



FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

João Gonçalo Ramalho de Almeida e Sousa

Contributions to the Automatic Detection of Laser Marks in Retinal Digital Fundus Images

Dissertation submitted to the University of Coimbra to fulfill the requirements of
the Master of Science in Biomedical Engineering program

Supervisors:

Eng. Carlos Manta Oliveira (Retmarker S.A.)

Prof. Dr. Luis Alberto da Silva Cruz (Department of Electrical and Computer
Engineering, Faculty of Sciences and Technology, University of Coimbra)

Coimbra, 2015

This work was developed in collaboration with:



Retmarker S.A. — Coimbra, Portugal



Instituto de Telecomunicações — Coimbra, Portugal

Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

"Science is what we have learned about how to keep from fooling ourselves."

Richard P. Feynman

À minha avó Fernanda e ao meu amigo Gustavo Moura

Acknowledgements

First of all, I would like to acknowledge AIBILI and Centro Cirúrgico de Coimbra, in the persons of their directors, respectively, Dr. José Cunha-Vaz and Dr. António Travassos, for kindly providing and authorizing the use of most of the images processed in this work. Without them this master thesis would not have been accomplished.

I would like to thank my colleagues and teammates at Retmarker S.A. for the excellent and enriching opportunity during this past year. João Diogo Ramos, Frederico Teles and specially Carlos Manta Oliveira, one of my supervisors, were the ones responsible for providing me this amazing first experience at a company environment. Adriano Galante, André Lemos, Guilherme Pereira, Henrique Soares and Jason Bolito were other colleagues at Retmarker that also gave me useful and precious advice to have success during this thesis work.

I am also grateful to have worked with Professor Luís Cruz from the Department of Electrical and Computer Engineering, who is the main supervisor of this work. Without his support and guidance, I would not be able to overcome the obstacles of this project and all the small but so important details. Also, my colleagues from the professor's laboratory Nuno Almeida, Nuno Carvalho, Pedro Rocha, Ricardo Rocha and Rui Peliteiro provided special support during the writing of this document.

There is no doubt that what I accomplished during this 5-year adventure is because I had the great pleasure of meeting and making new friends in Biomedical Engineering, most of them graduating at the same time, that helped me to grow up a lot.

None of my accomplishments would make sense without the presence of my lifetime friends from Coimbra. André Tomé, Manuel Pinheiro, Tomás Resendes, José Pedro Marcelino, David Carvalho and João André Branquinho will always have a mark in everything I do.

Last, but not least, my deepest thanks to my family, who is and always will be the number one supporter. The example set by my parents Luís and Teresa, my aunt Guida, my sisters Inês and Joana and my brother Luís made me who I am today.

Contents

Acknowledgements	vii
Contents	ix
Abstract	xiii
Resumo	xv
List of Figures	xvii
List of Tables	xxi
Abbreviations	xxv
1 Introduction	1
1.1 Problem Contextualization and Motivation	1
1.2 Document Overview	5
2 Background on Ophthalmology and Eye-related Pathologies	7
2.1 Eye Anatomy and Physiology	7
2.2 Retinal Images	9
2.3 Eye Pathologies	10
2.3.1 Cataract	10
2.3.2 Glaucoma	11
2.3.3 Age-related Macular Degeneration	11
2.3.4 Diabetic Retinopathy	12
2.3.4.1 Proliferative Retinopathy	14
2.3.4.2 Diabetic Macular Edema	15
2.4 Treatment using Laser Photocoagulation	16
2.4.1 Panretinal Photocoagulation	17
2.4.2 Macular Edema Photocoagulation	19
3 Laser Marks Detection State-of-the-Art	21
3.1 Classification Problems in Image Processing	21
3.2 Detection of Macular Photocoagulation Scars With Confocal Infrared Reflection Imaging	23
3.3 Detection of Laser Marks in Retinal Images	25
3.4 Laser Marks Detection From Fundus Images	27

4	Automatic Laser Mark Detection Algorithm	29
4.1	Materials	30
4.1.1	Public Datasets	31
4.1.2	Proprietary Datasets	32
4.2	Image Pre-Processing	33
4.2.1	Retinal Image Resizing and Cropping	34
4.2.2	Vascular Network and Optic Disc Segmentation	35
4.2.3	Uneven Illumination Correction and Contrast Enhancement	36
4.2.4	Green Channel Separation and Noise Reduction	38
4.2.5	Kernels and Constants	39
4.3	Candidate Regions Extraction	42
4.3.1	Circular Hough Transform	42
4.3.2	Frangi Vesselness Filter	44
4.3.3	Laser Mark Segmentation	47
4.4	Features Computation	49
4.4.1	Geometrical Descriptors	50
4.4.2	Texture Descriptors	51
4.4.3	Spatial Distribution Descriptors	53
4.4.4	Intensity-based Descriptors	56
4.5	Classification	56
4.5.1	Decision Tree	57
4.5.2	Random Forest	58
4.5.3	Threshold-based classification	59
4.6	Features Selection	59
4.6.1	Selection using Information Gain	63
4.6.2	Selection using Gain Ratio	65
5	Results	69
5.1	Performance Evaluation	69
5.1.1	Decision Tree	71
5.1.2	Random Forest with 5 Trees	75
5.1.3	Random Forest with 50 Trees	79
5.1.4	Random Forest with 500 Trees	82
5.1.5	Simple threshold-based classifier	86
5.2	Computation Time	89
6	Discussion	93
7	Conclusion and Future Work	99
A	Performance and Information Theory Measures	103
B	Datasets	107
C	Pre-processing Examples	111

D Misclassified Retinal Images	117
References	126

Abstract

Diabetic retinopathy is the most frequent and serious complication of diabetes mellitus that affects vision. In advanced stages its progress can be delayed with recourse to laser photocoagulation treatments which destruct retinal tissue alleviating retinal hypoxia conditions and thus slowing the neovascularization processes typical of diabetic retinopathies. Currently, screening programs rely on automatic algorithms implemented to detect signs of diabetic retinopathy in patients. Unfortunately, these systems can malfunction if the patients screened have already undergone laser photocoagulation treatments.

This work, proposed by Retmarker S.A., presents a digital retinal fundus image tree-based classifier which uses features computed from candidate regions extracted from an input image to decide if that image shows evidence of previous photocoagulation laser treatments. A scientific paper describing the proposed algorithm and providing two new public accessible datasets of retinal images containing retinal images with laser marks was also prepared in order to submit to a scientific journal in the near future.

Three state-of-the-art image processing segmentation algorithms were combined to identify candidate laser marks, then features based on geometrical, texture, spatial distribution and intensity descriptors were calculated. The final decision about the presence of laser marks was obtained using specifically trained classifiers. A simple threshold-based classifier as well as four different tree-based classifiers, all of them using the image-based features, were tried and the ones providing the best classification performance were kept as the choice classifiers.

The performance of each algorithm was obtained after training and testing using retinal images with positive (containing laser marks) and negative cases from eight public datasets and three proprietary datasets, making a total of 233 images with laser marks and 2401 images without laser marks. A classifier based on a decision tree proved to be the best solution for this application with a sensitivity of 88.1% and a specificity of 98.9%. Further aspects and results of this novel contribution are presented and discussed in detail in this document.

Finally, some considerations on possible improvements and future developments, such as the optimization of the implemented algorithm in order to reduce the computation time, are made to close this thesis dissertation.

Keywords: *diabetic retinopathy, digital retinal fundus image, photocoagulation treatment, laser marks, medical image processing, feature extraction, classification*

Resumo

A retinopatia diabética é a complicação mais grave e frequente da diabetes mellitus que afecta a visão. Em estados avançados, o seu progresso pode ser travado com recurso a tratamentos de fotocoagulação que usam laser e que destroem tecido da retina. Actualmente, existem programas de rastreio que usam métodos automáticos implementados que detectam sinais de retinopatia diabética nos doentes. Contudo, estes sistemas podem funcionar indevidamente perante doentes que já foram tratados.

Este trabalho, no âmbito do projecto de tese proposto pela empresa Retmarker S.A., apresenta um classificador de retinografias baseado em árvores, que usa características calculadas a partir de regiões candidatas que por sua vez são extraídas a partir da imagem da retinografia, e decide se essa imagem contém indícios de tratamentos prévios de fotocoagulação por laser. Adicionalmente, resultou deste trabalho um artigo científico, para ser submetido a revista da especialidade num futuro próximo, que descreve o método de detecção de marcas laser proposto, assim como dois conjuntos de imagens retinográficas com marcas laser abertos ao uso pela comunidade científica.

Três algoritmos estado-da-arte de segmentação em processamento de imagem foram combinados com o intuito de detectar candidatos a marcas laser. Resultam destes algoritmos vários parâmetros baseados em descritores geométricos, texturais, de distribuição espacial e de intensidade que funcionam como elementos característicos das imagens e que são as entradas dos classificadores testados. Para efectuar a detecção de marcas laser foram testados um classificador simples baseado em limiar e quatro classificadores em árvore, todos eles recebendo os parâmetros característicos das imagens como entradas, tendo os melhores sido escolhidos para um estudo mais aprofundado e escolhido como a solução de classificação a adoptar.

A performance de cada algoritmo foi avaliada após a respectiva fase de treino e de teste, usando imagens de retinografia com casos positivos (imagens com marcas laser) e com casos negativos provenientes de oito datasets públicos e de três datasets proprietários, num total de 233 imagens com marcas laser e de 2401 imagens sem marcas laser. A árvore de decisão demonstrou ser a melhor solução para o problema da classificação para esta aplicação, tendo sido conseguida uma sensibilidade de 88.1% e uma especificidade de 98.9%. Os capítulos que se seguem descrevem em maior detalhe a abordagem seguida, fornecendo informação sobre os métodos estudados e apresentando também mais resultados.

Finalmente, algumas considerações sobre possíveis melhorias e ideias para trabalhos futuros, tais como optimização e redução do tempo de cálculo do algoritmo implementado, concluem esta dissertação.

Palavras-Chave: *retinopatia diabética, retinografia digital do fundo da retina, fotocoagulação, marcas laser, processamento de imagem médica, extracção de parâmetros característicos, classificação*

List of Figures

1.1	Number of people with diabetes by IDF Region, 2013.	2
1.2	Overview of this work’s problem.	4
2.1	Anatomy of the Eye.	8
2.2	Histology of the Retina.	8
2.3	Most frequent retinal imaging modalities.	9
2.4	Example of a photographic protocol.	10
2.5	Example of a retinography with MAs and haemorrhages.	13
2.6	Overview of the progression of DR.	14
2.7	Example of a retinography with PR.	14
2.8	Example of a retinography with DME.	15
2.9	Examples of PRP.	18
2.10	Examples of fluorescein angiographies after MEP.	20
3.1	Flowchart of the laser marks detection algorithm.	25
4.1	Examples of ”Laser” and ”No Laser” retinal images.	29
4.2	Flowchart of the proposed algorithm.	30
4.3	Flowchart of the pre-processing step.	34
4.4	Region of Interest (ROI) and cropped version of the original retinal image.	35
4.5	Vascular Network and OD masks of the original retinal image.	36
4.6	L^* color component of the retinal image.	37
4.7	L^* color component after uneven illumination correction (L^*_{final}).	38
4.8	Retinal Image after adaptive histogram equalization in the $L^*a^*b^*$ color space.	38
4.9	Final image resulting from the pre-processing step.	39
4.10	Brief explanation of the CHT algorithm.	43
4.11	Example of an accumulator array containing a peak corresponding to a detected center of a circle resulting from CHT.	43
4.12	Binary image of the detected laser mark candidates using Circular Hough Transform.	44
4.13	Possible patterns in 2D and 3D, depending on the value of the eigenvalues.	46
4.14	Binary image of the detected laser mark candidates using Frangi Vesselness Filter.	47
4.15	Binary image of the detected laser mark candidates using Laser Mark Segmentation.	48
4.16	Original image with all the detected candidate regions drawn.	49
4.17	Convex Hull of the candidate regions.	54

4.18	Decision Tree built using the training set of attributes and respective values of the weather a Saturday morning.	57
4.19	Information Gain (red) and Gain Ratio (blue) measures of each feature after thresholding at $t = 0.1$	62
5.1	Decision Tree model built after training the classifier with the dataset (S+EONMA).	72
5.2	ROC Curve after thresholding the output of the DT classifier with 5-fold cross validation.	72
5.3	ROC Curve obtained by thresholding the output of the trained DT classifier tested on 9 datasets.	75
5.4	ROC Curve obtained by thresholding the output of the RF5 classifier with 5-fold cross validation.	76
5.5	ROC Curve obtained by thresholding the output of the trained RF5 classifier tested on 9 datasets.	78
5.6	ROC Curve after thresholding the output of the RF50 classifier with 5-fold cross validation.	79
5.7	ROC Curve obtained by thresholding the output of the trained RF50 classifier tested on 9 datasets.	82
5.8	ROC Curve after thresholding the output of the RF500 classifier with 5-fold cross validation.	83
5.9	ROC Curve obtained by thresholding the output of the trained RF500 classifier tested on 9 datasets.	86
5.10	ROC Curve after thresholding by the number of regions detected by the CHT algorithm.	87
5.11	ROC Curve after thresholding by the number of regions detected by the FVF algorithm.	88
5.12	ROC Curve after thresholding by the number of regions detected by the LMS algorithm.	89
5.13	Plot of the Computation time per image against the total number of pixels.	91
B.1	Overview of Screening Dataset.	108
B.2	Overview of BAT Dataset.	108
B.3	Examples of images with laser marks from the Screening Dataset.	109
B.4	Examples of images without laser marks from the Screening Dataset.	110
B.5	Images of patient 08 from the BAT Dataset.	110
B.6	Examples of images with laser marks from the PDJD.	110
C.1	Pre-processing outputs and candidates extraction of image 0002.jpg from the Screening Dataset.	113
C.2	Pre-processing outputs and candidates extraction of image 0132.jpg from the Screening Dataset.	114
C.3	Pre-processing outputs and candidates extraction of image 0198.jpg from the Screening Dataset.	115
C.4	Retinal image 0010.jpg from the S dataset.	116
C.5	Retinal image 0189.jpg from the S dataset.	116
C.6	Retinal image 04_A1.jpg from the BAT dataset.	116
C.7	Images from the EOEX dataset.	116

D.1	Pre-processing outputs and candidates extraction of image 0219.jpg from the S Dataset.	117
D.2	Pre-processing outputs and candidates extraction of image 0225.jpg from the S Dataset.	118
D.3	Pre-processing outputs and candidates extraction of image 0419.jpg from the S Dataset.	119
D.4	Pre-processing outputs and candidates extraction of image 06_B1.jpg from the BAT Dataset.	120
D.5	Pre-processing outputs and candidates extraction of image 07_B2.jpg from the BAT Dataset.	121
D.6	Pre-processing outputs and candidates extraction of image 15.jpg from the FAZD Dataset.	122
D.7	Pre-processing outputs and candidates extraction of image 28.jpg from the FAZD Dataset.	123
D.8	Pre-processing outputs and candidates extraction of image 32.jpg from the FAZD Dataset.	124
D.9	Pre-processing outputs and candidates extraction of image 20051130_60691_0400_PP.jpg from the Messidor Dataset.	125
D.10	Pre-processing outputs and candidates extraction of image 20051214_51733_0100_PP.jpg from the Messidor Dataset.	125

List of Tables

3.1	Results of the examiners' answer for each imaging modality.	24
3.2	João Dias <i>et al.</i> laser mark detection algorithm performance.	26
3.3	Faraz Tahir <i>et al.</i> laser mark detection algorithm performance.	28
4.1	Implemented 21 texture features based on [84–86].	53
4.2	Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier DT.	64
4.3	Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier RF5.	64
4.4	Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier RF50.	65
4.5	Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier RF500.	65
4.6	Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier DT.	66
4.7	Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier RF5.	66
4.8	Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier RF50.	67
4.9	Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier RF500.	67
5.1	Confusion Matrix after a 5-fold Cross Validation using DT of the images from the dataset S+EONMA.	71
5.2	Confusion Matrix after testing the trained DT on M.	71
5.3	Confusion Matrix after testing the trained DT on EOMA.	73
5.4	Confusion Matrix after testing the trained DT on EOEX.	73
5.5	Confusion Matrix after testing the trained DT on EONEX.	73
5.6	Confusion Matrix after testing the trained DT on VBR.	73
5.7	Confusion Matrix after testing the trained DT on HP.	73
5.8	Confusion Matrix after testing the trained DT on FAZD.	73
5.9	Confusion Matrix after testing the trained DT on BAT.	73
5.10	Confusion Matrix after testing the trained DT on PDJD.	74
5.11	Algorithm's performance on each dataset using Decision Tree.	74
5.12	Performance stats using the trained Decision Tree averaged over all testing datasets.	74
5.13	Confusion Matrix after a 5-fold Cross Validation using RF5 of the images from the dataset S+EONMA.	75
5.14	Confusion Matrix after testing the trained RF5 on M.	76

5.15	Confusion Matrix after testing the trained RF5 on EOMA.	76
5.16	Confusion Matrix after testing the trained RF5 on EOEX.	76
5.17	Confusion Matrix after testing the trained RF5 on EONEX.	76
5.18	Confusion Matrix after testing the trained RF5 on VBR.	77
5.19	Confusion Matrix after testing the trained RF5 on HP.	77
5.20	Confusion Matrix after testing the trained RF5 on FAZD.	77
5.21	Confusion Matrix after testing the trained RF5 on BAT.	77
5.22	Confusion Matrix after testing the trained RF5 on PDJD.	77
5.23	Algorithm's performance on each dataset using Random Forest with 5 Trees.	78
5.24	Performance stats using the trained Random Forest with 5 Trees averaged over all testing datasets.	78
5.25	Confusion Matrix after a 5-fold Cross Validation using RF50 of the images from the dataset S+EONMA.	79
5.26	Confusion Matrix after testing the trained RF50 on M.	80
5.27	Confusion Matrix after testing the trained RF50 on EOMA.	80
5.28	Confusion Matrix after testing the trained RF50 on EOEX.	80
5.29	Confusion Matrix after testing the trained RF50 on EONEX.	80
5.30	Confusion Matrix after testing the trained RF50 on VBR.	80
5.31	Confusion Matrix after testing the trained RF50 on HP.	80
5.32	Confusion Matrix after testing the trained RF50 on FAZD.	80
5.33	Confusion Matrix after testing the trained RF50 on BAT.	81
5.34	Confusion Matrix after testing the trained RF50 on PDJD.	81
5.35	Algorithm's performance on each dataset using Random Forest with 50 Trees.	81
5.36	Performance stats using the trained Random Forest with 50 Trees averaged over all testing datasets.	82
5.37	Confusion Matrix after a 5-fold Cross Validation using RF500 of the im- ages from the dataset S+EONMA.	82
5.38	Confusion Matrix after testing the trained RF500 on M.	83
5.39	Confusion Matrix after testing the trained RF500 on EOMA.	83
5.40	Confusion Matrix after testing the trained RF500 on EOEX.	83
5.41	Confusion Matrix after testing the trained RF500 on EONEX.	84
5.42	Confusion Matrix after testing the trained RF500 on VBR.	84
5.43	Confusion Matrix after testing the trained RF500 on HP.	84
5.44	Confusion Matrix after testing the trained RF500 on FAZD.	84
5.45	Confusion Matrix after testing the trained RF500 on BAT.	84
5.46	Confusion Matrix after testing the trained RF500 on PDJD.	84
5.47	Algorithm's performance on each dataset using Random Forest with 500 Trees.	85
5.48	Performance stats using the trained Random Forest with 500 Trees averaged over all testing datasets.	85
5.49	Confusion Matrix after applying the optimal threshold for the number of regions detected by the CHT algorithm ($number_CHT \geq 3$ means that the retinal image is "Laser").	87
5.50	Performance stats using the simple threshold-based classifier $number_CHT \geq 3$ averaged over all testing datasets.	87

5.51	Confusion Matrix after applying the optimal threshold for the number of regions detected by the FVF algorithm ($number_FVF \geq 3$ means that the retinal image is "Laser").	88
5.52	Performance stats using the simple threshold-based classifier $number_FVF \geq 3$ in every testing dataset.	88
5.53	Confusion Matrix after applying the optimal threshold for the number of regions detected by the LMS algorithm ($number_LMS \geq 5$ means that the retinal image is "Laser").	89
5.54	Performance stats using the simple threshold-based classifier $number_LMS \geq 5$ in every testing dataset.	89
5.55	Computation time per image of the algorithm for each resolution.	91
B.1	List of images not considered from Messidor Dataset.	107
B.2	Details of Screening Dataset.	107
B.3	Details of BAT Dataset.	109

Abbreviations

AIBILI	A ssociation for I nnovation and B iomedical Research on L ight and I mage
AMD	A ge-related M acular D egeneration
ARS	A dministração R egional de S aúde
AUC	A rea u nder the R OC c urve
BAT	B efore and A fter T reatment D ataset
CHT	C ircular H ough T ransform
DME	D iabetic M acular E dema
DR	D iabetic R etinopathy
DRS	D iabetic R etinopathy S tudy
DT	D ecision T ree
EOEX	e -ophtha EX
EOMA	e -ophtha MA
EONEX	e -ophtha N o EX
EONMA	e -ophtha N o MA
ETDRS	E arly T reatment D iabetic R etinopathy S tudy
EX	E xudate
FAZD	F oveal A vascular Z one D etection D ataset
FOV	F ield o f V iew
FVF	F rangi V esselness F ilter
GIS	G eographic I nformation S ystem
GPC	G eneric P arameter C lassification
GR	G ain R atio
HP	H ealthy P eople D ataset
IDF	I nternational D iabetes F ederation
IG	I nformation G ain

LMS	L aser M ark S egmentation
M	M essidor Dataset
MA	M icroaneurysm
MEP	M acular E dema P hotocoagulation
NPV	N egative P redictive V alue
OCT	O ptical C oherence T omography
OD	O ptic D isc
PDJD	P roprietary D ataset J oão D ias
PPV	P ositive P redictive V alue
PR	P roliferative R etinopathy
PRP	P anretinal P hotocoagulation
RF5	R andom F orest with 5 T rees
RF50	R andom F orest with 50 T rees
RF500	R andom F orest with 500 T rees
ROC	R eceiving O perator C haracteristic
ROI	R egion of I nterest
S	S creening Dataset
SPC	S tructural P arameter C lassification
SVM	S upport V ector M achine
VBR	V essel- B ased R egistration Dataset
VEGF	V ascular E ndothelial G rowth F actor
WHO	W orld H ealth O rganization

Chapter 1

Introduction

1.1 Problem Contextualization and Motivation

Diabetes Mellitus is one of the diseases with higher prevalence amongst the world's population, currently. As a matter of fact, the International Diabetes Federation (IDF) registered 382 million cases of diabetes in 2013 and predicted that this number will rise to 592 million by 2035. In 2013, there were 5.1 million deaths caused by diabetes, which means that every seven seconds a person dies from this systemic disease. IDF's Atlas also states that 80% of people with diabetes live in low and middle-income countries and at least USD 548 billion dollars were spent in health expenditure in 2013 [1].

Eyesight tends to deteriorate as part of the ageing process. However, there are also some specific eye conditions that may cause vision loss: cataracts, glaucoma, age-related macular degeneration and diabetic retinopathy, which are the most common causes of sight reduction. As a result, these eye conditions are the priority of the VISION 2020 - a global project launched by World Health Organization (WHO) in 1999, whose purpose is to research mechanisms to prevent and treat these ocular diseases by the year of 2020 [2].

Diabetic Retinopathy (DR) is the most frequent and most serious complication of diabetes that afflicts vision. It is a sight-threatening and chronic ocular disease that results from retinal blood vessels damage. It affects up to 80% of all patients who have had diabetes for 10 years or more [3] and it is the leading cause of blindness amongst the

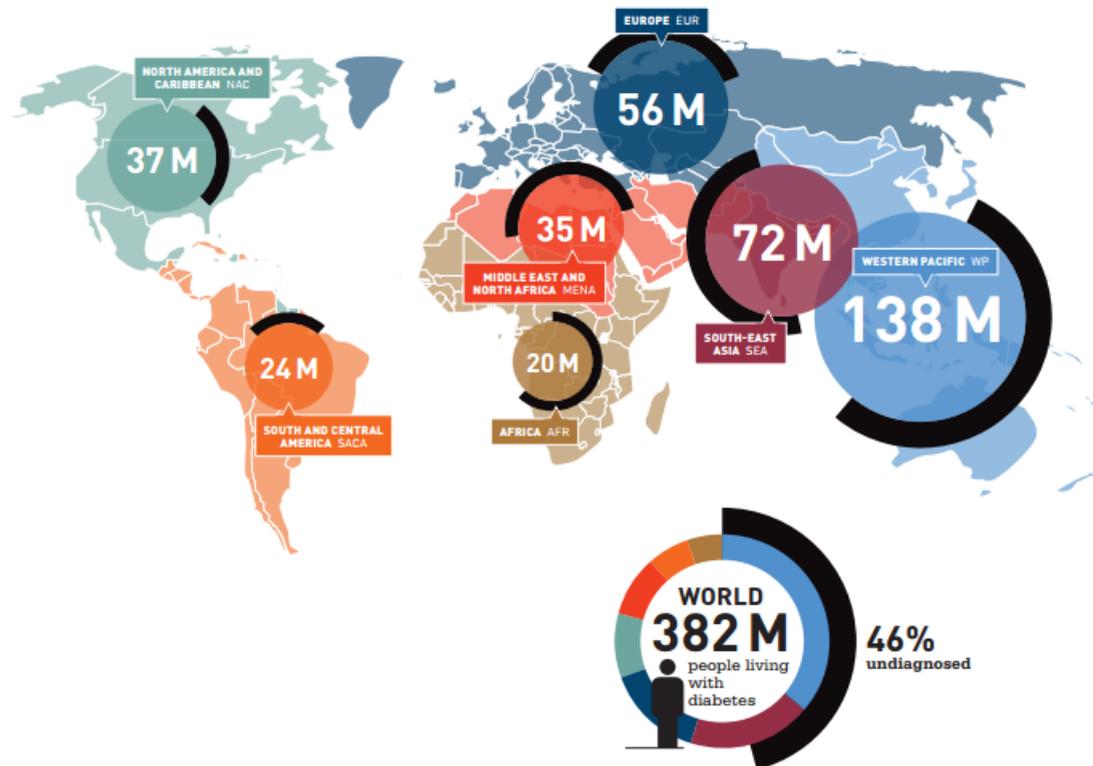


Figure 1.1: Number of people with diabetes by IDF Region, 2013 [1].

working-age population. Laser treatment and vitrectomy (still very controversial in scientific community) reduce the risk of blindness in patients with DR by more than 90%, if this condition is detected early [4].

It is recommended that each diabetic patient participates in a DR Screening Program annually by taking digital retinal fundus photographs and having them analysed and graded by an ophthalmologist or optometrist [5, 6]. However, this is a slow and difficult process because the number of persons with potential risk to develop DR is much higher than the total number of specialists. There are around 300 000 optmetrists [7] and 206 812 ophthalmologists [8] worldwide, which is equivalent to 30 ophthalmologists per million.

Taking into account the total number of people with diabetes and the percentage of this population that develops DR, it is clear that a huge number of people must be periodically seen by specialists in order to detect and track the existence of early lesions and signs of DR.

Meetings, strategies and recommendations to implement cost-effective DR Screening Programs have been multiplying these last years [9, 10]. The main objective of these

cost-effective programs is to reduce the workload and the burden of the specialists and ensure a high coverage (reaching at least 80% of the target population) in a short period of time [11].

Several studies of DR mass screening programs in different countries concluded that the implementation of these initiatives in remote rural areas, where the availability of healthcare professionals is scarce, is possible and effective if the correct procedure is followed. In these screening sessions conducted in remote areas, only a photographer and a grader are required at the local where the screening program is held. The photographer captures, stores the retinal images of the patients, which are then sent to the grader who observes and identifies those with signs of DR. The cases detected with lesions need more medical attention and hence are referred to an ophthalmologist by the grader. Moreover, in these studies, around 20% of the participants were detected with DR, so the ophthalmologists' burden reduction is huge [12–14].

However, the screening programs discussed that are based on telemedicine technologies still require that graders and specialists observe and analyze an enormous amount of retinal images. Since there are 30 ophthalmologists per million people with diabetes, automated grading algorithms have been recently developed to cover this issue [6]. Some of these algorithms are shortly described and reviewed by Dawn Sim *et al.* [15].

Retmarker[®] is one of these automated assessment algorithms. It is a product developed by Retmarker S.A. and its main purpose is to decrease the human grading workload and improve and simplify this healthcare system [16]. This automated algorithm detects lesions related to DR such as microaneurysms (MAs) and exudates (EXs) and it is integrated in a DR Screening Program in the Central Region of Portugal, coordinated by Administração Regional de Saúde do Centro (ARS Centro) and Association for Innovation and Biomedical Research on Light (AIBILI).

In these screening initiatives, such as the Portuguese example, each person with diabetes mellitus is called for screening and sometimes patients that already underwent laser surgery show up. This situation occurs because these patients do not know or cannot recall having been treated and there are even cases where they are aware of the fact that they were treated, but are called to participate in the screening.

Such scenario poses 2 problems: one concerning the coordination and management of the program and the other one concerning technical issues related to the operation of the automatic diagnostic procedures. First, there is a risk that the patient might be scheduled for further treatment again. Also, the detection and removal of these patients would avoid unnecessary processing, which would result in a more reduced workload. In the DR screening program in the Central Region of Portugal, there were 219 subjects with signs of treatment out of 13305 during 2014, which is equal to 1.65% of the participants. It is important to note that not only will the diabetic population grow, but also will the total number of DR-treated patients rise. Second, the treatment usually leaves behind scars produced by laser incidence on the retina and this fact causes misbehaviours in automatic diagnostic algorithms used in Telemedicine. For example, these laser marks can interfere negatively by being identified as MAs or as EXs or as other signs of eye pathology by the automatic algorithms, resulting in false detections.

It is therefore important to be able to automatically detect laser marks that result from photocoagulation treatment on retinal fundus images, which is the main objective of this thesis. By doing so, not only treated patients are removed from the screening programs and directed to adequate follow-up, but also a filtering step is performed before the processing of the images to detect DR lesions. Figure 1.2 shows how the developed automatic laser mark detector in this thesis should be implemented in a DR screening program. Moreover, there are no publicly available datasets of retinal images with photocoagulation scars, so it is also of great importance to provide the scientific community with images containing these special characteristics, thus contribution for the development of algorithms with similar aims as ours.

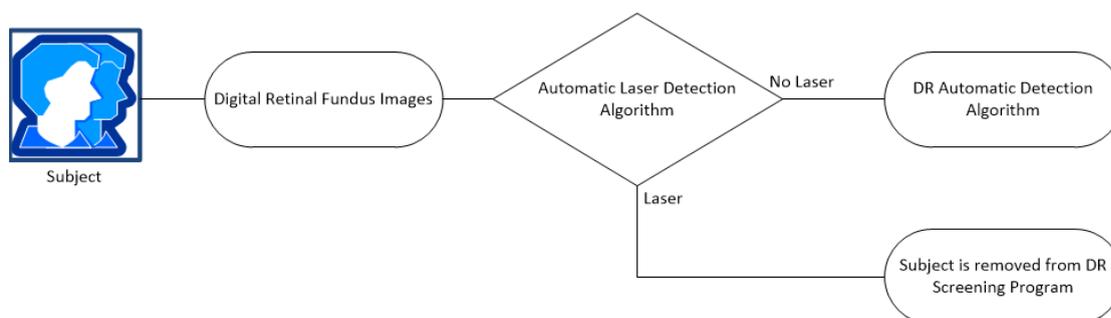


Figure 1.2: Overview of this work's problem.

1.2 Document Overview

This thesis is structured in seven chapters plus four appendixes.

Chapter 2: Background on Ophthalmology and Eye-related Pathologies reviews the the foundations on eye anatomy, eye physiology and retinal images. Also, the different eye pathologies and respective treatments are explained to provide a clarifying background so that the reader understands the problem stated in chapter 1.

Chapter 3: Laser Marks Detection State-of-the-Art summarizes previous research based on image processing techniques and lists the previous work developed on this classification problem and the respective results and conclusions.

Chapter 4: Automatic Laser Mark Detection Algorithm describes the content and the details of the datasets used in this work and explains the proposed algorithm step-by-step.

Chapter 5: Results presents the performance of the algorithm explained in chapter 4 when applied to the images from the datasets described in the same chapter.

Chapter 6: Discussion analyses and discusses the results presented in chapter 5.

Chapter 7: Conclusion and Future Work evaluates the results of the work, in confront with its objectives and provides suggestions for future work.

Appendix A: Performance and Information Theory Measures details some mathematical parameters used in this thesis.

Appendix B: Datasets contains images and further details of some datasets described in chapter 4.

Appendix C: Pre-processing Examples illustrates the image processing steps of three different retinal images.

Appendix D: Misclassified Retinal Images shows examples of retinal images without laser marks that were classified as having laser marks.

Chapter 2

Background on Ophthalmology and Eye-related Pathologies

2.1 Eye Anatomy and Physiology

The human eye transduces energy into nerve impulses. This energy is the visible light and it is a range of energy of the electromagnetic spectrum with wavelength between 400nm and 700nm. The eye is structured in three main anatomical structures (Figure 2.1): the outer layers' lenses of the eyeball, the retina and the Optic Disc (OD) [17].

Considering the light path of the eye, the sclera comes first, which is a coat of connective tissue continuous with the cornea and it can be seen externally as the white of the eyes. Light passes through the cornea to enter the anterior chamber and then passes through the pupil, which is surrounded by a pigmented muscle known as the iris. This anatomical element functions like a diaphragm of a camera: the aperture (the pupil) is widened or narrowed to admit more or less light. The light then passes through the lens, a crystalline and biconvex structure placed between posterior and vitreous chambers. The aqueous humor is a fluid that provides nutrients to the avascular lens and cornea and it fills the anterior chamber and the posterior chamber, that is located between the iris and the lens. Finally, the light suffers a refraction in the lens and reaches the retina after passing through a thick substance called vitreous humor [17]. Figure 2.1 shows all these parts in greater detail.

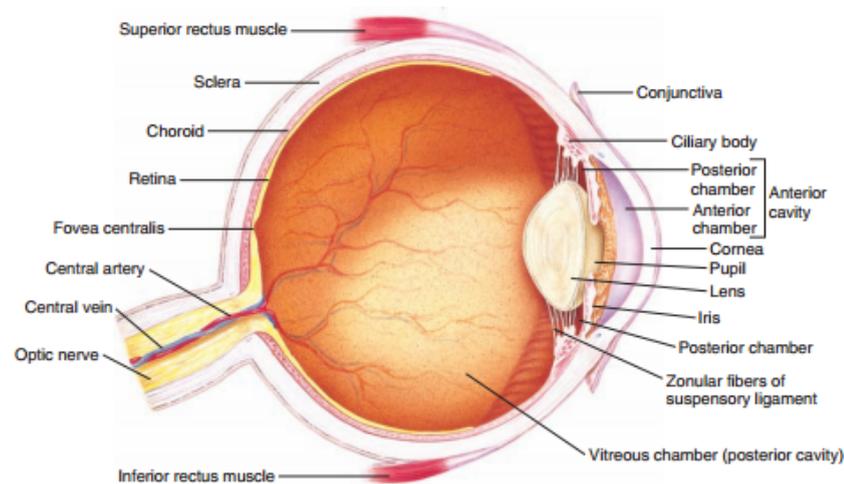


Figure 2.1: Anatomy of the Eye [17].

The retina is a neural layer that contains photoreceptors, bipolar and ganglion cells. The photoreceptors translate the received light into electric impulses and the bipolar and ganglion cells are activated by the photoreceptors to carry the message to a region where all the perceptual information is gathered: the optic disc. The information is then transferred to the brain by several neuron fibers that constitute the optic nerve. In the external layer of the retina, there are two types of photoreceptor neurons with different shapes: the cones and rods. The cones are light sensitive receptors with a conic shape responsible for colour vision and most of them are in the fovea (central region of the retina - macula). On the other hand, the rods are thin and long and have a bigger diameter than the few cones located on the peripheral area of the retina. The rods are responsible for the black and white vision [17, 18].

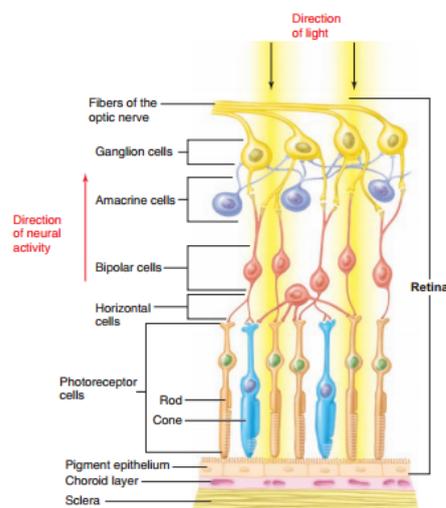


Figure 2.2: Histology of the Retina [17].

2.2 Retinal Images

There is a wide range of medical diagnostic exams in the ophthalmology area and with the evolution of science, it is likely that there will be more and some of them will be further developed. For instance, there are two imaging modalities that are the most popular: digital fundus imaging and optical coherence tomography (OCT).

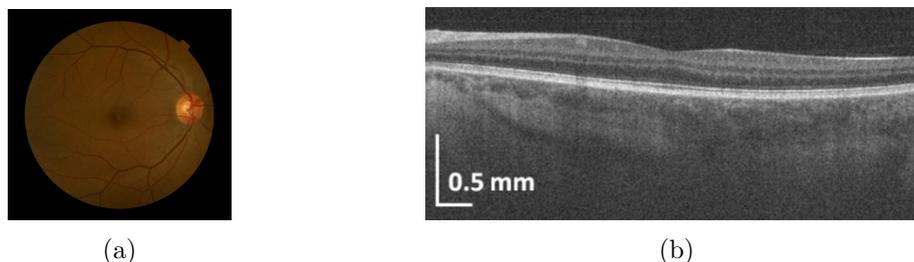


Figure 2.3: Most frequent retinal imaging modalities: (a) Digital fundus image of a right eye [19] (b) OCT Image of the macular region [20].

Digital retinal colour fundus imaging is the most frequently used in screening initiatives. It is not an invasive technique, unlike angiography with intravenous fluorescein which is too invasive, although it provides images with good quality for diagnostic purposes [21]. Moreover, medical image analysis is currently a research area that attracts both scientists and physicians to develop computational tools and systems that support and assist the diagnosis process. This kind of technology is easy to use and very compatible with digital fundus images [22]. Also, this imaging modality is convenient to use with subjects who do not have any sight-threatening disease, because unnecessary invasive interventions are avoided, the time spent capturing the image is short and the subject can resume his or her normal activities comfortably without any problems. These facts make screening programs based on colour fundus photography more cost-effective and a popular choice [23].

In this work, different datasets are used that are made up of images captured using different protocols. For example, in the ongoing DR screening in the centre region of Portugal, the images are from right and left eyes. For each eye, a photo is taken centered on the optical disc and nasal region (field 1), as presented in Figure 2.4a and another photo is taken centered on the macula (field 2), as shown in Figure 2.4b [24].

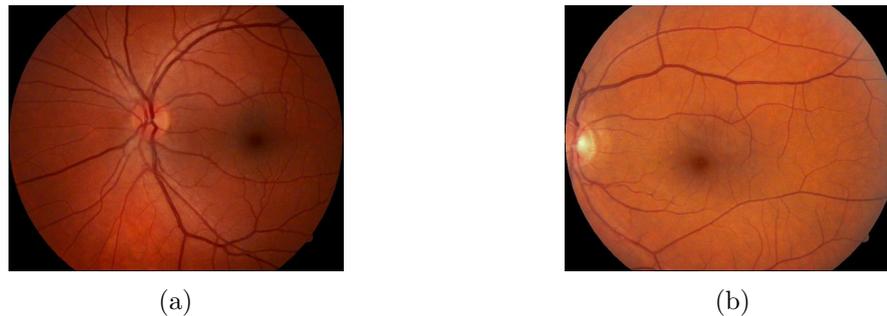


Figure 2.4: Example of a photographic protocol: (a) Field 1 image of a left eye [25] (b) Field 2 image of a left eye [25].

2.3 Eye Pathologies

The eye is composed of several layers of different type of cells. Like any other biological tissue, it is vulnerable to damage and to disease.

There are several pathologies related to failures of the eye's optical arrangement. Some of them result from a deficiency of the focal distance (hyperopia, myopia) of the lens and/or from geometrical imperfections of the eye (astigmatism). Others result from damages of the supporting structures: strabismus is caused by weakness of the muscles that control the movements of the eyeball and conjunctivitis result from an inflammation of the sclera's surface.

In this section, only the four sight-threatening diseases considered as priorities by the VISION 2020, namely cataract, glaucoma, age-related macular degeneration (AMD) and diabetic retinopathy, will be covered in detail.

2.3.1 Cataract

Cataract, which according to 2005 numbers affect 16 million people worldwide, is caused by the opacification of the lens. It is undoubtedly one of the main causes for vision loss and blindness. Age is the principal risk factor, but there are others such as genetic composition, exposure to ultraviolet light and diabetes. It is no surprise that several patients that show up in DR screenings have cataracts [26] thus justifying the widespread use of automatic cataract detection algorithms [27].

2.3.2 Glaucoma

Glaucoma is a pathology that results from a set of diseases such as high intra-ocular pressure and damage to the OD. It affects peripheral vision and it has no cure but its early diagnosis and treatment can prevent the vision loss.

There are two main types of Glaucoma: Primary Open Angle (POAG) and Angle Closure Glaucoma (ACG). POAG is the most common form of Glaucoma accounting for at least 90% of all cases. With POAG, the main problem originates inside the drainage canals of the eye, which results in a high intra-ocular pressure. There are essentially four changes to the retinal structures when Glaucoma occurs [28]:

- **OD variance** - The OD, also known as the blind spot due to lack of receptors, is easily visible on retinal images as a bright elliptic region. If this structure becomes larger, it means that more optic nerve fibres disappeared, so the risk to develop Glaucoma is higher.
- **Neuroretinal Rim Loss Determination** - Although it does not happen with every patient, the positions of the Nasal and Temporal regions are changed and as a consequence the usual pattern of the Neuroretinal Rim is not followed.
- **Retinal Nerve Fibre Layer defects** - If detected, it serves as the earliest sign of Glaucoma. Normally, this layer appears as bright bundle striations unevenly distributed, but in the presence of Glaucoma it appears as a wedge-shaped dark area.
- **Peripapillary Atrophy (PPA)** - It is one of the most important risk factors and its progression may lead to haemorrhage of the OD. PPA is the degeneration of the retinal pigment epithelial layer and photoreceptors in the region surrounding the OD.

2.3.3 Age-related Macular Degeneration

It is estimated that in many countries, more than 20% of the population has age-related macular degeneration, which is a progressive chronic disease of the central retina and a leading cause of blindness worldwide. Despite having this name, older age is not the

only cause as smoking, nutritional factors, genetic markers and cardiovascular diseases are also risk factors.

In early stage of AMD, symptoms are usually inexistent but in retinal images it is possible to observe yellow drusens underneath the retinal pigment epithelium. In advanced AMD, patients might develop neovascular (wet) AMD or geographic atrophy (late dry). When neovascular AMD occurs, there is a rapid visual loss and patients describe sudden worsening of central vision with distortion (metamorphopsia) or a dark patch in the central vision (scotoma) or both. This condition is also characterised and identified by haemorrhages and intraretinal fluid. When geographic atrophy occurs, the progression of the vision loss is much slower over many years. In retinal images, it is easily identified when a sharply demarcated area of depigmentation is seen near the fovea.

Three decades ago, the Macular Photocoagulation was the technique used for treatment. Although this laser photocoagulation procedure was able to reduce long-term severe visual loss, there was a 40% risk of immediate moderate visual loss. This is an issue that will be discussed in more detail on section 2.4. Currently, the use of vascular endothelial growth factor (VEGF) inhibitors (ranibizumab and bevacizumab) is the most common treatment strategy for this disease. Although it is not an easy process because patients need a monthly injection of anti-VEGF, there is a very good recovery in vision acuity even for the patients with advanced AMD.

Screening initiatives are also useful in order to monitor the disease progression, which is similar to what happens with DR. AMD can be detected with automatic diagnostic algorithms, but the performance (sensitivity of 75%) is not as good as the automatic algorithms used to detect DR [29].

2.3.4 Diabetic Retinopathy

Diabetic retinopathy is a set of lesions on the retina caused by complications that accompany diabetes mellitus. It is the leading cause of preventable blindness amongst working age population. During the first 20 years of diabetes, nearly all patients with type 1 diabetes develop DR, while more than 60% of patients with type 2 diabetes have DR. It is clear that type 1 diabetic patients are much more prone to develop DR than type 2 diabetic patients.

Some of the risk factors are ageing (the principal), genetic variation, poor control of blood sugar, high blood pressure and obesity. In other words, if a diabetic person does not have a healthy lifestyle and does not properly control the disease, it is very likely to develop DR. As for the ageing, it is an uncontrollable factor and for that reason DR screening programs play an important role to detect early signs of retinal lesions and to prevent sight loss. According to the American Diabetes Association, approximately 18.3% of Americans aged 60 and older have diabetes, because the prevalence of this disease also increases with age. Therefore, not only the working age population is of concern, but also the elderly population because a high percentage has diabetes and, as a consequence, the incidence of DR is also very high.

Nonproliferative DR's presence is detected in retinal images when MAs and small haemorrhages are visible (Figure 2.5). These initial lesions are focal and located within the inner nuclear layer of the retina and hyperglycemia is sufficient to serve as a trigger to this pathology, due to excessive transport of glucose near the cells surrounding the deep retinal vascular network. The rupture of MAs and small vessels may lead to the formation of intraretinal haemorrhages. Another characteristic of nonproliferative DR is the presence of hard exudates, which are extracellular accumulations of lipids, proteins and lipoproteins resulting from leakage from abnormal vessels.



Figure 2.5: Example of a retinography with MAs and haemorrhages [30].

The existence of exudates indicate a more severe stage of nonproliferative DR. If further progression of DR is verified, there is a risk of developing complications of DR, which are sight-threatening situations. The complications are the Proliferative Retinopathy and the Diabetic Macular Edema, discussed on subsections 2.3.4.1 and 2.3.4.2, respectively. A severe DR patient can develop either one of these complications or both [3, 30, 31]. Figure 2.6 is a scheme that describes the possible progressions of DR.

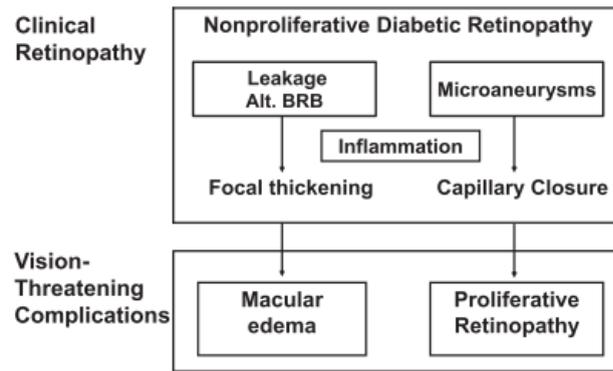


Figure 2.6: Overview of the progression of DR [30].

2.3.4.1 Proliferative Retinopathy

The first sign of Proliferative Retinopathy (PR) is the observation of retinal areas with capillary nonperfusion, also known as capillary closure (Figure 2.7). Vasoregression is the initial step of a series of vascular damages.

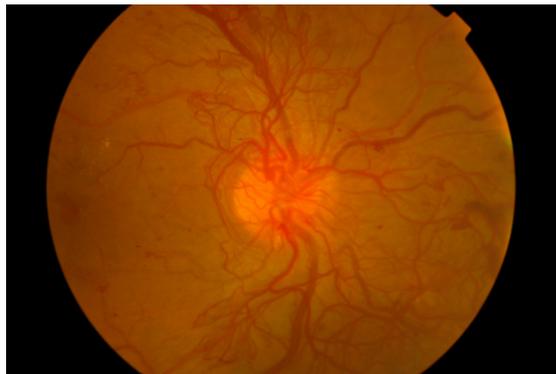


Figure 2.7: Example of a retinography with PR [30].

As the name itself suggests, PR is the growth of new vessels that occurs when the damaged blood vessels are not able to properly nourish the retina with oxygen and nutrients. These new vessels arise from the OD but are very fragile and their presence leads to retraction of the vitreous humor. This pulling effect and consequent shrinkage of the vitreous humor results in vitreous haemorrhage which cause vision loss and permanent low vision.

It is important to highlight that PR is not specific to DR, as it also occurs in other retinal vascular diseases such as sickle cell disease and retinal vein occlusion. Besides the neovascularization in the OD region, if PR is DR-specific, there is growth of new vessels from the remaining perfused vessels in the posterior pole [3, 30, 32].

2.3.4.2 Diabetic Macular Edema

Diabetic Macular Edema (DME) is the largest cause of sight loss in diabetes because it affects central vision (fovea). Although it is a complication of DR, DME can occur during any stage of DR, but it is more likely to accompany PR. It is a fact that about half of those with PR also have DME. Also, DME affects particularly older type 2 diabetic patients.

Diabetic Macular Edema is a thickening of the macula due to accumulation and swelling of fluid in the central macular area by the retinal pigment epithelium resulting from fluid leakage because of the changes in the blood-retinal barrier, which can lead to scarcity of oxygen (ischaemia). Due to the swelling, the retinal tissue volume in the macular area increases and may provoke changes in this region's position (Figure 2.8). For that reason, DME is easier to detect using OCT.

Although it is a deeply studied and well-known disease, it is not always easy to classify the severity and the stage of the DME. Some guidelines were established to solve this issue and hence DME is considered clinically significant under two circumstances [3, 30, 32]:

- Thickening of the retina within 500 μm of the center of the fovea, only observable in OCT.
- Presence of EXs within 500 μm of the center of the fovea, which can be seen in digital colour fundus images.



Figure 2.8: Example of a retinography with DME [32].

2.4 Treatment using Laser Photocoagulation

Laser photocoagulation is an intervention that is commonly used to treat DR, in which energy in the form of light is applied to the retina with the purpose of preserving vision but not reversing it, although it causes some damage to the photoreceptors. It is indeed one more reason why it is so important that patients with sight-threatening diseases such as DR be identified before irreversible visual loss occurs [33, 34].

”It is the second-most common eye procedure after cataract extraction, and yet little has changed in laser design over the last 35 years until recently” [35]. Using laser to treat DR has been demonstrated to be a very successful and effective procedure throughout the years, as confirmed by the Diabetic Retinopathy Study (DRS) [36] and Early Treatment Diabetic Retinopathy Study (ETDRS) [37].

More recently, VEGF inhibitors and steroids have been used alone to treat complications of DR. The use of these substances has been proven to be more effective and produce a larger gain in vision than laser photocoagulation. However, there are two disadvantages when using these compounds: they have a short-effect duration and there is a significant increase in intraocular pressure and, as a consequence, increase the risk of incidence of cataract. Therefore, anti-VEGF and steroids have been used in low dosages as a supplement to laser photocoagulation. The idea of this combination is to enhance the effectiveness of the traditional treatment procedure [38, 39].

The photocoagulation procedure using laser pulses is either controlled manually or semi-automatically by the laser operator. In either case, the scars left on the retinal pigment have special visual characteristics. After an observation of a number of such images from the DR screening coordinated by ARS, it can be concluded that the laser mark patterns usually have the following properties:

- They are not randomly distributed over the retina,
- They tend to occur in clusters,
- They occur in periphery regions, far from the fovea and from the OD so that the central vision is not affected,
- They present a dark or bright appearance and their colour can be yellow or green.

- They usually have a circular or asymmetric shape.

It can be assumed that the observed characteristics in these images are universal since there is no reason to believe that using the same methodologies during retinal imaging capture in different places produces different results.

Even though it is possible to describe some generally occurring laser spot patterns, after observation of several retinographies, it is very important to consider the treatment parameters. For example, the use of pattern or single spot, type of pattern, spot size, power, burn duration, number of burns per session and retinal coagulation are variables that contribute to a wide range of possible laser marks size, shape, colour and distribution [35].

The treatment applied varies depending on the complication of DR. As stated in subsection 2.3.4, the most severe stages of DR are PR and DME. So there are two types of laser treatment, one for each of the mentioned DR complications, described in subsections 2.4.1 and 2.4.2, respectively. The following subsections present information about the treatments' procedure and the formation of the laser marks.

2.4.1 Panretinal Photocoagulation

Panretinal Photocoagulation (PRP) is the most proven and accepted treatment for PR. This treatment is recommended when neovascularization on the OD is greater than one-third of its diameter or when there are haemorrhages resulting from new vessels [40].

Since the 1980s until 2006, the conventional argon lasers were the most frequently used in clinical practices. Patients did not have to go to the hospital, because the slit-lamp-delivered laser system through a wide angle contact lens was simple enough to deliver treatment in clinical centers. Only those who were unable to cooperate needed to go to the hospital to be treated while under anesthesia. After 2006, Pattern Scan Laser (PASCAL) started to be used more frequently (Figure 2.9). It is a semi-automatic system containing a 532 nm frequency-doubled (Nd:YAG) solid-state laser. Even though the type of laser design and respective pulse frequency changed, the clinical procedure did not change much [35, 40, 41].

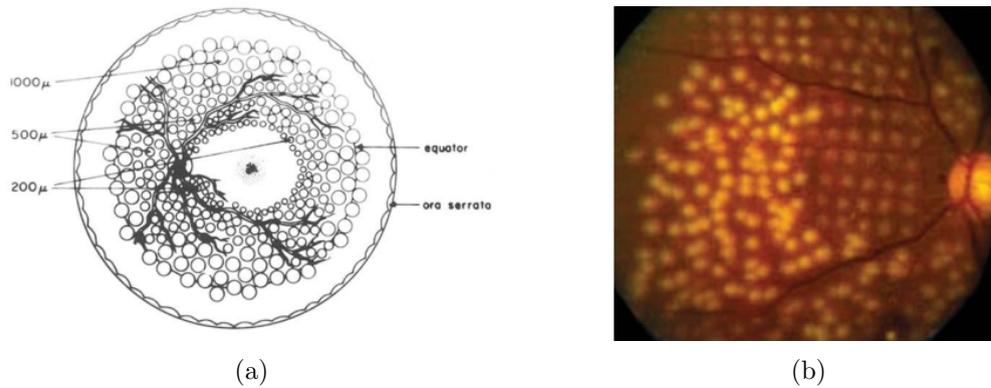


Figure 2.9: Examples of PRP: (a) Pattern of PRP with argon laser [40] (b) PRP with combined argon laser and PASCAL photocoagulation [35].

Panretinal Photocoagulation treatment involves applying a large number of large and intense burns (spot sizes range from $100\ \mu\text{m}$ to $500\ \mu\text{m}$) with pulse durations from 20 to 500 milliseconds. Muqit *et al.* proved that laser mark visibility and spot size increase in a linear relationship for both power and pulse duration. The number of spots delivered to the peripheral retina may vary between 1000 and 3000 depending on the severity of the PR, which means it depends on the nonperfused zones and on the size and extent of neovascularization. The main purpose of PRP is to force the regression of the new vessels' growth in order to delay the progression and limit the damage to the OD and macula and hence preserve the central vision. As a result, there is an improved tissue oxygenation of the peripheral retina [33, 35, 40, 41].

Normally, a PRP treatment session takes about 30 minutes. During a treatment session, the eye must be kept still to ensure accurate delivering of the laser pulses. Moreover, the patient cannot see at all from the treated eye or eyes immediately after treatment. It is a normal reaction and it tends to fade over the following days or weeks.

Doft *et al.* studied the differences between single and multiple treatment sessions of PRP. They concluded that there were no major differences in terms of the effect on visual acuity and DR risk factors. However, the risk for side-effects with single session is much higher. Besides becoming very sensitive to light and having a reduced peripheral vision, the patient may have exudates and may potentially develop DME. In case that situation occurs, PRP must be fractioned over several visits to avoid or minimize worsening of DME [33, 42].

It is possible to verify a positive response to the treatment after a short period of time,

but PRP only achieves full effect after several months or years. For example, in 1985, Little *et al.* performed a study where they did long-term follow-up and repeated PRP when needed on 86 males and 122 females in a total of 351 eyes. After 5 to 12 years follow-up, they concluded that 75% of the total number of eyes had 20/100 or better visual acuity and only 8% had less than 20/400 [40]. This confirms that the reduction of the risk of severe vision loss is verified in more than 90% of the cases.

2.4.2 Macular Edema Photocoagulation

Macular Edema Photocoagulation (MEP) is the treatment of choice for DME. It has some similarities to PRP, specially in terms of equipment. In fact, the conventional argon laser inserted in the slit-lamp system and PASCAL are also used in DME. Moreover, subthreshold diode laser has been used more recently for this type of treatment and it produces very small and almost invisible scars on the retina. However, the clinical procedure is slightly different from PRP.

Diabetic Macular Edema occurs in the macular area, so there is a huge risk that central vision is compromised. For that reason, this treatment is much more difficult and delicate to apply than PRP and it demands more precision, because a slight mistake by the laser operator can displace the laser beam and directly aim to the fovea, causing a potential visual loss. This is a very unlikely occurrence, because anaesthesia is administered, the eye is kept still and the laser operator warns the patient several times to avoid looking directly at the laser beam before the treatment. In MEP therapy, the patient also experiences bright flashes but no pain is felt, since it occurs in the macular area and consequently no special precautions are needed after the photocoagulation. This treatment takes only about 10 minutes and the side-effects are fewer: some reduction in colour vision and presence of small dark spots in central vision might be perceived by the patient in the initial moments after the treatment [33].

Macular Edema Photocoagulation takes the form of focal (Figure 2.10a) or grid (Figure 2.10b) laser photocoagulation, depending on the distribution of the fluid leakages that come from MAs and/or EXs. If the number of leakages is small and they are located in a small area of the thickened retina, the focal laser is more adequate. If there is diffused leakage spread over great part of the area of the thickened retina, the grid laser is the most used. The settings of the laser used and consequent laser scars produced are

different from the settings used in PRP. The number of spots left is only around 200 and the size of each one ranges from $50\ \mu\text{m}$ to $100\ \mu\text{m}$. The duration of each pulse is 100 milliseconds and the power of the laser varies between 70-150 mW, which is much lower than the one used in PRP. If the linear relationship between laser mark visibility plus spot size and power plus pulse duration is taken into account, the combination of all these parameters result in smaller, fewer and much more difficult to observe and detect laser marks compared to the laser marks left by PRP [33, 41, 43].

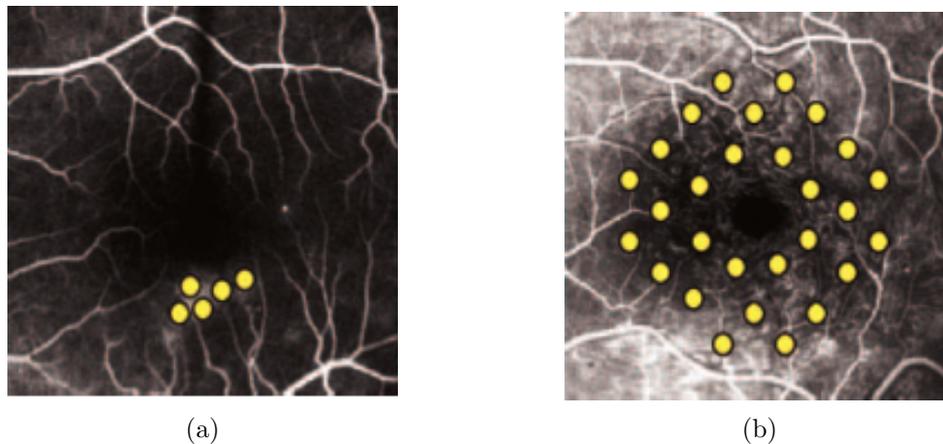


Figure 2.10: Examples of fluorescein angiographies after MEP: (a) Focal Laser Photocoagulation [33] (b) Grid Laser Photocoagulation [33].

The ETDRS performed a study in 1987 [37] where they treated patients with DME using focal and grid photocoagulation and after a three year follow-up, MEP reduced the chance of moderate vision loss in 50% of the cases. Moreover, 12% of the MEP treated eyes and 24% of the control group (non-treated patients) experienced significant visual loss. These results reveal the advantages and effectiveness of MEP. However, 12% of unsuccess is a percentage that exceeds what should happen in a clinical environment. This study was the trigger that initiated the use of anti-VEGFs and steroids to treat DME. Although it is possible to use these substances for both PR and DME, as discussed in subsection 2.3.4.2, their use in clinical studies increased substantially in DME treatment due to the MEP's lower performance in terms of vision loss reduction [37, 39, 44].

Chapter 3

Laser Marks Detection

State-of-the-Art

As mentioned before, classifying retinal images as showing the results of laser treatments or not is important to avoid automatic retinal image processing malfunctions and unnecessary processing. It is a complex task because the detection of laser scars must be efficient and have a very good performance, given that these structures can have different shapes, sizes, colours and brightness. Section 3.1 briefly presents and describes some examples of classification problems where state-of-the-art image segmentation algorithms are used and have applications in other fields besides medical imaging.

Very few papers were produced concerning the detection of scars left by laser photocoagulation in retinal images. A bibliography search uncovered three works: one about laser mark recognition by trained ophthalmologists in different retinal imaging modalities [45] (Section 3.2) and two other describing automatic laser marks detection algorithms [46, 47] (Sections 3.3 and 3.4).

3.1 Classification Problems in Image Processing

Detecting and classifying objects and structures is a problem not only in medical imaging but also in many other image processing application areas. In section 2.4, it was stated that laser marks may present a circular or an asymmetric shape, a bright or a

dark appearance and have different colours. A somewhat similar problem is grain detection and size measurement through visual means, for which several methods have been proposed and published. These methods are also used to identify structures with shapes, colours and brightness similar to the ones described previously in the context of the problem addressed in this thesis. Taking this fact into account, before searching and studying about classification methods, the research in this thesis started with finding segmentation algorithms such as grain detectors in geological images, with the purpose of adapting these algorithms to the identification of laser marks in retinal images.

Kingshuk Choudhury *et al.* [48] proposed an automated grain detection based on a seeded region growing algorithm. The accuracy of this boundary detection was validated by comparison with hand-drawn benchmark image datasets. Although it had a good performance, the author noted that it was not entirely automatic because it required some user intervention. In fact, for the region growing algorithm to work correctly, it is necessary to manually set an initial point (seed) that initiates the algorithm, which means that the user needs to point the region where he or she wants the grain to be detected. Therefore, this is a good segmentation algorithm, but it needs input from the user, which even if small, makes the method not appropriate for this work.

Yingkui Li *et al.* [49] proposed a grain detection applied in the context of a Geographic Information System (GIS). The approach used in this work is different than the previous one, because the priority was to completely avoid the user's intervention. The purpose was to extract the centroids and textural features from geological images and save this information to a database in order to be used as a reference for other works. This algorithm applies a noise reduction and then uses an edge detector. It had a performance accuracy slightly above 70%, but it only worked on geological images exclusively containing grains and no other structure in the images.

Chauris *et al.* [50] proposed a very interesting approach to detect objects with circular shapes. The authors of this work developed a robust classification algorithm using the circlet transform. It is a simple and efficient detector that operates directly on the image gradient using Fast Fourier transforms. It proved to have a very good accuracy in images from different fields: ophthalmology (OD's detection), astronomy (crater detection) and oceanography (detection of eddies).

Kaichang Di *et al.* [51] pointed out in 2014 that most grain and crater detectors were very specific for some particular images. Therefore, in order to add robustness, a machine learning approach to detect craters was proposed and tested in images with different resolution and geological surfaces. The algorithm starts with the detection of the craters with a local analysis using the Circular Hough Transform, extracts Haar-like and local binary pattern features and then a boosting classifier is applied to obtain the final result. The best performance achieved had a sensitivity of 90% and a specificity of 79%.

There are several segmentation algorithms used in retinal images. For example, the algorithms proposed in [52] and [53] are interesting and detect structures that resemble to laser marks. Sekhar *et al.* [52] used the Circular Hough Transform to locate the OD, which is a circular and bright object like most laser marks, only larger in size. Salem *et al.* [53] proposed a segmentation of retinal blood vessels based on the analysis of the Hessian matrix but the authors stated that the Hessian matrix approach can also be used to detect blob-like structures, which is a common characteristic of laser marks.

The examples presented in this section clearly show that a machine learning approach is essential to add robustness and to allow the classification algorithms to be automatic and have good performances, but the task of choosing the best image processing techniques to accurately detect structures in images is demanding.

3.2 Detection of Macular Photocoagulation Scars With Confocal Infrared Reflection Imaging

Macular Edema Photocoagulation using argon laser is a technique that causes some damage to the retina, leaving very few and small scars on the retina that can barely be seen. In clinical practices, it is important to know whether a patient had MEP treatment in previous occasions.

Kotoula *et al.* performed a comparative study with the purpose of evaluating the diagnostic reliability of four imaging modalities in detecting laser scars provoked by MEP. This work analyses the accuracy of ophthalmologists in identifying patients with DME using four different imaging modalities: biomicroscopic fundus imaging, digital colour fundus imaging, fluorescein angiography and infrared reflection imaging.

The authors examined 56 eyes of 56 patients with diabetes mellitus type 2 who had significant DME. From these 56 eyes, 34 of them needed to be treated again and the other 22 demanded treatment for the first time. Every subject that had undergone treatment had MEP applied by a laser operator 12 to 16 weeks before image capture.

Three retina specialists examined these images blindly and independently. On the day of biomicroscopic fundus imaging examination, one of these three was randomly selected to perform a direct examination of the source eye/patient. As for the other three types of images, 56 images of each were given to the three specialists in randomized order. Each specialist had then to answer the following question: Has the patient undergone a MEP? All the answers were then sorted according to the imaging technique in order to obtain the results presented on Table 3.1.

Imaging Modality	Classifications for images with Laser Treatment (n=34)	Classifications for images without Laser Treatment(n=22)	Sensitivity ¹	Specificity ¹
Digital colour fundus	20 yes, 14 no	4 yes, 18 no	59%	81.8%
Biomicroscopic fundus	24 yes, 10 no	0 yes, 22 no	70.6%	86.4%
Fluorescein angiography	31 yes, 3 no	0 yes, 22 no	91.2%	100%
Infrared reflection	33 yes, 1 no	0 yes, 22 no	97%	100%

Table 3.1: Results of the examiners' answer for each imaging modality. Adapted from Table 1 and Table 2 of Kotoula *et al.*[45]

Sensitivity refers to the number of patients with laser marks correctly classified over the total of 34 patients and specificity refers to the number of patients without laser marks correctly classified over the total of 22 patients.

The authors concluded that biomicroscopic and digital colour fundus images have limited value in clinical use when it comes to direct human based detection of laser marks left by MEP. In fact, for both of these types of images the human based classification resulted in low sensitivity performance, which means that the human based detection of laser marks is very poor. On the other hand, fluorescein angiography and infrared reflection proved to be the best options for clinical application. The preference is given to the infrared reflection, because it is the method with the best performance and it is a noninvasive technique, unlike fluorescein angiography.

The main conclusion to retain for this thesis' work is that even ophthalmologists are not able to accurately detect laser marks produced by MEP using digital fundus images, hence it is clear that this thesis' classification problem is challenging.

¹The detailed description about sensitivity and specificity can be found on Appendix A

3.3 Detection of Laser Marks in Retinal Images

João Dias *et al.* proposed another solution to the problem of detecting laser marks in retinal fundus images. The authors of this paper recognised and described the impact and importance of automatically flagging patients that had undergone laser treatments based on images captured during DR Screening Programs.

The proposed laser mark detection algorithm combines the results obtained by two different approaches: Generic Parameter Classification (GPC) and Structural Parameter Classification (SPC). Each one of these approaches return a binary classification and if both outputs are positive, then the retinal image is classified as containing laser marks, otherwise it is considered an image of a non-treated patient, as described in Figure 3.1.

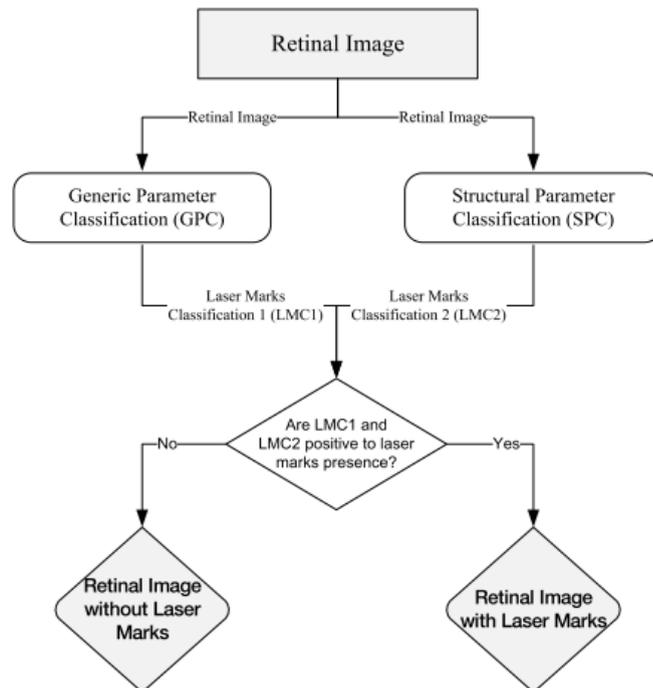


Figure 3.1: Flowchart of the laser marks detection algorithm [46].

The concept behind GPC is very similar to that used in an Image Quality Assessment method proposed in [54]. The fourteen image-related measures used the Quality Assessment work are the input of the GPC. In fact, colour, focus, contrast and illumination are parameters that change and vary between images with laser marks and without them, so these features can distinguish these two classes. Thus, the same methodology is used, but the classes are different. A dataset with 40 retinal images with laser marks and 176 images without laser marks was used with a 4-fold cross validation where 75% of the

dataset was used for training and 25% was used for testing, in order to build the binary classifier.

Structural Parameter Classification has as inputs four features: three of them are related to colour information (CM1, CM2 and CM3) and the other is a single measure (SM) related to the structural information. The colour information is obtained through histogram projection on three specific colourmaps constructed through statistical analysis of 101 retinal images with laser marks and 141 retinal images without laser marks. The colourmaps are *eye-fundus*, *bright marks* and *dark marks* and they are used to perform colour indexing in order to compute CM1, CM2 and CM3, respectively. SM is a value obtained after applying several image processing steps on the retinal image. First, a sobel operator is applied to compute a gradient map of the greyscaled image. Then, five erosion operators with different structuring elements are applied and summed to obtain a binary image. A morphological opening is used to remove small objects from the binary image and finally SM is calculated by computing the percentage of the total of foreground pixels that are within laser marks. The retinal images and the validation process used in GPC are the same as the ones used in SPC. The only difference is that CM1, CM2, CM3 and SM are inputs to a Feed Forward Neural Network.

To evaluate the performance of this algorithm (Table 3.2), the authors used a public dataset containing 1200 retinal images without laser marks called Messidor [55]. Also, 996 retinal images were collected for the algorithm's performance evaluation from the ongoing DR screening program in the centre region of Portugal, with 101 of those showing laser marks.

Dataset	Classification Accuracy
Laser Treatment (n=101)	63.37%
No Laser Treatment (n=2095)	99.90%

Table 3.2: João Dias *et al.* laser mark detection algorithm performance. Adapted from Table 1 of [46].

Considering the images classified with laser marks as the positive case, then Sensitivity = 63.37% and Specificity = 99.90%. In the final remarks, the authors state that this result is noteworthy given the complexity of the classification problem. However, this is the first paper about this topic and hence some further work to improve the algorithm is suggested, such as using or adding spatial dispersion metrics and different classifiers and decision methods.

The most outstanding aspect of this work is the fact that specificity is close to 100%, which means that almost no retinal image without laser marks is incorrectly classified. This is of great importance, because it is not clinically acceptable to wrongly identify a patient as having undergone treatment, when he or she can potentially develop DR, because it has not been treated yet. However, the sensitivity value is not very high and the diversity of the dataset used is poor in terms of resolution, models of fundus cameras and even types of laser marks. For instance, after inspecting the dataset, it was found that the images with laser photocoagulation scars were captured with the same model of camera and all of them had a significant number of large-sized PRP-related laser marks. These facts may indicate that the algorithm proposed in [46] is in fact detecting camera and type of mark specific features. Since this is an algorithm to detect visual objects and structures on retinal images, visual spatially based segmentation techniques should play a major role.

3.4 Laser Marks Detection From Fundus Images

Faraz Tahir *et al.* work is another published solution proposed to the problem of automatically detecting laser marks. Its authors state that laser scars hinder other automatic systems and that treated patients represent unnecessary workload for graders in DR Screening Programs.

The proposed algorithm is divided into four steps: pre-processing, candidate region identification, feature computation and classification.

Pre-processing: Retinal images are often noisy and lack proper illumination. To reduce the effect of these problems, all the processing is performed only on the green channel, which usually has better contrast. Then, a circular averaging filter is applied to blur the image and remove the noise and then adaptive histogram equalization is applied to enhance the contrast and make the laser marks fairly visible. The uneven illumination is corrected by estimating the background illumination and subtracting it to the image (top-hat filtering).

Candidate Region: The purpose of this step is to identify all possible laser marks, even if that leads to the detection of some undesirable regions that are false detections.

A binary image identifying the masked candidate regions is computed by thresholding the previous stage output, with a small-valued threshold.

Feature Computation: For each candidate region identified in the previous step, 10 features are computed: Compactness, Max hue, Max saturation, Standard deviation of saturation, Intensity mean, Intensity max, Intensity standard deviation, Mean Red channel, Max Red channel, Max Green channel.

Classification: In this final step, the false regions are removed from the candidate regions set. Then, a Minimum Distance classifier is used to decide whether the retinal image has laser marks or not, using the 10 features computed as input.

To evaluate the performance of the proposed algorithm (Table 3.3), 51 retinal images with laser marks and 329 retinal images without laser marks were used. These images were taken with a *Topcon* mydriatic camera, have a resolution of 1504 by 1296 pixels and were classified by ophthalmologists. The dataset also contains some images from normal patients and others with different levels of retinal disease but without any laser treatment.

Dataset	Classification Accuracy
Laser Treatment (n=51)	94%
No Laser Treatment (n=329)	97%

Table 3.3: Faraz Tahir *et al.* laser mark detection algorithm performance. Adapted from Table 3 of [47].

The results presented indicate that the proposed algorithm has a really good performance and is quite acceptable for a clinical environment, despite of the fact that Specificity is not 100%. Although the performance results provided in the paper are remarkable, the number of images containing laser marks is small and there is no explanation about the method used to choose the threshold used to obtain the maximum possible laser marks (even if there are undesired false detections), there is little explanation on how this value is computed to obtain the binary image.

Chapter 4

Automatic Laser Mark Detection Algorithm

As stated in the introduction, the purpose of the implemented algorithm is to detect if the patient had already undergone treatment in order to avoid malfunctioning of further processing steps, unnecessary processing and to reduce the burden of the photographer and of the grader relieving them of the cumbersome direct observation and analysis of all the images captured during screening sessions. The implemented laser mark detection algorithm returns two possible outputs: "Laser" or "No Laser" (as illustrated in the two cases of Figure 4.1).

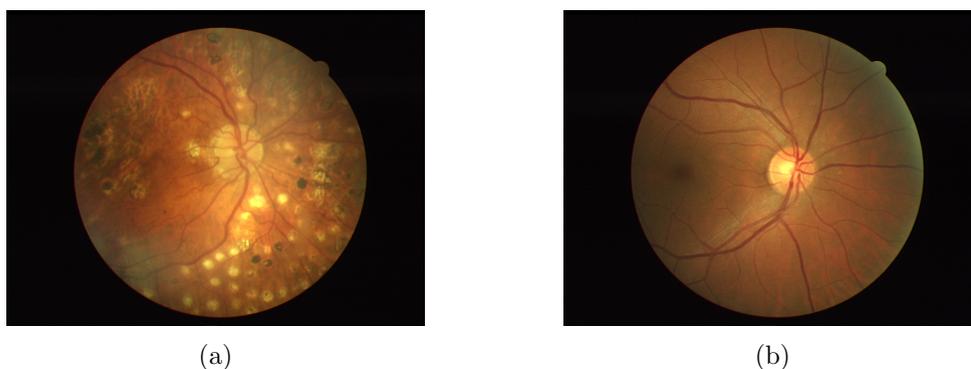


Figure 4.1: Examples of (a) "Laser" and (b) "No Laser" retinal images.

The approach followed in this thesis follows the procedure typical for this type of image classification problem and involves pre-processing steps, segmentation techniques, features computation and a final binary classification. Figure 4.2 summarizes the workflow of this algorithm and this chapter describes in detail each step represented. This chapter

also includes a section at the beginning (Section 4.1) characterising the image datasets used in this thesis.

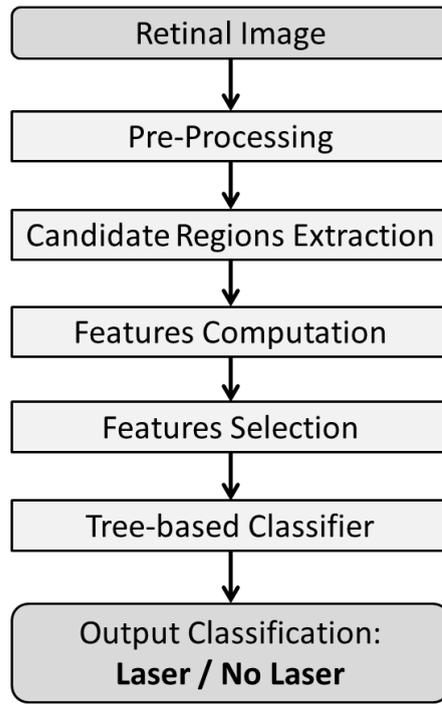


Figure 4.2: Flowchart of the proposed algorithm.

The retinal image from Figure 4.1a is the file named 0087.jpg from the Screening Dataset (described in section 4.1) and it is used as the main example throughout chapter 4. Other examples can be found on Appendix C.

4.1 Materials

In order to develop and test the performance of the proposed algorithm, eight public datasets and three proprietary datasets were used.¹

¹For more detailed information, please consult Appendix B

4.1.1 Public Datasets

The following eight datasets result from other studies and they are available online. All these datasets are labeled as "No Laser".

Messidor (M) [55]: 1200 retinal images without laser marks from three ophthalmology research departments in France. All of the images were captured using a colour video 3CCD camera on a *Topcon TRC NW6 non-mydratic retinograph* with a 45° field of view (FOV). The images have a resolution of 1444 by 960 pixels, 2240 by 1488 pixels or 2304 by 1536 pixels. For this work, 13 images from this dataset were not used as they presented laser scars, according to image processing experts from Retmarker S.A. These 13 images are listed on Table B.1.

e-ophtha MA (EOMA) [56]: 148 images with a total of 1306 MAs. The images have a resolution of 1440 by 960 pixels, 1504 by 1000 pixels, 2048 by 1360 pixels or 2544 by 1696 pixels. These images are part of ANR-TECSAN-TELEOPHTA project funded by the French Research Agency.

e-ophtha No MA (EONMA) [56]: 233 images of healthy patients without any signs of DR and treatment. The images have a resolution of 1440 by 960 pixels, 1504 by 1000 pixels, 2048 by 1360 pixels or 2544 by 1696 pixels. These images are part of ANR-TECSAN-TELEOPHTA project funded by the French Research Agency.

e-ophtha EX (EOEX) [56, 57]: 47 images with a total of 12,278 EXs. The images have a resolution of 1440 by 960 pixels, 1504 by 1000 pixels, 2048 by 1360 pixels or 2544 by 1696 pixels. These images are part of ANR-TECSAN-TELEOPHTA project funded by the French Research Agency.

e-ophtha No EX (EONEX) [56, 57]: 35 exudate-free images and contain structures such as reflections and optical artifacts which can mislead automatic detection algorithms [57]. The images have a resolution of 1440 by 960 pixels, 2048 by 1360 pixels or 2544 by 1696 pixels. These images are part of ANR-TECSAN-TELEOPHTA project funded by the French Research Agency.

Vessel-Based Registration (VBR) [58]: 22 retinal images without laser marks. These images were acquired with *Topcon 3D OCT-1000 Camera* and have a resolution of 1200 by 1143 pixels. Although this public set consists of 22 pairs of images (17

macular and 5 prepapillary) from different patients, each of these pairs contains one OCT image and one colour fundus images. This last imaging modality mentioned is the one used in this work.

50 Healthy People (HP) [19]: 100 retinal images without laser marks. These colour fundus photographs come from left and right eyes of 50 healthy volunteers and have a resolution of 1612 by 1536 pixels.

Foveal Avascular Zone Detection (FAZD) [25]: 60 retinal images without laser marks. This dataset includes 25 retinal images of healthy subjects and 35 retinal images of patients with DR. Each image has a resolution of 720 by 576 pixels. It is important to note that on this study's website it is stated that there are 30 images of patients with DR. However, their download link contains 35 images with DR lesions.

4.1.2 Proprietary Datasets

Screening (S): 203 retinal images classified by optometrists as having laser marks from an ongoing DR screening program in the centre region of Portugal, managed by ARS Centro. All of the images are non-mydratic and have a 45° FOV and they were acquired between the 7th week of 2014 and the 3rd week of 2015. There are 26 images captured using *Nidek AFC-330 Retinal Camera* and with a resolution of 1920 by 1920 pixels. The remaining 177 images were captured using *Canon CR6-45NM Retinal Camera* and have a resolution of 768 by 584 pixels.

It also includes 419 retinal images without laser marks from the same screening program in the same time interval. There are 68 images that were captured using *Nidek AFC-330 Retinal Camera*: 18 have a resolution of 2448 by 2448 and 50 have a resolution of 1920 by 1920 pixels. There is one image captured using *CSO Cobra Retinal Camera* with 60°x45° FOV and a resolution of 1624 by 1232 pixels. The remaining 350 images were captured using *Canon CR6-45NM Retinal Camera* and have a resolution of 768 by 584 pixels. Amongst the 419 images without laser marks, six were graded by ophthalmologists with MAs, six were graded with EXs and 20 were commented as having drusens (AMD-related lesions).

Each image corresponds to a different patient, so the total number of patients that contributed to the constitution of this dataset is equal to the total number of images: 622. The Table B.2 and Figure B.1 are a summary of the structure of this dataset.

Before and After Treatment (BAT): 34 retinal images with laser marks and 15 without laser marks. The images were kindly provided by Centro Cirúrgico de Coimbra (CCC). This dataset is composed by images from 9 patients. Each patient contains retinal fundus images from before and after treatment. The images with the letter "A" in their filenames are labeled as "Laser" and the images with the letter "B" in their filenames are labeled as "No Laser". The Table B.3 and Figure B.2 are a summary of the structure of this dataset.

Proprietary Dataset João Dias (PDJD) [46]: 101 retinal images with laser marks from a DR screening program in the centre region of Portugal before 2013, managed by ARS Centro. All of the images are non-mydratic and have a 45° FOV and were captured using *Canon CR6-45NM Retinal Camera* and have a resolution of 768 by 584 pixels. This dataset contains images selected by image processing experts and they were not classified by any optometrist or ophthalmologist.

4.2 Image Pre-Processing

Before analysing the content of the retinal images, each image is subjected to a pre-processing step. This is one of the most important phases because the photographic protocol, the image acquisition technique and the camera itself may not allow to clearly identify all the different anatomical structures of the retina. For these reasons, getting a "clean" version of the original retinal image is a huge progress towards having an efficient segmentation that allows a good performance of the algorithm in the end. Figure 4.3 shows the sequence of techniques applied in this pre-processing step.

The following subsections will cover the different operations represented in Figure 4.3. Subsection 4.2.5 was added to explain how the sizes of the kernels used throughout the pre-processing and following sections were chosen.

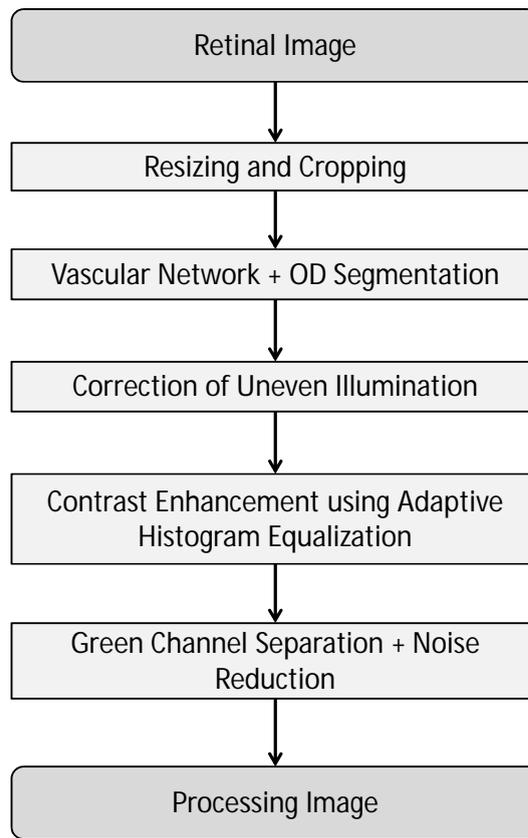


Figure 4.3: Flowchart of the pre-processing step.

4.2.1 Retinal Image Resizing and Cropping

Image resizing is performed only in case the image diagonal is smaller than 400 pixels or larger than 1500 pixels. When these conditions are met, this resizing step changes the number of rows and columns so that the corresponding diagonal is within the 500 and 1400 pixels, using a bicubic interpolation.

Then, the resized image is cropped circularly leaving only the central region of interest (ROI) of the retinal image. This operation uses a mask similar to the one shown in Figure 4.4a. The result of this cropping operation (application of the circular ROI mask on the retinal image) is shown in Figure 4.4b.

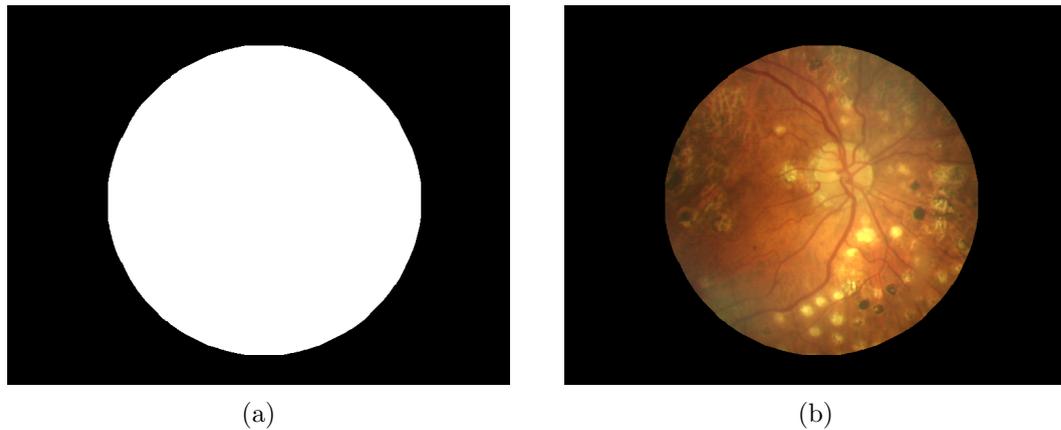


Figure 4.4: (a) Region of Interest (ROI) and (b) cropped version of the original retinal image.

4.2.2 Vascular Network and Optic Disc Segmentation

In any digital retinal fundus image, there are at least three identifiable anatomical structures: the Vascular Network, the Optic Disc and the Fovea. Other structures and lesions can be seen in a retinal image depending on the patient's condition, but these three are common to every person.

During a treatment session, the operator cannot aim the laser to any of these anatomical structures. As it was already explained in chapter 2, if the patient looks directly to the beam, the central vision (fovea) is immediately compromised. Also, the laser beam can be used to stop and destroy growing blood vessels, but does not leaves scars in the vascular network. Moreover, photocoagulation treatment is not applied to the OD and no scars will occur in its vicinity, for clinical reasons.

Therefore, since there will not be any laser marks on the regions co-located with the vascular tree, OD and fovea, these regions can be excluded from further processing performed to identify candidate laser marks.

A mask for the Vascular Network and a mask for the OD were obtained and the respective binary images shown in Figure 4.5 will be useful for further processing, specially for extracting Candidate Regions (Section 4.3).

The extracted vascular network is used to remove false detections during Candidate Regions Extraction. Since laser marks are not found on top of the vascular tree or in its vicinity, as a first step the vessels in the vascular tree binary mask are enlarged through the application of a dilation operation as described in Equation 4.1

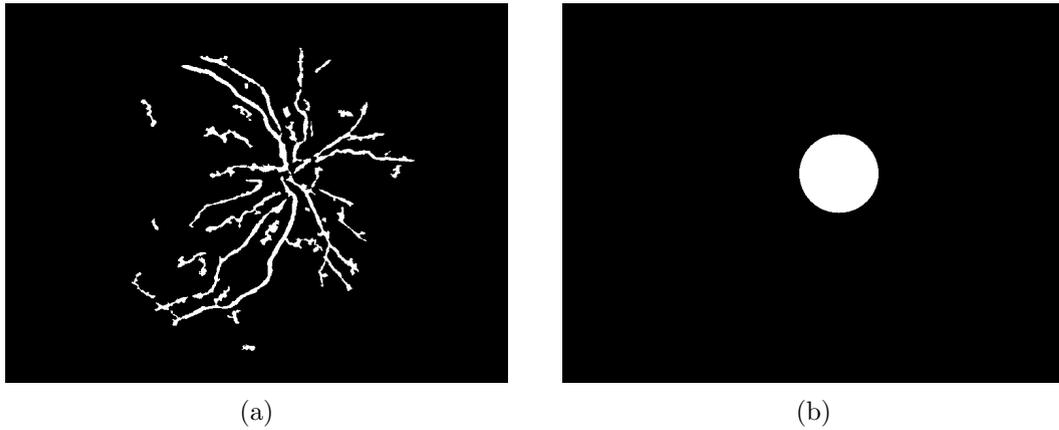


Figure 4.5: (a) Vascular Network mask and (b) OD mask of the original retinal image.

$$vessel_{final_mask} = vessel_{original_mask} \oplus k_{vessel} \quad (4.1)$$

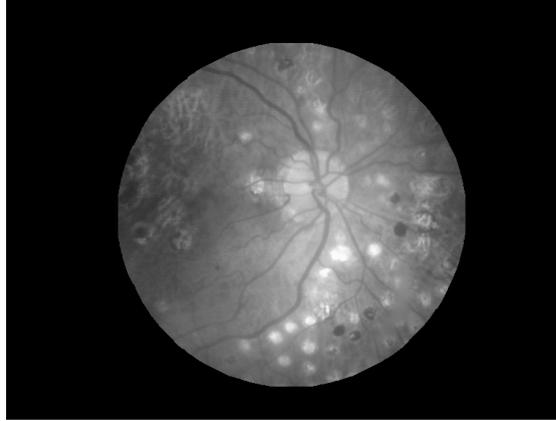
where $vessel_{original_mask}$ is the vascular network extracted, \oplus represents dilation and k_{vessel} represents the circle-shaped structuring element. The size of this kernel is defined in subsection 4.2.5).

The fovea was not extracted because not only it is a very small point (very few pixels), but also it was not easily detectable or present on every image used in this work, since some field 1 retinal images did not cover it. The blood vessels extraction is based on the Contourlet Transform [59]. The OD identification method is based on [60].

4.2.3 Uneven Illumination Correction and Contrast Enhancement

Poor illumination of the retina and other problems occurring during the image capture can result in fundus images with large spatial variations in the locally averaged luminances. This is a common problem that brings additional difficulties to the identification of the anatomical and pathology-related structures of the retinal image. There are several state-of-the-art techniques to correct this illumination issue such as top-hat filtering [61, 62], homomorphic filtering [62] and changing the Value channel in HSV colorspace [63].

In this work, the retinal image is converted to the $L^*a^*b^*$ colorspace [64] in order to correct the uneven illumination by changing the L^* color component (Figure 4.6). Each value of L^* represents the brightness of the corresponding pixel.

Figure 4.6: L^* color component of the retinal image.

Firstly, a local means image (L_{lm}) is computed from the L^* component (Equation 4.3). This image was obtained by performing a 2D convolution between L^* and the kernel k_{lum} (Equation 4.2, where J is a square matrix of ones). The size of this kernel is c_{lum} and the value of each element is $\frac{1}{c_{lum}^2}$. For the retinal images with a resolution of 584 by 768 pixels, c_{lum} value was set to 28 after observing the result of the uneven illumination correction with c_{lum} varying between 20 and 40.

$$k_{lum} = \frac{1}{c_{lum}^2} \times J_{c_{lum}} \quad (4.2)$$

$$L_{lm} = L^*(x, y) * k_{lum}(x, y) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} L^*(n_1, n_2) \cdot k_{lum}(x - n_1, y - n_2) \quad (4.3)$$

In other words, L_{lm} is a map of luminance values where each pixel is equal to the mean of L^* in the pixel's neighbourhood c_{lum} by c_{lum} and it is normalised to $[0,1]$. Then, L^* is divided by L_{lm} plus a constant in order to obtain the final L^* color component. The purpose of this mathematical expression (Equation 4.4) is to reduce the luminance of the brightest areas (higher values in L_{lm}) and increase the luminance in the darkest areas (lower values in L_{lm}). The value of the constant depends on the mean value of L^* and it was determined empirically by observation: if the mean is lower than 0.32, the constant is 0.45, otherwise this constant is equal to 0.3. The uneven illumination problem is solved by replacing L^* with L^*_{final} (Figure 4.7).

$$\begin{cases} L^*_{final} = \frac{L^*}{L_{lm} + 0.45}, & \text{mean}_{L^*} < 0.32 \\ L^*_{final} = \frac{L^*}{L_{lm} + 0.3}, & \text{mean}_{L^*} > 0.32 \end{cases} \quad (4.4)$$

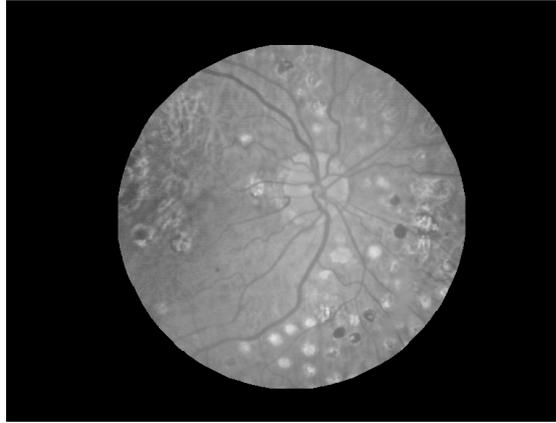


Figure 4.7: L^* color component after uneven illumination correction (L^*_{final}).

Most laser marks present a bright appearance and to take advantage of this characteristic an adaptive histogram equalization step is applied to the illumination corrected L^* channel, improving the contrast between the laser marks and the background and enhancing the brightness. The result of this color-domain image enhancement is shown in Figure 4.8.

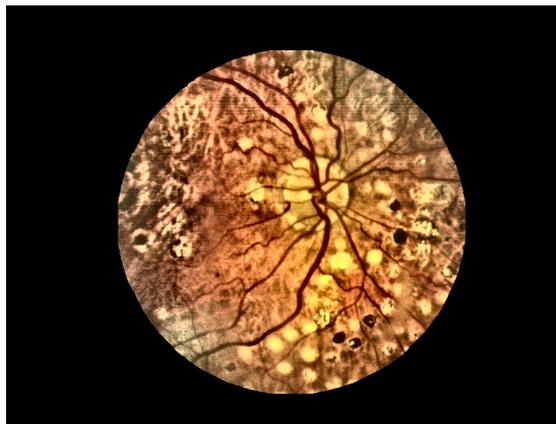


Figure 4.8: Retinal Image after Adaptive Histogram Equalization in the $L^*a^*b^*$ colorspace.

4.2.4 Green Channel Separation and Noise Reduction

The image is then converted back to the RGB colorspace. At this point, several channels are extracted and saved because they will be useful to calculate some features. The Red channel from the RGB colorspace, the Hue and the Saturation channels from the HSV colorspace are obtained in order to be used for intensity-based descriptors in subsection 4.4.4.

Further processing is done using only the information from the green channel. Not only there are bright laser marks, but also there are several with a dark appearance. However, most of these dark laser marks are green, so extracting the green channel results in an image where the intensity of the pixels containing the laser marks, whether they are bright or green, is very high. On the other hand, the pixels of the vascular network have a very low intensity, which means that the green channel shows the best contrast between anatomical structures and background [65].

An additional filtering of the green channel is then performed using a 5x5 median filter (for the retinal images with a resolution of 584x768) to reduce the noise. The result of this latest operation is represented in Figure 4.9 which shows a gray-level encoding of the green-channel image obtained after median filtering.

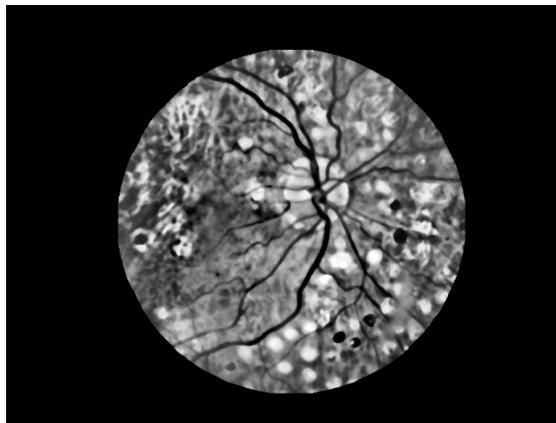


Figure 4.9: Final image resulting from the pre-processing step.

4.2.5 Kernels and Constants

Some morphological operations and mathematical expressions in image processing require choosing some parameter values and kernel sizes. In this case, the main objective is to develop an automatic and robust algorithm, so a criterion is necessary to set these values automatically because we desire a method that requires as little human intervention as possible.

The sizes of the filter kernels have a considerable impact on the output of the image processing operations. It is therefore very important to take into consideration the ROI area, the FOV angle, the brightness and the size of the anatomical structures of the retinal images.

The first experiments of the proposed algorithm were performed on several images with a resolution of 584x768 from the Screening Dataset and some constant values and kernel sizes were determined for these images. However, after testing the algorithm with these same values on other datasets, the results were far from acceptable. For this reason, it was decided to adopt a scaling approach such that some constants and parameters, such as kernel sizes, are scaled according to the major characteristics of the image, such as ROI and OD size. This scaling uses a parameter called *scaling_constant* defined shortly (Equation 4.5).

After carefully inspecting a wide range of retinal images, it is possible to conclude that:

- There is a wide variety of different resolutions, so it is important to obtain measures independent to resolution. In other words, relative distances and sizes are the measures used in this work.
- A lower FOV angle usually leads to a higher ROI area, so the anatomical structures are larger, if the resolution is kept constant. In another perspective, it is similar to the effect of a zoom-in command.
- Laser marks are typically about 20%-50% of the OD's size, so this anatomical structure is a good term of comparison. The application of these relative measurements is not a novelty in ophthalmology, since it is common to use the OD's size as the measurement unit.

The images with a resolution of 584x768 were used as the reference. Every ROI of these images have the same area, so does the ($ROI_diameter = 480$) pixels. Most of the segmented ODs have a $OD_diameter = 110$ pixels and the observed ODs vary between 100 and 120 pixels. Taking the stated observations into account, for each image, $ROI_diameter$ and $OD_diameter$ are calculated and then divided by the reference (denominator of Equation 4.5) in order to obtain *scaling_constant*.

$$scaling_constant = \frac{ROI_diameter - OD_diameter}{480 - 110} = \frac{ROI_diameter - OD_diameter}{370} \quad (4.5)$$

Let us take c_{lum} as the example to explain the application of *scaling_constant*. For the images with a resolution of 584x768, $c_{lum} = 28$. However, in order to adapt this parameter to different retinal images, this luminance constant was changed to $c_{lum} = 28 \times \textit{scaling_constant}$, where 28 was obtained after a series of empirical experiences. The same happens to every other kernel size used in this thesis except for $k_{vessels}$, which is slightly different. Most kernel sizes were determined after several experiments in images with a resolution of 584x768, for example by varying a certain kernel size in an interval of values and selecting the value that produced better results. The value of $k_{vessels}$ was obtained according to the description in [57]: the maximal width of vessels is estimated to be $\frac{ROI_diameter}{74}$. The purpose of the filter used to enlarge the vessels is to avoid detections of laser marks in the areas nearby the vasculature, the size of the kernel size was determined to be the double of the maximal width of the vessels, hence $k_{vessels} = 2 \times \frac{ROI_diameter}{74} \times \textit{scaling_constant}$.

To clarify how the Equation 4.5 was obtained, let us take the images from the EOEX dataset for example. After the resizing step, every image of this dataset has a resolution of 1024x683 pixels. As a consequence, the anatomical structures are larger and therefore a larger kernel is needed. This part of the problem is solved by the fact that *ROI_diameter* is greater than the diameter of 480 pixels from the reference ROI, so initially it was set that $\textit{scaling_constant} = \frac{ROI_diameter}{480}$. However, there is still an issue to consider: the use of this initially proposed *scaling_constant* in images with low ROI size from the S dataset and in images with high resolution from the BAT dataset was making the c_{lum} and the minimum radius used in the segmentation algorithms (Section 4.3) so high that the small-sized laser marks were not being detected as candidates, so the task of detecting images "Laser" became more difficult. For this reason, *OD_diameter* is subtracted from *ROI_diameter* with the purpose to slightly lower this constant, so that more laser marks can be detected even if some false detections also occur. Also, retinal images with a large OD also have thick vessels, so if a high *scaling_constant* value is used, the masking operation using the binary image containing the blood vessels (Figure 4.5a) after applying the morphological operation in equation 4.1 may incorrectly lead to the removal of laser marks detected by the segmentation algorithms. In these situations it is also important to reduce the value of the constant, hence the subtraction operation solves this problem. Some retinal images illustrating this issue can be found on Appendix C.

4.3 Candidate Regions Extraction

As a first step to obtain the features that will be used to classify the retinal images, the proposed algorithm identifies small patches of pixels which represent potential laser scars. The detection of these candidates is performed using three different algorithms: Circular Hough Transform (CHT), Frangi Vesselness Filter (FVF) [66] and an algorithm proposed in [67], which will be referred as Laser Mark Segmentation (LMS). Each of these algorithms will be discussed in detail in the following subsections. In the end, the outputs of each of these three segmentation algorithms are merged to obtain the entire set of candidate laser marks.

4.3.1 Circular Hough Transform

Hough Transform is a very popular method introduced by Paul Hough and used in Computational Vision that efficiently identifies lines and circles in images. This transform has several variants and each of them has applications in various fields, specially in automated digital image analysis for shape detection [68].

One of these variants is the Circular Hough Transform which was firstly introduced by Duda and is used to detect circular shapes in digital images. It is very used in medical imaging, for example to detect welded joints in radiographies [69] and even to detect MAs [70, 71] and the OD [60] in retinal images. Since in many cases the laser beam leaves circle-like scars, CHT is one of the best operators one can use to detect laser marks, also due to its robustness in the presence of noise, occlusion and varying illumination.

Circular Hough Transform is a technique equivalent to a convolution between the image and a circle operator. Firstly, an edge detector is applied to the processing image in order to obtain a binary image space of edge points. Each edge point in the image space contributes a circumference of radius R to a parameter space or accumulator array. In other words, for each pixel of the binary image with edge points, a circumference with a desired R is drawn in the parameter space. Figure 4.10 represents schematically the concept of this step of the algorithm [70, 72].

If enough circles with a certain radius R drawn in the parameter space intersect in the same point, it can be concluded that a circumference with that radius R is found at

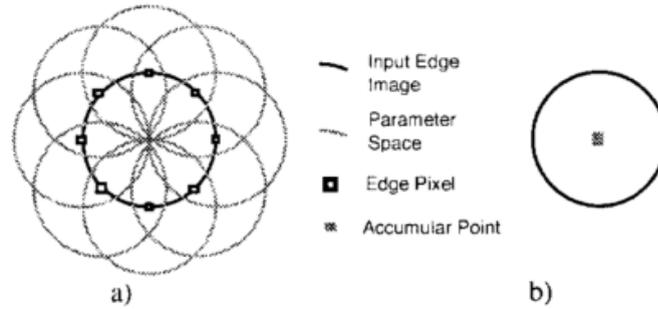


Figure 4.10: Brief explanation of the CHT algorithm: a) Contribution of edge points to the accumulator space b) Edge point contribution to a single accumulator point [72].

the pixel (x_0, y_0) , which is the center of the found circumference. Therefore, the task of detecting shapes in the image space is simplified to peak finding in an accumulator array, as shown in Figure 4.11 [69, 70, 72].

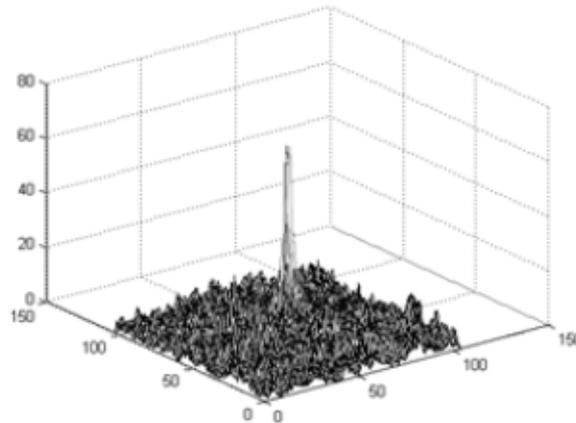


Figure 4.11: Example of an accumulator array containing a peak corresponding to a detected center of a circle resulting from CHT [70].

In the proposed algorithm, the accumulator array is normalised to $[0,1]$ and the threshold value to consider a pixel as the center of a circumference is set to 0.225. The interval of radius values used during CHT varies between 6 and 30 pixels for images with a resolution of 584x768.

Obtaining the binary image with the edge points is a defining phase that has a considerable impact on the final output of the CHT algorithm. Sobel and canny edge detectors are the most popular state-of-the-art techniques and the last one mentioned is usually used in CHT. The canny edge detector smooths the processing image with a gaussian filter to reduce noise and then computes gradients using sobel operators to find the edges where the change in grayscale intensity is maximum. A non-maximum suppression is

carried out to preserve all local maxima and finally a double threshold is applied to get the final binary image with the edge points.

However, the authors in [73] and [74] pointed out two problems: although the gaussian filter removes the noise, it also smoothes and weakens the important edges which increases false edges detection and the two thresholds used are set manually. These problems were also verified when applying the traditional CHT in these retinal images. As a consequence, Gao Jie and Liu Ning proposed the use of a bilateral filter [75] instead of gaussian filter and the use of Otsu's method for thresholding [76]. These two suggestions were applied in the proposed algorithm and proved to be very effective. In fact, the bilateral filter not only smooths the image but also sharply detects and preserves the edges, reducing the probability of false edges.

The output of the proposed CHT is a vector for each laser mark detected containing the pixel coordinates of the center of the detected circle, the corresponding radius and the value of the center in the accumulator array, which will be called l_{CHT} from now on because it indicates how similar to a circle the detected laser mark is. Figure 4.12 shows a binary image with circles that represent the output of the CHT algorithm.

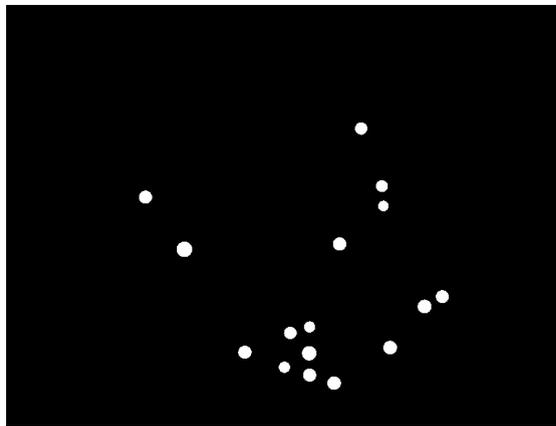


Figure 4.12: Binary image of the detected laser mark candidates using Circular Hough Transform.

4.3.2 Frangi Vesselness Filter

Alejandro Frangi *et al.* [66] developed a filter that measures vesselness of medical images based on the eigenvalues of the Hessian of the image. They tested their algorithm on Digital Subtraction Angiography (DSA) and Magnetic Resonance Angiography (MRA)

images, proving that the Frangi Vesselness Filter shows very good noise and background suppression and vessel enhancement in both 2D and 3D images.

The algorithm proposed by Frangi has been widely used in different imaging modalities. For example, FVF can be used in 3D lung vessel segmentation [77], in coronary artery segmentation in Computed Tomography (CT) images [78, 79] and even in vessels enhancement in digital colour fundus images [53, 80, 81].

The FVF is based on a multiscale vesselness approach that detects tubular, ridge and blob-like objects in medical images. In fact, when FVF is applied on retinal images with photocoagulation scars, not only the vascular network is detected but also the laser marks that have a blob-like shape. For this reason, FVF is another of the three operators used to identify potential laser marks. After applying the FVF to the retinal image, a binary image containing the blood vessels and the laser marks is obtained and since a mask with the vascular network is obtained for each image during pre-processing (Figure 4.5a), the vessels resulting from FVF are removed, keeping solely the laser marks as the final output.

Frangi Vesselness Filter searches for geometrical structures which can be regarded as tubular and blob-like and these can appear in different sizes. Therefore, a multiscale approach is used, so this is an iterative algorithm that merges the vesselness measure computed for each scale size σ . For each iteration, the retinal image is convolved with a Gaussian filter with a kernel of size σ . Then, the Hessian matrix is calculated, according to Equation 4.6, where I is the processing image.

$$H(I) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \quad (4.6)$$

From the Hessian matrix, the eigenvalues λ_1 and λ_2 are calculated with $|\lambda_1| < |\lambda_2|$. The idea behind eigenvalue analysis of the Hessian is to extract the principal directions in which the local second order structure of the image can be decomposed. Therefore, low λ_1 values occur when the structures are tubular and high λ_1 values occur when the structures are blob-like. Also, if $\lambda_2 < 0$, the structures are bright, otherwise the objects have a dark appearance. Figure 4.13 is Table 1 reproduced from [66] that summarizes the type of objects detected as a function of the Hessian's eigenvalues.

2D		3D			orientation pattern
λ_1	λ_2	λ_1	λ_2	λ_3	
N	N	N	N	N	noisy, no preferred direction
		L	L	H-	plate-like structure (bright)
		L	L	H+	plate-like structure (dark)
L	H-	L	H-	H-	tubular structure (bright)
L	H+	L	H+	H+	tubular structure (dark)
H-	H-	H-	H-	H-	blob-like structure (bright)
H+	H+	H+	H+	H+	blob-like structure (dark)

Figure 4.13: Possible patterns in 2D and 3D, depending on the value of the eigenvalues.

N - Null, L - Low, H - High. It is a copy from Table 1 in [66].

For 2D images, Frangi originally proposed a vesselness measure for each scale V_σ , according to Equation 4.7

$$V_\sigma = \begin{cases} 0, & \lambda_2 > 0 \\ \exp(-\frac{R_b^2}{2\beta^2})(1 - \exp(-\frac{S^2}{2c^2})), & \lambda_2 < 0 \end{cases} \quad (4.7)$$

where $R_b = \frac{\lambda_1}{\lambda_2}$ is the blobness measure and $S = \sqrt{\lambda_1^2 + \lambda_2^2}$ is the Hessian's matrix norm, also known as Frobenius norm. If the value of this norm is low, it means that V_σ is low in the background pixels, where no structure is present and the eigenvalues are small for the lack of contrast. The parameters β and c control the sensitivity of FVF to the measures R_b and S respectively. The authors in [82] suggest that $\beta = 0.5$, $c =$ half of the maximum value of the Frobenius norm and that the set of σ values to be used is $\{1, 3, 5, 7, 9\}$. In this work, c and the set of σ values used are the same as in [82], but β is equal to 0.25 because using the recommended value resulted in the detection of noisy pixels, so the value of this parameter had to be lowered in order to avoid this situation.

In the proposed algorithm, there is another slight but very important change compared to the original FVF. Since the processing image results from the green channel of the original RGB image, the vessels have the lowest intensity so they have a dark appearance. As a consequence, the condition used for $V_\sigma = 0$ in Equation 4.7 is $\lambda_2 > 0$. Another reason for changing this condition is because the CHT already searches for bright objects, so it is important to have an algorithm that detects the dark laser marks, which is the role of the FVF used in the proposed algorithm.

FVF returns a grayscale image where the presence high intensity pixels means a potential vessel or blob-like structure. A threshold value is computed using Otsu's method in order to get a binary image. Afterwards, the pixels belonging to the OD and Vascular Network masks are removed. Also, an area opening operation is performed to remove the small regions detected, so regions with an area lower than 100 pixels for images with a resolution of 584x768 are removed to obtain the output binary image of the FVF algorithm, such as the one shown in Figure 4.14.



Figure 4.14: Binary image of the detected laser mark candidates using Frangi Vesselness Filter.

4.3.3 Laser Mark Segmentation

The third and last segmentation algorithm used was based on a dissertation by Sohini Roy Chowdbury from University of Minnesota titled "Automated Segmentation and Pathology Detection in Ophthalmic Images" [67] to complete her PhD degree. In this dissertation, several segmentation algorithms are proposed: automated vessel segmentation, automated OD segmentation, automated non-proliferative DR detection, automated PR detection and automated segmentation of OCT images. The automated PR detection is proposed in chapter 6 of Sohini's work and one of the steps is the detection of laser scars and fibrosis described in section 6.2.1 [67].

The author suggests an initial pre-processing step not very different to the one described in section 4.2: the green channel from each retinal image is extracted and the pixel intensities are scaled in the range [0,1] followed by contrast enhancement.

A morphological operation - erosion - is performed on the image (I) using a circular structuring element and then an image reconstruction is performed on I , using the output

of the erosion operation as the mask. Finally, the reconstructed image is subtracted from I and subjected to contrast enhancement in order to obtain the final image that will be used in the LMS segmentation (I_{LMS}). The circular structuring element used by Sohini has a radius of 15 pixels for images with a resolution of 500x500. In the proposed algorithm, the structuring element also has a radius of 15 pixels for a resolution of 584x768.

For each retinal image, a threshold value is computed and applied to I_{LMS} using Otsu's method. Then, the pixels superimposing the Vascular Network and the OD are ignored. Finally, two discriminating parameters are computed in order to select which regions of I_{LMS} represent potential laser marks. These parameters are:

- Roundedness (f_1) - it is the ratio of the major to minor axis length of the region. Most laser marks have a circular shape, as it was stated when describing the CHT algorithm. Therefore, $f_1 \geq 1$ and $f_1 = 1$ for regions that have a circular shape.
- Solidity (f_2) - it is the ratio of the area to the convex hull bordering the region. It is a parameter with a range of $[0,1]$ and it is high for regions that are convex, which is characteristic for laser scars.

A candidate region is considered a potential laser mark if $1 < f_1 < 2.5$ and $f_2 > 0.8$. Moreover, regions with an area lower than 100 pixels for images with a resolution of 584x768 are removed, similar to what was done during FVF algorithm. The output binary image of LMS is represented in Figure 4.15.



Figure 4.15: Binary image of the detected laser mark candidates using Laser Mark Segmentation.

After applying CHT, FVF and LMS algorithms, the regions detected with each algorithm are merged in order to obtain the final set of candidate regions. The potential laser marks drawn in Figure 4.16 are the starting point to compute the features that will be used for the final classification.

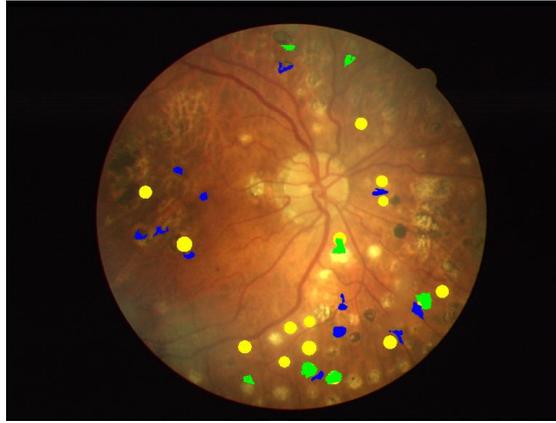


Figure 4.16: Original retinal image with all the detected candidate regions drawn: CHT - yellow, FVF - blue, LMS - green.

4.4 Features Computation

After identifying the regions candidate to be laser marks, a total of 65 features are calculated for each retinal image in order to obtain a vector containing 65 values that will be used with a classifier to determine if the image is "Laser" or "No Laser". The features are divided in 4 categories:

- **Geometrical Descriptors** - a set of 12 features that includes the number of candidate regions, their total area, average radius and likelihoods to a circle for the segmentation algorithms described in Section 4.3.
- **Texture Descriptors** - a set of 27 features that represent an analysis of texture statistics. This is the only category where the features are computed directly from the processing image and not from the candidate regions. There are six texture descriptors introduced by Gonzalez and Woods [83] and the remaining 21 result from an implementation of 11 Haralick features [84] plus 10 features based on [85, 86].
- **Spatial Distribution Descriptors** - a set of 10 features that describe the distribution of the candidate regions in the retinal image. In short words, these features

represent the distances between the laser marks and describe if they are dispersed, clustered or randomly distributed in the ROI of the retinal image.

- **Intensity-based Descriptors** - a set of 16 features related to the intensity values of the candidate regions on each channel saved during pre-processing (red, hue and saturation) plus the processing image (green channel with median filter). The computation of these features was inspired by Tahir Fazar's work [47].

4.4.1 Geometrical Descriptors

The features in this category depend on general characteristics of the detected candidate laser marks.

number_CHT, *number_FVF* and *number_LMS* are the number of regions detected by the Circular Hough Transform, by the Frangi Vesselness Filter and by the Laser Mark Segmentation algorithms, respectively, after exclusion of false marks overlapping the OD and the Vascular Network.

area_CHT is the total number of pixels detected by the Circular Hough Transform divided by the number of pixels of the OD. In this category of features, the OD is used as the term of comparison. *area_FVF* and *area_LMS* are calculated in the same way as *area_CHT* but for the blobs detected by the Frangi Filter and for the marks detected by the Laser Mark Segmentation.

Moreover, two likelihood features are calculated. For CHT, one of the outputs already mentioned is the value of the accumulator array at the position of the center of each circle detected. This value indicates the likelihood l_{CHT} that the detected region is a circle. Therefore, *likelihood_CHT* is the mean of the likelihoods l_{CHT} of all circles detected by the CHT algorithm. For the FVF detector, the likelihood of each blob is $l_{FVF}=1-ecc$, where ecc is the eccentricity value computed for each blob and if $ecc=0$, it means that the corresponding blob is a perfect circle. The value of *likelihood_FVF* is the average of the likelihoods l_{FVF} of all blobs detected by the FVF algorithm. The individual likelihoods l_{CHT} and l_{FVF} take values in the interval $[0,1]$ with 1 indicating a perfect circle.

Two other features are calculated based on the likelihood measures. These two features are weighted areas, which are the sum of the products between the number of pixels of

each region (area) and their respective likelihood parameter divided by the number of pixels of the OD, as described in Equation 4.8.

$$weighted_area = \frac{\sum_{i=1}^n l_i \pi R_i^2}{OD} \quad (4.8)$$

Therefore, *weighted_area_CHT* is the weighted area of the circles detected by the CHT, where n is *number_CHT*, l_i is the l_{CHT} and R_i is the radius of each circle and *weighted_area_FVF* is the weighted area of the blobs detected by the FVF, where n is *number_FVF*, l_i is the l_{FVF} and R_i is the radius of each blob.

To complete the set of 12 geometrical features, two features related with the radius of the regions are computed. For each potential laser mark, the radius R_i is extracted and divided by the radius of the OD. The same procedure is performed for all the other regions and then the mean and variance of those values are calculated to obtain *radius_mean* and *radius_var*, respectively. These are the only two features of this category where there is no discrimination between the three segmentation algorithms.

The normalization reflected in the divisions by the number of pixels in the OD and by the radius of the OD performed during these features computations is fundamental to make the algorithm invariant to changes in image resolution.

4.4.2 Texture Descriptors

The features in this category describe the retinal image texture based on the intensity histogram of the image in the ROI. Haralick stated that "too small a subimage region will not have enough textural information to separate image categories of interest, while a large subimage region may have objects belonging to several different categories" [84]. This means that an image containing only the laser marks lacks textural information and hence it does not provide enough conditions to compute features that allow a good classification performance. On the other hand, the ROI of the retinal image is better suited, although every image contains the same two structures: the OD and the retinal blood vessels. However, images where the retina is covered with PRP scars have a very distinguishable texture in comparison to the normal ones.

Gonzalez and Woods [83] proposed six statistical texture measures:

- ***avg_green_level*** - a measure of average intensity. It is described as

$$avg_green_level = \sum_{i=1}^{L-1} z_i p(z_i) \quad (4.9)$$

where z_i is a random variable indicating intensity, $p(z_i)$ is the normalized histogram of the intensity levels of the processing image and L is the number of possible intensity levels.

- ***avg_contrast*** - a measure of average contrast, which is the standard deviation of the image. In Equations 4.10 and 4.12, m is the *avg_green_level*.

$$avg_contrast = \sqrt{\sum_{i=1}^{L-1} (z_i - m)^2 p(z_i)} \quad (4.10)$$

- ***smoothness*** - a measure of the relative smoothness of the intensity in the ROI of the retinal image. This feature is 0 for an image of constant intensities and 1 if the intensity values vary drastically.

$$smoothness = 1 - \frac{1}{1 + \sigma^2} \quad (4.11)$$

- ***skewness*** - a measure of the third moment of the histogram of the processing image's intensity levels. This feature is equal to 0 for symmetric histograms, positive by histograms skewed to the right (about the mean) and negative for histograms skewed to the left.

$$skewness = \sum_{i=1}^{L-1} (z_i - m)^3 p(z_i) \quad (4.12)$$

- ***uniformity*** - a measure that is maximum when all green levels are equal. It takes values in the interval [0,1].

$$uniformity = \sum_{i=1}^{L-1} p^2(z_i) \quad (4.13)$$

- ***entropy*** - it is a statistical measure of randomness of the processing image.

$$entropy = - \sum_{i=1}^{L-1} p(z_i) \log_2 p(z_i) \quad (4.14)$$

Table 4.1 contains the names of the remaining 21 features implemented [87] and the reference of the work on which they were based [84–86]. These features were computed from the Green Level Co-occurrence Matrix (GLCM). Most of these statistical texture measures (11 features) were introduced by Haralick in 1973, giving birth to the so called and famous Haralick features [84].

Texture Descriptor	Reference
<i>autocorrelation</i>	[85]
<i>contrast</i>	[84, 85]
<i>correlation</i>	[84, 85]
<i>cluster_prominence</i>	[85]
<i>cluster_shade</i>	[85]
<i>dissimilarity</i>	[85]
<i>energy</i>	[84, 85]
<i>entropy_glcm</i>	[85]
<i>inv_diff</i>	[86]
<i>homogeneity</i>	[85]
<i>max_prob</i>	[85]
<i>variance</i>	[84]
<i>sum_avg</i>	[84]
<i>sum_var</i>	[84]
<i>sum_entropy</i>	[84]
<i>diff_var</i>	[84]
<i>diff_entropy</i>	[84]
<i>info_corr_1</i>	[84]
<i>info_corr_2</i>	[84]
<i>inv_diff_norm</i>	[86]
<i>inv_diff_moment_norm</i>	[86]

Table 4.1: Implemented 21 texture features based on [84–86].

4.4.3 Spatial Distribution Descriptors

The ten features in this category are representative of spatial descriptive statistics. It is reasonable in this work to estimate these statistical parameters that describe spatial patterns because laser marks can be regarded as a set of points/regions in a plane (retinal image), so they can be represented as gridded data. The comparative values used to normalize the computed features in this category are related to the ROI (*ROI_diameter* and *ROI_area*, which is the total number of pixels in the ROI of the retinal image).

First, the binary image containing the candidates in Figure 4.16 is used as input to obtain the Convex Hull, illustrated in Figure 4.17. Three features are computed using the Convex Hull [88, 89]:

- *hull* - it is the total number of pixels detected by the three segmentation algorithms divided by the number of pixels inside the Convex Hull. This value varies between $[0,1]$ and if it is near to 1, it means that the laser marks are clustered.
- *hull_area* - it is the number of pixels inside the Convex Hull divided by *ROI_area*. This value represents the portion of the retinal image occupied by the Convex Hull of the laser marks.
- *point_density* - it is the number of candidate regions divided by *hull_area*. This value represents the number of laser marks in function of the size of the Convex Hull. For few laser marks, if the size of the hull is large, then *point_density* will be low, so it means that the laser scars are more dispersed.

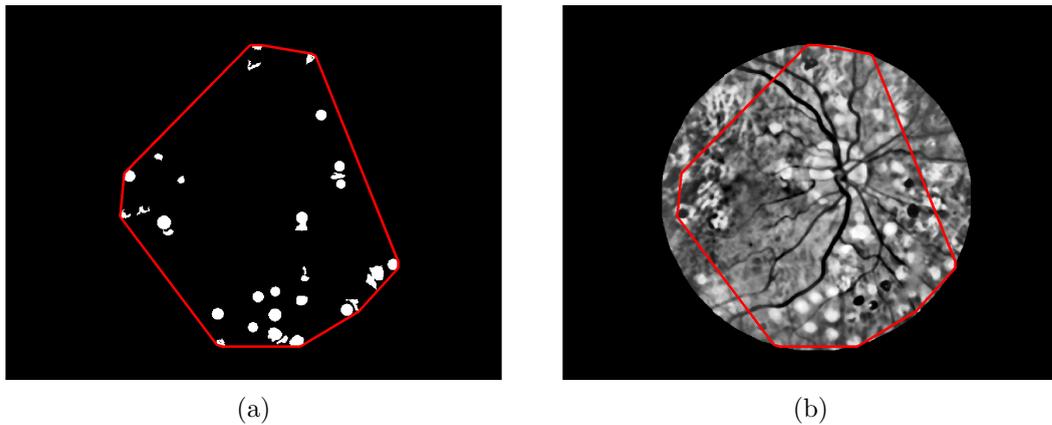


Figure 4.17: Convex Hull of the candidate regions (a) drawn in the binary image (b) drawn in the processing retinal image.

To compute the remaining features, the centroid of each region is obtained and the respective coordinates are stored. Also, it is very important to note that for the following features, if there are less than three candidate regions, their values are set to zero. Three simple measures of spatial dispersion can be defined by using the covariance matrix of the coordinates of the centroids: the *trace*, the *determinant* and the *max_eigen* (largest eigenvalue) of the covariance matrix are three features that describe the dispersion of the laser marks in the ROI of the retinal image, using only the set of centroids [90]. All these three values are divided by *ROI_area* to normalize.

The analysis of the distances between the potential laser marks is also a good approach to study the spatial pattern. A candidate region very distant from all the others has a high probability of being an outlier and hence the probability of not being a laser mark is high. Taking this premise into consideration, *m_dist_total* and *v_dist_total* are two features based on the distances between each region and its closest neighbour. In order to obtain these two features, the distance matrix (D) is computed using the matrix containing the coordinates of the centroids. Then, the minimum value in each column i of the distance matrix D (D_i) is determined, divided by *ROI_diameter* and saved in a vector. The average of the elements of this vector is *m_dist_total* and the variance is *v_dist_total* as described respectively in Equations 4.15 and 4.16 where i is the index of the candidate region and n is the total number of candidate regions.

$$m_dist_total = \frac{\sum_{i=1}^n \frac{\min\{D_i\}}{ROI_diameter}}{n} \quad (4.15)$$

$$v_dist_total = \frac{\sum_{i=1}^n \left(\frac{\min\{D_i\}}{ROI_diameter} - m_dist_total \right)^2}{n - 1} \quad (4.16)$$

The two features left to describe are Moran's I Observed Value (*moran_i*) and Moran's I Statistic (*moran_null_hipot*) [91, 92], two indicators of the spatial autocorrelation of the given set of centroids. The practical examples and applications of these two measures are presented in [93].

Moran's I Observed Value varies between [-1,1]. If it is equal to -1, it indicates perfect dispersion of the points; if it is equal to 1, it means perfect autocorrelation of the set of points (clustered); if it is equal to 0, it indicates the presence of a random spatial pattern.

Moran's I Statistic is a statistical test that represents the confidence of *moran_i* value, so *moran_null_hipot* is a binary feature. The null hypothesis states that the attribute being analyzed, which is the set of the calculated centroids, is randomly distributed in the ROI. If the p-value is statistically significant at the 5% level, the null hypothesis can be rejected and *moran_null_hipot* =1. On the other hand, if the p-value is not statistically significant at the 5% level, the value of *moran_null_hipot* is null.

4.4.4 Intensity-based Descriptors

As mentioned in the beginning of this section, the 16 features of this category are inspired on Faraz's work [47] and they describe the intensity values of the potential laser marks on four different image planes.

The binary image of the candidates present in Figure 4.16 is used to mask the processing image, the red channel of the retinal image in the RGB colorspace, the Hue dimension of the retinal image in the HSV colorspace and the Saturation dimension of the retinal image in the HSV colorspace, resulting in four different images: *green*, *red*, *hue* and *sat*.

Four features are computed from each of these images normalized to [0,1]. In each image, the average intensity value and respective variance are calculated for each laser mark and stored in the arrays *mean_intensities* and *var_intensities*. Then:

- The average of *mean_intensities* is extracted and the result are the features *mean_laser_green*, *mean_laser_red*, *mean_laser_hue* and *mean_laser_sat*.
- The average of *var_intensities* is extracted and the result are the features *var_laser_green*, *var_laser_red*, *var_laser_hue* and *var_laser_sat*.
- The maximum value of *mean_intensities* is extracted and the result are the features *max_laser_green*, *max_laser_red*, *max_laser_hue* and *max_laser_sat*.
- The variance of *var_intensities* is extracted and the result are the features *variance_var_green*, *variance_var_red*, *variance_var_hue*, *variance_var_sat*,

4.5 Classification

A classifier is an algorithm that receives data containing descriptive information as input and returns a final decision for each instance of the input data. The classifiers used in this work were the Decision Tree (pruned C4.5) [94] and Random Forests [95]. Also, a classifier based on a simple threshold method was applied. (Subsection 4.5.3). Tree-based classifiers are knowledge-based systems that are very popularly used to solve difficult and complex real-world problems. They are characterised for being accurate,

robust, with high execution speed, reliable and producing descriptive outputs that can be interpreted by the user, using a set of features as inputs. Not only do these properties make the tree-based classifiers well-known in the scientific community, but also they are very simple to use and to implement because the user only needs to provide learning data, such as the 65 features computed previously, in order to train and build the classifier and generate a representation of the acquired knowledge [94, 96]. The subsections 4.5.1 and 4.5.2 briefly describe the theoretical concepts of the Decision Tree and of the Random Forest, respectively.

4.5.1 Decision Tree

Ross Quinlan introduced an algorithm (ID3) based on a family of learning systems that are characterized by their representation of acquired knowledge, such as decision trees [96]. ID3 is a supervised learning algorithm that builds a classification rule expressed as a decision tree from a given set of examples (features) and the resulting tree is used to receive new instances and make a decision for each instance of the input [97].

Quinlan explains the generation of a decision tree by giving a practical example: according to the weather of a Saturday morning, should an unspecified activity be positive (P) or negative (N)? The attributes used to describe the weather are the outlook (sunny, overcast, rain), temperature (cool, mild, hot), humidity (high, normal), windy (true, false). The author used the attributes and respective values in Table 1 in [96] as the training set and obtained the tree shown in Figure 4.18.

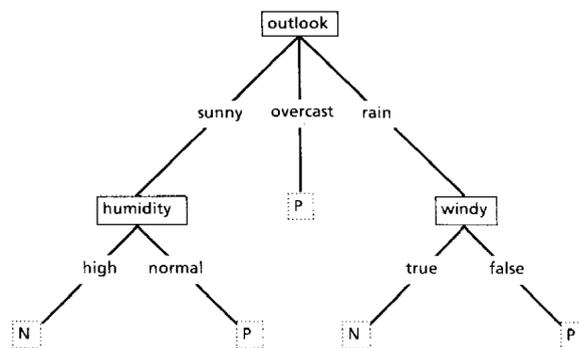


Figure 4.18: Decision Tree built using the training set of attributes and respective values of the weather a Saturday morning. This figure corresponds to Figure 2 in [96].

For example, an instance with the following parameters ($outlook=rain$, $windy=true$) will be classified as N by the tree in Figure 4.18. During training phase, each node of the

decision tree is obtained by calculating a gain for each attribute and the one with the highest gain is the attribute selected for the corresponding node. This measure is the Information Gain¹ and it is based on the entropy of each attribute. The ID3 algorithm stops after every attribute are used or after every instance of the training set can be correctly classified by the tree model in construction [96].

There are some limitations to the ID3 algorithm: it only works with nominal attributes, there is a risk of overfitting which leads to a loss of predictive accuracy and it is too sensitive to attributes with a wide range of values which contributes to a biased tree model. To overcome these issues, Quinlan *et al.* [98] developed latter another algorithm using decision tree - C4.5. The latter algorithm is compatible with continuous variables and instead of Information Gain it uses Gain Ratio¹ for attribute selection at each node. Overfitting can be avoided by adding a stopping criterion during training phase, but in this specific case the decision tree is built first and then some of the structure of the tree is removed (pruning) [98].

4.5.2 Random Forest

Leo Breiman [95] introduced the concept of Random Forest, an ensemble learning method that is based on the building and aggregation of several decision trees. Random Forests are accurate classifiers and regressors that were originally developed correct the overfitting risk of decision trees. The main criterion of random forests is the minimization of a generalization error. This error converges to a limit as the number of decision trees in the forest becomes large because of the Law of Large Numbers and it depends on the strength of the individual trees and on the correlation between them.

At the training phase, a random selection of features from the training set is performed to split each node for each tree of the forest. In other words, each set of randomly selected features is generated to govern the growth of each tree, which are not pruned, in the ensemble. The purpose of this randomness approach is to minimize the correlation between trees in order to improve the accuracy of the random forest.

Out-of-bag estimates are calculated during training with the purpose of monitoring the generalization error. This means that before the random feature selection one-third of

¹For more details about the Information Gain and Gain Ratio, please consult Appendix A

the instances are left out so that the error rates computed from the growing trees using this portion of instances overestimate the current/real error rate. This bagging strategy allows the enhancement of the accuracy when random features are used and it gives ongoing estimates of the generalization error as well as the strength and correlation.

In the end, after the generation of a certain number of trees chosen by the user, a voting process is carried to select the most frequent class and make the final decision [95].

4.5.3 Threshold-based classification

The threshold-based classifier is a very simple approach and easy to implement. The simplicity of this method is due to the fact that the final decision is directly extrapolated from very few features, which are the candidate regions detected for each segmentation algorithm in this work. For instance, a threshold is applied to the number of candidates detected by one of the three segmentation algorithms. The number of candidates was varied and the condition applied was that if a retinal image presented an equal or greater number of candidates than the threshold, that image would be classified as "Laser". The number of candidates for which this simple classifier had the smallest classification error was considered the optimal threshold and its computation can be interpreted as the training step.

4.6 Features Selection

Feature selection is a fundamental problem in many areas because it reduces the dimensionality of feature space, removes redundant, irrelevant and noisy data. For this specific problem, all features might be important but usually only a small subset of features is relevant. "Until recently, classification tasks with more than 50 attributes were considered to have a high dimensionality" [99]. Therefore, the purpose of this section is to identify which minimum size subset of the 65 features is more suitable for performing a fast and accurate final classification [100, 101].

The software used for feature selection, classifiers' training and testing was the WEKA framework [102]. The dataset used to train the classifiers is a union of the S and EONMA datasets, resulting in 203 retinal images "Laser" and 652 retinal images "No

Laser”. These two datasets were chosen because their images have different resolution and were captured by different cameras so that the selected features are more likely to represent the diverse type of images encountered in real life. This is one of the most crucial steps to certify that the algorithm is robust and at the same time has a wide variety of datasets for testing.

There are two types of feature selection methods: Filter and Wrapper methods. The filter method uses statistical characteristics of the data to evaluate and returns a score of variable importance for each feature. It is independent of any learning algorithm and it is computationally simple. This method ranks and selects the most relevant features that can be provided as input to different classifiers. The inconvenience of filter method is that the resulting features are usually redundant. The wrapper method leads to a better performance than filter methods because it is a scheme-dependent attribute evaluator, which means it is optimized for the classification algorithm to be used, although it needs a high computational effort. Wrappers are a type of feature selection that calculate a score for a subset of features by making use of a classifier that uses only the features from that subset and return a predictive performance. In other words, this features selection method selects several subsets, then it uses the classification algorithm itself to measure the predictive quality of each subset of features and the one which provides the best performance is considered the best set of features to be used with that specific classifier. Choosing all possible subsets means a huge consumption of computational time, so a search method is usually coupled to the wrapper with the purpose to reduce the complexity without compromising the performance [99, 101, 103].

In this work, it was used a mixture of these two methods similar to what was proposed in [99]. First, an attribute evaluator based on information theory is used to rank all 65 features according to their relevance score. Then a threshold is applied to remove the least important features. The output of this filter method serves as input to a wrapper method with a forward selection as the searching method, which will eliminate the redundant features and keep only the most important ones.

Several approaches to the variable selection problem using information theoretic criteria have been proposed. Information Gain (IG) and Gain Ratio (GR)² are scheme-independent attribute evaluators that rely on empirical estimates of the mutual information between each variable and the target or class [104]. Information Gain is based on entropy, which is a common information theory measure and it scores the system's unpredictability. One inconvenience of IG is that it is biased in favor of features with a wide range of values, even if they are not more informative. Gain Ratio is a modification of IG to correct its bias by dividing IG by the intrinsic information of a split. This is an approach similar to a normalisation, so GR decreases as intrinsic information gets larger [100, 103, 105].

The purpose of using one of these filters is to discard irrelevant features from the feature vector because they introduce noise in the data and may overshadow the important information present. Therefore, using the training dataset (S+EONMA), two feature ranks were obtained: one using IG as attribute evaluator and the other one using GR. For each rank, the filter method was completed by applying a threshold $t = 0.1$. This threshold operation using IG resulted in 40 out of 65 features remaining, while for GR 43 out of 65 features followed the criteria. Figure 4.19 shows the features that had a value greater than 0.1 for at least one of the filters (IG or GR) and their respective scores for each filter method. Some bars are absent in the figure because for one of the information measures the score was lower than 0.1 and therefore set to zero. A high score of a feature for a certain filter means that the corresponding feature is important when using that filter method. Figure 4.19 confirms that the importance of each feature is different depending on the filter method applied.

The subsets that resulted from the filtering method were used as input for the wrapper evaluation. The wrapper used is a WEKA-implementation with default settings that evaluates subsets using a classifier determined by the user and performs an inner stratified 5-fold cross-validation on the training dataset. Sequential Forward Selection is the hill-climbing search technique used in the wrapper. It starts with the empty subset, it evaluates all possible single-feature extension of the current subset and the feature that leads to the best accuracy of the current classifier is added permanently. Usually the search terminates when no single-feature extension improves the current accuracy. This poses a problem: is the optimal value a local or an absolute maxima? To solve

²For more about the information theory measures, please consult Appendix A

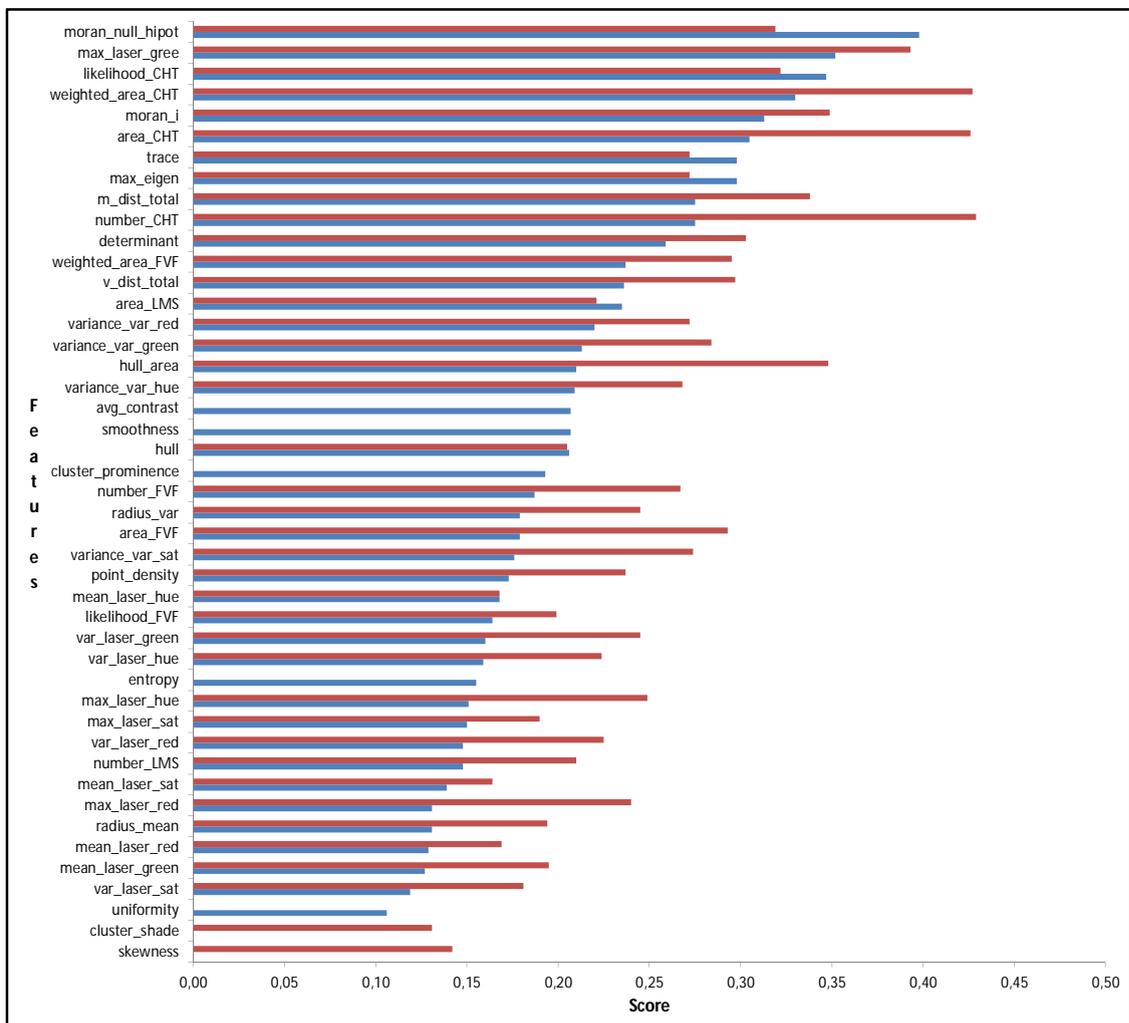


Figure 4.19: Information Gain (red) and Gain Ratio (blue) measures of each feature after thresholding at $t = 0.1$. The features are ordered in descending order of GR measure.

this problem, a bi-directional search was proposed by the scientific community, but in this work a different approach is followed. The classical forward selection is used, but a stop criteria of 15 nodes is added: this means that the search only terminates after consecutively adding 15 features without improving the accuracy, instead of stopping right after the first attempt without improving the performance [99].

The classifiers used in this wrapper method were: pruned C4.5 Decision Tree (DT), which is the J48 classifier on WEKA, Random Forest with 5 trees (RF5), Random Forest with 50 trees (RF50), Random Forest with 500 trees (RF500) and the training dataset used was (S+EONMA).

The following subsections describe the results of this features' selection phase. Subsection 4.6.1 presents the accuracy of the wrapper method for each of the four tree-based classifiers after filtering the initial 65 features with the IG evaluator thresholded at $t = 0.1$. Subsection 4.6.2 shows the results of the wrapper method for each of the 4 tree-based classifiers after filtering the initial 65 features using the GR evaluator thresholded at $t = 0.1$.

4.6.1 Selection using Information Gain

A filter using the IG measure is applied to the 65 features contained in the images from the dataset (S+EONMA) resulting in 40 features. With the remaining 40 features, the wrapper method is applied in the training dataset (S+EONMA) using four different tree-based classifiers.

Table 4.2 reveals the 13 most suitable features to be used in a Decision Tree classifier and the respective accuracy score, after IG filtering followed by the wrapper method.

Table 4.3 lists the four most suitable features to be used in a Random Forest with 5 trees classifier and the respective accuracy score, after IG filtering followed by the wrapper method.

Table 4.4 tabulates the 17 most suitable features to be used in a Random Forest with 50 trees classifier and the respective accuracy score, after IG filtering followed by the wrapper method.

Best Subset: IG-DT

weighted_area_CHT
 area_CHT
 hull_area
 likelihood_CHT
 determinant
 v_dist_total
 weighted_area_FVF
 trace
 var_laser_green
 var_laser_red
 number_LMS
 mean_laser_green
 mean_laser_hue

Accuracy: 92.8%

Table 4.2: Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier DT.

Best Subset: IG-RF5

number_CHT
 weighted_area_FVF
 moran_null_hipot
 determinant

Accuracy: 92.2%

Table 4.3: Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier RF5.

Table 4.5 itemizes the six most suitable features to be used in a Random Forest with 500 trees classifier and the respective accuracy score, after IG filtering followed by the wrapper method.

Best Subset: IG-RF50

number_CHT
 weighted_area_CHT
 area_CHT
 m_dist_total
 moran_null_hipot
 determinant
 weighted_area_FVF
 trace
 max_eigen
 variance_var_red
 variance_var_hue
 number_FVF
 var_laser_green
 max_laser_red
 var_laser_hue
 mean_laser_green
 mean_laser_sat

Accuracy: 93.2%

Table 4.4: Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier RF50.

Best Subset: IG-RF500

number_CHT
 weighted_area_CHT
 area_CHT
 moran_i
 likelihood_CHT
 moran_null_hipot

Accuracy: 92.4%

Table 4.5: Best subset of features and respective accuracy after applying the IG filter and using wrapper evaluation with the classifier RF500.

4.6.2 Selection using Gain Ratio

A filter using the GR measure is applied to the 65 features contained in the dataset (S+EONMA) resulting in 43 features. With the remaining 43 features, the wrapper method is applied in the training dataset (S+EONMA) using four different tree-based classifiers.

Table 4.6 enumerates the eight most suitable features to be used in a Decision Tree classifier and the respective accuracy score, after GR filtering followed by the wrapper

method.

Best Subset: GR-DT
weighted_area_CHT
moran_i
area_CHT
weighted_area_FVF
variance_var_red
hull_area
hull
mean_laser_hue
Accuracy: 92.6%

Table 4.6: Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier DT.

Table 4.7 shows the 15 most suitable features to be used in a Random Forest with 5 trees classifier and the respective accuracy score, after GR filtering followed by the wrapper method.

Best Subset: GR-RF5
moran_null_hipot
max_laser_green
weighted_area_CHT
moran_i
number_CHT
determinant
variance_var_green
variance_var_hue
hull
number_FVF
mean_laser_hue
mean_laser_sat
mean_laser_red
mean_laser_green
var_laser_sat
Accuracy: 92.7%

Table 4.7: Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier RF5.

Table 4.8 displays the 14 most suitable features to be used in a Random Forest with 50 trees classifier and the respective accuracy score, after GR filtering followed by the wrapper method.

Best Subset: GR-RF50

moran_null_hipot
 likelihood_CHT
 weighted_area_CHT
 area_CHT
 max_eigen
 number_CHT
 m_dist_total
 determinant
 variance_var_red
 max_laser_sat
 mean_laser_sat
 max_laser_red
 mean_laser_red
 mean_laser_green

Accuracy: 93.3%

Table 4.8: Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier RF50.

Table 4.9 presents the 17 most suitable features to be used in a Random Forest with 500 trees classifier and the respective accuracy score, after GR filtering followed by the wrapper method.

Best Subset: GR-RF500

moran_null_hipot
 likelihood_CHT
 weighted_area_CHT
 moran_i
 area_CHT
 trace
 max_eigen
 number_CHT
 m_dist_total
 determinant
 area_LMS
 variance_var_red
 variance_var_hue
 point_density
 var_laser_hue
 number_LMS
 mean_laser_sat

Accuracy: 93.0%

Table 4.9: Best subset of features and respective accuracy after applying the GR filter and using wrapper evaluation with the classifier RF500.

Chapter 5

Results

The content in this chapter presents the classification results for the automatic laser mark detection algorithm addressing four different tree-based classifiers: Decision Tree (Subsection 5.1.1), Random Forest with 5 Trees (Subsection 5.1.2), Random Forest with 50 Trees (Subsection 5.1.3) and Random Forest with 500 Trees (Subsection 5.1.4). Also, some results of a simple threshold-based classification are presented (Subsection 5.1.5). The latter method and respective results are listed to provide a comparison anchor in the form of a very simple naive classifier. The evaluation of each classifier was performed by computing the confusion matrix for each tested dataset and by calculating two widely known classification performance measures for each tested dataset, sensitivity and specificity. Furthermore, this chapter contains an assessment of the time consumption of the proposed algorithm for each retinal image (Section 5.2).

5.1 Performance Evaluation

In section 4.6 two approaches were followed to select the features to use in the decision trees and forests. One of these approaches was to apply a filter using Information Gain and then performing a wrapper method using the four classifiers. The other was similar but uses the Gain Ratio as filter. From both methods, the best subset of features and respective accuracy for each classifier was obtained. This chapter lists results for the best case of each type of classifier.

Using the DT classifier, 13 features were selected with a classification accuracy of 92.8% after applying the IG filter, while 8 features were selected with a classification accuracy of 92.6% after applying the GR filter. Therefore, only the case with application of IG filter is considered in this chapter, since it showed a better accuracy. A similar procedure was taken with the other three classifiers. For instance, the 15 features selected using RF5 after GR filtering (92.7% against 92.2% with IG filtering), the 14 features selected using RF50 after GR filtering (93.3% against 93.2% with IG filtering) and the 17 features selected using RF500 after GR filtering (93% against 92.4% with IG filtering) are the other three classifiers and features sets considered in this chapter.

The first four subsections of this chapter have a standard structure. First, a stratified 5-fold cross validation is performed on the dataset (S+EONMA) using the corresponding classifier with the selected features and the performance measures are presented along with the Receiving Operator Characteristic (ROC) Curve, which is the True Positive Rate in function of the False Positive Rate. True Positive Rate is the equivalent of Sensitivity and False Positive Rate = $\frac{FP}{FP+TN} = 1 - Specificity$. Then, the classifier is trained using the dataset (S+EONMA) and tested on the remaining datasets, thus obtaining the performance for each dataset. In the end, every testing datasets (except the datasets S and EONMA used for training) are aggregated in order to obtain the final performance measures of the implemented algorithm, which are the accuracy, the sensitivity, the specificity, the positive predictive value (PPV), the negative predictive value (NPV)¹ and the area under the ROC curve (AUC) [106].

The ROC curves on WEKA are generated by plotting the True Positive Rate in function of the False Positive Rate for different probabilities that the classifier assigns to the negative class. For instance, each point of the curve corresponds to setting a threshold on the probability assigned by the classifier to the negative class. The performance measures presented on Table 5.1 occur when this probability is equal to 0.5. This method of obtaining the ROC curve is applicable to subsections 5.1.1, 5.1.2, 5.1.3 and 5.1.4.

¹For further information about the performance measures, please consult Appendix A

5.1.1 Decision Tree

As described in Section 4.6.1, 13 attributes were chosen during features selection phase using a Decision Tree wrapper method after a IG filtering. Table 5.1 shows the results of a stratified 5-fold cross validation using a Decision Tree classifier with the selected 13 features on the retinal images of the dataset (S+EONMA).

Dataset: (S+EONMA) 5-fold CV	Laser	No Laser
Laser	155 (76.4%)	48 (23.6%)
No Laser	16 (2.5%)	636 (97.5%)

Table 5.1: Confusion Matrix after a 5-fold Cross Validation using DT of the images from the datasets S+EONMA.

Training the pruned C4.5 classifier using the dataset (S+EONMA) resulted in a Decision Tree model with 15 leaves and containing only 8 of the selected 13 features used to build the classifier (Figure 5.1).

Considering the class "Laser" as the positive case and the class "No Laser" as the negative case, a ROC curve was obtained by thresholding the output of the built model of the Decision Tree after the stratified 5-fold cross validation (Figure 5.2). The area under ROC was also computed and has a value of $AUC_{(S+EONMA)}=0.885$.

The DT classifier was trained using the dataset (S+EONMA) and was tested on the remaining 9 datasets of this work. The performance measures were calculated for each dataset: M (Table 5.2), EOMA (Table 5.3), EOEX (Table 5.4), EONEX (Table 5.5), VBR (Table 5.6), HP (Table 5.7), FAZD (Table 5.8), BAT (Table 5.9), PDJD (Table 5.10). The BAT dataset contains retinal images from 9 patients and after presenting the testing results of this dataset (Table 5.9), an evaluation is performed concerning the number of patients correctly detected. A patient correctly detected contains at least one "Laser" image correctly classified and cannot have any "No Laser" image classified as having laser marks.

Dataset: M	Laser	No Laser
Laser	-	-
No Laser	12 (1%)	1175 (99%)

Table 5.2: Confusion Matrix after testing the trained DT on M.

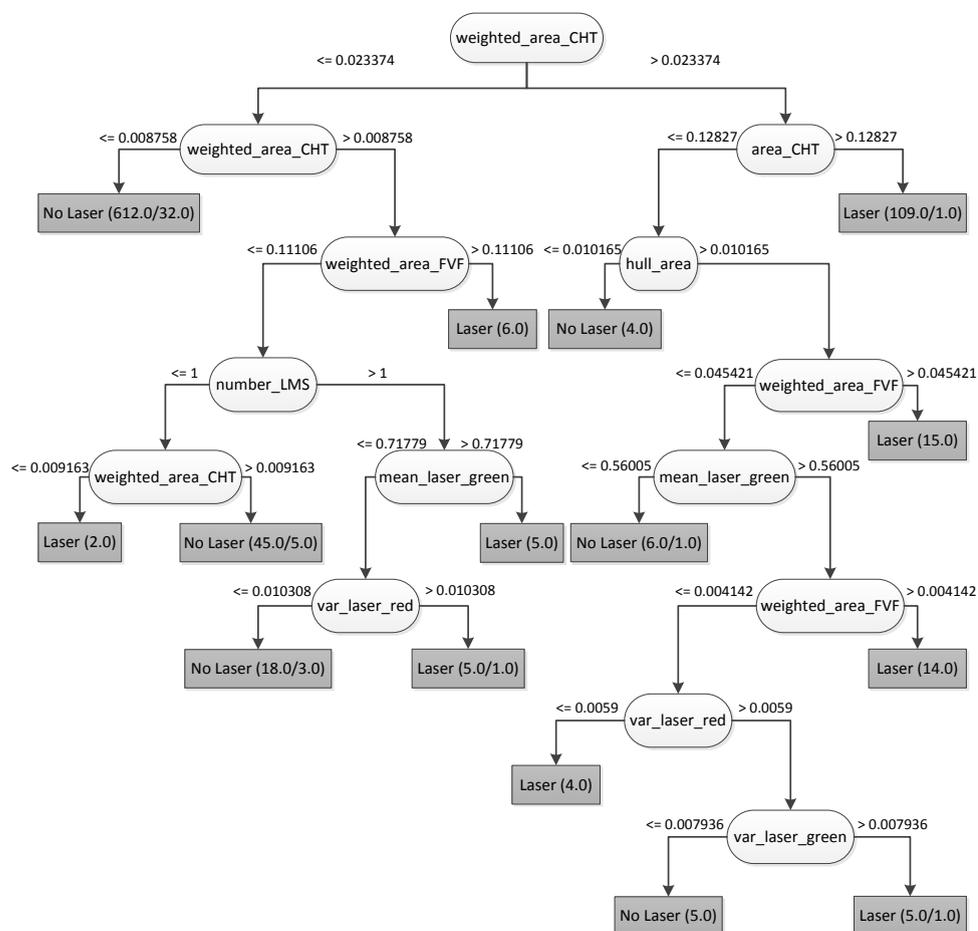


Figure 5.1: Decision Tree model built after training the classifier with the dataset (S+EONMA).

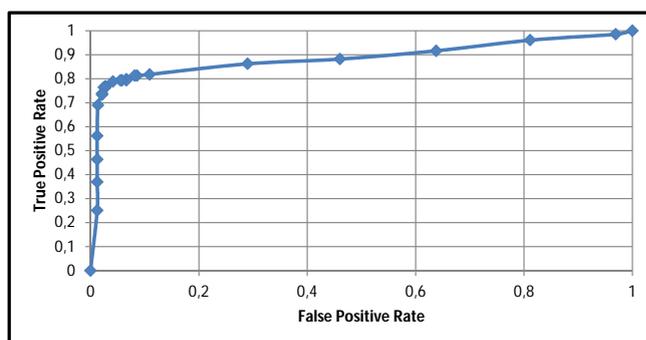


Figure 5.2: ROC Curve after thresholding the output of the DT classifier with 5-fold cross validation.

Dataset: EOMA	Laser	No Laser
Laser	-	-
No Laser	0 (0%)	148 (100%)

Table 5.3: Confusion Matrix after testing the trained DT on EOMA.

Dataset: EOEX	Laser	No Laser
Laser	-	-
No Laser	0 (0%)	47 (100%)

Table 5.4: Confusion Matrix after testing the trained DT on EOEX.

Dataset: EONEX	Laser	No Laser
Laser	-	-
No Laser	1 (2.9%)	34 (97.1%)

Table 5.5: Confusion Matrix after testing the trained DT on ENOEX.

Dataset: VBR	Laser	No Laser
Laser	-	-
No Laser	0 (0%)	22 (100%)

Table 5.6: Confusion Matrix after testing the trained DT on VBR.

Dataset: HP	Laser	No Laser
Laser	-	-
No Laser	1 (1%)	99 (99%)

Table 5.7: Confusion Matrix after testing the trained DT on HP.

Dataset: FAZD	Laser	No Laser
Laser	-	-
No Laser	3 (5%)	57 (95%)

Table 5.8: Confusion Matrix after testing the trained DT on FAZD.

Dataset: BAT	Laser	No Laser
Laser	21 (28.2%)	13 (61.8%)
No Laser	1 (6.7%)	14 (93.3%)

Table 5.9: Confusion Matrix after testing the trained DT on BAT.

The retinal image 07_B2.jpg was the only one without laser marks that was classified as "Laser". As for the 13 misclassified "Laser" images, at least one retinal image "Laser" per patient was correctly classified by the algorithm, except for patient 09 where none of the 3 "Laser" images was detected. Therefore, only patients 07 and 09 were not

correctly classified by the implemented algorithm, so 7 out of 9 patients were accurately detected using the DT classifier.

Dataset: PDJD	Laser	No Laser
Laser	98 (97%)	3 (3%)
No Laser	-	-

Table 5.10: Confusion Matrix after testing the trained DT on PDJD.

Table 5.11 presents the DT classifier sensitivity and specificity values of the 5-fold cross validation using the dataset (S+EONMA) plus the 9 tested datasets.

Dataset	Sensitivity	Specificity
S+EONMA	76.4%	97.5%
M	-	99%
EOMA	-	100%
EOEX	-	100%
EONEX	-	97.1%
VBR	-	100%
HP	-	99%
FAZD	-	95%
BAT	61.8%	93.3%
PDJD	97%	-

Table 5.11: Algorithm's performance on each dataset using Decision Tree.

The retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749 images: 135 images "Laser" and 1614 images "No Laser". This merged dataset was used to test the DT classifier and the results are shown in Table 5.12.

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity	PPV	NPV	AUC
98.06%	1.94%	88.1%	98.9%	0.869	0.99	0.946

Table 5.12: Performance stats using the trained Decision Tree averaged over all testing datasets.

Considering the class "Laser" as the positive case and the class "No Laser" as the negative case, a ROC curve was obtained by thresholding the output of the implemented binary DT classifier after testing it on the merged dataset containing 1749 images (Figure 5.3). The area under ROC was also computed and has a value of $AUC_{Test}=0.946$.

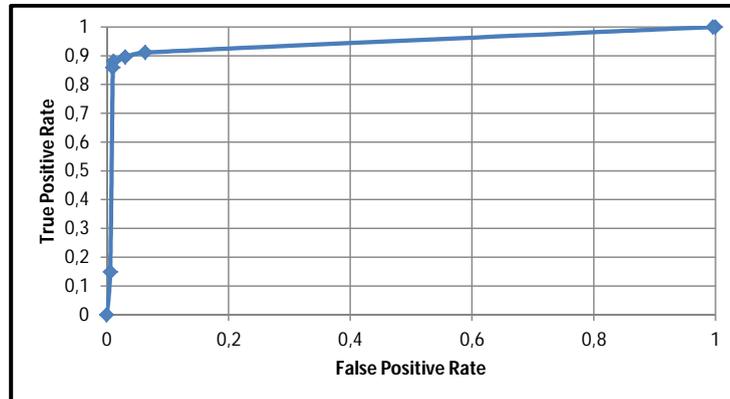


Figure 5.3: ROC Curve obtained by thresholding the output of the trained DT classifier tested on 9 datasets.

5.1.2 Random Forest with 5 Trees

As described before, 15 attributes were chosen during features selection phase using a Random Forest with 5 Trees classifier wrapper method after a GR filtering. Table 5.13 shows the results of a stratified 5-fold cross validation using a Random Forest with 5 Trees classifier with the selected 15 features on the retinal images of the dataset (S+EONMA).

Dataset: (S+EONMA) 5-fold CV	Laser	No Laser
Laser	152 (74.9%)	51 (25.1%)
No Laser	33 (5.1%)	619 (94.9%)

Table 5.13: Confusion Matrix after a 5-fold Cross Validation using RF5 of the images from the datasets S+EONMA.

A ROC curve was obtained by thresholding the output of the RF5 classifier after the stratified 5-fold cross validation (Figure 5.2). The area under ROC was also computed and has a value of $AUC_{(S+EONMA)}=0.9$.

The RF5 classifier was trained using the dataset (S+EONMA) and was tested on the remaining 9 datasets of this work. The performance measures were calculated for each dataset: M (Table 5.14), EOMA (Table 5.15), EOEX (Table 5.16), EONEX (Table 5.17), VBR (Table 5.18), HP (Table 5.19), FAZD (Table 5.20), BAT (Table 5.21), PDJD (Table 5.22). As described in subsection 5.1.1, an evaluation is performed concerning

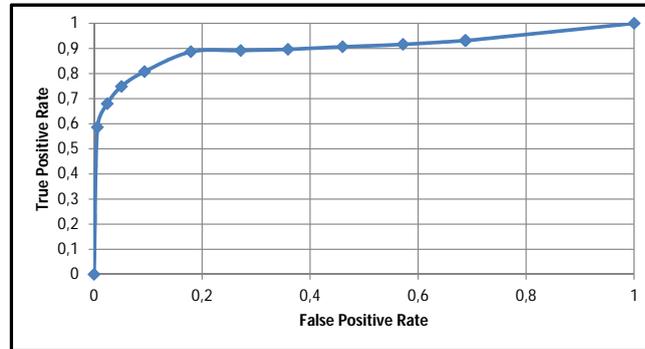


Figure 5.4: ROC Curve obtained by thresholding the output of the RF5 classifier with 5-fold cross validation.

the number of patients from the BAT dataset (Table 5.9) correctly detected. A patient correctly detected contains at least one "Laser" image correctly classified and cannot have any "No Laser" image classified as having laser marks.

Dataset: M	Laser	No Laser
Laser	-	-
No Laser	17 (1.4%)	1170 (98.6%)

Table 5.14: Confusion Matrix after testing the trained RF5 on M.

Dataset: EOMA	Laser	No Laser
Laser	-	-
No Laser	0 (0%)	148 (100%)

Table 5.15: Confusion Matrix after testing the trained RF5 on EOMA.

Dataset: EOEX	Laser	No Laser
Laser	-	-
No Laser	0 (0%)	48 (100%)

Table 5.16: Confusion Matrix after testing the trained RF5 on EOEX.

Dataset: EONEX	Laser	No Laser
Laser	-	-
No Laser	1 (2.9%)	34 (97.1%)

Table 5.17: Confusion Matrix after testing the trained RF5 on ENOEX.

Dataset: VBR	Laser	No Laser
Laser	-	-
No Laser	1 (4.5%)	21 (95.5%)

Table 5.18: Confusion Matrix after testing the trained RF5 on VBR.

Dataset: HP	Laser	No Laser
Laser	-	-
No Laser	4 (4%)	96 (96%)

Table 5.19: Confusion Matrix and after testing the trained RF5 on HP.

Dataset: FAZD	Laser	No Laser
Laser	-	-
No Laser	4 (6.7%)	56 (93.3%)

Table 5.20: Confusion Matrix after testing the trained RF5 on FAZD.

Dataset: BAT	Laser	No Laser
Laser	23 (67.6%)	11 (32.4%)
No Laser	3 (20%)	12 (80%)

Table 5.21: Confusion Matrix after testing the trained RF5 on BAT.

After inspecting the images from the BAT dataset that were incorrectly classified, it was verified that 06_B1.jpg, 07_B2.jpg and 09_B2.jpg are retinal images without laser marks that were classified as "Laser". As a consequence, patients 06, 07 and 09 were not correctly detected. As for the 11 misclassified "Laser" images, at least one retinal image "Laser" per patient was correctly classified by the algorithm. Therefore, 6 out of 9 patients were accurately detected using the RF5 classifier.

Dataset: PDJD	Laser	No Laser
Laser	97 (96%)	4 (4%)
No Laser	-	-

Table 5.22: Confusion Matrix after testing the trained RF5 on PDJD.

Table 5.23 presents the RF5 classifier sensitivity and specificity values of the 5-fold cross validation using the dataset (S+EONMA) and of the 9 tested datasets.

According to the standard structure of this chapter, the retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749

Dataset	Sensitivity	Specificity
S+EONMA	74.9%	94.9%
M	-	98.6%
EOMA	-	100%
EOEX	-	100%
EONEX	-	97.1%
VBR	-	95.5%
HP	-	96%
FAZD	-	93.3%
BAT	67.6%	80%
PDJD	96%	-

Table 5.23: Algorithm's performance on each dataset using Random Forest with 5 Trees.

images. This merged dataset was used to test the RF5 classifier and the results are shown in Table 5.24.

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity	PPV	NPV	AUC
97.43%	2.57%	88.9%	98.1%	0.8	0.991	0.955

Table 5.24: Performance stats using the trained Random Forest with 5 Trees averaged over all testing datasets.

A ROC curve was obtained by thresholding the output of the implemented binary classifier RF5 after testing it on the merged dataset containing 1749 images (Figure 5.5). The area under ROC was also computed and has a value of $AUC_{Test}=0.955$.

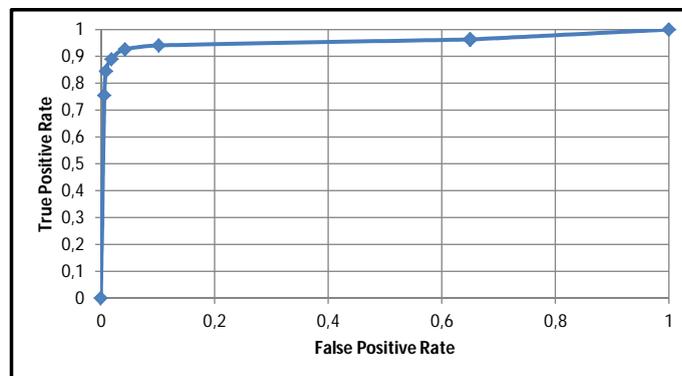


Figure 5.5: ROC Curve obtained by thresholding the output of the trained RF5 classifier tested on 9 datasets.

5.1.3 Random Forest with 50 Trees

As described in Section 4.6.2, 14 attributes were chosen during features selection phase using a Random Forest with 50 Trees classifier wrapper method after a GR filtering. Table 5.25 shows the results of a stratified 5-fold cross validation using a Random Forest with 50 Trees classifier with the selected 14 features on the retinal images of the dataset (S+EONMA).

Dataset: (S+EONMA) 5-fold CV	Laser	No Laser
Laser	156 (76.8%)	47 (23.2%)
No Laser	22 (3.4%)	630 (96.6%)

Table 5.25: Confusion Matrix after a 5-fold Cross Validation using RF50 of the images from the datasets S+EONMA.

Once again, considering the class "Laser" as the positive case and the class "No Laser" as the negative case, a ROC curve was obtained by thresholding the output of the implemented binary classifier RF50 after the stratified 5-fold cross validation (Figure 5.6). The area under ROC was also computed and has a value of $AUC_{(S+EONMA)}=0.918$.

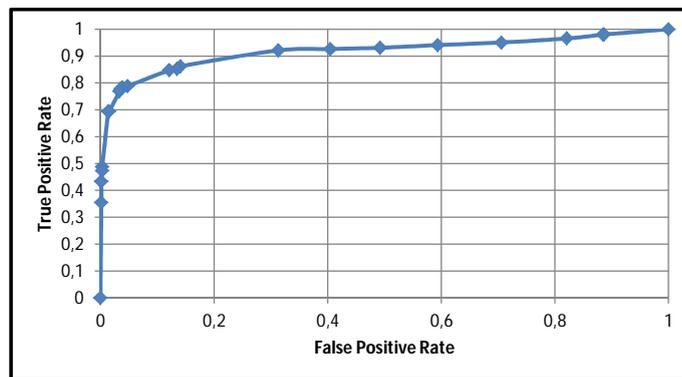


Figure 5.6: ROC Curve after thresholding the output of the RF50 classifier with 5-fold cross validation.

The RF50 classifier was trained using the dataset (S+EONMA) and was tested on the remaining 9 datasets of this work. The performance measures were calculated for each dataset: M (Table 5.26), EOMA (Table 5.27), EOEX (Table 5.28), EONEX (Table 5.29), VBR (Table 5.30), HP (Table 5.31), FAZD (Table 5.32), BAT (Table 5.33), PDJD (Table 5.34). Moreover, an evaluation is performed concerning the number of patients

from the BAT dataset (Table 5.33) correctly detected. A patient correctly detected contains at least one "Laser" image correctly classified and cannot have any "No Laser" image classified as having laser marks.

Dataset: M	Laser	No Laser
Laser	-	-
No Laser	24 (2%)	1163 (98%)

Table 5.26: Confusion Matrix after testing the trained RF50 on M.

Dataset: EOMA	Laser	No Laser
Laser	-	-
No Laser	1 (0.7%)	147 (99.3%)

Table 5.27: Confusion Matrix after testing the trained RF50 on EOMA.

Dataset: EOEX	Laser	No Laser
Laser	-	-
No Laser	1 (2.1%)	46 (97.9%)

Table 5.28: Confusion Matrix after testing the trained RF50 on EOEX.

Dataset: EONEX	Laser	No Laser
Laser	-	-
No Laser	1 (2.9%)	34 (97.1%)

Table 5.29: Confusion Matrix after testing the trained RF50 on ENOEX.

Dataset: VBR	Laser	No Laser
Laser	0	0
No Laser	0 (0%)	22 (100%)

Table 5.30: Confusion Matrix after testing the trained RF50 on VBR.

Dataset: HP	Laser	No Laser
Laser	-	-
No Laser	1 (1%)	99 (99%)

Table 5.31: Confusion Matrix after testing the trained RF50 on HP.

Dataset: FAZD	Laser	No Laser
Laser	0	0
No Laser	4 (6.7%)	56 (93.3%)

Table 5.32: Confusion Matrix after testing the trained RF50 on FAZD.

Dataset: BAT	Laser	No Laser
Laser	24 (70.6%)	10 (29.4%)
No Laser	4 (26.7%)	11 (73.3%)

Table 5.33: Confusion Matrix after testing the trained RF50 on BAT.

The images 01_B2.jpg, 05_B1.jpg 06_B1.jpg and 07_B2.jpg from the BAT dataset are the retinal images "No Laser" incorrectly classified. For that reason, patients were not correctly detected by the implemented algorithm. As for the 10 misclassified "Laser" images, at least one retinal image "Laser" per patient was correctly detected by the algorithm. Therefore, 5 out of 9 patients were accurately detected using the RF50 classifier.

Dataset: PDJD	Laser	No Laser
Laser	94 (93.1%)	7 (6.9%)
No Laser	-	-

Table 5.34: Confusion Matrix after testing the trained RF50 on PDJD.

Table 5.35 presents the RF50 classifier sensitivity and specificity values of the 5-fold cross validation using the dataset (S+EONMA) plus the 9 tested datasets.

Dataset	Sensitivity	Specificity
S+EONMA	76.8%	96.6%
M	-	98%
EOMA	-	99.3%
EOEX	-	97.9%
EONEX	-	97.1%
VBR	-	100%
HP	-	99%
FAZD	-	93.33%
BAT	70.6%	73.3%
PDJD	93.1%	-

Table 5.35: Algorithm's performance on each dataset using Random Forest with 50 Trees.

The retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749 images. This merged dataset was used to test the RF50 classifier and the results are shown in Table 5.36.

The ROC curve was obtained by thresholding the output of the implemented classifier RF50 after testing it on the merged dataset containing 1749 images (Figure 5.7). The area under ROC was also computed and has a value of $AUC_{Test}=0.959$.

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity	PPV	NPV	AUC
96.97%	3.03%	87.4%	97.8%	0.766	0.989	0.969

Table 5.36: Performance stats using the trained Random Forest with 50 Trees averaged over all testing datasets.

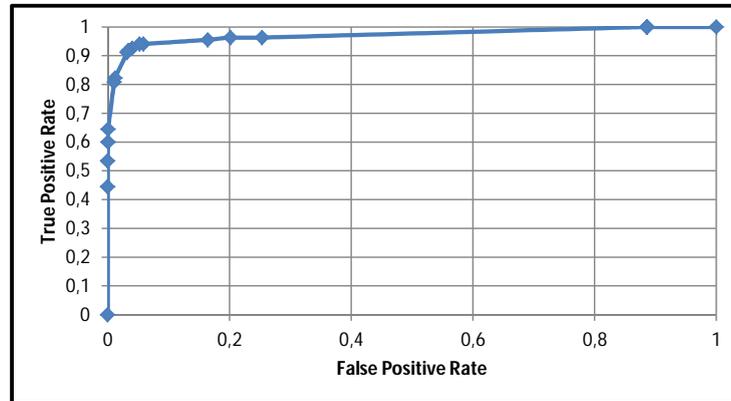


Figure 5.7: ROC Curve obtained by thresholding the output of the trained RF50 classifier tested on 9 datasets.

5.1.4 Random Forest with 500 Trees

According to the results in 4.6.2, 17 features were chosen during the selection phase using a Random Forest with 500 Trees classifier wrapper method after a GR filtering. Table 5.37 shows the results of a stratified 5-fold cross validation using a Random Forest with 500 Trees classifier with the selected 17 features on the retinal images of the dataset (S+EONMA).

Dataset: (S+EONMA) 5-fold CV	Laser	No Laser
Laser	157 (22.7%)	46 (77.3%)
No Laser	17 (2.6%)	635 (97.4%)

Table 5.37: Confusion Matrix after a 5-fold Cross Validation using RF500 of the images from the datasets S+EONMA.

The ROC curve in Figure 5.8 was obtained by thresholding the output of the implemented classifier RF500 after the stratified 5-fold cross validation. The area under ROC was also computed and has a value of $AUC_{(S+EONMA)}=0.934$.

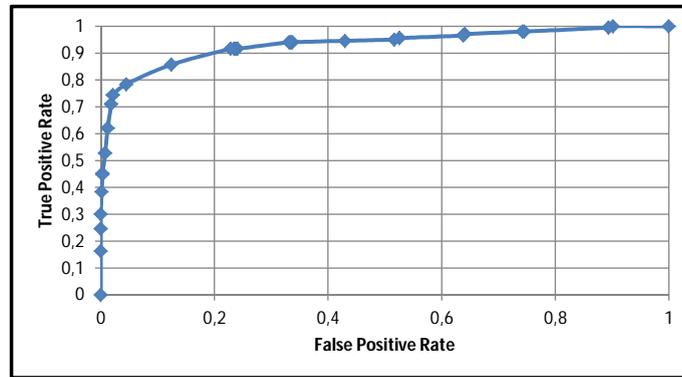


Figure 5.8: ROC Curve after thresholding the output of the RF500 classifier with 5-fold cross validation.

The RF500 classifier was trained using the dataset (S+EONMA) and then tested on the remaining 9 datasets of this work. The performance measures were calculated for each dataset: M (Table 5.38), EOMA (Table 5.39), EOEX (Table 5.40), EONEX (Table 5.41), VBR (Table 5.42), HP (Table 5.43), FAZD (Table 5.44), BAT (Table 5.45), PDJD (Table 5.46). After presenting the testing results of the BAT dataset (Table 5.45), an evaluation is performed concerning the number of patients correctly detected. A patient correctly detected contains at least one "Laser" image correctly classified and cannot have any "No Laser" image classified as having laser marks.

Dataset: M	Laser	No Laser
Laser	-	-
No Laser	32 (2.7%)	1155 (97.3%)

Table 5.38: Confusion Matrix after testing the trained RF500 on M.

Dataset: EOMA	Laser	No Laser
Laser	-	-
No Laser	2 (1.4%)	146 (98.6%)

Table 5.39: Confusion Matrix after testing the trained RF500 on EOMA.

Dataset: EOEX	Laser	No Laser
Laser	-	-
No Laser	1 (2.1%)	46 (97.9%)

Table 5.40: Confusion Matrix after testing the trained RF500 on EOEX.

Dataset: EONEX	Laser	No Laser
Laser	-	-
No Laser	1 (2.9%)	34 (97.1%)

Table 5.41: Confusion Matrix after testing the trained RF500 on EONEX.

Dataset: VBR	Laser	No Laser
Laser	-	-
No Laser	0 (0%)	22 (100%)

Table 5.42: Confusion Matrix after testing the trained RF500 on VBR.

Dataset: HP	Laser	No Laser
Laser	-	-
No Laser	1 (1%)	99 (99%)

Table 5.43: Confusion Matrix after testing the trained RF500 on HP.

Dataset: FAZD	Laser	No Laser
Laser	-	-
No Laser	4 (6.7%)	56 (93.3%)

Table 5.44: Confusion Matrix after testing the trained RF500 on FAZD.

Dataset: BAT	Laser	No Laser
Laser	23 (67.6%)	11 (32.4%)
No Laser	2 (13.3%)	13 (86.7%)

Table 5.45: Confusion Matrix after testing the trained RF500 on BAT.

In this case, 05_B1.jpg and 07_B2.jpg are the retinal images without laser marks that were classified as "Laser". As a result, patients 05 and 07 are not correctly classified by the algorithm. As for the 11 misclassified "Laser" images, at least one retinal image "Laser" per patient was correctly classified by the algorithm, except for patient 09 where none of the 3 "Laser" images was detected. Therefore, 6 out of 9 patients were accurately classified using the RF500 classifier.

Dataset: PDJD	Laser	No Laser
Laser	92 (91.1%)	9 (8.9%)
No Laser	-	-

Table 5.46: Confusion Matrix after testing the trained RF500 on PDJD.

Table 5.47 presents the RF500 classifier sensitivity and specificity values of the 5-fold cross validation using the dataset (S+EONMA) plus the 9 tested datasets.

Dataset	Sensitivity	Specificity
S+EONMA	77.3%	97.4%
M	-	97.3%
EOMA	-	98.6%
EOEX	-	97.9%
EONEX	-	97.1%
VBR	-	100%
HP	-	99%
FAZD	-	93.33%
BAT	67.6%	86.7%
PDJD	91.1%	-

Table 5.47: Algorithm's performance on each dataset using Random Forest with 500 Trees.

Once again, the retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749 images. This merged dataset was used to test the RF500 classifier and the results are shown in Table 5.48.

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity	PPV	NPV	AUC
96.40%	3.60%	85.2%	97.3%	0.728	0.987	0.963

Table 5.48: Performance stats using the trained Random Forest with 500 Trees averaged over all testing datasets.

The ROC curve was obtained by thresholding the output of the binary classifier RF500 after testing it on the merged dataset (Figure 5.9). The area under ROC was also computed and has a value of $AUC_{Test}=0.963$.

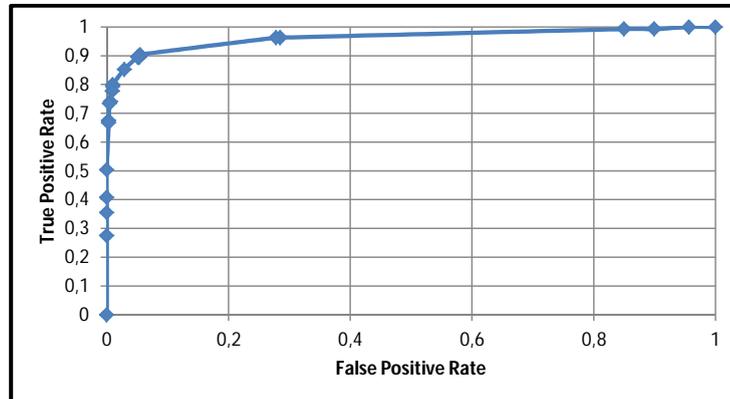


Figure 5.9: ROC Curve obtained by thresholding the output of the trained RF500 classifier tested on 9 datasets.

5.1.5 Simple threshold-based classifier

A threshold-based classifier was built using each segmentation algorithm's detected candidates, according to the description in subsection 4.5.3. The dataset (S+EONMA) was used for training in order to obtain the optimal threshold and the testing phase was the application of the selected threshold to the images from the merged dataset containing 1749 retinal images.

Figure 5.10 shows a ROC curve that was obtained by thresholding the number of candidates detected by the CHT algorithm and calculating the True Positive Rate and the False Positive Rate for each threshold value, using the (S+EONMA) dataset. The values used for thresholding vary between $[0,9]$ and the condition used was greater than or equal.

The best accuracy value (smallest classification error) was 91.46% and it occurred when $number_CHT \geq 3$ (optimal threshold). Table 5.49 shows the matrix confusion after applying the condition previously stated on the (S+EONMA) dataset.

According to the standard structure of this section, the retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749 images. This merged dataset was used to test the built threshold-based classifier and the results are shown in Table 5.50.

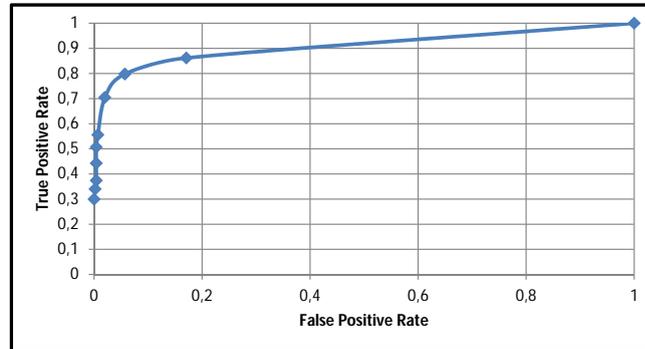


Figure 5.10: ROC Curve after thresholding by the number of regions detected by the CHT algorithm.

Dataset: (S+EONMA)	Laser	No Laser
Laser	143 (70.4%)	60 (29.6%)
No Laser	13 (2%)	639 (98%)

Table 5.49: Confusion Matrix after applying the optimal threshold for the number of regions detected by the CHT algorithm ($number_CHT \geq 3$ means that the retinal image is "Laser").

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity
96.34%	3.66%	79.3%	97.8%

Table 5.50: Performance stats using the simple threshold-based classifier $number_CHT \geq 3$ averaged over all testing datasets.

Figure 5.11 shows a ROC curve that was obtained by thresholding the number of candidates detected by the FVF algorithm and calculating the True Positive Rate and the False Positive Rate for each threshold value, using the (S+EONMA) dataset. The values used for thresholding vary between $[0,9]$ and the condition used was greater than or equal.

The best accuracy value (smallest classification error) was 86.31% and it occurred when $number_FVF \geq 3$ (optimal threshold). Table 5.51 shows the matrix confusion after applying the condition previously stated on the (S+EONMA) dataset.

According to the standard structure of this section, the retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749

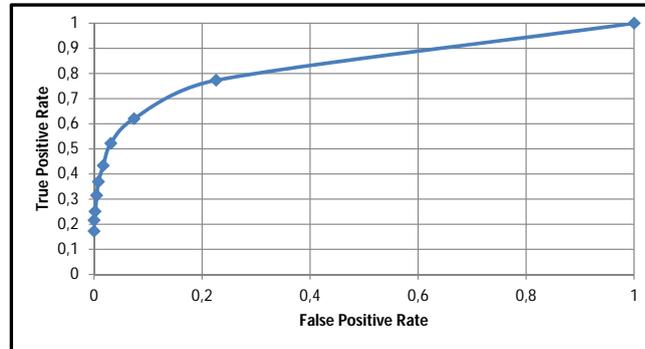


Figure 5.11: ROC Curve after thresholding by the number of regions detected by the FVF algorithm.

Dataset: (S+EONMA)	Laser	No Laser
Laser	106 (52.2%)	97 (47.8%)
No Laser	20 (3.1%)	632 (96.9%)

Table 5.51: Confusion Matrix after applying the optimal threshold for the number of regions detected by the FVF algorithm ($number_FVF \geq 3$ means that the retinal image is "Laser").

images. This merged dataset was used to test the built threshold-based classifier and the results are shown in Table 5.52.

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity
95.03%	4.97%	78.5%	96.4%

Table 5.52: Performance stats using the simple threshold-based classifier $number_FVF \geq 3$ in every testing dataset.

Figure 5.12 shows a ROC curve that was obtained by thresholding the number of candidates detected by the LMS algorithm and calculating the True Positive Rate and the False Positive Rate for each threshold value, using the (S+EONMA) dataset. The values used for thresholding vary between [0,9] and the condition used was greater than or equal.

The best accuracy value (smallest classification error) was 83.97% and it occurred when $number_LMS \geq 5$ (optimal threshold). Table 5.53 shows the matrix confusion after applying the condition previously stated on the (S+EONMA) dataset.

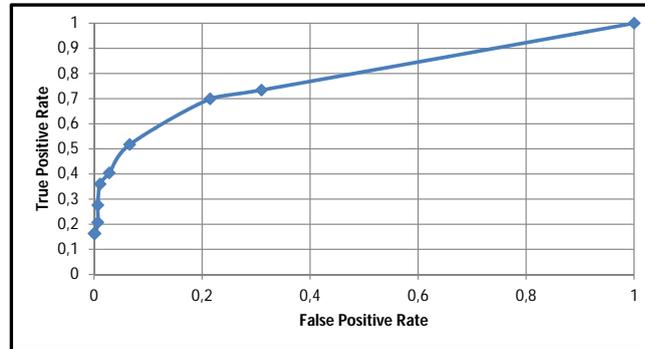


Figure 5.12: ROC Curve after thresholding by the number of regions detected by the LMS algorithm.

Dataset: (S+EONMA)	Laser	No Laser
Laser	73 (36%)	130 (64%)
No Laser	7 (1.1%)	645 (98.9%)

Table 5.53: Confusion Matrix after applying the optimal threshold for the number of regions detected by the LMS algorithm ($number_LMS \geq 5$ means that the retinal image is "Laser").

According to the standard structure of this section, the retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749 images. This merged dataset was used to test the built threshold-based classifier and the results are shown in Table 5.54.

Correctly Classified	Incorrectly Classified	Sensitivity	Specificity
95.25%	4.75%	49.6%	99.1%

Table 5.54: Performance stats using the simple threshold-based classifier $number_LMS \geq 5$ in every testing dataset.

5.2 Computation Time

An important characteristic of the algorithm is its typical computation time per retinal image. This is specially true in the cases where one wants to do laser mark detections in real-time to allow online applications.

Training the tree-based classifiers using WEKA consumes computation time. For instance, the time taken was:

- 0.14 seconds to train the Decision Tree using the dataset (S+EONMA) with the selected 13 features.
- 0.16 seconds to train the Random Forest with 5 Trees using the dataset (S+EONMA) with the selected 15 features.
- 0.28 seconds to train the Random Forest with 50 Trees using the dataset (S+EONMA) with the selected 14 features.
- 3.24 seconds to train the Random Forest with 500 Trees using the dataset (S+EONMA) with the selected 17 features.

However, these computation costs are incurred only once during classifier training and so can be disregarded and not included in the following analysis presented for the computational complexity of the classification.

In summary, having the classifier built with the selected features, each retinal image has to go through four time-consuming phases until the proposed algorithm makes the final decision: pre-processing, candidate regions extraction, features computation and the test performed by the built classifier.

The resizing step plus the extraction of the ROI, OD and the Vascular Network are performed by Retmarker[®] taking on average about 9 seconds per image with a resolution of 768x584. For the remaining steps of the pre-processing phase plus the candidate regions extraction plus the computation of all 65 features, the average processing times per image for each resolution are presented in Table 5.55. Figure 5.13 shows that for a higher total number of pixels, the computation time per image increases. Note that the resolution values presented are obtained after the resizing step.

For example, each retinal image with a resolution of 768x584 takes a total of 27.95 seconds (9s+18.95s) to process and obtain all the 65 features. Then each image will be tested by the built classifier, which is almost instantaneous, so it can be assumed that for each retinal image with the previously mentioned resolution the proposed algorithm takes approximately a total of 27.95 seconds to make the final decision. However, if we

Image Resolution	Number of Pixels	Computation Time per Image (s)
720x576	414720	21.93
768x584	448512	18.95
768x768	589824	28.65
1024x680	696320	28.68
1024x681	697344	28.52
1024x683	699392	28.73
1024x687	703488	32.12
1024x729	746496	39.11
806x768	619008	28.54
782x768	600576	30.56
937x768	719616	42.89
938x768	720384	38.63
948x768	728064	41.85

Table 5.55: Computation time per image of the algorithm for each resolution.

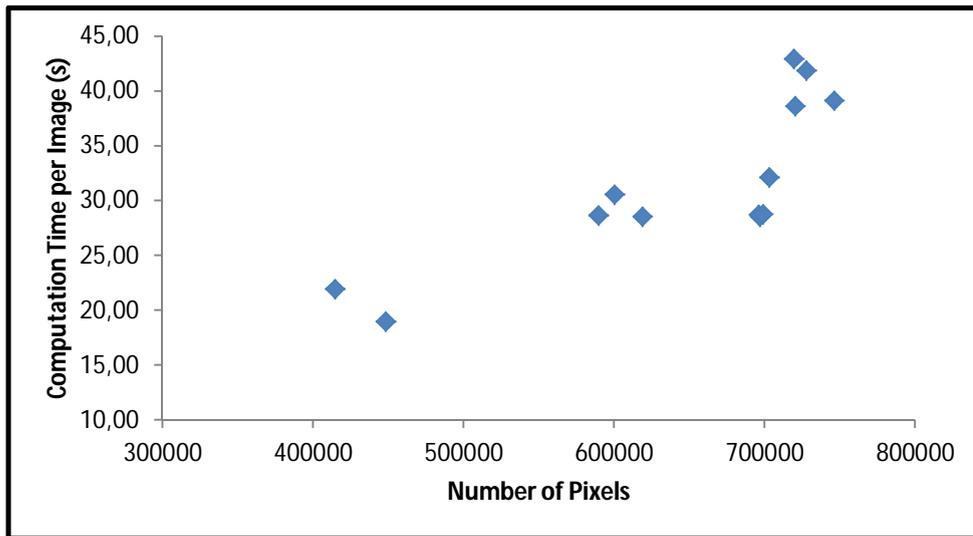


Figure 5.13: Plot of the Computation time per image against the total number of pixels.

consider the proposed algorithm integrated within the Retmarker software package, the 9 seconds are not contemplated as additional computation time since the pre-processing step of Retmarker's algorithm is applicable to all images, hence the effective computation time would be the one presented in table 5.55.

The processing time for each retinal image was measured using a non-optimized MATLAB®

implementation of the proposed using a machine with an Intel[®] Xeon I5520 processor with a speed of 2.27 GHz and equipped with 24 GB of RAM in each node.

Chapter 6

Discussion

Each of the four tree-based classifiers has its own best subset of features and the performance measures were obtained in chapter 5 after testing the algorithm with the different classifiers in 9 different datasets. Overall, for the 5-fold cross validation with (S+EONMA) dataset, the sensitivity value was higher than 74.8% and the specificity value was higher than 94.8% and for the merged dataset with 1749 images the sensitivity value was higher than 85.1% and the specificity value was higher than 97.2%. The RF500 classifier had the best accuracy during the 5-fold cross validation since it correctly classified 92.6% of the retinal images of (S+EONMA). The Decision Tree had the best accuracy while testing the remaining 1749 images because 98.1% of those images were correctly classified. It is also interesting to analyze the results in chapter 5 class-by-class. For instance, RF50 and DT were both the classifiers that correctly detected most images "Laser", having correctly detected 274 out of a total of 338 images "Laser", which means an overall sensitivity of 81.1%. Regarding the class "No Laser", DT was by far the best classifier, once it presented the best specificity value for every single dataset. Therefore, DT correctly detected 2232 out of a total of 2266 "No Laser" images, having an overall specificity of 98.5%. Moreover, a particular evaluation was performed using the BAT dataset for each classifier. BAT is a special proprietary dataset because it contains images from 9 patients before and after the treatment. It is therefore important that the implemented algorithm distinguishes the different retinal images for each patient. Although RF500 was the one that correctly classified most images from the BAT dataset (36 of 49), the DT was the classifier that detected more

accurately the patients (7 of 9), considering that a single false positive or no true positive detected for each patient means a failed classification for that patient. A brief note for the simple threshold-based classifier: the decision method that used the number of regions detected by the CHT algorithm as the criteria was the one with the best performance amongst this type of classification. As expected, this threshold approach had the lowest classification accuracy (96,34%) compared to the tree-based classifiers.

Regarding the performance of the four implemented classifiers, there is no doubt that the Decision Tree showed the best results when comparing to the results obtained with the Random Forests. In fact, it had the best sensitivity, specificity and accuracy values. Also, it is the simplest classifier to implement, the generated model is very easy to interpret (Figure 5.1) and its training time is only 0.14 seconds, which means that it allows the possibility of retraining the classifier more frequently after gathering more retinal images. Not only it is the best classifier for its accuracy and simplicity, but also it is the most adequate for clinical practice. The proposed algorithm with DT classifier is the best option to integrate a DR screening program, because it is the one with best specificity values which is the most important requirement since it is not acceptable to remove a patient for having laser marks, when in fact that same patient as not undergone any kind of treatment.

The proposed automatic system relies on 2 main steps: image processing and classification. The importance of the image processing step is to retrieve the features from any retinal image that are used for the final classification. The testing results on the datasets described in this thesis confirm that the proposed algorithm is robust to images with different resolution, FOV angle and captured by different cameras. For instance, this fact can be confirmed by observing every specificity values in chapter 5 which do not vary very much. However, a considerable difference can be observed in the sensitivity values between datasets S, BAT and PDJD. Any classifier can detect between 60% and 80% of the retinal images "Laser" from S and BAT datasets, while for PDJD the percentage of images "Laser" detected is above 90%. The source images from S and BAT datasets are labeled and identified by optometrists during screening initiatives and by ophthalmologists in clinical environments, respectively, during a certain period of time. On the other hand, the images from PDJD (which are from the same retinal camera *Canon* used to capture the images in S) were carefully selected by image processing experts and all the chosen images very clearly show PRP scars, whereas images where

the marks were not clear would not be selected, *i.e.* it is less realistic. It is important to remind that this was the first work and experiment in laser marks detection.

Relative to the features selected, there are some noteworthy remarks. First, none of the 27 texture descriptors was selected in any of the filter+wrapper methods. Most of these texture features are removed during filtering, whether by using information gain or gain ratio. Nevertheless, this observation means that texture statistical measures do not contribute to detect laser marks at all. The accuracy of a wrapper method is a prediction of the classification quality of the corresponding classifier with the selected best subset. In subsection 4.6, every attribute subset chosen had at least geometrical and spatial distribution descriptors. Indeed, features related to the CHT (number_CHT, area_CHT, weighted_area_CHT, likelihood_CHT) and to the FVF (number_FVF, area_FVF, weighted_area_FVF, likelihood_FVF) segmentation algorithms were very frequent. Moreover, features related to spatial dispersion (max_eigen, determinant and trace), convex hull of the candidate regions (hull and hull_area) and autocorrelation (morani and moran_null_hipot) were shown to be almost always included in the best subsets and consequently be important for the final classification. Not less important are the intensity-related features proposed by Fazar [47]. Although these features do not appear in every selected subset, they have a very important role on the Decision Tree. Despite the fact that four spatial distribution features are used to train the Decision Tree, only hull_area is part of the built model, while mean_laser_green, var_laser_green and var_laser_red are the intensity-related features used in the terminal nodes of the built Decision Tree (Figure 5.1).

As for the tree-based classifiers used, the Random Forest is an ensemble learning method that operates by constructing several decision trees (in this work it is used with 5, 50 and 500 trees) at training time that uses a random subset of features at each node split. Originally, random forests were developed to correct the tendency of decision trees to overfit, but it is interesting to verify in the results that the decision tree proved to be the best classifier to be used in this application.

As it was already stated, in a clinical context it is more important to classify correctly the non-treated patients. However, the algorithm's specificity is not 100% for any of the classifiers. Therefore, there are retinal images "No Laser" where the features extracted from the candidates are similar to the standard set of features of an image "Laser".

The candidates falsely detected are the trigger that mislead the tree-based classifiers and, as a consequence, incorrect outputs are obtained. Some examples of retinal images that were misclassified by the Decision Tree and the Random Forests can be found on Appendix D. For example, images 06_B1.jpg (Figure D.4) and 07_B2.jpg (Figure D.5) represent very severe PR and a huge number of haemorrhages and small EXs can be seen. These structures promote differences of contrast that are segmented as candidates and the largest haemorrhages are detected as blobs by the FVF algorithm. Image 0219.jpg (Figure D.1) from the screening dataset contains some artifacts: the round bubbles resulting from the fact that the lens is not cleaned are detected as circles by the CHT. Images 0225.jpg (Figure D.2) and 0419.jpg (Figure D.3) represent thin retinas, which results in highly illuminated images with bright spots that are detected as candidates. Furthermore, some candidates in image 0419.jpg are detected near small branches of blood vessels, which means that the vessel extractor is also not 100% efficient. As for the images 15.jpg (Figure D.6), 28.jpg (Figure D.7) and 32.jpg (Figure D.8) from the FAZD dataset, they all presented EXs with a huge size and since these are bright structures, the probability of false detection by the CHT algorithm is very high. Moreover, 32.jpg, as well as image 20051214_51733_0100_PP.jpg (Figure D.10) from the Messidor dataset, presents some failures in the vascular network detection. Images 20051130_60691_0400_PP.jpg (Figure D.9) and 20051214_51733_0100_PP.jpg from the Messidor dataset are also thin retinas, so the candidate region extraction is not the most desirable.

Figure 5.13 confirms that a retinal image with a higher resolution results in a longer processing step of the image. It is not possible to state that it is a linear relation because there are images with a similar total number of pixels that have quite different processing times. For example, the images with a resolution of 1024x680, 1024x681, 1024x683 and 1024x687 (696320, 697344, 699392, 703488 pixels respectively) are “No Laser” and take about 30 seconds, due to the fact that not many candidates are detected and analyzed. On the other hand, the images with a resolution of 806x768 and 782x768 (619008 and 600576 pixels respectively) are “Laser” and take about 30 seconds and the images with a resolution of 97x768, 938x768 and 948x768 (719616, 720384 and 728064 pixels respectively) are “Laser” and take about 40 seconds, because many regions are extracted and processed to calculate most features.

In the screening program coordinated by ARS Centro, the automatic DR detector takes

about 90 seconds per retinal image, using MATLAB. The proposed laser mark detector takes between 18 and 42 seconds depending on the resolution. A great portion of processing time is spent on the LMS algorithm, on the bilateral filter of the CHT, on removing the false detections located on the vessels and on the OD and on computing all 65 features, which would not happen in an integration with a screening program because it was already confirmed that only a small subset of features is sufficient to obtain a very good performance.

Chapter 7

Conclusion and Future Work

This thesis contribution consists of a robust automatic detection algorithm for signs of treatment on digital retinal fundus images and preparation of a paper describing the algorithm created and providing the datasets *Screening* and *Before and After Treatment* for public access. In order to do so, the steps followed were:

- Studying the problem, reviewing the scientific bibliography about DR complications and the different type of treatment practices and researching previous work on this topic.
- Gathering retinal fundus images datasets with and without laser marks.
- Developing the algorithm by organising it in three parts: extracting candidates in each retinal image, computing features from the candidates and then using a tree-based classifier to label the image as "Laser" or "No Laser".
- Evaluating the performance of the algorithm on the gathered datasets.

Using the described automated algorithm it is possible to identify a considerable amount of images with laser marks, a noteworthy result considering the complexity of the task even for a human expert.

The simple threshold-based classifier had a poor classification performance, proving that using solely the number of detected regions is insufficient for this classification problem. Although it is an extremely simple approach, the accuracy of the final decision is worse.

After testing the algorithm with three Random Forest classifiers and with a Decision Tree classifier, the last one proved to be more effective, as confirmed in chapter 6. As a matter of fact, using DT with the 13 features listed in Table 4.2 resulted in an overall sensitivity of 81.1% and in an overall specificity of 98.5%. This final result indicates that the proposed robust algorithm detects a reasonable quantity of retinal images with laser marks. However, even if a specificity of 98.5% seems to be a high measure value, it is still not enough to implement the automatic system in a screening program because it is much less tolerable to remove patients that have not undergone under treatment whether they are healthy or not, even if it is a small percentage.

Taking the results mentioned before into account, the performance of the implemented algorithm is comparable to that of Kotoula *et al.* [45], of João Dias *et al.* [46] and of Faraz *et al.* [47]. The first referred work consisted in a manual grading of three ophthalmologists where they had to identify images with signs of Macular Edema Photocoagulation, which are much smaller and much more difficult to detect. Nevertheless, the performance by the ophthalmologists when observing digital colour retinal fundus images was very poor: 58% sensitivity and 81.1% specificity. Although Kotoulo's work only considers MEP-related scars and the proposed algorithm does not discriminate the type of photocoagulation scar, it is clear that the classification algorithm has a better performance than the ophthalmologists themselves. João Dias proposed the first automatic laser mark detection algorithm. The specificity of 99.9% achieved is outstanding and for that reason it is much more resistant to false positives than the algorithm proposed in this thesis (98.5%). However, there is a remarkable difference in the capacity of detecting laser marks: while João Dias' algorithm only captures 63.37% of the treated patients, the current algorithm detects 97% of those same patients. The approach used in [46] is based on quality parameters and therefore only the most evident cases are detected. Fazar's consists in a classical image processing classification problem that follows a very similar approach to the one used in this work. It uses different image processing techniques, different features and a different classifier (minimum distance). The paper [47] reports a sensitivity of 94% and a specificity of 97%, but no comparison can be made for the same images, since the dataset is not public. It is important to note that they included images with MAs and EXs in the "No Laser" category, so at first it is possible to conclude that this algorithm is more vulnerable to false positives than the one proposed in this thesis. As for the retinal images "Laser", the content and criteria

is completely unknown, hence they might have chosen manually the most evident ones, like in João Dias work.

The value of this work is in the fact that it uses simple, effective and robust to noise and resolution segmentation algorithms like the Circular Hough Transform and the Frangi Filter plus a very simple to interpret and to implement tree-based classifier in order to accurately solve a quite complex problem - detecting structures (laser marks) with variable shape, size and colour. The other novelty are the type of features used: texture descriptors proved to have no importance in this specific classification; features derived directly from the candidate regions properties are the most relevant (specially the ones related to CHT); spatial distribution descriptors are in fact crucial to determine if the segmented candidates have a regular or dispersed pattern characteristic of patients with laser photocoagulation scars; it was confirmed that the intensity-based features proposed by Faraz are important and effective to classify images with laser marks.

The automatic laser mark detection algorithm described in this work should be subjected to further work in order to improve the performance of the current algorithm. The priority should be focusing in alternatives that allow less detection of false positives. It is true that it is important to detect laser marks in the retinal images "Laser" where they were not detected, but more important than that is to reduce the number of candidate regions in retinal images without treatment applied, so that specificity is closer to 100% and therefore applicable in clinical context.

One possibility is finding a criteria to remove false detected candidate regions. For example, regions detected nearby the OD and the blood vessels are removed, so identifying other characteristics that would indicate that the region detected is not a laser mark (such as large EXs or bright spots resulting from thin retinas as described in chapter 6) would be one valid approach. Besides, although the vessel extractor is quite effective, it would be important to improve it in order to avoid detections in blood vessels not segmented. Another interesting approach would be researching and possibly adding another blob detector that makes use of the Hessian matrix. Moreover, finding more features is a key to add more relevant features to the classifier in order to get a better performance. According to the remarks written in this document, there is no need to research for texture descriptors and finding new spatial distribution descriptors are potentially the best type of features to add to this work. For example, geographical and

econometrics tools contain many statistical measures that describe the distribution of points in a 2D plane (for example, Ripley K is a good measure of point dispersion). Finally, there is a very wide variety of classifiers. Only simple tree-based classifiers are used in this work, so it is possible that using other more complex classifiers such as naive bayes, SVM, neural networks or even fuzzy logic may increase the accuracy.

Furthermore, in order to implement the proposed algorithm in the Retmarker family of products, the computation time should be reduced. Although the computation time increases for higher resolutions, it is possible to fasten the algorithm by optimizing the script in MATLAB and by converting it to a compiled language, such as C++. Also, the algorithm spends a great percentage of its time on candidate region extraction, specially on the LMS and on the bilateral filter of the CHT algorithm that together take approximately 14 seconds in retinal images “Laser” with a resolution of 768x584.

In conclusion, the contribution of this thesis was to propose an automatic algorithm with a very promising performance, as confirmed by the results, that detects treatment signs of diabetic retinopathy in digital colour fundus images by combining the results of three state-of-the-art segmentation algorithms adapted to this context, extracting and selecting the most important features and adding a decision tree classifier. There was an additional effort towards publishing two datasets containing retinal images with laser marks that serve as benchmarks for future work developed on this complex task even for a human expert.

Appendix A

Performance and Information Theory Measures

Sensitivity and **Specificity** are two parameters that measure the quality of a classifier, its ability to correctly classify instances and are computed according to Equations A.1 and A.2. The positive case is when a retinal image is classified as having laser marks.

$$\textit{Sensitivity} = \frac{TP}{TP + FN} \quad (\text{A.1})$$

$$\textit{Specificity} = \frac{TN}{TN + FP} \quad (\text{A.2})$$

where

- TP are true positives, meaning the number of retinal images with laser marks correctly classified.
- TN are true negatives, meaning the number of retinal images without laser marks correctly classified.
- FP are false positives, meaning the number of retinal images without laser marks classified as "Laser".
- FN are false negatives, meaning the number of retinal images with laser marks classified as "No Laser".

Positive Predictive Value and **Negative Predictive Value** are two measures that indicate the probability of an instance being correctly classified. Positive Predictive Value is the percentage of retinal images classified as "Laser" that actually present laser marks (Equation A.3) and Negative Predictive Value is the percentage of retinal images classified as "No Laser" that actually do not present signs of laser treatment (Equation A.4).

$$\text{Positive Predictive Value} = \frac{TP}{TP + FP} \quad (\text{A.3})$$

$$\text{Negative Predictive Value} = \frac{TN}{TN + FN} \quad (\text{A.4})$$

Information Gain an attribute evaluator based on entropy, a common information theory measure. Let S be a set that is composed by s data samples with m classes, in this case $m = 2$. The entropy, which is the expected information necessary to classify a sample is given by

$$H(S) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (\text{A.5})$$

where p_i is the probability that a random sample belongs to class C_i and is equal to $\frac{s_i}{s}$.

Considering that feature F has n distinct values and that s_{ij} is the number of samples of class C_i in a subset S_j that has the samples in S that have a value f_j of F . Then, the entropy based on the partitioning into subsets by F is given by

$$H(F) = - \sum_{j=1}^n H(S) \frac{s_{1j} + s_{2j} + \dots + s_{mj}}{s} \quad (\text{A.6})$$

Therefore, the Information Gain of feature F in relation to both classes "Laser" and "No Laser" is the difference between the two entropies mentioned above, as described in Equation A.7 [100, 103].

$$IG(F) = H(S) - H(F) \quad (\text{A.7})$$

Gain Ratio corrects the issue involving the bias of the Information Gain attribute evaluator. It normalises the Information Gain measure of a feature F by taking the intrinsic information into account. Intrinsic information (Equation A.8) is the information generated by splitting the training dataset S into n partitions corresponding to n outcomes of a test on the feature F .

$$SI_S(F) = - \sum_{j=1}^n \frac{|S_j|}{|S|} \log_2 \left(\frac{|S_j|}{|S|} \right) \quad (\text{A.8})$$

Therefore, the Gain Ratio measure of a feature is the Information Gain measure of that feature divided by the intrinsic information, as described in Equation A.9. Gain Ratio decreases as intrinsic information gets larger [103, 105].

$$GR(F) = \frac{IG(F)}{SI_S(F)} \quad (\text{A.9})$$

Appendix B

Datasets

Messidor ID
20051020_44843_0100_PP.png
20051020_54209_0100_PP.png
20051021_40377_0100_PP.png
20051021_40450_0100_PP.png
20051021_59459_0100_PP.png
20051021_59504_0100_PP.png
20051205_31396_0400_PP.png
20051212_41432_0400_PP.png
20051213_62648_0100_PP.png
20051214_40912_0100_PP.png
20051214_41582_0100_PP.png
20051214_56821_0100_PP.png
20051214_57940_0100_PP.png

Table B.1: List of images not considered from Messidor Dataset.

Screening Dataset Folder	Files
Laser	0001.jpg to 0203.jpg
No pathologies	0204.jpg to 0622.jpg excluding the images in Drusens, Exudates and Microaneurysms folders
Drusens	0212.jpg, 0284.jpg, 0352.jpg, 0354.jpg, 0355.jpg, 0383.jpg, 0391.jpg, 0424.jpg, 0453.jpg, 0460.jpg, 0467.jpg, 0473.jpg, 0488.jpg, 0505.jpg, 0532.jpg, 0540.jpg, 0548.jpg, 0577.jpg, 0608.jpg, 0612.jpg
Exudates	0204.jpg, 0220.jpg, 0233.jpg, 0267.jpg, 0348.jpg, 0382.jpg
Microaneurysms	0237.jpg, 0340.jpg, 0360.jpg, 0379.jpg, 0513.jpg, 0544.jpg

Table B.2: Summary of Screening Dataset.

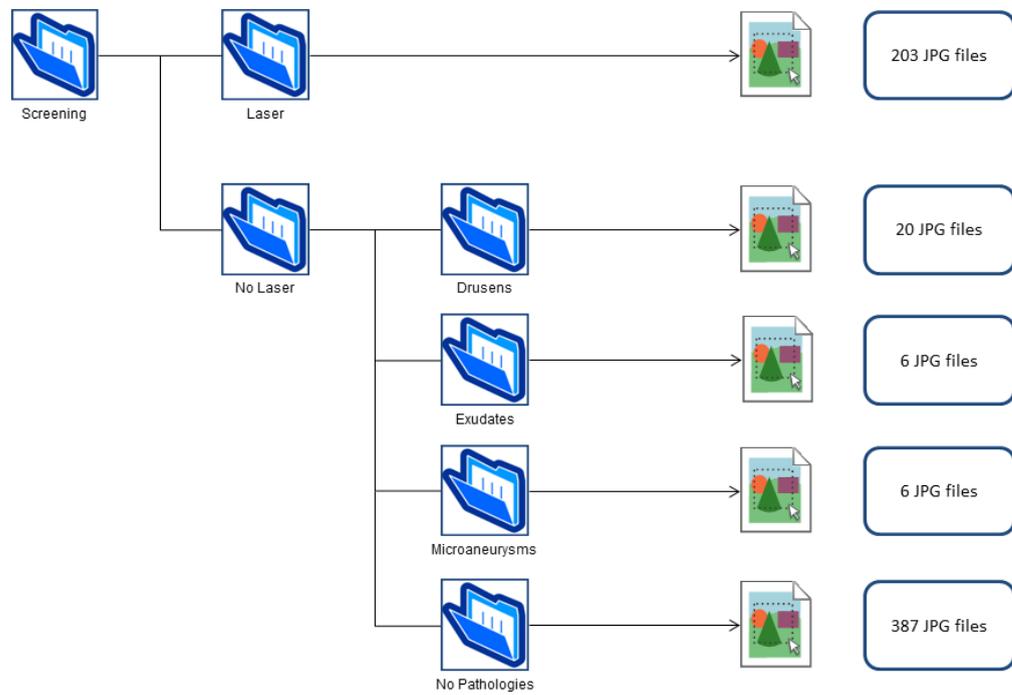


Figure B.1: Overview of Screening Dataset.

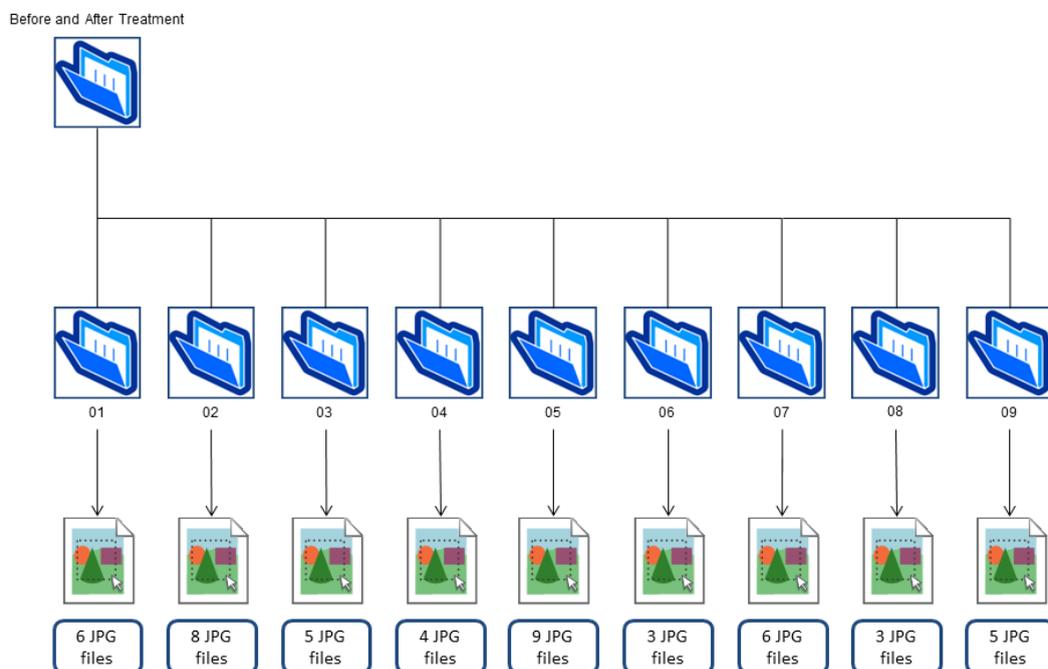


Figure B.2: Overview of BAT Dataset.

Patient	# Images "No Laser"	# Images "Laser"
01	2	4
02	2	6
03	2	3
04	1	3
05	1	8
06	2	1
07	2	4
08	1	2
09	2	3

Table B.3: Details of BAT Dataset.

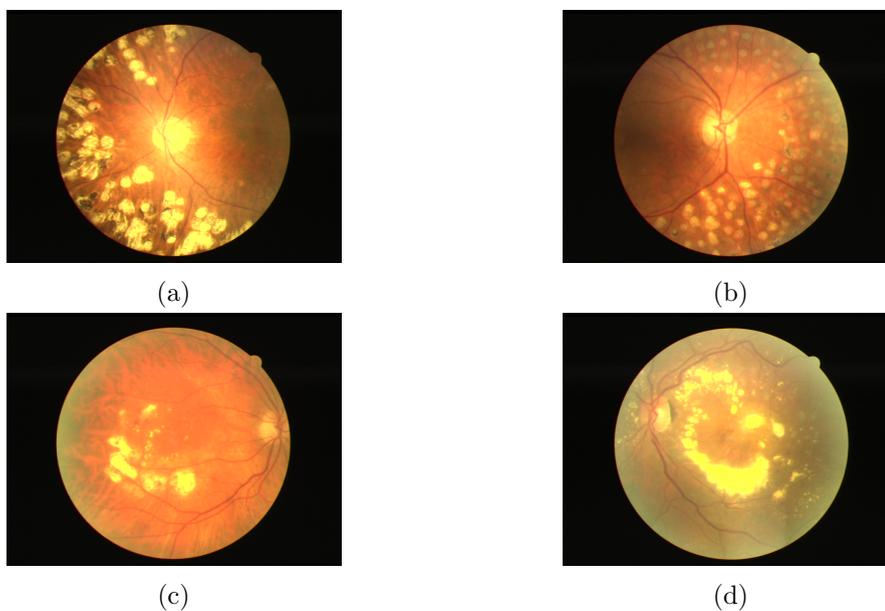


Figure B.3: Examples of images with laser marks from the Screening Dataset: (a),(b) Retinal images of patients treated with PRP (c),(d) Retinal images of patients treated with MEP.

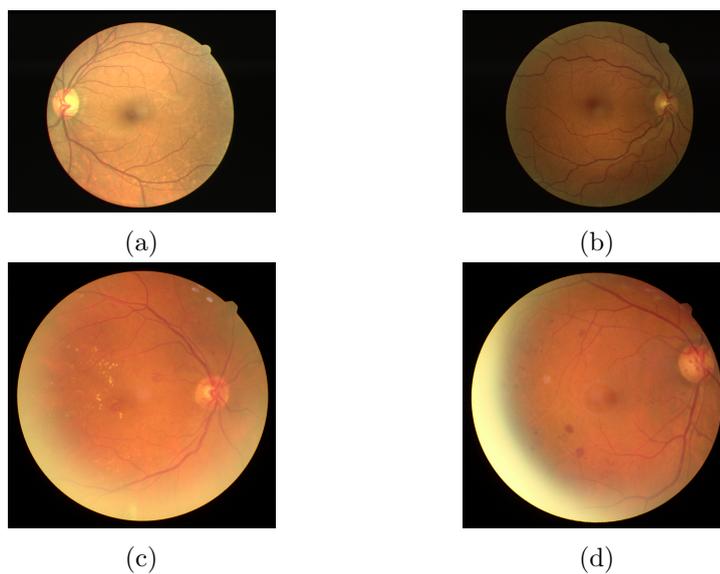


Figure B.4: Examples of images without laser marks from the Screening Dataset: (a) Retinal image of a patient with drusens (b) Retinal image of a healthy person (c) Retinal image of a patient with EXs (d) Retinal image of a patient with MAs.

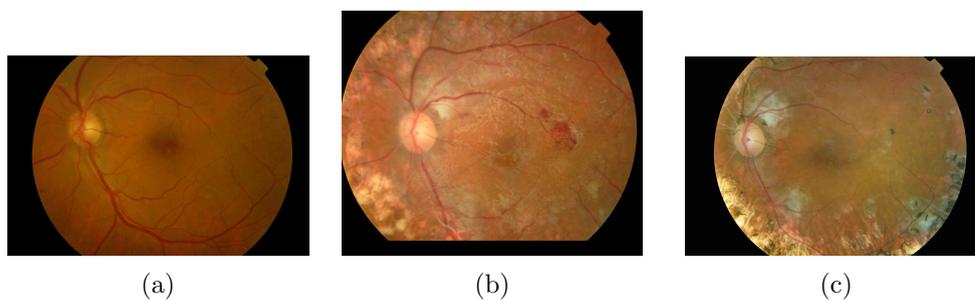


Figure B.5: Images of patient 08 from the BAT Dataset: (a) Retinal image before treatment (healthy) (b),(c) Retinal images after PRP.

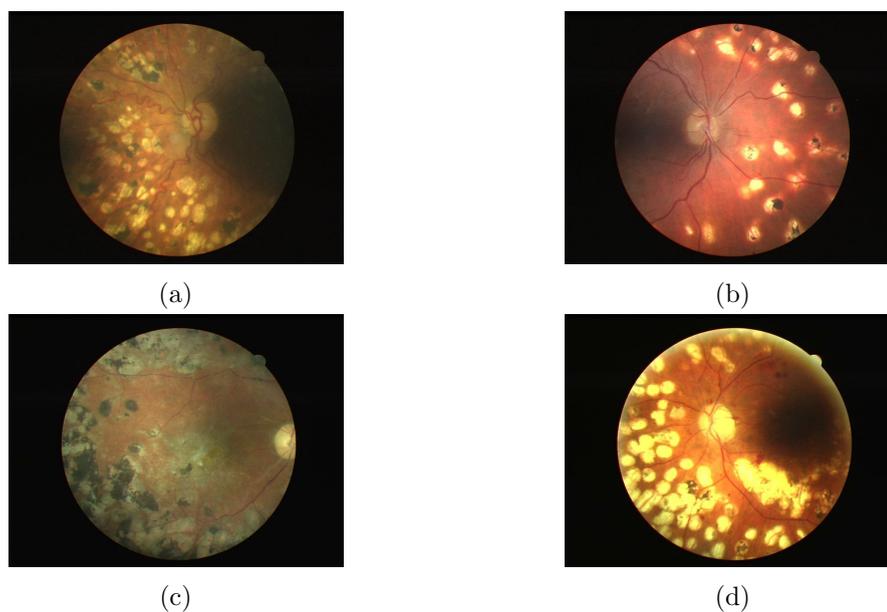


Figure B.6: Examples of images with laser marks from the PDJD.

Appendix C

Pre-processing Examples

This section of the Appendix presents the results obtained during Pre-processing (Subsection 4.2) and Candidate Regions Extraction (Subsection 4.3) of three other retinal images with laser marks besides the one used in the main text: 0002.jpg, 0132.jpg and 0198.jpg. Also, some examples of retinal images that support the explanation of the *scaling_constant* expression in section 4.2.5 are given at the end of this section.

Figure C.4 represents the retinal image 0010.jpg ($ROI_diameter = 447$ and $OD_diameter = 112$) from the Screening dataset with the candidates to laser marks drawn after using $scaling_constant = \frac{ROI_diameter}{480}$ (Subfigure C.4b) and after using $scaling_constant = \frac{ROI_diameter - OD_diameter}{370}$ (Subfigure C.4c). Even though there are few false detections, subfigure C.4c shows more laser marks detected.

Figure C.5 represents the retinal image 0189.jpg ($ROI_diameter = 479$ and $OD_diameter = 118$) from the Screening dataset with the candidates to laser marks drawn after using $scaling_constant = \frac{ROI_diameter}{480}$ (Subfigure C.5b) and after using $scaling_constant = \frac{ROI_diameter - OD_diameter}{370}$ (Subfigure C.5c). Even though there are few false detections, subfigure C.5c shows more laser marks detected.

Figure C.6 represents the retinal image 04_A1.jpg ($ROI_diameter = 836$ and $OD_diameter = 128$) from the BAT dataset with the candidates to laser marks drawn after using $scaling_constant = \frac{ROI_diameter}{480}$ (Subfigure C.6b) and after using $scaling_constant = \frac{ROI_diameter - OD_diameter}{370}$ (Subfigure C.6c). There are no laser marks detected in subfigure C.6b. On the contrary, the retinal image in subfigure C.6c presents two laser marks detected.

The retinal images from the EOEX dataset displayed in figure C.7 have the same $ROI_diameter = 698$. However, the OD's size are very different from each other. The retinal image in subfigure C.7a has $OD_diameter = 142$, while the retinal image in subfigure C.7b has $OD_diameter = 187$ (it can be observed that the vessels are also larger).

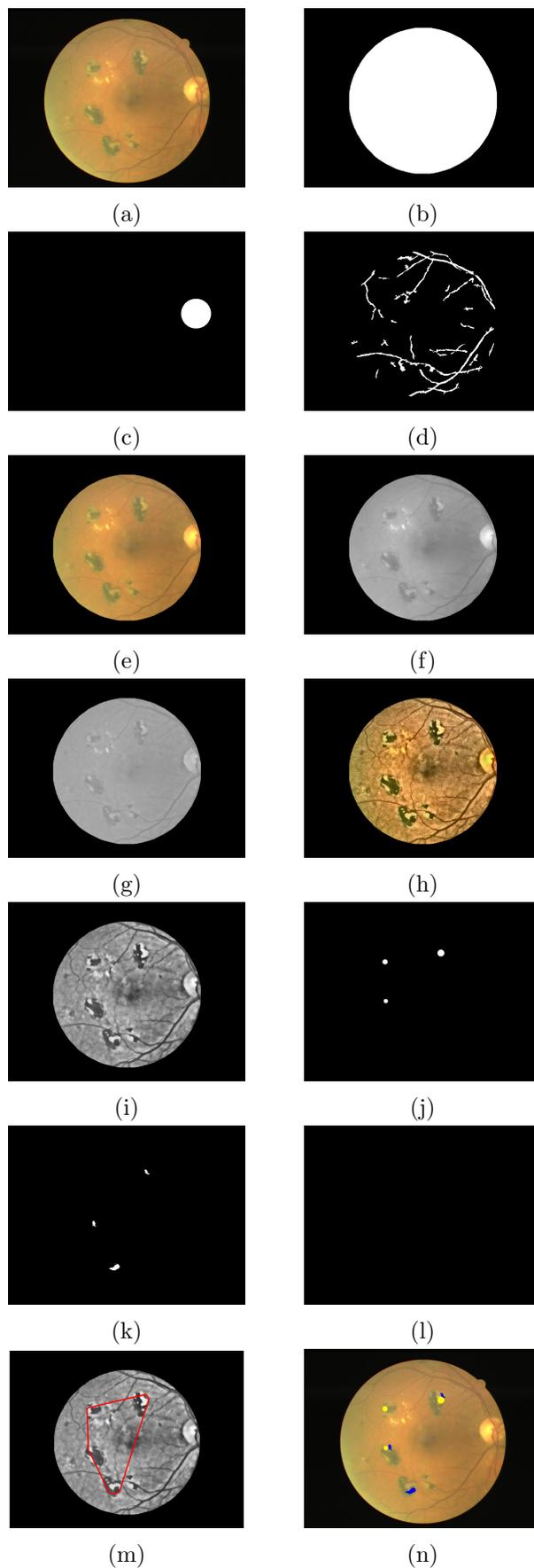


Figure C.1: Pre-processing outputs and candidates extraction of image 0002.jpg from the Screening Dataset: a) Original image b) ROI c) OD d) Vascular Network e) Image cropped f) L^* dimension before uneven illumination correction g) L^* dimension after uneven illumination correction h) After adaptive histogram equalization i) Processing image j) CHT binary output k) FVF binary output l) LMS binary output m) Convex Hull n) Original image with candidates drawn - yellow=CHT; blue=FVF.

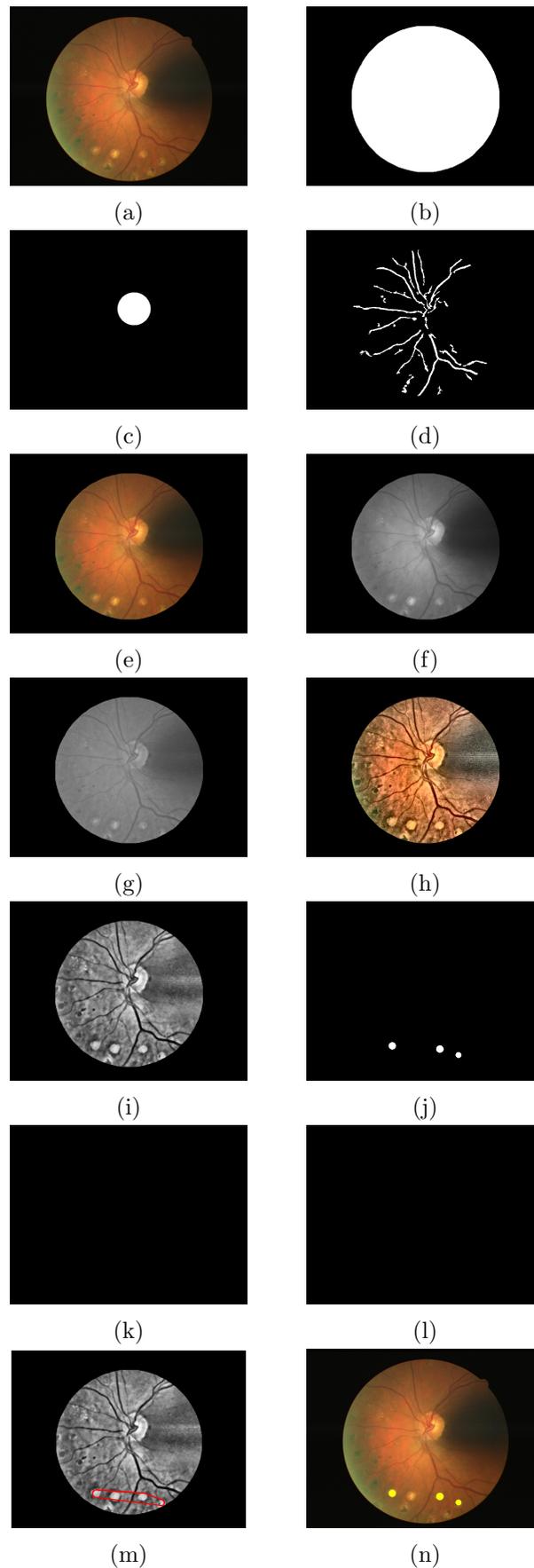


Figure C.2: Pre-processing outputs and candidates extraction of image 0132.jpg from the Screening Dataset: a) Original image b) ROI c) OD d) Vascular Network e) Image cropped f) L^* dimension before uneven illumination correction g) L^* dimension after uneven illumination correction h) After adaptive histogram equalization i) Processing image j) CHT binary output k) FVF binary output l) LMS binary output m) Convex Hull n) Original image with candidates drawn - yellow=CHT.

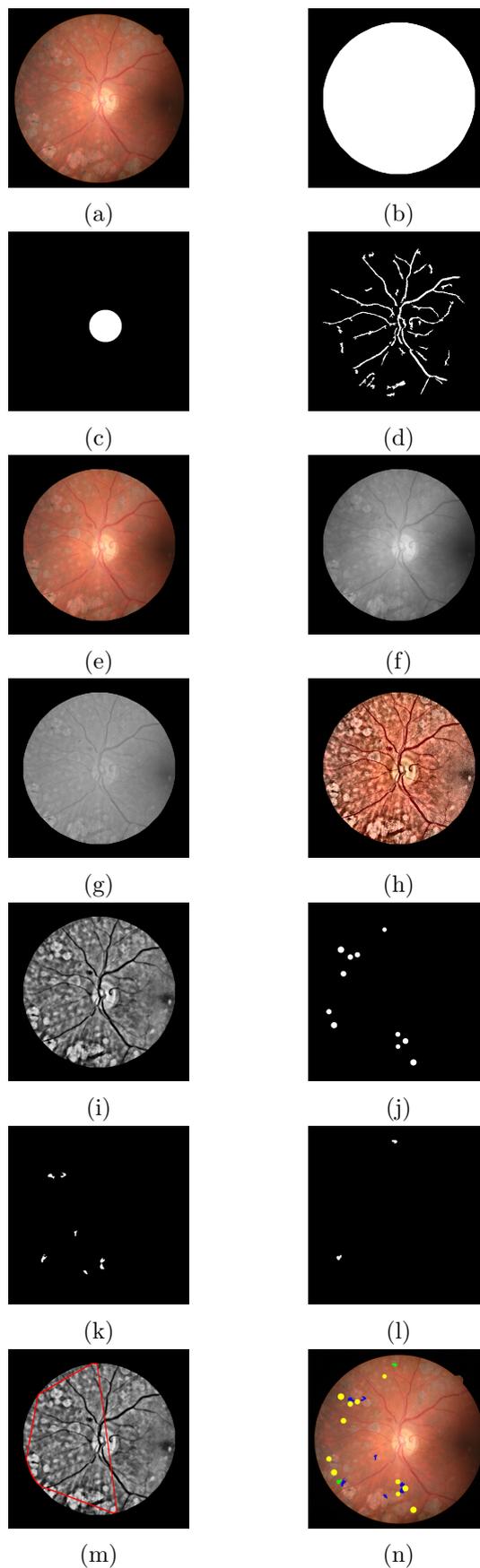


Figure C.3: Pre-processing outputs and candidates extraction of image 0198.jpg from the Screening Dataset: a) Original image b) ROI c) OD d) Vascular Network e) Image cropped f) L^* dimension before uneven illumination correction g) L^* dimension after uneven illumination correction h) After adaptive histogram equalization i) Processing image j) CHT binary output k) FVF binary output l) LMS binary output m) Convex Hull n) Original image with candidates drawn - yellow=CHT; blue=FVF; green=LMS.

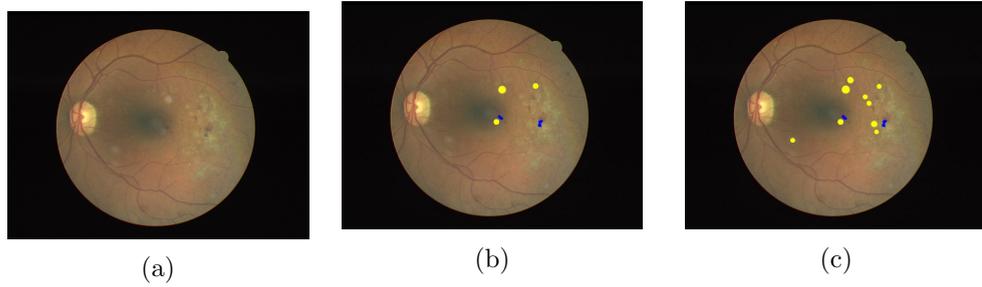


Figure C.4: Retinal image 0189.jpg from the S dataset: (a) original (b) candidates detected using $scaling_constant = \frac{ROI_diameter}{480}$ (c) candidates detected using $scaling_constant = \frac{ROI_diameter - OD_diameter}{370}$

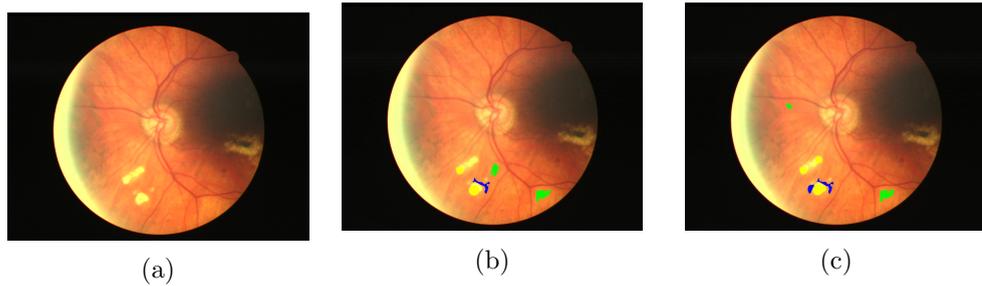


Figure C.5: Retinal image 0189.jpg from the S dataset: (a) original (b) candidates detected using $scaling_constant = \frac{ROI_diameter}{480}$ (c) candidates detected using $scaling_constant = \frac{ROI_diameter - OD_diameter}{370}$

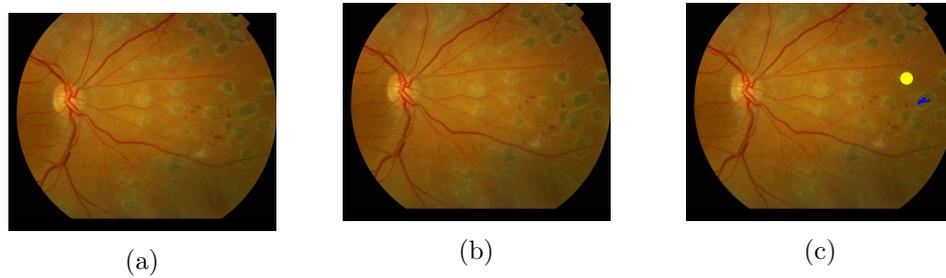


Figure C.6: Retinal image 04_A1.jpg from the BAT dataset: (a) original (b) candidates detected using $scaling_constant = \frac{ROI_diameter}{480}$ (c) candidates detected using $scaling_constant = \frac{ROI_diameter - OD_diameter}{370}$

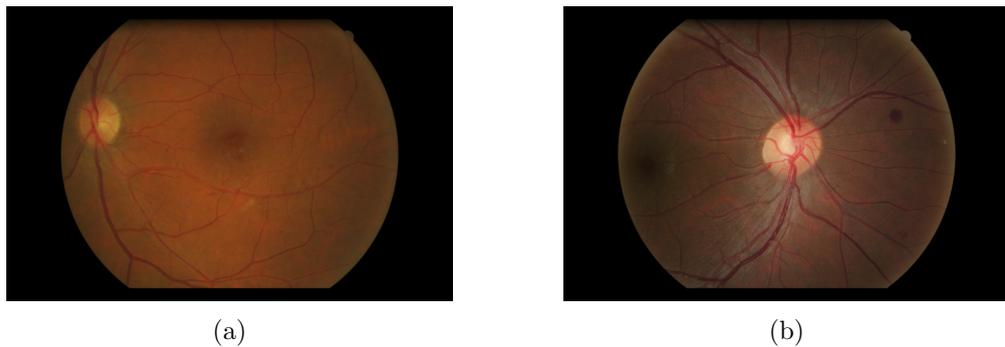


Figure C.7: Images from the EOEX dataset: (a) C0007248.jpg (b) C0007240.jpg

Appendix D

Misclassified Retinal Images

This section of the Appendix presents some results obtained during Pre-processing (Subsection 4.2) and Candidate Regions Extraction (Subsection 4.3) of 10 retinal images "No Laser" that were classified as "Laser". Of these 10 images, 3 are from the S dataset, 2 are from the BAT dataset, 2 are from the Messidor and 3 are from the FAZD dataset.

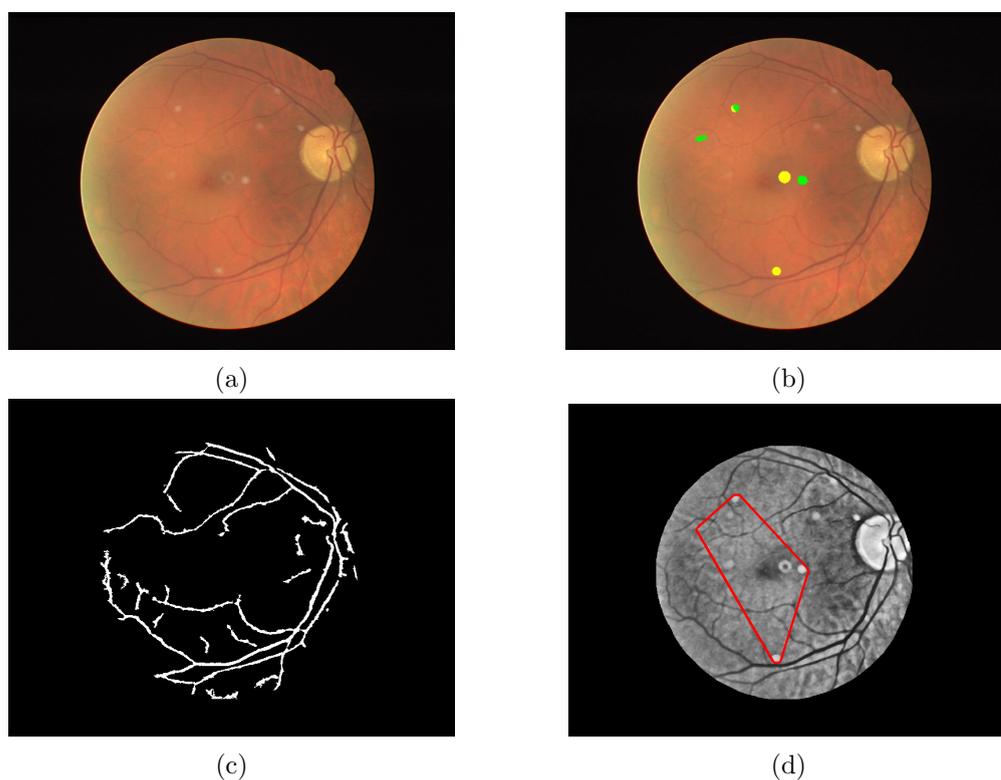


Figure D.1: Pre-processing outputs and candidates extraction of image 0219.jpg from the S Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; green=LMS c) Vascular Network d) Convex Hull.

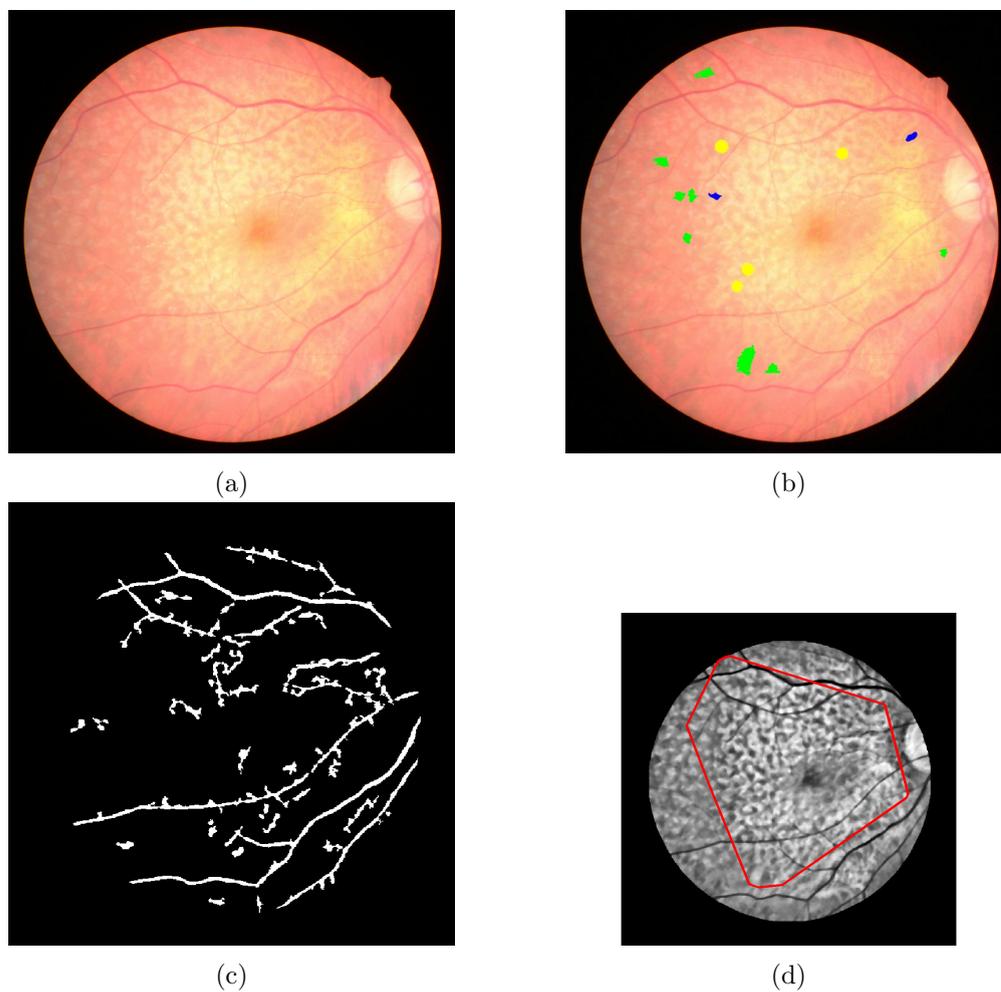


Figure D.2: Pre-processing outputs and candidates extraction of image 0225.jpg from the S Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF; green=LMS c) Vascular Network d) Convex Hull.

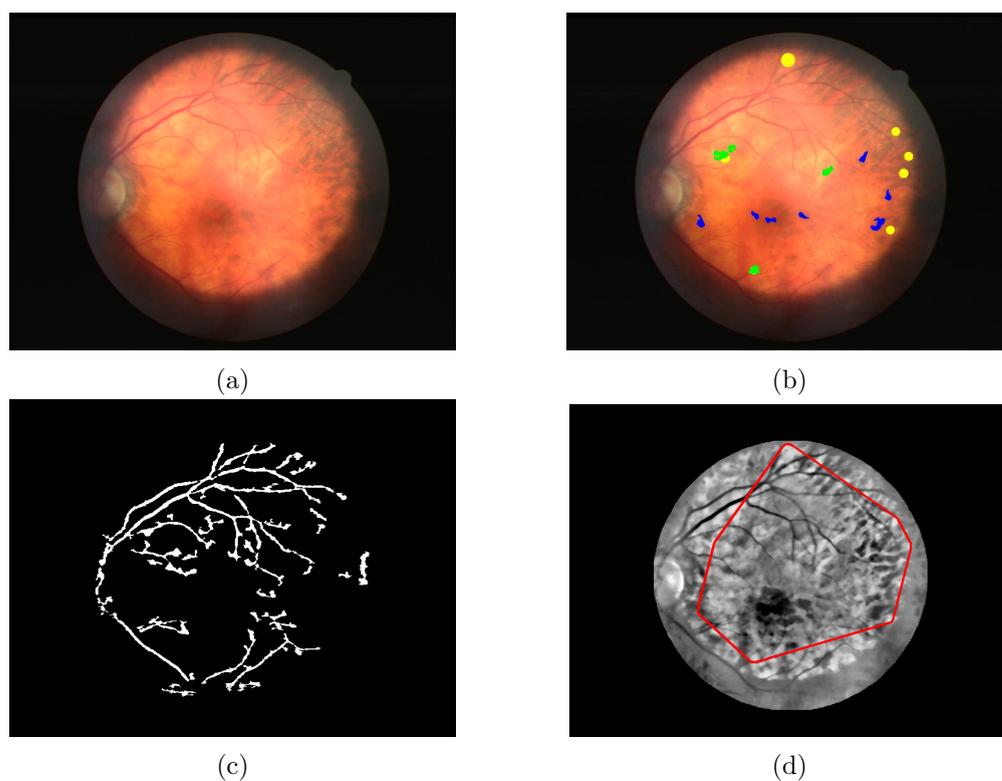


Figure D.3: Pre-processing outputs and candidates extraction of image 0419.jpg from the S Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF; green=LMS c) Vascular Network d) Convex Hull.

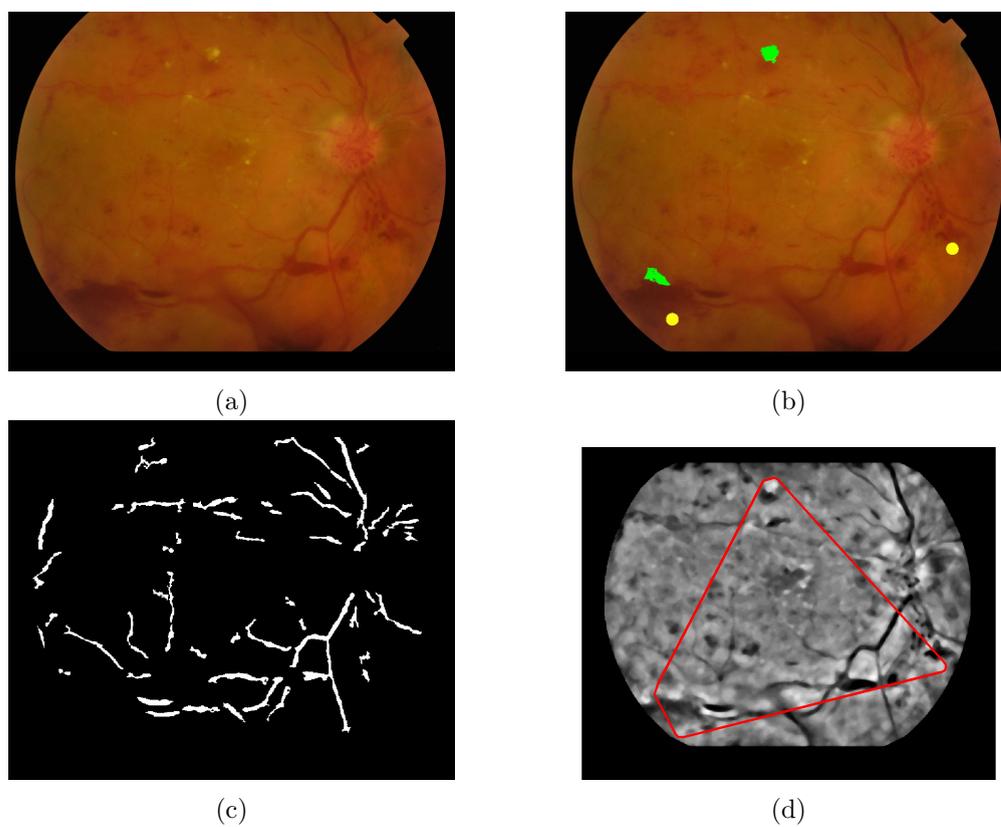


Figure D.4: Pre-processing outputs and candidates extraction of image 06_B1.jpg from the BAT Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; green=LMS c) Vascular Network d) Convex Hull.

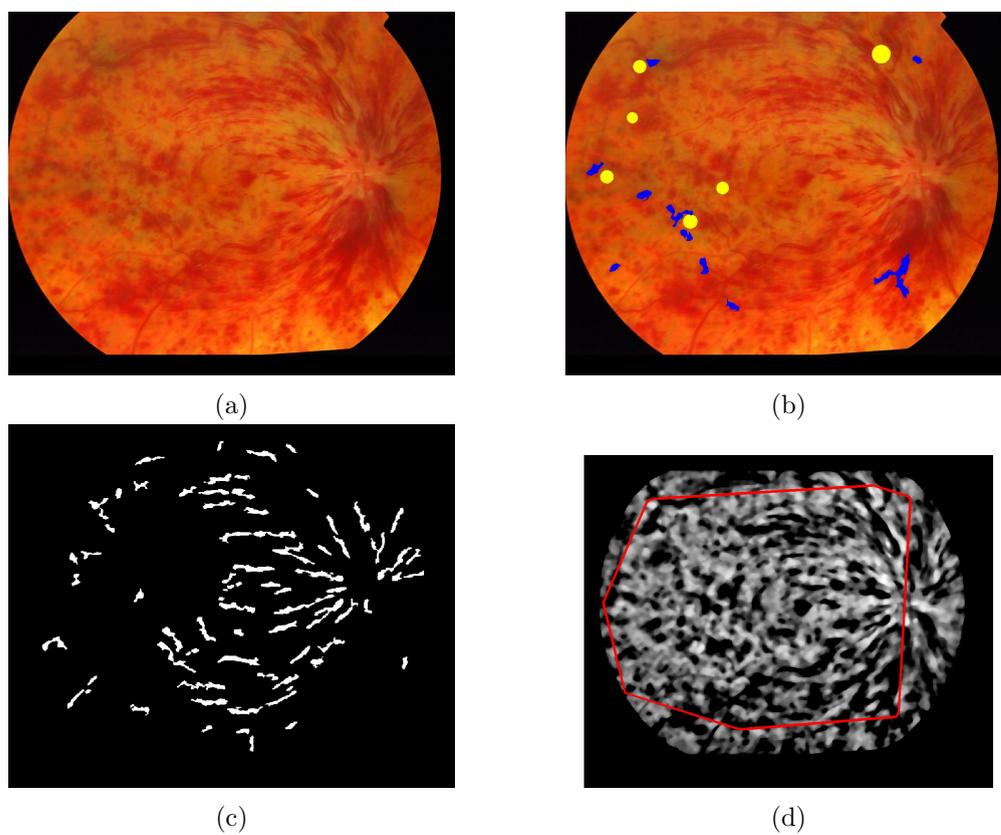


Figure D.5: Pre-processing outputs and candidates extraction of image 07_B2.jpg from the BAT Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF c) Vascular Network d) Convex Hull.

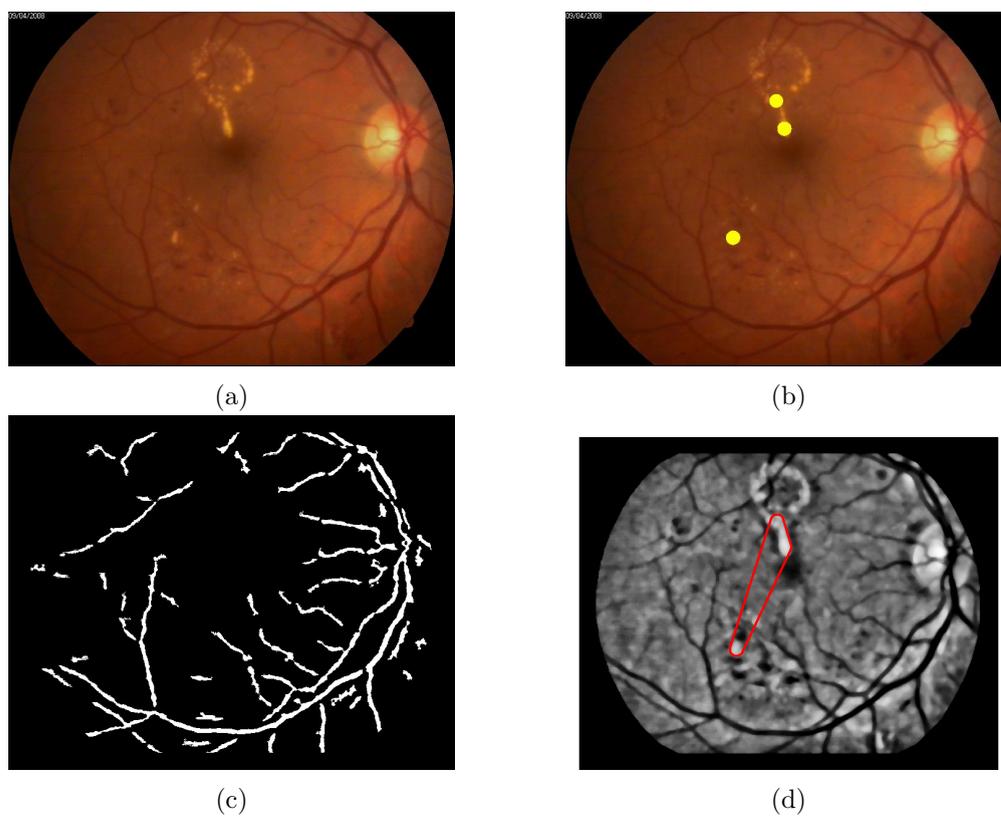


Figure D.6: Pre-processing outputs and candidates extraction of image 15.jpg from the FAZD Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT c) Vascular Network d) Convex Hull.

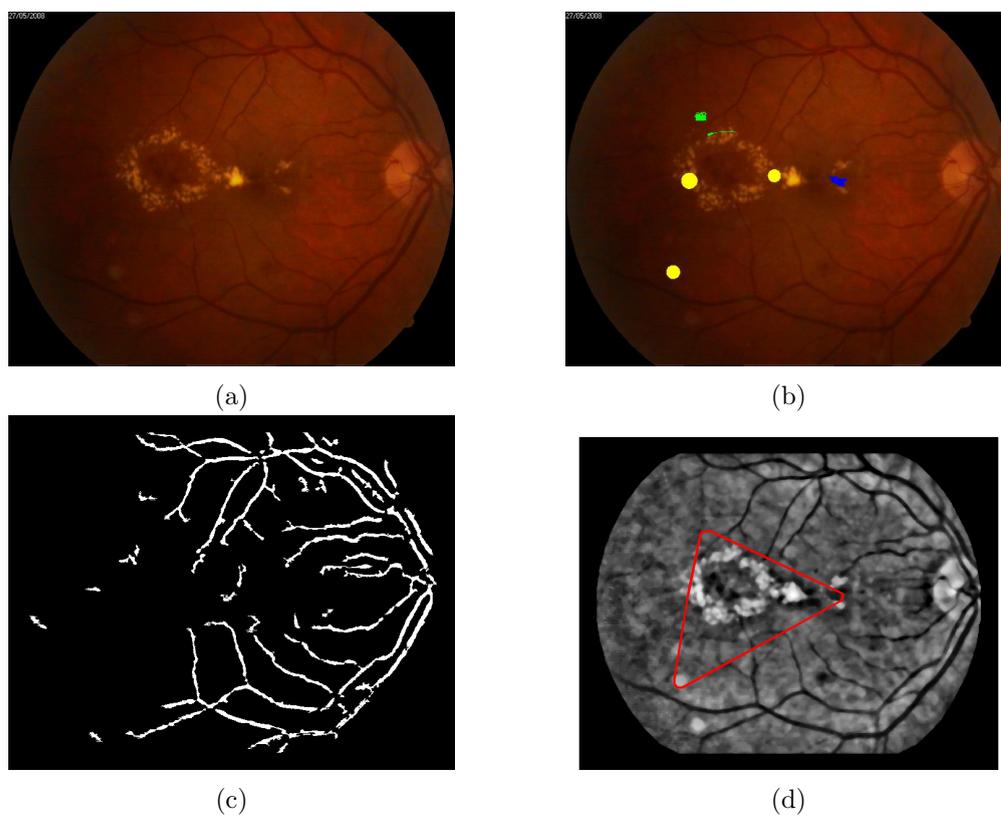


Figure D.7: Pre-processing outputs and candidates extraction of image 28.jpg from the FAZD Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF; green=LMS c) Vascular Network d) Convex Hull.

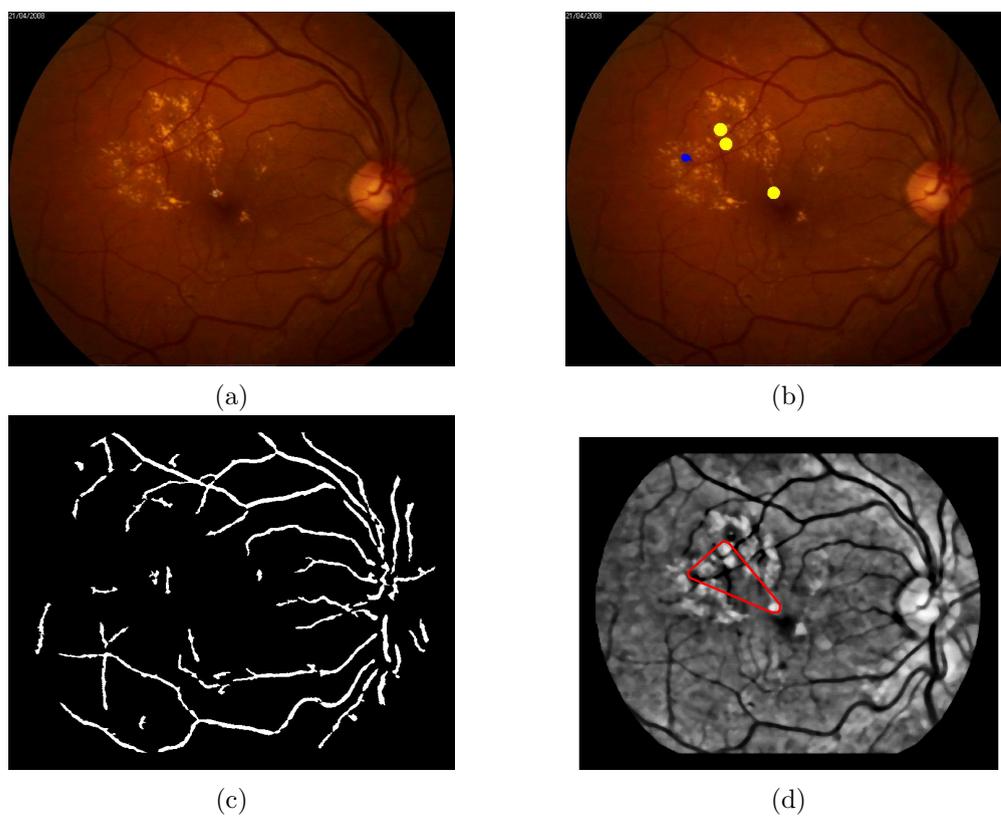


Figure D.8: Pre-processing outputs and candidates extraction of image 32.jpg from the FAZD Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF c) Vascular Network d) Convex Hull.

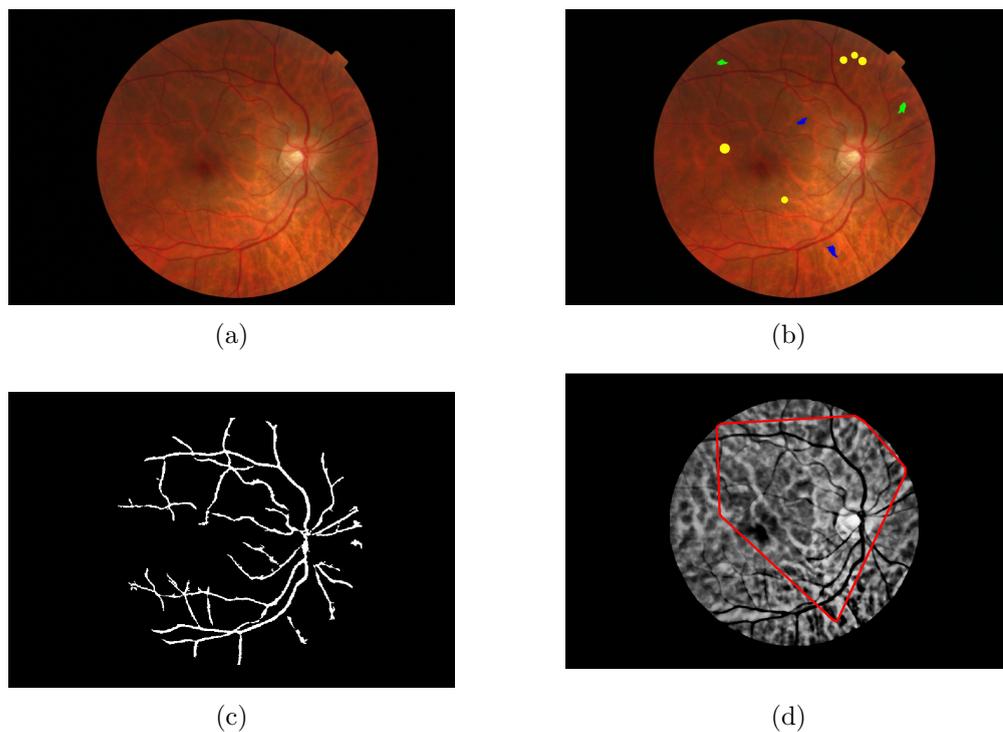


Figure D.9: Pre-processing outputs and candidates extraction of image 20051130_60691_0400_PP.jpg from the Messidor Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF; green=LMS c) Vascular Network d) Convex Hull.

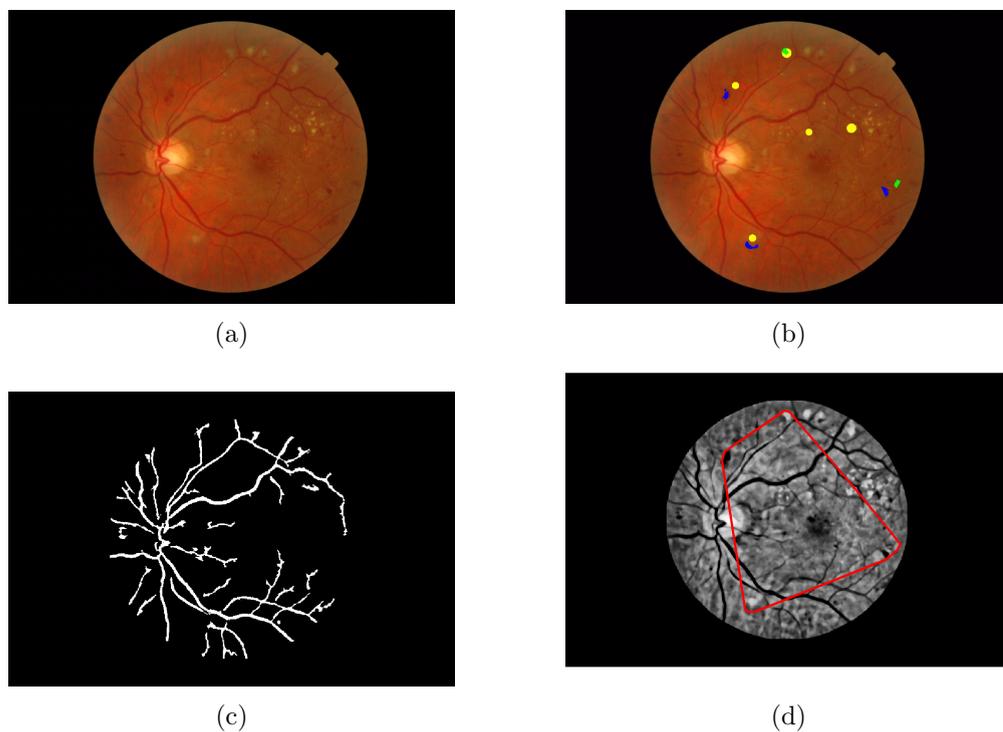


Figure D.10: Pre-processing outputs and candidates extraction of image 20051214_51733_0100_PP.jpg from the Messidor Dataset: a) Original image b) Original image with candidates drawn - yellow=CHT; blue=FVF; green=LMS c) Vascular Network d) Convex Hull.

References

- [1] I. D. Federation, *IDF Diabetes Atlas, 6th edn.* Brussels, Belgium: International Diabetes Federation, 2013. [Online]. Available: <http://www.idf.org/diabetesatlas>
- [2] T. I. A. for the Prevention of Blindness. Project vision 2020 - global facts. [Online]. Available: <http://www.iapb.org/vision-2020/global-facts>
- [3] A. Rohilla and S. Yadav, "A Review on Diabetic Retinopathy," *International Journal of Pharmacology Research*, vol. 3, no. 2, pp. 53–56, 2013.
- [4] S. Garg and R. M. Davis, "Diabetic retinopathy screening update," *Clinical Diabetes*, vol. 27, no. 4, pp. 140–145, 2009.
- [5] UK National Screening Committee, "Essential elements in developing a Diabetic Retinopathy screening programme - Workbook 4.3," Tech. Rep., 2009.
- [6] A. M. Syed, M. U. Akbar, M. U. Akram, and J. Fatima, "Automated Laser Mark Segmentation from Colored Retinal Images," *IEEE 17th International Multi-Topic Conference (INMIC), Karachi, Pakistan*, pp. 282–286, December 2014.
- [7] W. C. of Optometry. Member benefits. Accessed: 2015-06-28. [Online]. Available: <http://www.worldoptometry.org/en/membership2/member-benefits.cfm>
- [8] I. C. of Ophthalmology. Number of ophthalmologists in practice and training worldwide. Accessed: 2015-06-28. [Online]. Available: <http://www.icoph.org/ophthalmologists-worldwide.html>
- [9] L. P. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F. L. Ferris III, and R. Klein, "Diabetic Retinopathy," *Diabetes Care*, vol. 25, pp. 590–593, Mar. 2012. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3121116&tool=pmcentrez&rendertype=abstract>

- [10] S. meeting to EASDec, “Screening for Diabetic Retinopathy in Europe – Progress since 2005,” Tech. Rep. May, 2011.
- [11] K. Viswanath and D. D. M. McGavin, “Diabetic Retinopathy : Clinical Findings and Management,” *Community Eye Health*, 2003.
- [12] R. Raman, S. Mahajan, P. K. Rani, S. Agarwal, and T. Sharma, “eHEALTH INTERNATIONAL Telehealth Program for Diabetic Retinopathy in Rural South India : A Pilot Study,” 2005. [Online]. Available: <http://www.ehealthinternational.net>
- [13] J. E. Villena, C. a. Yoshiyama, J. E. Sánchez, N. L. Hilario, and L. M. Merin, “Prevalence of diabetic retinopathy in Peruvian patients with type 2 diabetes: results of a hospital-based retinal telescreening program.” *Revista panamericana de salud pública = Pan American journal of public health*, vol. 30, no. 5, pp. 408–14, 2011. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22262266>
- [14] K. a. Goatman, S. Philip, a. D. Fleming, R. D. Harvey, K. K. Swa, C. Styles, M. Black, G. Sell, N. Lee, P. F. Sharp, and J. a. Olson, “External quality assurance for image grading in the Scottish Diabetic Retinopathy Screening Programme,” *Diabetic Medicine*, vol. 29, no. 6, pp. 776–783, 2012.
- [15] D. a. Sim, P. a. Keane, A. Tufail, C. a. Egan, L. P. Aiello, and P. S. Silva, “Automated Retinal Image Analysis for Diabetic Retinopathy in Telemedicine,” *Current Diabetes Reports*, vol. 15, no. 3, 2015. [Online]. Available: <http://link.springer.com/10.1007/s11892-015-0577-6>
- [16] L. Ribeiro, C. M. Oliveira, C. Neves, J. a. D. Ramos, H. Ferreira, and J. Cunha-Vaz, “Screening for Diabetic Retinopathy in the Central Region of Portugal. Added Value of Automated ‘Disease/No Disease’ Grading.” *Ophthalmologica. Journal international d’ophtalmologie. International journal of ophthalmology. Zeitschrift fur Augenheilkunde*, Nov. 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/25427567>
- [17] S. I. Fox, *Human Physiology*, 12th ed. McGraw-Hill, 2011, pages 286-294.
- [18] J. E. H. Arthur C. Guyton, *Tratado de Fisiologia Médica*, tradução - 11th ed. Elsevier, 2006, pages 626-628.

- [19] T. Mahmudi, K. R., and H. Rabbani, "Comparison of macular OCTs in right and left eyes of normal people," *Proc. SPIE, Medical Imaging: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 9038, pp. 15–20, 2014. [Online]. Available: <https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1>
- [20] I. Grulkowski, J. J. Liu, B. Potsaid, V. Jayaraman, C. D. Lu, J. Jiang, A. E. Cable, J. S. Duker, and J. G. Fujimoto, "Retinal, anterior segment and full eye imaging using ultrahigh speed swept source OCT with vertical-cavity surface emitting lasers," *Biomedical Optics Express*, vol. 3, no. 11, pp. 2733–2752, 2012.
- [21] A. Ben-Haroush, Y. Yogeve, and M. Hod, "Epidemiology of gestational diabetes mellitus and its association with Type 2 diabetes." *Diabetic medicine : a journal of the British Diabetic Association*, vol. 21, no. 2, pp. 103–113, 2004.
- [22] O. Faust, R. Acharya U, E. Y. K. Ng, K.-H. Ng, and J. S. Suri, "Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review." *Journal of medical systems*, vol. 36, no. 1, pp. 145–57, Feb. 2012. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20703740>
- [23] G. a. Williams, I. U. Scott, J. a. Haller, A. M. Maguire, D. Marcus, and H. R. McDonald, "Single-field fundus photography for diabetic retinopathy screening: A report by the American Academy of Ophthalmology," *Ophthalmology*, vol. 111, no. 5, pp. 1055–1062, 2004.
- [24] C. M. Oliveira, L. M. Cristóvão, M. L. Ribeiro, and J. R. F. Abreu, "Improved automated screening of diabetic retinopathy." *Ophthalmologica. Journal international d'ophtalmologie. International journal of ophthalmology. Zeitschrift für Augenheilkunde*, vol. 226, no. 4, pp. 191–7, Jan. 2011. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21865671>
- [25] S. H. M. Alipour, H. Rabbani, M. Akhlaghi, A. M. Dehnavi, and S. H. Javanmard, "Analysis of foveal avascular zone for grading of diabetic retinopathy severity based on curvelet transform," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 250, no. 11, pp. 1607–1614, 2012. [Online]. Available: <https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1/colour-fundus-images-of-healthy-persons--patients-with-diabetic-retinopathy>

- [26] P. a. Asbell, I. Dualan, J. Mindel, D. Brocks, M. Ahmad, and S. Epstein, "Age-related cataract," *Lancet*, vol. 365, no. 9459, pp. 599–609, 2005.
- [27] L. Guo, J.-J. Yang, L. Peng, J. Li, and Q. Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis," *Computers in Industry*, vol. 69, pp. 72–80, may 2015. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0166361514001754>
- [28] M. S. Haleem, L. Han, J. van Hemert, and B. Li, "Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: a review." *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 37, no. 7-8, pp. 581–96, 2013. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24139134>
- [29] L. S. Lim, P. Mitchell, J. M. Seddon, F. G. Holz, and T. Y. Wong, "Age-related macular degeneration," *The Lancet*, vol. 379, no. 9827, pp. 1728–1738, 2012. [Online]. Available: [http://dx.doi.org/10.1016/S0140-6736\(12\)60282-7](http://dx.doi.org/10.1016/S0140-6736(12)60282-7)
- [30] J. Cunha-Vaz, L. Ribeiro, and C. Lobo, "Phenotypes and biomarkers of diabetic retinopathy." *Progress in retinal and eye research*, vol. 41, pp. 90–111, Jul. 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24680929>
- [31] D. Flanagan, "Screening for diabetic retinopathy." *The Practitioner*, vol. 238, no. 1534, pp. 37–42, 1994.
- [32] "Diabetic retinopathy (DR): Management and referral," *Community Eye Health Journal*, vol. 24, no. 75, pp. 12–13, 2011.
- [33] J. G. F. Dowler, "Laser management of diabetic retinopathy," *Journal of the Royal Society of Medicine*, vol. 96, no. 6, pp. 277–279, June 2003. [Online]. Available: <http://jrsm.rsmjournals.com/cgi/doi/10.1258/jrsm.96.6.277>
- [34] E. Jr, M. Michelessi, and G. Virgili, "Laser photocoagulation for proliferative diabetic retinopathy (Review)," *Cochrane Database Syst Rev.*, no. 11, November 2014.
- [35] A. G. Salman, "Pascal laser versus conventional laser for treatment of diabetic retinopathy." *Saudi journal of ophthalmology : official journal of the Saudi Ophthalmological Society*, vol. 25, no. 2, pp. 175–9, Apr. 2011.

- [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3729574&tool=pmcentrez&rendertype=abstract>
- [36] S. Riaskoff, "Photocoagulation treatment of proliferative diabetic retinopathy." *Bulletin de la Societe belge d'ophtalmologie*, vol. 197, no. 7, pp. 9–17, 1981. [Online]. Available: [http://dx.doi.org/10.1016/S0161-6420\(81\)34978-1](http://dx.doi.org/10.1016/S0161-6420(81)34978-1)
- [37] E. Treatment and D. Retinopathy, "Treatment Techniques and Clinical Guidelines for Photocoagulation of Diabetic Macular Edema," *Ophthalmology*, vol. 94, no. 7, pp. 761–774, Jul. 1987. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0161642087335274>
- [38] P. Osaadon, X. J. Fagan, T. Lifshitz, and J. Levy, "A review of anti-VEGF agents for proliferative diabetic retinopathy." *Eye (London, England)*, vol. 28, no. 5, pp. 510–20, 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24525867>
- [39] P. Hooper, M. C. Boucher, A. Cruess, K. G. Dawson, W. Delperio, M. Greve, V. Kozousek, W.-C. Lam, and D. a. L. Maberley, "Canadian Ophthalmological Society evidence-based clinical practice guidelines for the management of diabetic retinopathy." *Canadian journal of ophthalmology. Journal canadien d'ophtalmologie*, vol. 47, no. 2 Suppl, pp. S1–30, S31–54, Apr. 2012. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22632804>
- [40] H. L. Little, "Treatment of Proliferative Diabetic Retinopathy," *Ophthalmology*, vol. 92, no. 2, pp. 279–283, Feb. 1985. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0161642085340599>
- [41] M. M. K. Muqit, J. Denniss, V. Nourrit, G. R. Marcellino, D. B. Henson, I. Schiessl, and P. E. Stanga, "Spatial and spectral imaging of retinal laser photocoagulation burns." *Investigative ophthalmology & visual science*, vol. 52, no. 2, pp. 994–1002, Feb. 2011. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20861493>
- [42] B. H. Doft and G. W. Blankenship, "Single versus Multiple Treatment Sessions of Argon Laser Panretinal Photocoagulation for Proliferative Diabetic Retinopathy," *Ophthalmology*, vol. 89, no. 7, pp. 772–779, Jul. 1982. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S016164208234734X>

- [43] S. N. Lee, J. Chhablani, C. K. Chan, H. Wang, G. Barteselli, S. El-Emam, M. L. Gomez, I. Kozak, L. Cheng, and W. R. Freeman, "Characterization of microaneurysm closure after focal laser photocoagulation in diabetic macular edema." *American journal of ophthalmology*, vol. 155, no. 5, pp. 905–12, May 2013. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3627831&tool=pmcentrez&rendertype=abstract>
- [44] J. Arévalo, C. Fernández, a.J. Mendoza, R. García, and F. Arévalo, "Intravitreal triamcinolone combined with grid laser photocoagulation for patients with cystoid macular edema and advanced diabetic retinopathy: Pilot study," *Archivos de la Sociedad Española de Oftalmología (English Edition)*, vol. 88, no. 10, pp. 373–379, Oct. 2013. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S2173579413001588>
- [45] M. G. Kotoula, D. Z. Chatzoulis, S. Tahmitzi, and E. E. Tsironi, "Detection of macular photocoagulation scars with confocal infrared reflection imaging." *Ophthalmic surgery, lasers & imaging : the official journal of the International Society for Imaging in the Eye*, vol. 40, no. 4, pp. 385–388, 2009.
- [46] J. Miguel, P. Dias, C. M. Oliveira, A. Luís, and S. Cruz, "Detection of Laser Marks in Retinal Images," *2013 IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS), Porto, Portugal*, pp. 532–533, June 2013.
- [47] F. Tahir, M. U. Akram, M. Abbass, and A. A. Khan, "Laser Marks Detection From Fundus Images," *14th International Conference on Hybrid Intelligent Systems (HIS), Kuwait*, pp. 147–151, December 2014.
- [48] P. A. M. Kingsuk Roy Choudhury and K. F. Mulchrone, "Automated grain boundary detection by casrg," *Journal of Structural Geology*, no. 28, pp. 363–375, 2006.
- [49] C. M. O. Yingkui Li and Y. Guo, "Gis-based detection of grain boundaries," *Journal of Structural Geology*, no. 30, pp. 431–443, 2008.
- [50] H. Chauris, I. Karoui, P. Garreau, H. Wackernagel, P. Craneguy, and L. Bertino, "The circllet transform: A robust tool for detecting features with circular shapes," *Computers & Geosciences*, vol. 37, no. 3, pp. 331–342, Mar. 2011. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0098300410002529>

- [51] Z. Y. Y. S. Kaichang Di, Wei Li and Y. Liu, "A machine learning approach to crater detection from topographic data," *Advances in Space Research*, 2014.
- [52] W. A.-N. S. Sekhar and A. K. Nandi, "Automated localisation of retinal optic disk using hough transform," *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1577–1580, May 2008.
- [53] N. M. Salem, S. a. Salem, and A. K. Nandi, "Segmentation of retinal blood vessels based on analysis of the hessian matrix and clustering algorithm," *European Signal Processing Conference, Poznan, Poland*, pp. 428–432, 2007.
- [54] J. M. Pires Dias, C. M. Oliveira, and L. a. da Silva Cruz, "Retinal image quality assessment using generic image quality indicators," *Information Fusion*, vol. 19, pp. 73–90, Sep. 2014. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S1566253512000656>
- [55] T. V. Program. Messidor dataset. Accessed: 2015-06-28. [Online]. Available: <http://messidor.crihan.fr/download-en.php>
- [56] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, a. Erginay, B. Laÿ, and a. Chabouis, "TeleOphta: Machine learning and image processing methods for teleophthalmology," *IRBM - Elsevier*, vol. 34, no. 2, pp. 196–203, 2013. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S1959031813000237>
- [57] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Laÿ, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, P. Massin, A. Chabouis, Z. Victor, and A. Erginay, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Medical Image Analysis*, vol. 18, no. 7, pp. 1026–1043, Oct. 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24972380>
- [58] M. Golabbakhsh and H. Rabbani, "Vessel-based registration of fundus and optical coherence tomography projection images of retina using a quadratic registration model," *IET Image Processing*, vol. 7, no. 8, pp. 768–776, 2013. [Online]. Available: <https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1/vessel-reg-oct-fundus>

- [59] U. of Minho. Minh n. do: Software. Access: 2015-07-31. [Online]. Available: <http://www.ifp.illinois.edu/~minhdo/software/>
- [60] J. Pinão and C. M. Oliveira, "Fovea and Optic Disc Detection in Retinal Images with Visible Lesions," *Technological Innovation for Value Creation - Third IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing, Electrical and Industrial Systems, Costa da Caparica, Portugal*, vol. 372, pp. 543–552, February.
- [61] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Prentice Hall, 2002, page 593.
- [62] W. Wen-cheng and C. Xiao-jun, "A Segmentation Method for Uneven Illumination Particle Images," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 5, no. 4, pp. 1284–1289, February 2013.
- [63] P. G. Cavalcanti, J. Scharcanski, and C. B. O. Lopes, "Shading attenuation in human skin color images," *Advances in Visual Computing*, pp. 190–198, 2010.
- [64] N. A. Ibraheem, M. M. Hasan, R. Z. Khan, and P. K. Mishra, "Understanding Color Models : A Review," *ARPJ Journal of Science and Technology*, vol. 2, no. 3, pp. 265–275, 2012.
- [65] H. F. Jelinek and M. J. Cree, *Automated Image Detection of Retinal Pathologies*. CRC Press, Taylor and Francis Group, 2010, page 235.
- [66] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. a. Viergever, "Multiscale vessel enhancement filtering," *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98. Lecture Notes in Computer Science*, vol. 1496, pp. 130–137, 1998.
- [67] S. Roy Chowdbury, "Automated Segmentation and Pathology Detection in Ophthalmic Images," *Dissertation to obtain PhD Degree*, pp. 133–138, 2014.
- [68] P. Mukhopadhyay and B. B. Chaudhuri, "A survey of Hough Transform," *Pattern Recognition*, vol. 48, no. 3, pp. 993–1010, Mar. 2015. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0031320314003446>
- [69] H. Yazid, H. Yazid, M. Harun, S. Mohd, a. Aziz Mohamed, M. Rizon, and S. Sayuti, "Circular discontinuities detection in welded joints using Circular Hough Transform," *NDT and E International*, vol. 40, no. 8, pp. 594–601, 2007.

- [70] S. A. Amiri, H. Hassanpour, M. Shahiri, and R. Ghaderi, "Detection of Microaneurysms in Retinal Angiography Images Using the Circular Hough Transform," *Journal of Advances in Computer Research*, vol. 3, no. 1, pp. 1–12, 2012.
- [71] R. J. Winder, P. J. Morrow, I. N. McRitchie, J. R. Bailie, and P. M. Hart, "Algorithms for digital image processing in diabetic retinopathy." *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 33, no. 8, pp. 608–22, Dec. 2009. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19616920>
- [72] D. J. Krcbywn and T. J. Atherton, "Circle Detection using hough transform filters," *Image (Rochester, N.Y.)*, pp. 370–374, 1995.
- [73] M. Kumar, A. K. Jaiswal, and R. Saxena, "Performance Analysis of Adaptive Canny Edge Detector Using Bilateral Filter," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 1, no. 4, pp. 930–936, June 2013.
- [74] G. Jie and L. Ning, "An Improved Adaptive Threshold Canny Edge Detection Algorithm," *2012 International Conference on Computer Science and Electronics Engineering*, pp. 164–168, Mar. 2012. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6187852>
- [75] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pp. 839–846, 1998. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=710815>
- [76] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS*, vol. 20, no. 1, pp. 62–66, January 1979.
- [77] D. Jimenez-Carretero, A. Santos, S. Kerkstra, R. D. Rudyanto, and M. J. Ledesma-Carbayo, "3D Frangi-based lung vessel enhancement filter penalizing airways," *Proceedings - International Symposium on Biomedical Imaging*, pp. 926–929, 2013.
- [78] J. Deng and X. Xie, "3D INTERACTIVE CORONARY ARTERY SEGMENTATION USING RANDOM FORESTS AND MARKOV RANDOM FIELD OPTIMIZATION Peninsular Radiology Academy , Plymouth University Schools of

- Medicine & Dentistry Plymouth Hospitals NHS Trust , United Kingdom Department of Computer S,” pp. 942–946, 2014.
- [79] Y. Zheng, “Machine learning based vesselness measurement for coronary artery segmentation in cardiac CT volumes,” *SPIE Medical Imaging 2011: Image Processing*, pp. 79 621K–79 621K–12, 2011. [Online]. Available: [http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.877233\delimiter"026E30F\\$http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1349868](http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.877233\delimiter)
- [80] M. Sofka and C. V. Stewart, “Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures,” *IEEE Transactions on Medical Imaging*, vol. 25, no. 12, pp. 1531–1546, 2006.
- [81] R. M. Rangayyan and S. Mukhopadhyay, “Detection of Blood Vessels in Retinal Fundus Images,” *Computer Science Journal of Moldova*, vol. 22, no. 2, pp. 155–185, 2014.
- [82] M. S. Drew, A. Alsam, A. Madooei, and M. H. Brill, “Fast Colour Vesselness,” *19th Color and Imaging Conference Final Program and Proceedings*, pp. 146–151.
- [83] R. E. W. Rafael C. Gonzalez and S. L. Eddins, *Digital Image Processing using MATLAB*. Pearson, 2004, pages 464-467.
- [84] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 610–621, 1973.
- [85] L. Soh and C. Tsatsoulis, “Texture analysis of sar sea ice imagery using gray level co-occurrence matrices,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, March 1999.
- [86] D. Clausi, “An analysis of co-occurrence texture statistics as a function of grey level quantization,” *Canadian Journal of Remote Sensing*, vol. 28, no. 1, pp. 45–62, 2002.
- [87] A. Uppuluri. Glcm texture features. Accessed: 2015-08-09. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/22187-glcm-texture-features>

- [88] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Transactions on Mathematical Software*, vol. 22, no. 4, pp. 469–483, 1996.
- [89] "Optimal set of features for accurate skin cancer diagnosis," *2014 IEEE International Conference on Image Processing (ICIP), Paris, France*, no. 3, pp. 2256–2260, October 2014.
- [90] W. H. Encyclopedia. Spatial descriptive statistics. Accessed: 2015-08-09. [Online]. Available: http://www.kobolibrary.com/articles/spatial_descriptive_statistics#Measures_of_spatial_dispersion
- [91] P. de Jong, C. Sprenger, and F. van Veen, "On extreme values of Moran's I and Geary's C," *Geographical Analysis*, vol. 16, no. 1, pp. 17–24, 1984.
- [92] W. O. University. Chapter 5 - pattern descriptors. Accessed: 2015-08-09. [Online]. Available: <http://www.wou.edu/las/physci/taylor/g492/avstat5.pdf>
- [93] R. Ferstl, "Spatial filtering with EViews and MATLAB," *Austrian J Stat*, vol. 36, no. 1, pp. 17–26, 2007. [Online]. Available: <http://www.stat.tugraz.at/AJS/ausg071/071Ferstl.pdf>
- [94] R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.
- [95] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [96] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.
- [97] B. Hssina, A. Merbouha, H. Ezzikouri, and M. Erritali, "A comparative study of decision tree ID3 and C4 . 5," *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 2, pp. 13–19.
- [98] R. Kohavi and R. Quinlan, "Decision tree discovery," in *HANDBOOK OF DATA MINING AND KNOWLEDGE DISCOVERY*. University Press, 1999, pp. 267–276.
- [99] M. Gütlein, E. Frank, M. Hall, and A. Karwath, "Large-scale attribute selection using wrappers," *2009 IEEE Symposium on Computational Intelligence and Data Mining, Nashville, USA*, pp. 332–339, April 2009.

- [100] J. Novakovic, "Using Information Gain Attribute Evaluation to Classify Sonar Targets," *17 th T elecommunication Forum*, pp. 1351–1354, 2009.
- [101] R. Genuer, J.-m. Poggi, C. Tuleau-malot, R. Genuer, J.-m. Poggi, C. T.-m. Variable, R. Genuer, J.-m. Poggi, and C. Tuleau-malot, "Variable selection using Random Forests," vol. 31, no. 14, pp. 2225–2236, 2012.
- [102] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: An update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009. [Online]. Available: <http://doi.acm.org/10.1145/1656274.1656278>
- [103] A. G. Karegowda, a. S. Manjunath, and M. a. Jayaram, "Comparative study of attribute selection using gain ration and correlation based feature selection," *International Journal of Information Technology and Knowledge Management*, vol. 2, no. 2, pp. 271–277, 2010.
- [104] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.
- [105] R. P. Priyadarsini and S. Sivakumari, "Gain ratio based feature selection method for privacy preservation," *ICTACT Journal of Soft Computing*, vol. 1, no. 4, pp. 201–205, April 2011.
- [106] A. P. S. S. C. Parikh, Rajul Mathai and R. Thomas, "Understanding and using sensitivity, specificity and predictive values," *Indian J Ophthalmol.*, vol. 56, no. 1, pp. 45–50, Jan-Feb 2008.