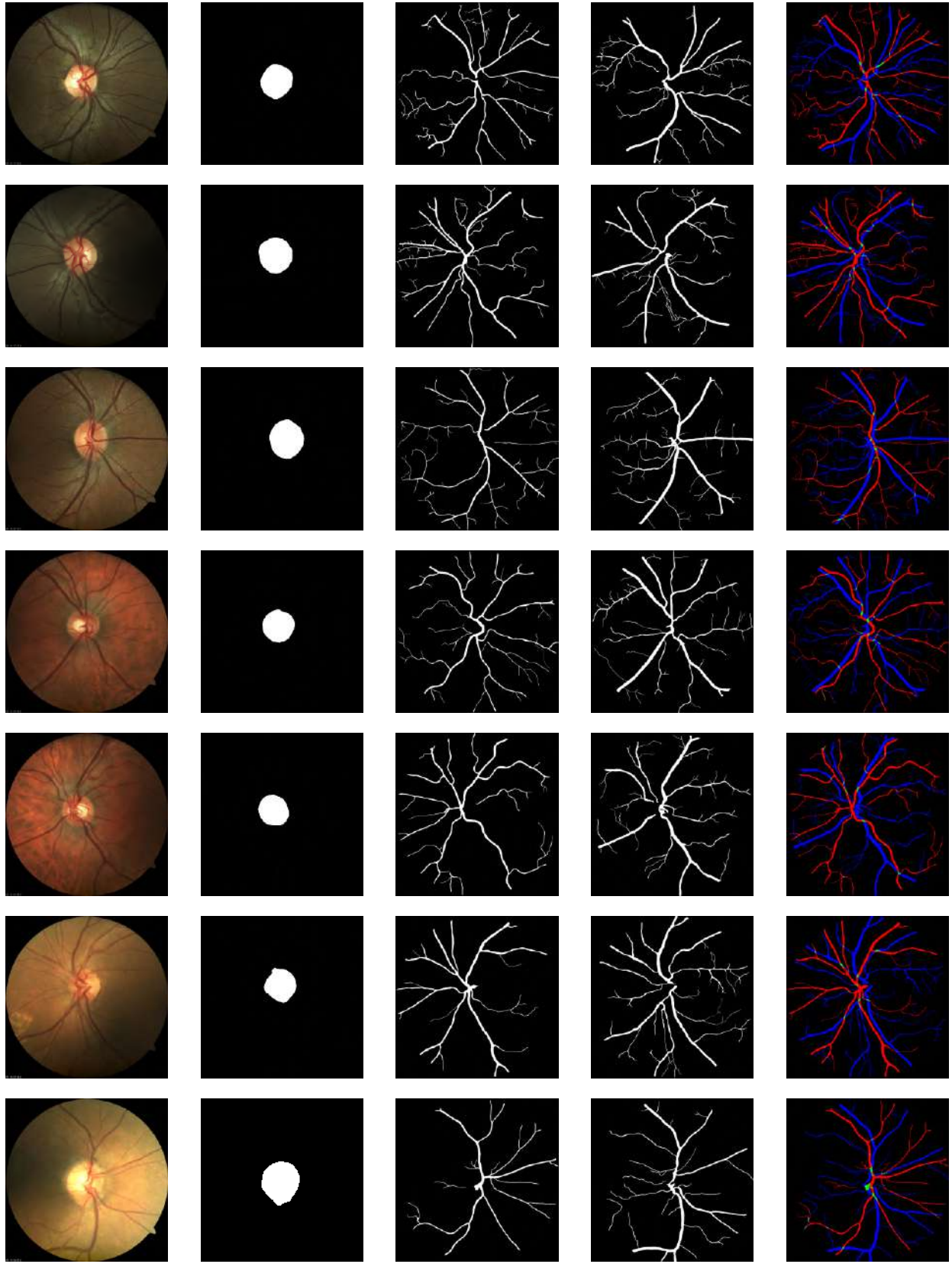
## E.1 Version 1

<https://github.com/mzamarron/INDEREB>

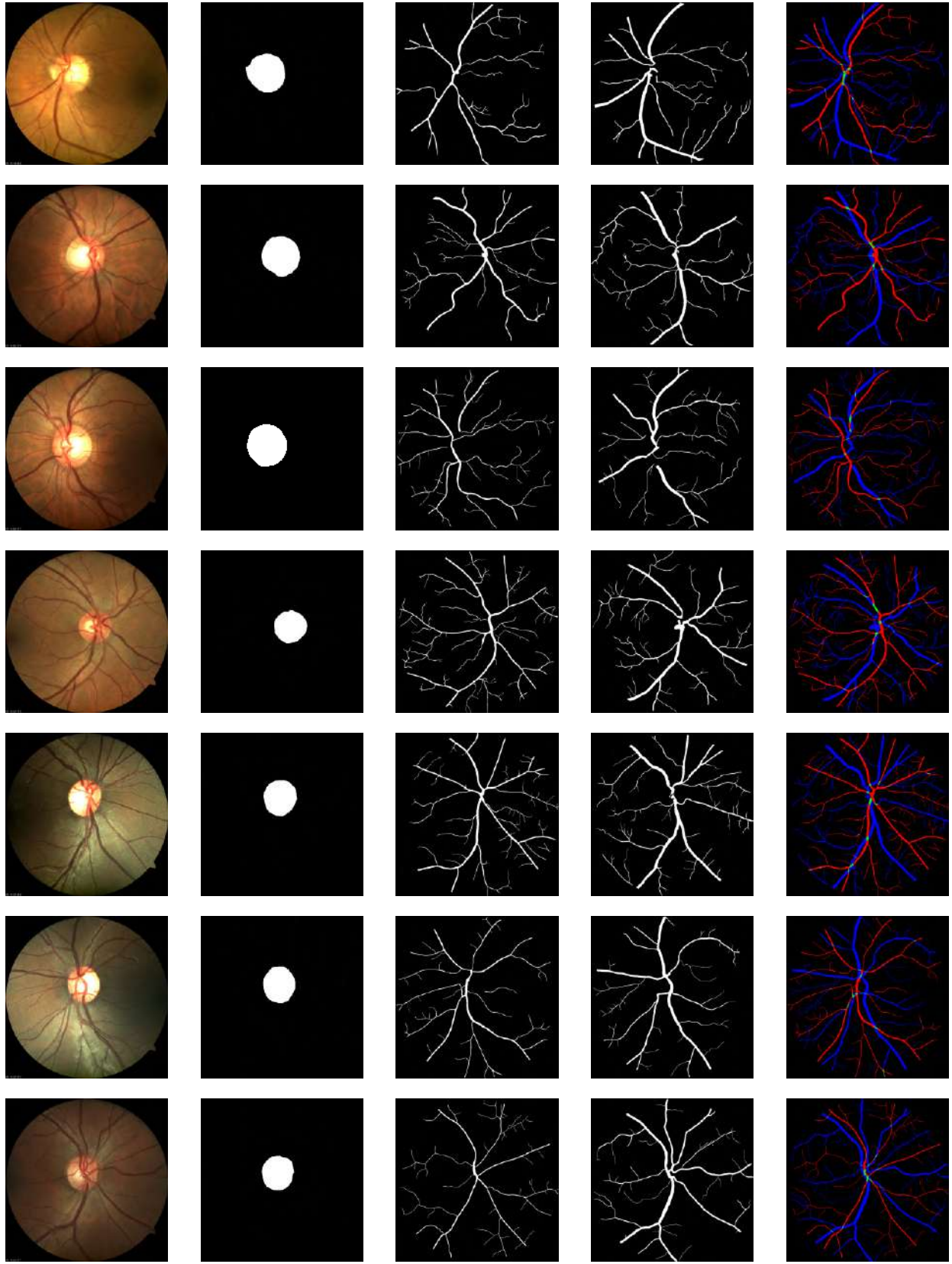


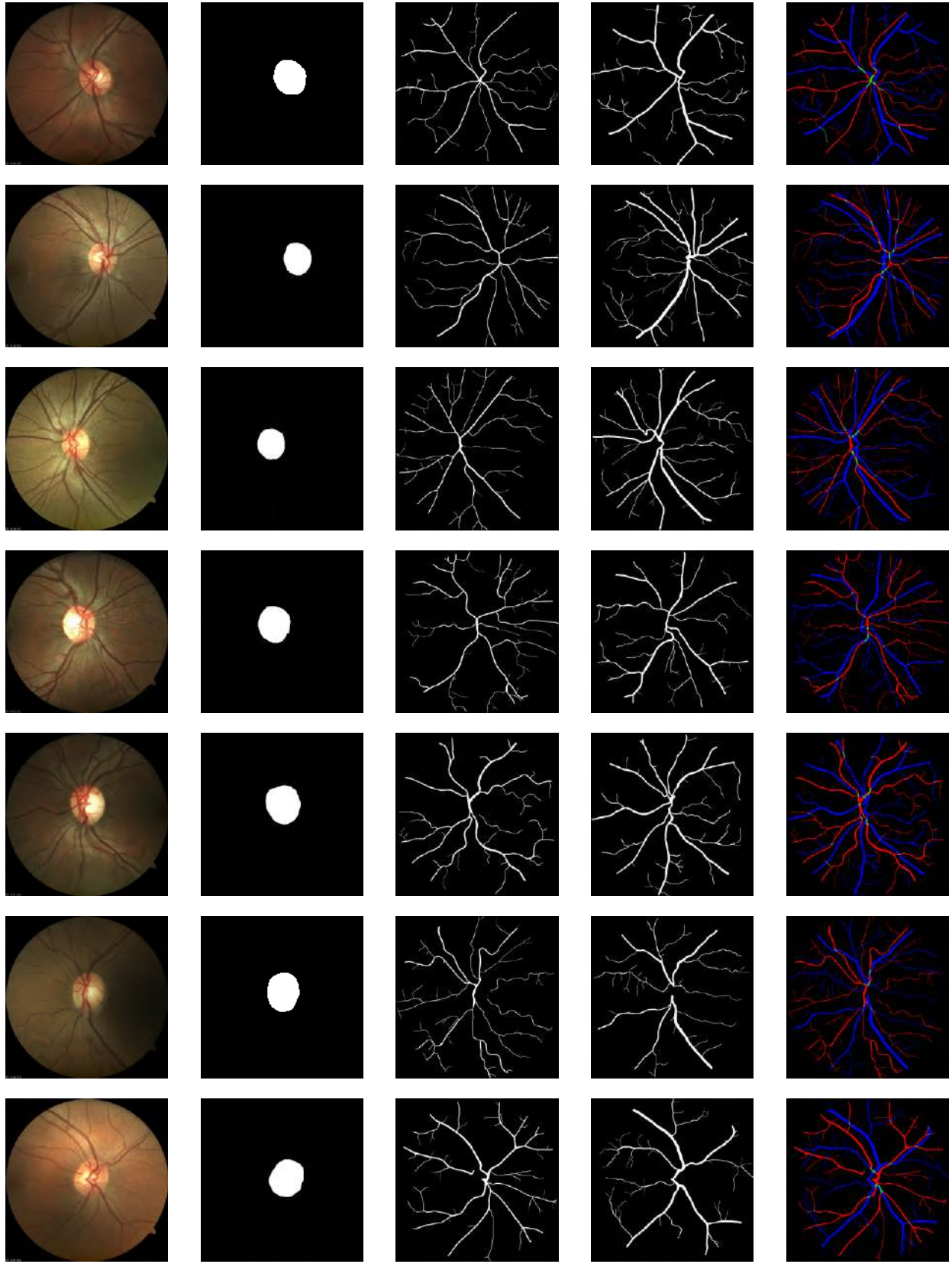


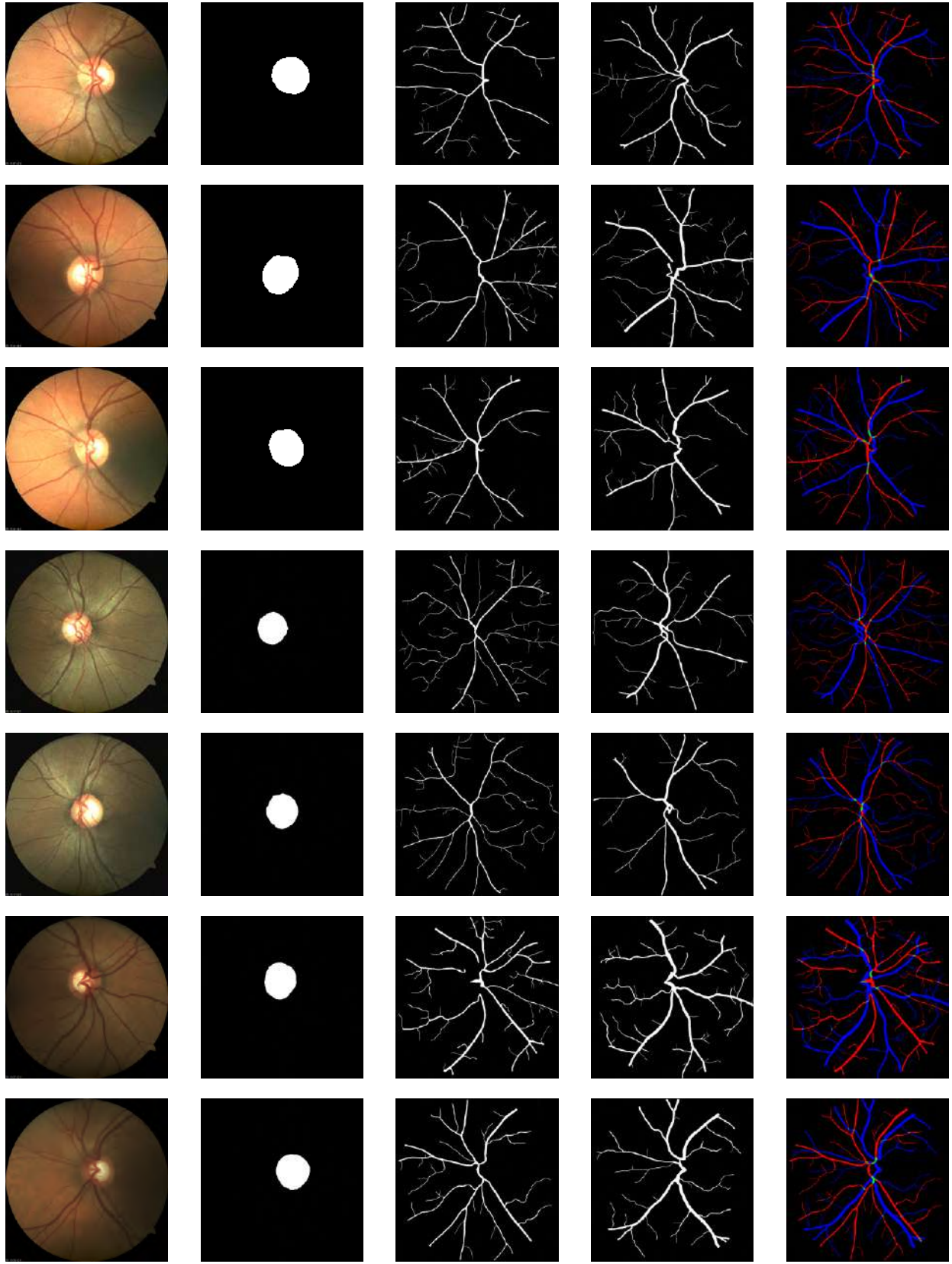






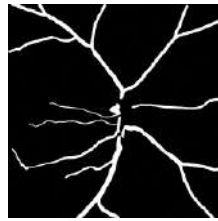
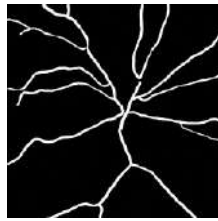
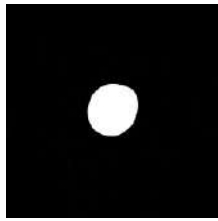






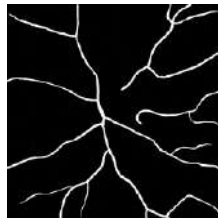
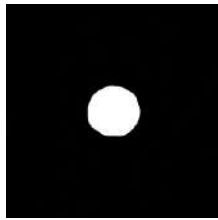
## E.2 Version 2

<https://github.com/mzamarron/INDEREBv2>



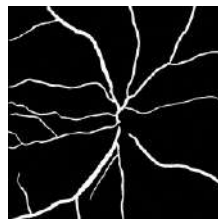
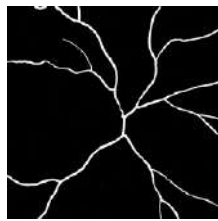
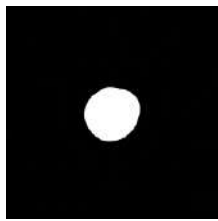
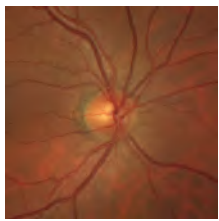
Systemic Hypertension : **No**

Expert AVR : **0.400**



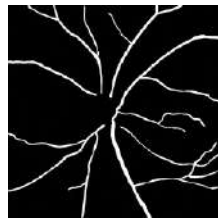
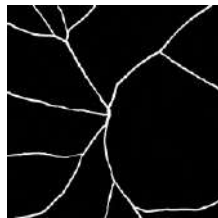
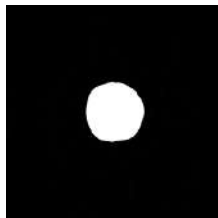
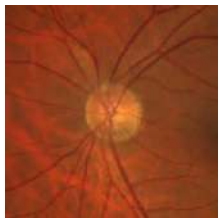
Systemic Hypertension : **No**

Expert AVR : **0.531**



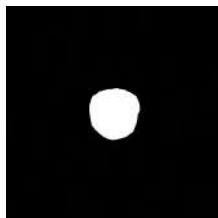
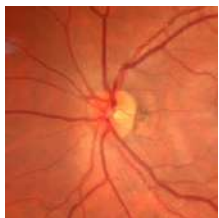
Systemic Hypertension : **No**

Expert AVR : **0.518**



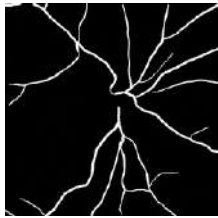
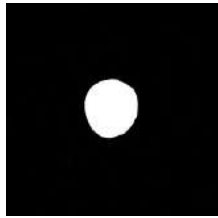
Systemic Hypertension : **Yes**

Expert AVR : **0.719**



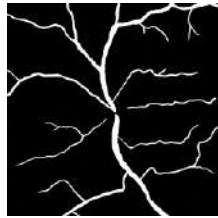
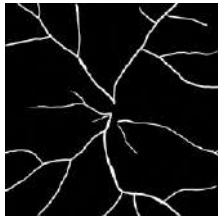
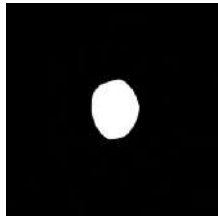
Systemic Hypertension : **Yes**

Expert AVR : **0.602**



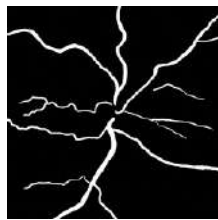
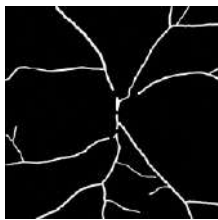
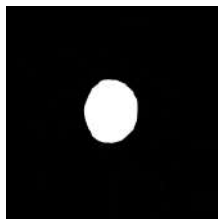
Systemic Hypertension : **Yes**

Expert AVR : **0.554**



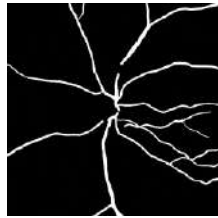
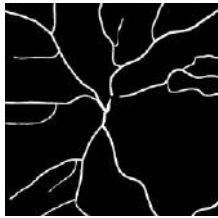
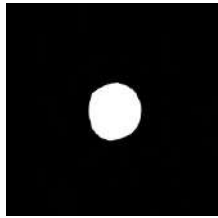
Systemic Hypertension : **Yes**

Expert AVR : **0.500**



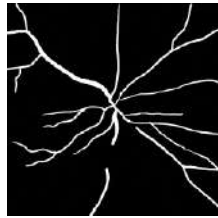
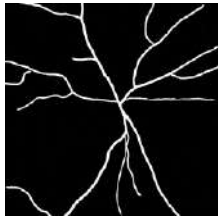
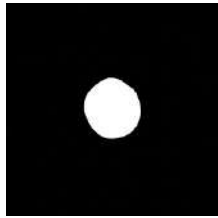
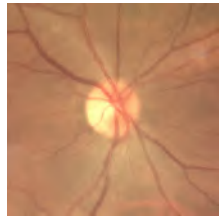
Systemic Hypertension : **Yes**

Expert AVR : **0.648**



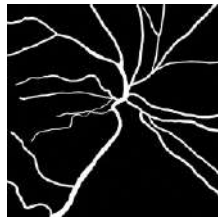
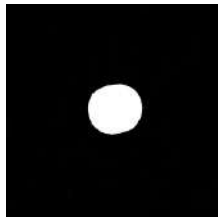
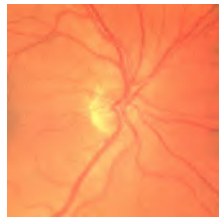
Systemic Hypertension : **Yes**

Expert AVR : **0.722**



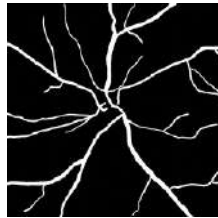
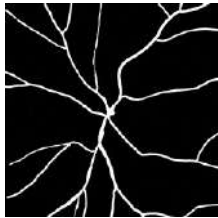
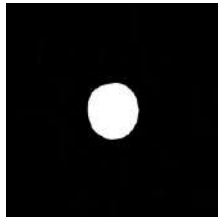
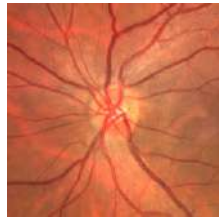
Systemic Hypertension : **Yes**

Expert AVR : **0.655**



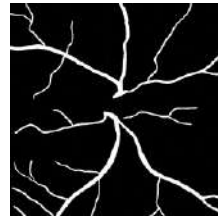
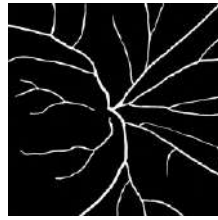
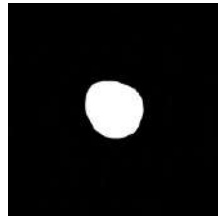
Systemic Hypertension : **Yes**

Expert AVR : **0.632**



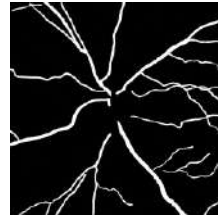
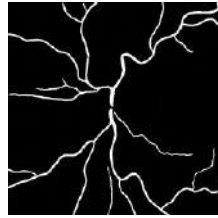
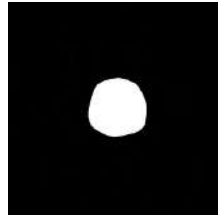
Systemic Hypertension : **Yes**

Expert AVR : **0.610**



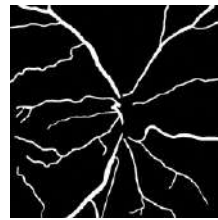
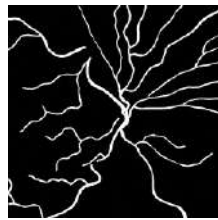
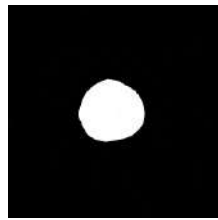
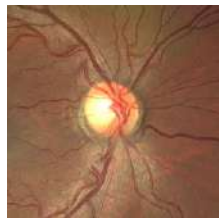
Systemic Hypertension : **Yes**

Expert AVR : **0.590**



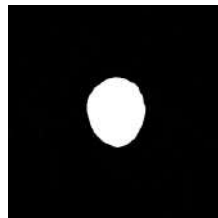
Systemic Hypertension : **Yes**

Expert AVR : **0.660**



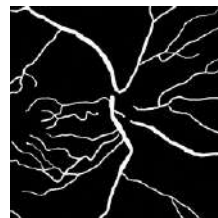
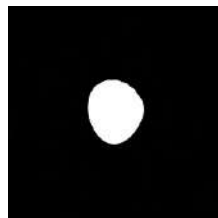
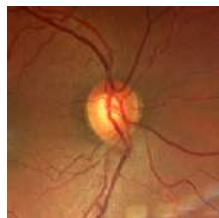
Systemic Hypertension : **Yes**

Expert AVR : **0.612**



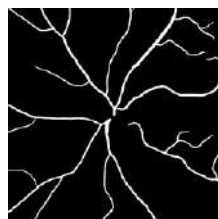
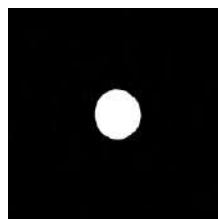
Systemic Hypertension : **No**

Expert AVR : **0.578**



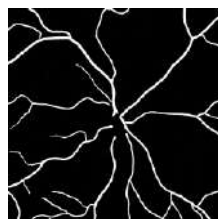
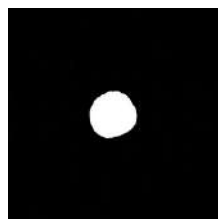
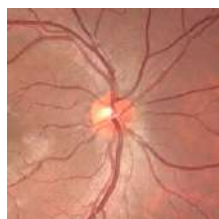
Systemic Hypertension : **No**

Expert AVR : **0.631**



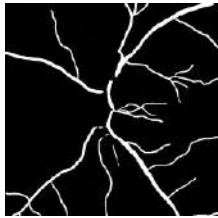
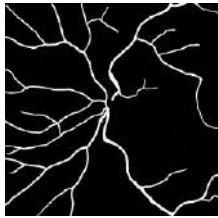
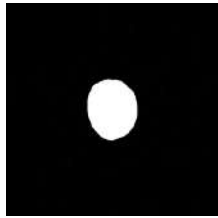
Systemic Hypertension : **No**

Expert AVR : **0.497**



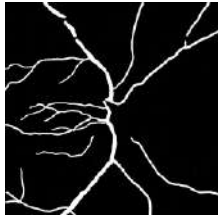
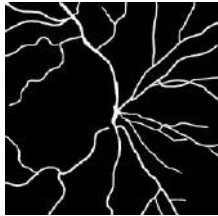
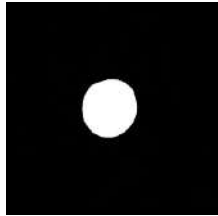
Systemic Hypertension : **No**

Expert AVR : **0.471**



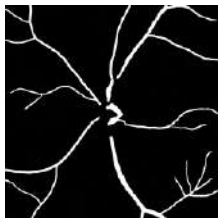
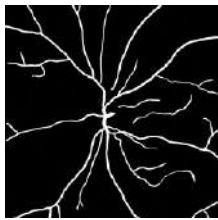
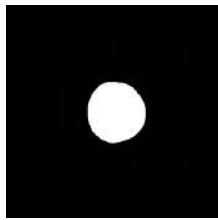
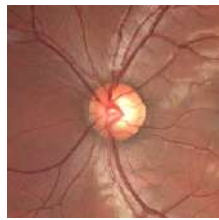
Systemic Hypertension : No

Expert AVR : 0.584



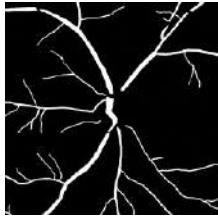
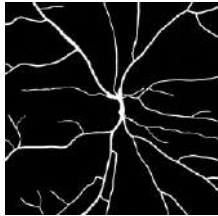
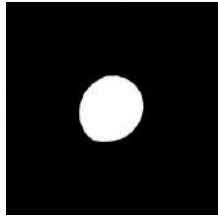
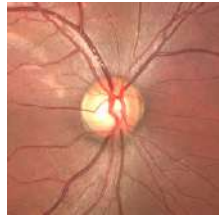
Systemic Hypertension : No

Expert AVR : 0.647



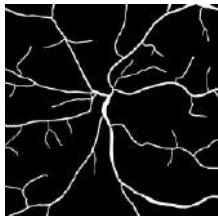
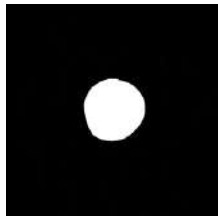
Systemic Hypertension : No

Expert AVR : 0.528



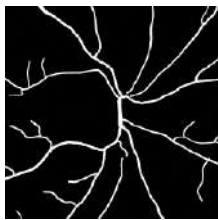
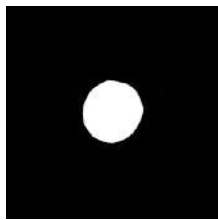
Systemic Hypertension : No

Expert AVR : 0.486



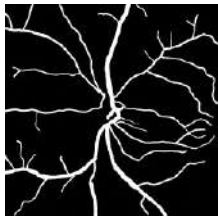
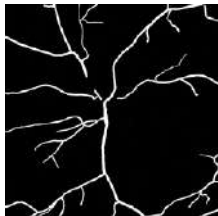
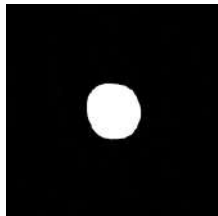
Systemic Hypertension : No

Expert AVR : 0.443



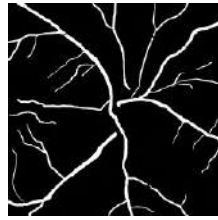
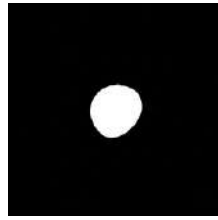
Systemic Hypertension : No

Expert AVR : 0.490



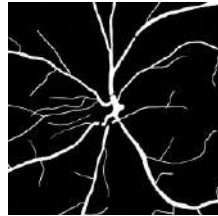
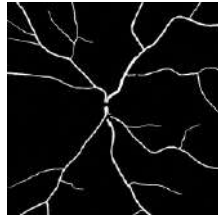
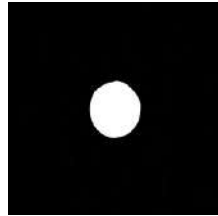
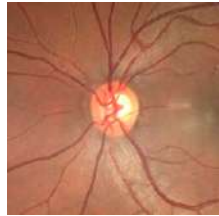
Systemic Hypertension : No

Expert AVR : 0.557



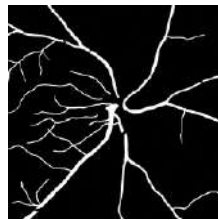
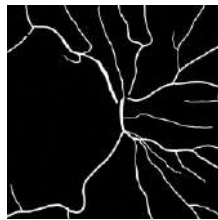
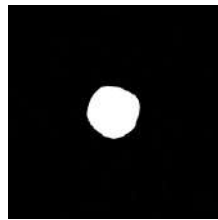
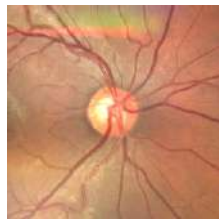
Systemic Hypertension : **No**

Expert AVR : **0.606**



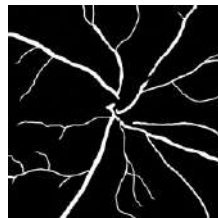
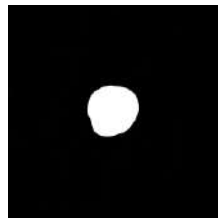
Systemic Hypertension : **No**

Expert AVR : **0.534**



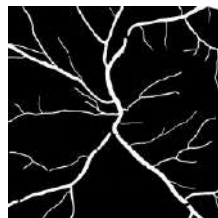
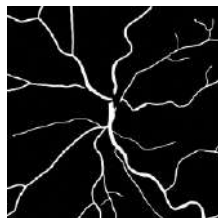
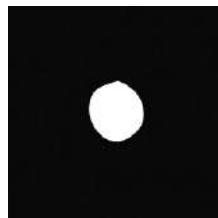
Systemic Hypertension : **No**

Expert AVR : **0.622**



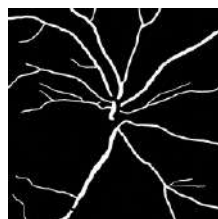
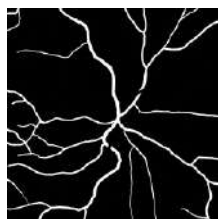
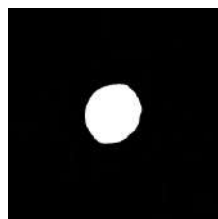
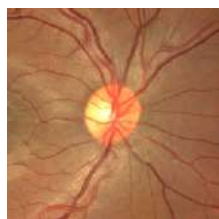
Systemic Hypertension : **No**

Expert AVR : **0.638**



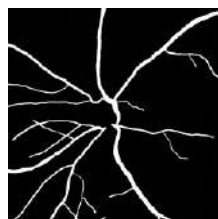
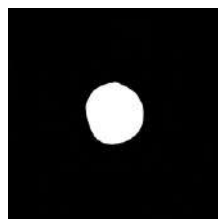
Systemic Hypertension : **No**

Expert AVR : **0.694**



Systemic Hypertension : **No**


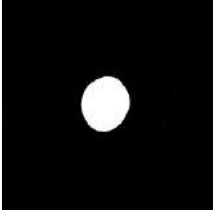

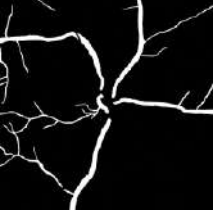

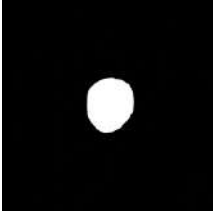

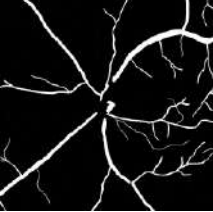

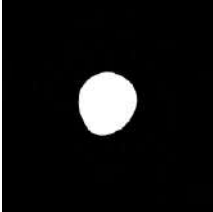
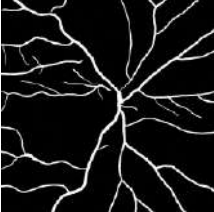

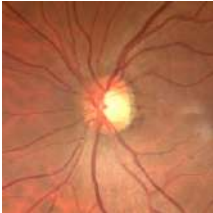
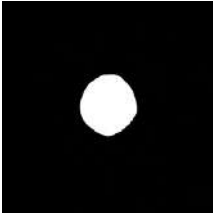

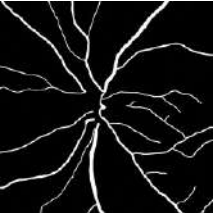
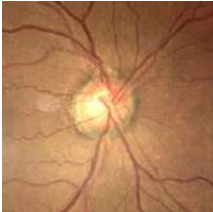
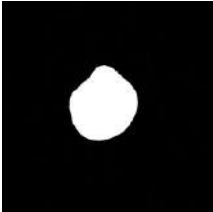

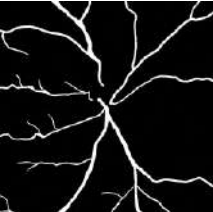
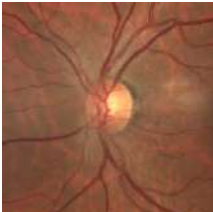
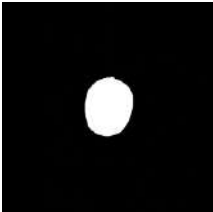



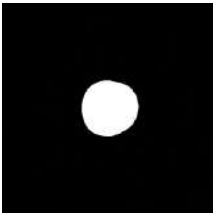

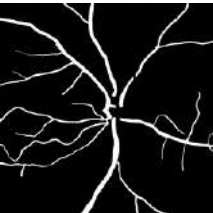
Expert AVR : **0.694**

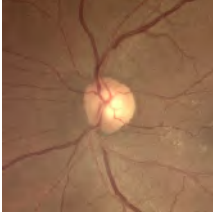
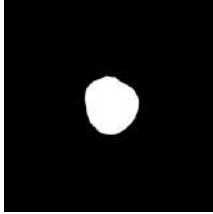
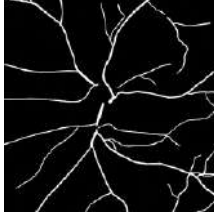


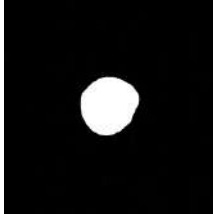
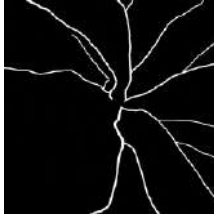


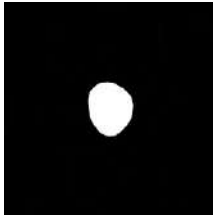
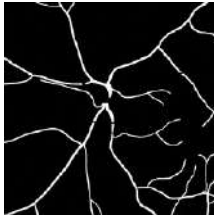


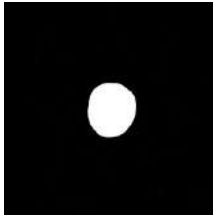
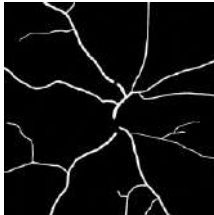
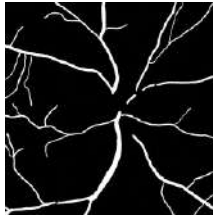

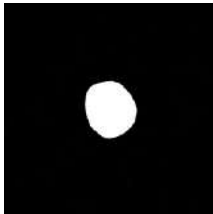
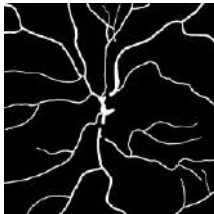


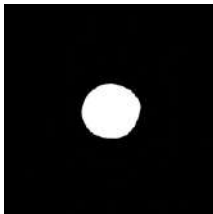
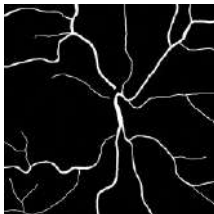
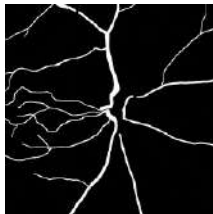

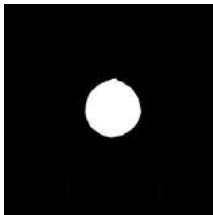
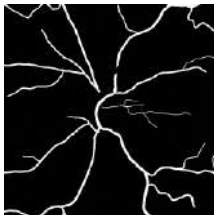




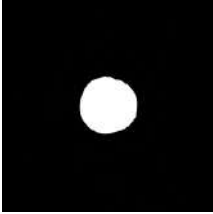
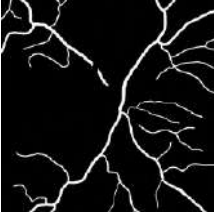
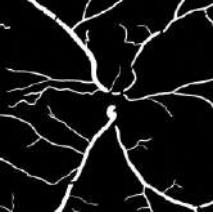
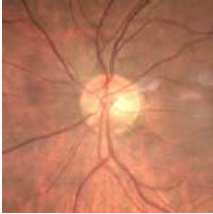
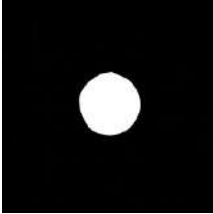


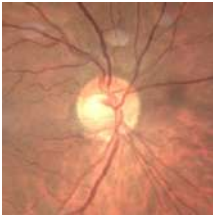
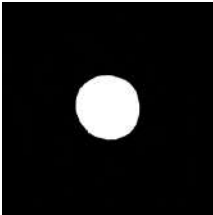
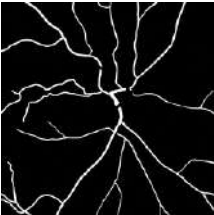


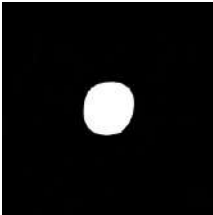

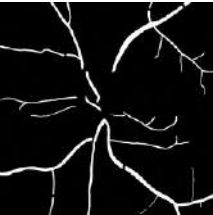

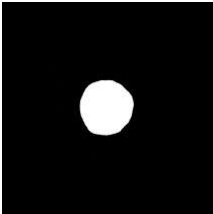
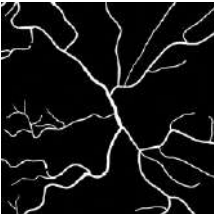
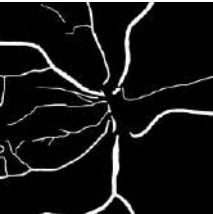

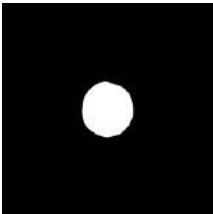



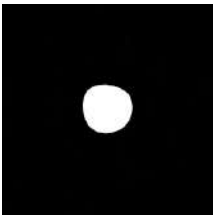

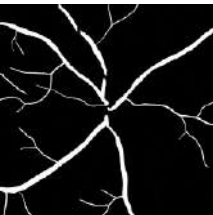
Systemic Hypertension : **No**

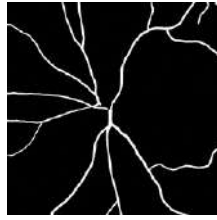
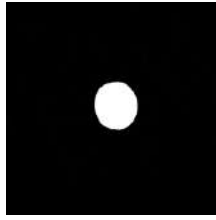
Expert AVR : **0.600**



				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.497</b>
				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.475</b>
				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.530</b>
				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.720</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.654</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.591</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.565</b>

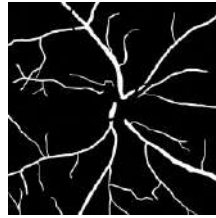
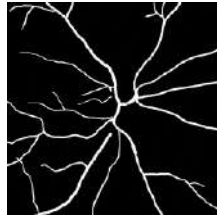
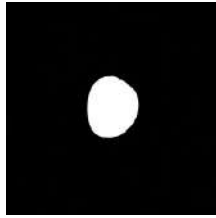
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.561</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.509</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.618</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.509</b>
				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.746</b>
				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.628</b>
				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.540</b>

				Systemic Hypertension : <b>No</b> Expert AVR : <b>0.595</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.673</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.627</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.607</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.532</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.554</b>
				Systemic Hypertension : <b>Yes</b> Expert AVR : <b>0.547</b>



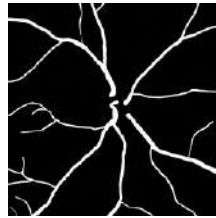
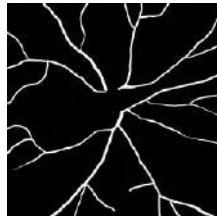
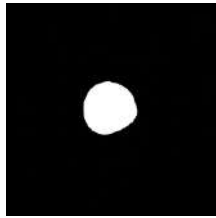
Systemic Hypertension : **Yes**

Expert AVR : **0.542**



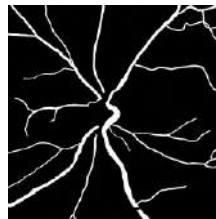
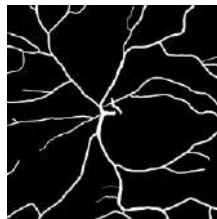
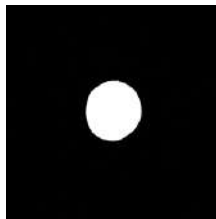
Systemic Hypertension : **No**

Expert AVR : **0.585**



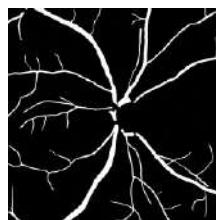
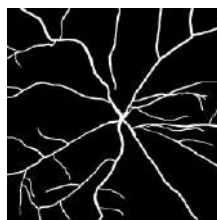
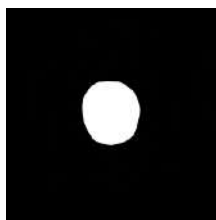
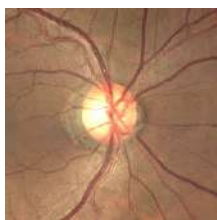
Systemic Hypertension : **Yes**

Expert AVR : **0.508**



Systemic Hypertension : **Yes**

Expert AVR : **0.528**



Systemic Hypertension : **Yes**

Expert AVR : **0.496**

---

# Bibliography

- [1] L. E. Downie, L. A. Hodgson, C. Dsylva, R. L. McIntosh, S. L. Rogers, P. Connell, and T. Y. Wong, “Hypertensive retinopathy: Comparing the Keith-Wagener-Barker to a simplified classification,” *Journal of Hypertension*, vol. 31, no. 5, pp. 960–965, 2013.
- [2] K. Suzuki, “Overview of deep learning in medical imaging,” *Radiological Physics and Technology*, no. June, 2017.
- [3] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [4] P. Hamet and J. Tremblay, “Artificial intelligence in medicine,” *Metabolism*, vol. 69, pp. S36–S40, 2017.
- [5] A. Ramesh, C. Kambhampati, J. R. Monson, and P. Drew, “Artificial intelligence in medicine.” *Annals of the Royal College of Surgeons of England*, vol. 86, no. 5, p. 334, 2004.
- [6] Y. Mintz and R. Brodie, “Introduction to artificial intelligence in medicine,” *Minimally Invasive Therapy & Allied Technologies*, vol. 28, no. 2, pp. 73–81, 2019.
- [7] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, “Artificial intelligence in retina,” *Progress in Retinal and Eye Research*, vol. 67, no. May, pp. 1–29, 2018. [Online]. Available: <https://doi.org/10.1016/j.preteyeres.2018.07.004>
- [8] F. Huang, B. Dashtbozorg, and B. M. Romeny, “Artery/vein classification using reflection features in retina fundus images,” *Machine Vision and Applications*, vol. 29, no. 1, pp. 23–34, 2018.
- [9] T. T. Nguyen and T. Y. Wong, “Retinal vascular changes and diabetic retinopathy,” *Current diabetes reports*, vol. 9, no. 4, pp. 277–283, 2009.
- [10] M. Tsukikawa and A. W. Stacey, “A review of hypertensive retinopathy and chorioretinopathy,” *Clinical optometry*, vol. 12, p. 67, 2020.
- [11] Q. Abbas and M. E. Ibrahim, “DenseHyper: an automatic recognition system for detection of hypertensive retinopathy using dense features transform and deep-residual learning,” *Multimedia Tools and Applications*, 2020.
- [12] M. Kiruthika, T. Swapna, K. C. Santhosh, and K. Peeyush, “Artery and vein classification for hypertensive retinopathy,” in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE, 2019, pp. 244–248.

- [13] M. Bhargava, M. Ikram, and T. Y. Wong, “How does hypertension affect your eyes?” *Journal of human hypertension*, vol. 26, no. 2, pp. 71–83, 2012.
- [14] S. Akbar, M. U. Akram, M. Sharif, A. Tariq, and U. ullah Yasin, “Arteriovenous ratio and papilledema based hybrid decision support system for detection and grading of hypertensive retinopathy,” *Computer methods and programs in biomedicine*, vol. 154, pp. 123–141, 2018.
- [15] WHO, “A global brief on hypertension: silent killer, global public health crisis,” [http://apps.who.int/iris/bitstream/10665/79059/1/WHO\\_DCO\\_WHD\\_2013.2\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/79059/1/WHO_DCO_WHD_2013.2_eng.pdf), 2013, online; accessed 12 October 2020.
- [16] M. Bhargava, M. K. Ikram, and T. Y. Wong, “How does hypertension affect your eyes,” *Journal of Human Hypertension*, vol. 26, no. 2, pp. 71–83, 2012.
- [17] N. Keith, “Some different types of essential hypertension: their course and prognosis,” *Am J Med Sci*, vol. 268, pp. 336–345, 1974.
- [18] A. Harjasouliha, V. Raiji, and J. M. Garcia Gonzalez, “Review of hypertensive retinopathy,” *Disease-a-Month*, vol. 63, no. 3, pp. 63–69, 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.disamonth.2016.10.002>
- [19] S. G. Vázquez, N. Barreira, M. G. Penedo, and M. Rodríguez-Blanco, “Reliable monitoring system for arteriovenous ratio computation,” *Computerized Medical Imaging and Graphics*, vol. 37, no. 5-6, pp. 337–345, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.compmedimag.2013.10.001>
- [20] R. Hemelings, B. Elen, I. Stalmans, K. Van Keer, P. De Boever, and M. B. Blaschko, “Artery–vein segmentation in fundus images using a fully convolutional network,” *Computerized Medical Imaging and Graphics*, vol. 76, p. 101636, 2019.
- [21] F. Girard, C. Kavalec, and F. Cheriet, “Joint segmentation and classification of retinal arteries/veins from fundus images,” *Artificial intelligence in medicine*, vol. 94, pp. 96–109, 2019.
- [22] V. Hemminki, M. Kähönen, M. T. Tuomisto, V. Turjanmaa, and H. Uusitalo, “Determination of retinal blood vessel diameters and arteriovenous ratios in systemic hypertension: Comparison of different calculation formulae,” *Graefe’s Archive for Clinical and Experimental Ophthalmology*, vol. 245, no. 1, pp. 8–17, 2007.
- [23] S. G. Vázquez, N. Barreira, and M. G. Penedo, “Automatic Arteriovenous Ratio Computation :,” pp. 563–570, 2012.
- [24] R. Heitmar, A. A. Kalitzeos, and V. Panesar, “Comparison of two formulas used to calculate summarized retinal vessel calibers,” *Optometry and Vision Science*, vol. 92, no. 11, pp. 1085–1091, 2015.
- [25] T. Y. Wong, M. D. Knudtson, R. Klein, B. E. Klein, S. M. Meuer, and L. D. Hubbard, “Computer-assisted measurement of retinal vessel diameters in the beaver dam eye study: methodology, correlation between eyes, and effect of refractive errors,” *Ophthalmology*, vol. 111, no. 6, pp. 1183–1190, 2004.
- [26] M. Ortega, N. Barreira, J. Novo, M. G. Penedo, A. Pose-Reino, and F. Gómez-Ulla, “Sirius: a web-based system for retinal image analysis,” *International journal of medical informatics*, vol. 79, no. 10, pp. 722–732, 2010.

- [27] C. Y. Cheung, D. Xu, C.-Y. Cheng, C. Sabanayagam, Y.-C. Tham, M. Yu, T. H. Rim, C. Y. Chai, B. Gopinath, P. Mitchell *et al.*, “A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre,” *Nature biomedical engineering*, vol. 5, no. 6, pp. 498–508, 2021.
- [28] I. N. de Salud Pública, “Hipertensión arterial sistémica (has) diagnóstico, tratamiento y prevención,” *Boletín de Práctica Médica Efectiva*, 2006.
- [29] M. Tsukikawa and A. W. Stacey, “A review of hypertensive retinopathy and chorioretinopathy,” *Clinical Optometry*, vol. 12, pp. 67–73, 2020.
- [30] S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, and V. Lakshminarayanan, “Ophthalmic diagnosis using deep learning with fundus images – A critical review,” *Artificial Intelligence in Medicine*, vol. 102, no. July 2019, p. 101758, 2020. [Online]. Available: <https://doi.org/10.1016/j.artmed.2019.101758>
- [31] B. K. Triwijoyo, B. S. Sabarguna, W. Budiharto, and E. Abdurachman, *Deep learning approach for classification of eye diseases based on color fundus images*. Elsevier Inc., 2020. [Online]. Available: <http://dx.doi.org/10.1016/B978-0-12-817440-1.00002-4>
- [32] M. Haenlein and A. Kaplan, “A brief history of artificial intelligence: On the past, present, and future of artificial intelligence,” *California management review*, vol. 61, no. 4, pp. 5–14, 2019.
- [33] N. Muthukrishnan, F. Maleki, K. Ovens, C. Reinhold, B. Forghani, and R. Forghani, “Brief history of artificial intelligence,” *Neuroimaging Clinics*, vol. 30, no. 4, pp. 393–399, 2020.
- [34] A. M. Turing and J. Haugeland, “Computing machinery and intelligence,” *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*, pp. 29–56, 1950.
- [35] L. J. Catania and E. Nicolitz, “Artificial Intelligence and Its Applications in Vision and Eye Care,” *Advances in Ophthalmology and Optometry*, pp. 1–18, 2018. [Online]. Available: <https://doi.org/10.1016/j.yaoo.2018.04.001>
- [36] P. Ongsulee, “Artificial intelligence, machine learning and deep learning,” in *2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE)*. IEEE, 2017, pp. 1–6.
- [37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [38] L. Deng and D. Yu, “Deep learning: methods and applications,” *Foundations and trends in signal processing*, vol. 7, no. 3–4, pp. 197–387, 2014.
- [39] D. Gupta, “A beginner’s guide to deep learning based semantic segmentation using keras,” Jun 2019. [Online]. Available: <https://divamgupta.com/image-segmentation/2019/06/06/deep-learning-semantic-segmentation-keras.html>
- [40] K. O. Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” pp. 1–11, 2015.
- [41] X.-X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, “U-net-based medical image segmentation,” *Journal of Healthcare Engineering*, vol. 2022, 2022.

- [42] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, “U-net and its variants for medical image segmentation: A review of theory and applications,” *IEEE Access*, 2021.
- [43] A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, “Color retinal image enhancement using clahe,” in *International Conference on ICT for Smart Society*. IEEE, 2013, pp. 1–3.
- [44] M. Nixon and A. Aguado, *Feature extraction and image processing for computer vision*. Academic press, 2019.
- [45] M. Bertalmío, *Vision models for high dynamic range and wide colour gamut imaging: techniques and applications*. Academic Press, 2019.
- [46] B. Dashtbozorg, A. M. Mendonça, and A. Campilho, “An automatic graph-based approach for artery/vein classification in retinal images,” *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1073–1083, 2013.
- [47] J. Morano, Á. S. Hervella, J. Novo, and J. Rouco, “Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images,” *Artificial Intelligence in Medicine*, vol. 118, p. 102116, 2021.
- [48] C. L. Srinidhi, P. Aparna, and J. Rajan, “Automated method for retinal artery/vein separation via graph search metaheuristic approach,” *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 2705–2718, 2019.
- [49] X. Xu, T. Tan, and F. Xu, “An improved u-net architecture for simultaneous arteriole and venule segmentation in fundus image,” in *Annual Conference on Medical Image Understanding and Analysis*. Springer, 2018, pp. 333–340.
- [50] W. Ma, S. Yu, K. Ma, J. Wang, X. Ding, and Y. Zheng, “Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 769–778.
- [51] H. Kang, Y. Gao, S. Guo, X. Xu, T. Li, and K. Wang, “Avnet: A retinal artery/vein classification network with category-attention weighted fusion,” *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105629, 2020.
- [52] A. Galdran, M. Meyer, P. Costa, A. Campilho *et al.*, “Uncertainty-aware artery/vein classification on retinal images,” in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 556–560.
- [53] J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken, “Ridge based vessel segmentation in color images of the retina,” *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004.
- [54] M. Niemeijer, X. Xu, A. V. Dumitrescu, P. Gupta, B. Van Ginneken, J. C. Folk, and M. D. Abramoff, “Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs,” *IEEE Transactions on medical imaging*, vol. 30, no. 11, pp. 1941–1950, 2011.
- [55] Q. Hu, M. D. Abramoff, and M. K. Garvin, “Automated construction of arterial and venous trees in retinal images,” *Journal of Medical Imaging*, vol. 2, no. 4, p. 044001, 2015.



- [56] S. Sharma, S. Sharma, and A. Athaiya, “Activation functions in neural networks,” *towards data science*, vol. 6, no. 12, pp. 310–316, 2017.
- [57] C. Banerjee, T. Mukherjee, and E. Pasilio Jr, “An empirical study on generalizations of the relu activation function,” in *Proceedings of the 2019 ACM Southeast Conference*, 2019, pp. 164–167.
- [58] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [59] S. Jadon, “A survey of loss functions for semantic segmentation,” in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, 2020, pp. 1–7.
- [60] N. Abraham and N. M. Khan, “A novel focal tversky loss function with improved attention u-net for lesion segmentation,” in *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*. IEEE, 2019, pp. 683–687.
- [61] J. Sklansky, “Finding the convex hull of a simple polygon,” *Pattern Recognition Letters*, vol. 1, no. 2, pp. 79–83, 1982.
- [62] “Streamlit docs.” [Online]. Available: <https://docs.streamlit.io/>
- [63] J. M. Bland and D. G. Altman, “Measuring agreement in method comparison studies,” *Statistical methods in medical research*, vol. 8, no. 2, pp. 135–160, 1999.
- [64] D. G. Altman and J. M. Bland, “Measurement in medicine: the analysis of method comparison studies,” *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 32, no. 3, pp. 307–317, 1983.
- [65] M. Ali, *PyCaret: An open source, low-code machine learning library in Python*, April 2020, pyCaret version 1.0. [Online]. Available: <https://www.pycaret.org>