UNIVERSIDADE D
COIMBRA

Vânia Sofia Bastos Silva

GLAUCOMA PREDICTION BASED ON 3D-OCT SPECKLE IMAGING

Thesis submitted to the University of Coimbra in fulfilment of the requirements of the Master's Degree in Biomedical Engineering under the scientific supervision of Ph.D. Danilo Andrade de Jesus, Professor João Manuel Rendeiro Cardoso and Ph.D. Pedro Guilherme da Cunha Leitão Dias Vaz

September of 2021

University of Coimbra

Integrated Master in Biomedical Engineering

Glaucoma prediction based on 3D-OCT speckle imaging

Vânia Sofia Bastos Silva

Thesis submitted to the Faculty of Sciences and Technology of the University of Coimbra in fulfillment of the requirements for the Master's Degree in Biomedical Engineering

Supervisors:
Ph.D. Danilo Andrade de Jesus
Professor João Manuel Rendeiro Cardoso
Ph.D. Pedro Guilherme da Cunha Leitão Dias Vaz

Coimbra, September 2021

Acknowledgements

First of all, I would like to thank the entire team behind this project, Dr. Danilo Jesus, Dr. Pedro Vaz, Dr. Luisa Sánchez, Prof. João Cardoso, Dr. Theo van Walsum, Dr. Stefan Klein, Prof. Dr. Ingeborg Stalmans, Dr. João Breda and Dr. Jan Van Eijgen, for all the help and support given throughout this project. Thank you to all the LIBPhys and BIGR team for receiving me, with a special thanks to my technical supervisors for all of the guidance, knowledge and tireless dedication given. Without everyone's expertise and advisory, this work would not have been possible.

I also thank the Laboratory for Advanced Computing of University of Coimbra for providing use of the supercomputer, crucial for the development of this project.

A very special thanks to my lab colleague and friend Rita, for your friendship and support during the last year, and for always being my right-hand during the most stressful times.

To António Sérgio, thank you for being by my side during all the good and the difficult moments. Thank you for the encouragement and for always believing in me.

Last, but not least, a big thank you to my family, in special my parents and sister Joana, for the endless love and support given throughout my life. You'll always be my greatest example. All that I am, I owe it to you.

Thank you all.

Abstract

Glaucoma is an insidious and unpredictable ophthalmic disease, with no symptoms at an early stage. Although glaucoma is the leading cause of irreversible blindness worldwide, the damage it causes to the eye is preventable, making an early diagnostic and close monitoring crucial.

Optical Coherence Tomography (OCT) is a low-coherence interferometry technique, widely used to retrieve relevant biomarkers for glaucoma diagnosis. OCT suffers from speckle, an interference phenomenon characterized by the appearance of granular patterns in the image, called speckle patterns. Although these patterns have been treated mainly as a source of noise, they are known to be also a carrier of information, and their analysis has many known biomedical applications. Nevertheless, changes in speckle of retinal OCT images due to glaucoma have yet to be studied.

In this project, five different methods to study and infer OCT speckle properties were implemented. These methods were then applied to glaucomatous OCT, in order to assess their potential to aid glaucoma diagnosis. The speckle-related features and other clinically relevant data (age, gender, retinal layers thickness) were used to develop a predictive model based on discriminative Event-Based modelling (DEBM): a data-driven technique to estimate the sequence in which biomarkers for a disease become abnormal. This model aims at not only predicting the presence of glaucoma, but also providing an interpretable insight to the disease and its processes. The data of 259 subjects from the Leuven Eye Study cohort was used for this matter.

The main results suggest that DEBM was able to successfully model glaucoma disease progression, with the higher performing model achieving an AUC of the overall separation between healthy and severe glaucoma of 0.984 ± 0.018. Additionally, speckle showed to have the potential to capture glaucoma derived microstructural changes in the retina, which occurred both in the inner and the outer retinal layers, although with different degrees. This information may potentiate the implementation of an automatic and reproducible way of estimating glaucoma

severity without relying on subjective measurements. However, further work is required to fully understand its potential and to attempt a more insightful interpretation of its physical meaning, as well as its feasibility to be introduced into glaucoma clinical practice.

Key-words: Optical Coherence Tomography, Retina, Speckle, Glaucoma, Image analysis, Disease progression

Resumo

Glaucoma é uma doença oftalmológica imprevisível e insidiosa, sem quaisquer sintomas numa fase inicial. Apesar de ser a principal causa de cegueira irreversível no mundo, os seus danos podem ser prevenidos, tornando crucial conseguir um diagnóstico precoce e manter uma monitorização cuidada da progressão da doença.

A tomografia de coerência ótica (OCT) é uma técnica de imagiologia baseada em interferometria de baixa coerência, frequentemente usada para extrair biomarcadores relevantes para o diagnóstico de glaucoma. O OCT sofre de speckle, um fenómeno de interferência caracterizado pelo aparecimento de padrões granulares na imagem, chamados padrões de speckle. Apesar destes padrões serem tratados maioritariamente como fonte de ruído, podem também ser uma fonte de informação, e a sua análise tem várias aplicações biomédicas conhecidas. Apesar disso, alterações nos padrões de speckle de imagens de OCT da retina de pacientes com glaucoma não foram ainda estudadas.

Neste projeto, foram implementados cinco métodos diferentes para analisar e inferir propriedades dos padrões de speckle. Estes métodos foram posteriormente aplicados a volumes de OCT de glaucoma, de forma a aferir o seu potencial para auxiliar no diagnóstico. Parâmetros do speckle e outros clinicamente relevantes (idade, género, espessura das camadas da retina) foram usadas para desenvolver um modelo preditivo baseado em discriminative Event-Based modelling (DEBM): uma técnica "data-driven" que estima a sequência na qual os biomarcadores para uma doença se tornam anormais. Este modelo tem como objetivo não só prever a presença de glaucoma, mas também proporcionar uma visão geral e interpretável dos processos que ocorrem durante a evolução da doença. Para este efeito, foram usados dados de 259 sujeitos do Leuven Eye Study cohort.

Os resultados principais sugerem que o DEBM foi capaz de modelar corretamente a progressão de glaucoma, obtendo para o modelo com o melhor desempenho uma AUC para a diferenciação entre indivíduos saudáveis e glaucomas severos de 0.984 ± 0.018. Adicionalmente, o speckle

mostrou ter o potencial de capturar diferenças na microestrutura da retina resultantes do glaucoma, as quais ocorrem tanto nas camadas internas da retina como nas mais externas, embora com diferentes níveis de importância. Esta informação pode auxiliar a implementação de um modelo automático e reproduzível que estime o grau de severidade da doença sem depender de parâmetros subjetivos. Contudo, é necessária mais investigação nesta área, de forma a compreender o potencial dos padrões de speckle e atingir uma melhor interpretação da sua significância física, assim como da viabilidade da sua introdução em prática clínica.

Palavras-chave: Tomografia de Coerência Ótica, Retina, Speckle, Glaucoma, Análise de imagem, Progressão de doença

Contents

List of Acronyms

AUC area under the receiver operating characteristics curve

cpRNFL circumpapillary retinal nerve fiber layer

DEBM discriminative event-based modeling

DFT Discrete Fourier transform

DPM Disease progression modelling

EBM Event-based modeling

EGS European Glaucoma Society

FD-OCT Fourier domain - optical coherence tomography

FPR false positive rate

GCC ganglion cell complex

GCIPL ganglion cell-Inner plexiform layer

GCL ganglion cell layer

GG Generalized Gamma

GMM Gaussian mixture model

ILM internal limiting membrane

INL inner nuclear layer

IOP intraocular pressure

IPL inner plexiform layer

IS/OS inner-segment/outer-segment

KDE Kernel density estimate

mRNFL macular retinal nerve fiber layer

NTG normal tension glaucoma

OCT Optical Coherence Tomography

OHT ocular hypertension

ONL outer nuclear layer

OPL outer plexiform layer

OPR outer photoreceptor

OSL outer segment layer

PACG primary angle closure glaucoma

PDF probability density function

POAG primary open-angle glaucoma

RGC retinal ganglion cells

RNFL retinal nerve fiber layer

ROC receiver operating characteristic

RPE retinal pigment epithelium

SD-OCT Spectral domain - optical coherence tomography

SGLDM spatial gray-level dependence matrices

SNR signal-to-noise ratio

TPR true positive rate

VF MD visual field mean deviation

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Glaucoma is a chronic ophthalmic disease, the leading cause of irreversible blindness worldwide. It affects a total of over 80 million people [1]. This number is thought to be an underestimation of reality, as studies estimate that around half of people with glaucoma are undiagnosed and unaware of their condition [2], making the exact number of affected people unknown.

Glaucoma has no known cure, and the damage it provokes can not be reverted. Nevertheless, it can be prevented if diagnosed at an early stage, stabilizing the disease progression. This makes an early diagnostic and close progression monitoring crucial to preserve patient's vision and quality of life. However, since glaucoma can remain asymptomatic until a very severe disease stage, an early diagnostic can be a challenging task, and a standard tool for the glaucoma diagnosis still does not exist. In addition, glaucoma diagnosis in clinical practise faces some challenges, such as the heterogeneity of the disease presentation, the subjectivity of clinicians judgment and inter-observer variability in the interpretation of clinical exams and parameters, as well as the need for patient cooperation for the correct assessment of visual defects, resulting in normal fluctuations and low repeatability of these metrics [3, 4]. The difficulty in establishing an objective method for the diagnosis has led to the growth of technology-aided research on this topic, with the ultimate goal of supplementing clinical judgment and assisting in the decision making regarding glaucoma diagnosis and monitoring [5].

Optical Coherence Tomography (OCT) is a commonly used imaging modality in the evaluation of glaucomatous structural damage, and it has been widely used in glaucoma research, as it provides objective and reproducible metrics for this assessment. As a coherent image modality, it is subject to speckle, a random interference phenomenon that creates granular pattern on the

image. Speckle is proved to contain information about the microstructure of the imaged tissue, smaller than the resolution of the image and imperceptible to the naked eye [6–9]. Speckle has many useful biomedical applications, and a number of works already developed show its potential in various fields [8, 10–27], including ophthalmology. Nevertheless, its interpretability remains a challenging task and signal-carrying speckle is still a research line in an early stage. Early changes in speckle patterns of retinal OCT images from glaucoma patients have, to the team's knowledge, never been studied and have not been considered in clinical practice yet. These patterns have the potential to add new information to the features currently used and improve glaucoma prediction methods, as well as provide more insight to the structure change associated with early development of the disease.

1.2 Objectives

The main goal of this project is the development of a disease progression model for glaucoma, based on micro and macrostructural features retrieved from OCT volumes. This model aims to make technology-aided diagnosis and disease severity classification more interpretable, as well as provide an insight of the disease timeline of glaucoma.

Additionally, this project aims to investigate the usability of speckle information from retinal OCT data for identifying glaucomatous damage, and potentially provide an added value to the diagnostic methods used currently in clinical practise.

For this purpose, a review of the developed methods to study and infer speckle properties from OCT data was performed, and the methods were then applied to 3D OCT data from healthy and glaucomatous eyes. The retrieved speckle-related features and macrostructural features were used, as well as other clinically relevant data that can be considered as confounders (age and gender). The predictive model is based on discriminative event-based modeling (DEBM): a data-driven technique to estimate the sequence in which biomarkers for a disease become abnormal. The data of 259 subjects from the Leuven Eye Study cohort was used, which comprises healthy controls as well as glaucoma patients with various disease severity levels.

Concluding, in this project, we expect to contribute to glaucoma prediction by developing a new model combining OCT-speckle information and other clinical and structural relevant parameters. We also provide some insight of the potential role of speckle information in understanding microstructural changes that happen within the retinal tissues at different stages of glaucoma (early/advanced stage), as well as assess the potential of speckle information to be

included in clinical practice, aiding the glaucoma screening/monitoring.

1.3 Project team

This project was developed within a collaboration between four research groups. The research team is presented in Table 1.1, with respective role and affiliation of each team member. This project also granted the possibility to integrate the Erasmus+ program, for a two month remote internship at Erasmus University Medical Center.

Table 1.1: Research team involved in the project.

Name	**Role**
Vânia Bastos Silva [1,2]	Master Student
Danilo Andrade de Jesus [2]	Technical Supervisor
João Cardoso [1]	Technical Supervisor
Pedro Guilherme Vaz [1]	Technical Supervisor
Luisa Sánchez Brea [2]	Technical Supervisor
Ingeborg Stalmans [3]	Clinical advisor
Jan van Eijgen [3]	Clinical advisor
João Barbosa Breda [3,4]	Clinical advisor
Stefan Klein [2]	Technical advisor
Theo van Walsum [2]	Technical advisor

[1] Laboratory for Instrumentation, Biomedical Engineering and Radiation Physics (LIBPhys-UC), Department of Physics, University of Coimbra, Coimbra, Portugal.

[2] Biomedical Imaging Group Rotterdam (BIGR), Department of Radiology & Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands.

[3] Research Group Ophthalmology, Department of Neurosciences, KU Leuven, Leuven, Belgium.

[4] São João Universitary Hospital Center (CHUSJ), Porto, Portugal.

1.4 Dissertation's content

This dissertation is organized into six core chapters.

1. **Introduction:** this chapter comprises the main goals and motivations of the project, as well as the research team involved in its development, and the list of scientific contributions.

2. **Theoretical Background:** in this chapter, a brief introduction to the theory behind this dissertation topics' is presented. First, a basic introduction to the humans' eye anatomy is presented, followed by an overview on glaucoma disease and its diagnosis, an introduction to the imaging modality used, OCT, a theoretical overview on speckle in OCT, and the current OCT biomarkers for glaucoma diagnosis. This chapter also includes the state of the art of both speckle analysis methods and disease progression modelling methods.

3. **Methods:** in this chapter, the methods used in the different steps of the project development are presented. The chapter begins with a description of the dataset, followed by the different steps taken, from the image processing methodology to an initial statistical analysis of features and implementation and evaluation of the predictive model.

4. **Results:** in this chapter, the results from the different steps of the methodology are presented. First, results of the statistical analysis of the features retrieved are shown, followed by the results of the experiments performed during the implementation of DEBM.

5. **Discussion:** critical interpretation and analysis of the results is presented in this chapter. The results are related with the main objectives of the project and with the current state of the art.

6. **Conclusions:** this chapter comprises the general conclusions and findings of this work, with a description of the future possible improvements and research directions.

1.5 Scientific dissemination

The work developed during this project resulted the following scientific contributions:

- V.B. Silva, D.A. Jesus, S. Klein, T. Walsum, J. Cardoso, L.S. Brea, P. Vaz. *"Signal-carrying speckle in Optical Coherence Tomography: a methodological review paper on biomedical applications"*, 2021 [28]. Submitted for publication.

- V.B. Silva, L.S. Brea, T. Walsum, S. Klein, J. Cardoso, R. Iskander, M. Danielewska, M. Kostyszak, J.B. Breda, J. Eijgen, I. Stalmans, D.A. Jesus, P.G. Vaz. *"Glaucoma prediction based on 3D-OCT speckle imaging"* [Conference session]. 8$^{\text{th}}$ Dutch Bio-Medical Engineering Conference, The Netherlands, January 28-29, 2021.

Chapter 2

Theoretical background

This section outlines the main theoretical topics regarding the work developed in this project. It starts by detailing the eye anatomy, namely the location and purpose of the principal structures. Then, a brief explanation about glaucoma disease is provided, including its effects and main biomarkers. After, Optical Coherence Tomography (OCT), the imaging modality used in this project, is introduced, followed by an introduction of the speckle effect theory. Moreover, details are presented about the main parameters extracted from OCT for glaucoma diagnosis. Lastly, the state of the art on speckle analysis and disease progression modelling is comprised.

2.1 Eye anatomy

The human eye is a complex and multifaceted organ, responsible for capturing, adjusting, and transforming light so the human brain can decipher what we see [29].

The inside of the human eye can be divided into three sections, called chambers: the anterior chamber (between the cornea and the iris), filled with aqueous humor, the posterior chamber (between iris and the crystalline lens) also filled with aqueous humor, and the vitreous body (between the lens and the retina), filled with vitreous humor [30]. These three chambers and their conforming structures are visible on Figure 2.1.

The wall of the eye globe is composed by three primary layers: external, intermediate, and internal layers [30]. The external, or outer, layer is composed by the sclera (commonly known as "the white" of the eye), which connects to the cornea in the anterior part. The cornea is transparent and approximately spherical, while the sclera is a dense, white and opaque fibrous tissue [32,33]. The external layer acts as a protective layer, maintaining the intraocular pressure (IOP) inside the eye [34].

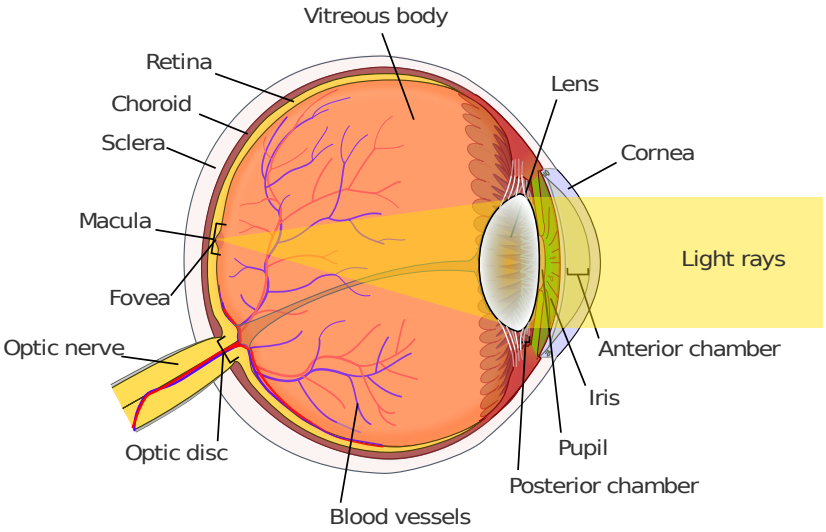

Figure 2.1: Diagram of the human eye. Adapted from public domain image [31].

The intermediate layer can also be called uveal tract. It is composed by the iris and ciliary body, in the anterior part, and the choroid in the posterior.

The internal layer is the retina, the multi-layered sensory part of the eye. The retina is an extension of the central nervous system and is connected to the brain by the optic nerve [34]. In the back of the eye, there is a region of the retina called macula. This region corresponds to the area surrounding the fovea, the retina's thinnest point, where the light entering the eye is focused. The macular region has the highest density of retinal ganglion cells and is vital for everyday visual function [35]. The location of the macular region is depicted in Figure 2.2.

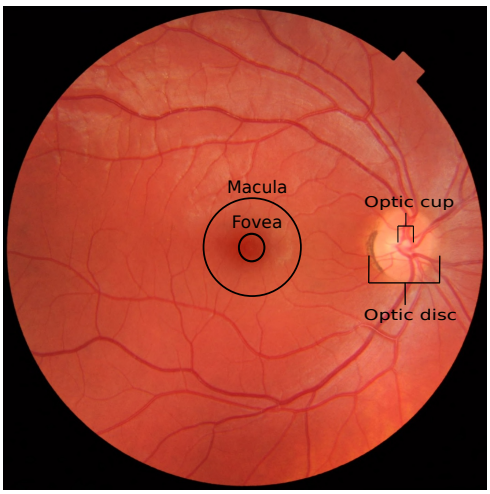

Figure 2.2: Image of the back of the eye, where the location of the macula, fovea, optic disc and optic cup are visible. Adapted from public domain image [36].

When light enters the eye, it first has contact with the cornea, that refracts the light, causing the image to converge on its way to the iris and pupil [37]. The pupil functions as an aperture that regulates the light that enters the eye, adjusted by the iris' ability to enlarge and shrink [38]. Then, the lens focuses the light through the humor vitreous, where it travels until it reaches the fovea. Here, the light is converted into electrical signals that are transmitted to the brain [39]. The trajectory of the light rays through the eye is also shown on Figure 2.1.

The retina, the anatomical region in which this project is focused, can be further divided into several distinguishable layers, with different properties and purposes. Different retinal layers can be affected by some ophthalmic diseases, like glaucoma [35], and non-ophtalmic diseases, like multiple sclerosis [40] or Alzheimer [41, 42]. Therefore, the study of their properties can unveil useful biomarkers for such diseases, with several clinical applications [43–45].

2.2 Glaucoma

Glaucoma is an insidious, unpredictable, and deceptive ophthalmic disease, affecting more than 80 million people worldwide. It represents the most common cause of irreversible blindness in the world, and its prevalence is expected to increase in the near future [1], making it a highly relevant topic for research.

Glaucoma patients may be asymptomatic or have minimal symptoms until a very severe stage of the disease, frequently resulting in delayed and missed diagnosis [46]. Around 50% of glaucoma patients are not diagnosed until advanced illness [47]. Since glaucomatous damage can be prevented if treated at an early stage, an early diagnostic and close progression monitoring is of the utmost importance to prevent further development of the disease and preserve vision. Because of this, the development of new methods to achieve an early diagnosis is in urge. Family history of glaucoma, older age, black race, hypertension, diabetes, cigarette smoking, use of systemic or topical corticosteroids, and high IOP, can all be considered risk factors for glaucoma, and are taken into account when evaluating suspect cases [46, 48, 49].

Pathologically, glaucoma can be characterized as a group of optic degenerative neuropathies, with a progressive loss of retinal ganglion cells (RGC), consequent structural changes in the optic nerve head and visual field defects [50]. Damage provoked in the optic nerve is usually characterized by a progressive thinning of the neuroretinal rim of the optic nerve head and loss of the retinal nerve fiber layer (RNFL) [51]. The progressive deterioration of the visual field usually begins in the midperiphery and progresses in a centripetal manner until only a central

or peripheral island of vision remains [46]. Because glaucoma is often asymmetrical between both eyes, and vision loss normally starts in the peripheral vision, the brain has the capacity to adjust to this loss, making the first symptoms very hard to notice and contributing to late diagnosis.

Several distinct types of glaucoma can be defined, including primary open-angle glaucoma (POAG), primary angle closure glaucoma (PACG) and normal tension glaucoma (NTG) [52]. The main difference between POAG and PACG is the iridocorneal angle (or anterior chamber angle, the part of the eye located between the cornea and iris), which is open in POAG and closed in PACG, blocking the aqueous fluid from draining out of the front of the eye [52].

POAG is the most incident type of glaucoma in the majority of populations [49]. Although its pathophysiology is not yet fully understood [53], it is usually associated with an increased level of IOP, causing mechanical stress and strain on the posterior structures of the eye, leading to damage to the optic nerve head and retinal ganglion cell death that characterizes glaucoma [54], depicted in Figure 2.3. Usually, no pain is associated with increased eye pressure, although sudden high levels of IOP might result in eye ball pain or headache.

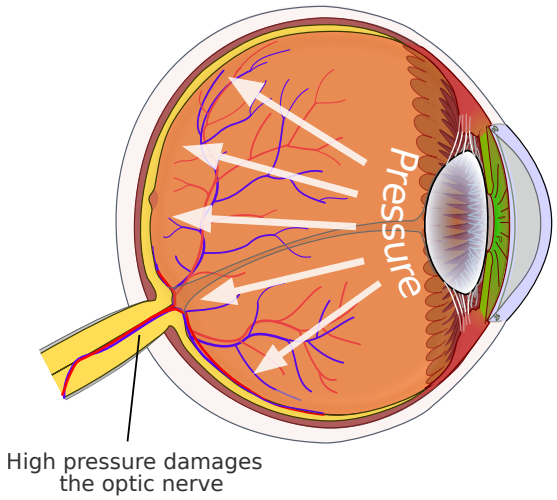

Figure 2.3: Visual representation of the effect of IOP in POAG. Adapted from public domain image [31].

Although high levels of IOP are frequently associated with glaucoma, in patients with NTG, optic nerve damage occurs with IOP levels within normal ranges (12-22 mmHg) [54], hampering even more the diagnosis. The remaining clinical characteristics and pathophysiology of NTG patients are very similar to those with POAG, being a type of open angle glaucoma as well [55]. The opposite can also happen, when a patient has elevated IOP in one or both eyes but there

is no clear evidence of optic nerve damage or visual field defects. This condition is designated by ocular hypertension (OHT).

In this project, subjects with POAG, NTG and OHT types of glaucoma were included.

Diagnosis and treatment

Risk assessment for glaucoma includes intensive treatment and monitoring for those who are vulnerable to vision loss, while avoiding over-treating and testing patients in cases when the disease is less likely to cause any damage. However, the limited understanding of the disease's pathophysiology, especially in early stages, prevents an effective risk assessment [56, 57].

Currently, glaucoma screening, diagnosis, and severity classification are based on clinicians' impression using the combination of IOP measurements, visual field and structural imaging parameters [58]. There is no single standard biomarker for the diagnosis of glaucoma, which often leads to suspect and misclassified cases. As glaucoma damages mainly the optic nerve area, monitoring of the optic nerve represents a crucial step to the diagnosis. The optic nerve damage appears histopathologically as ganglion cells death and nerve fiber loss with "cupping" of the optic nerve head (the central portion of the optic disc, also called optic cup, visible on Figure 2.2, becomes larger). These structural changes can be identified through eye imaging techniques, such as ophthalmoscopy or OCT, and have an important role in the diagnosis.

Current treatments for glaucoma focus on delaying the disease progression by reducing IOP, one of the main risk factors for this disease [59]. Even though other factors are involved, IOP is the only which can be pharmacomodulated [60]. However, in several patients, IOP-reduction is insufficient to significantly decrease disease progression, suggesting that other factors may be involved in the disease pathogenesis. In NTG patients, it is thought that vascular dysregulation of blood supply to the optic nerve might play a bigger role for the development and progression of the disease, with these patients often showing more signs of ocular and systemics vascular abnormalities [56, 59, 61, 62]. These abnormalities may further aggravate ganglion cell damage and vision loss.

2.3 Optical Coherence Tomography

OCT is an optical imaging modality based on low-coherence interferometry. It is a non-contact and non-invasive technique that provides *in vivo* cross-sectional images of microscopic structures. OCT has high longitudinal, lateral spatial and temporal resolutions and an extremely high sensitivity, making it an appealing technique to multiple areas in pre-clinical and clinical

research. Due to the transparency of ocular structures, ophthalmology is currently the main, and most common OCT application [63]. It is used as a key tool to analyze relevant biomarkers for some ophthalmic pathologies diagnosis, including glaucoma [64, 65]. Nevertheless, in the past decades, OCT applications have also emerged in other fields, such as dermatology [66], gastroenterology [67] and intra-arterial imaging [68].

OCT implementation is based on an interferometer. In a Michelson interferometer, light is emitted from a coherent source towards a beam splitter, and further split into two paths. The light is then reflected by two different mirrors and re-directed into the beam splitter, where it interferes. The light is then later detected by a photodetector.

Based on this, a basic OCT setup can be defined (Figure 2.4). A low-coherence light source emits light towards a beam splitter, which splits it into two different paths (the two arms of an OCT device). Light directed to the reference arm is reflected on a reference mirror and re-directed to the beam splitter, while the light directed to the sample arm is reflected by the multiple layers within the sample, and also re-directed to the beam splitter. Both reflected beams, after interference, will be detected by a photodetector [69]. Due to the broadband nature of the light, interference between the two light beams will only occur if the reference and the sample arm optical path lengths are matched to the coherence length of the light [69]. For this reason, the axial resolution of an OCT device is determined by the temporal coherence of the light source.

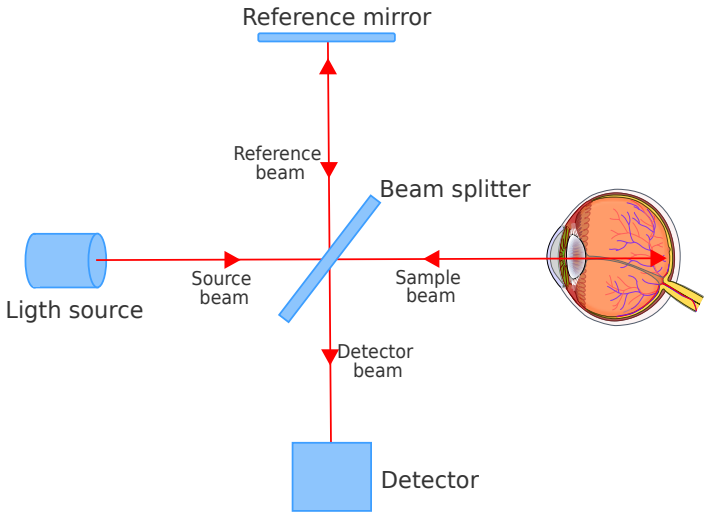

Figure 2.4: Diagram of the implementation of a basic OCT setup.

To obtain 3D information, an OCT device scans the object in both depth and transverse

directions [70]. At each point, an A-scan is produced from the time delay and intensity of the backscattered light, providing information about the properties of the sample as a function of depth at one position of the scanned beam. A B-scan, also called tomogram, is a cross-sectional image, produced by a sequence of neighboring A-scans. A B-scan is similar to a histology-like image, but obtained non-invasively. A 3D OCT volume can then be obtained by a sequence of B-scans (x-z scans). Figure 2.5 shows the difference between A-scan, B-scan and a 3D volume.

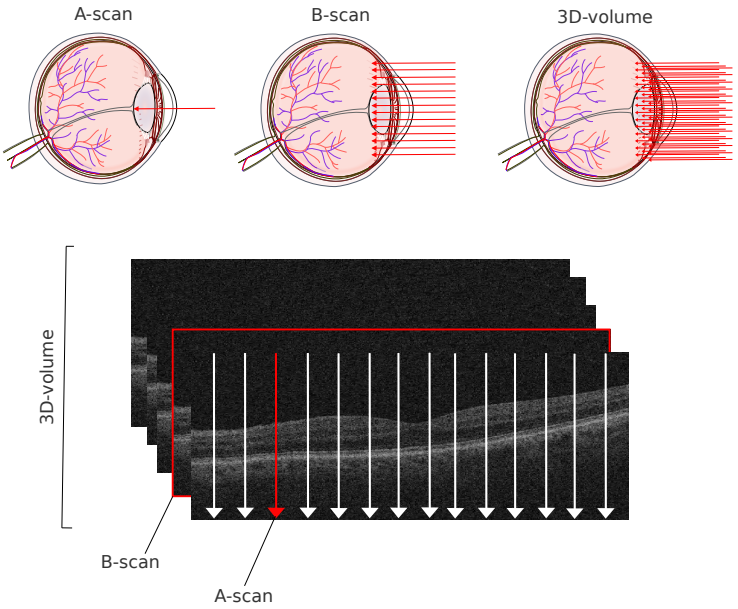

Figure 2.5: Diagram of the scan types in OCT.

Different OCT modalities have been developed over the last decades. The OCT modality used for data acquisition in this project is Spectral domain - optical coherence tomography (SD-OCT), a type of Fourier domain - optical coherence tomography (FD-OCT). FD-OCT is characterized by the absence of moving parts in the device, using a static reference mirror. Im SD-OCT, the depth of the sample is created from the Fourier analysis of the intensity spectrum across the returning light frequency range, also known as spectral interferogram.

2.4 Speckle in OCT

The speckle effect is an interference phenomenon that results from the reflection of coherent light by a rough surface, making it common to all coherent imaging modalities [25]. It is a scattering phenomenon that creates granular visual patterns composed by dark and bright spots, called speckle patterns.

The speckle phenomenon occurs in OCT, as it is a low-coherence interferometry imaging technique. In an OCT acquisition, the tissue is scanned by an optical beam, and most of the light is either refracted or scattered. The incident light scattered from different points in the sample will travel through different optical paths with varying lengths until it reaches the image plane. Because of this, the intensity of each specific point in the image plane will result from a destructive/constructive interference of all the backscattered light waves with different phases at that point. The distribution of the angles that constitute the scattered light will then depend on the optical refractive index of the imaged tissue. This phenomenon is illustrated in Figure 2.6.

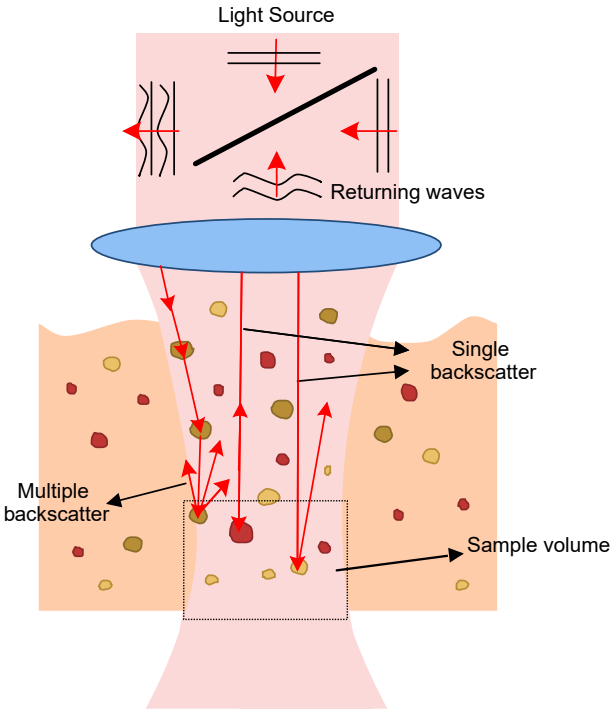

Figure 2.6: Process of speckle formation in OCT. Speckle patterns result from coherent superposition of multiple backscattered waves from particles in the sample volume.

In perfect conditions, with a completely polarized light source and a static medium, the speckle pattern resulting from the sample illumination is fully developed, which means that the amplitude and phase of the speckle field are statistically independent, and its speckle phase is uniformly distributed [71]. These fully developed patterns can be simulated mathematically [72], with results as presented in Figure 2.7. However, that is usually not the case in real applications.

Schmitt *et al.* [7] were among the first to discuss the origin of speckle in OCT, distinguishing between two types of speckle. The first one comes from the interference of multiple backscattering of the light beam inside and outside the sample. The second type is caused by delays of

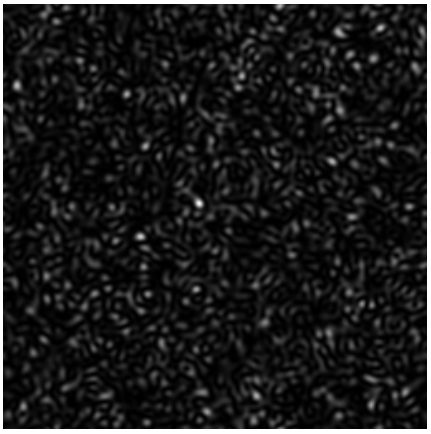

Figure 2.7: Synthetic speckle pattern generated with a speckle size of 4 pixels/speckle. Image size is 256×256 pixels. Speckle simulation algorithm is detailed in [72].

multiple forward scattering. The granular patterns visible in OCT scans (represented by image pixel intensity) represent speckle amplitude.

A number of works have interpreted speckle as a corrupting influence, *i.e.* a source of noise in OCT imaging. Speckle reduces the image quality, decreases the signal-to-noise ratio (SNR) and contrast, and makes boundaries between tissues less precise, which can hamper the extraction of image features. To prevent this, several methods to suppress and reduce speckle have been developed, including filtering [73], averaging between scans [74] or wavelet processing techniques [75]. Figure 2.8 shows a retinal OCT image before and after speckle reduction with a median filter.

Speckle patterns can change depending on different device-related parameters, such as the properties of the light source, the propagating beam, and the aperture of the detector. However, they are also known to change depending on the inner properties and structural organization of the tissues [8, 9], such as its optical refractive index. The latter indicates that speckle may contain relevant information regarding sub-resolution and structural properties of the tissues from which it is originated [6], as the optic refractive index depends on the biomechanics of the tissue, which depends on its inner properties and structural organization. Hence, it is expected that changes at the cellular level will affect this index and, consequently, the pattern of the scattered light. In fact, in the work presented by Schmitt *et al.* [7], speckle patterns in OCT are already mentioned as having a dual role, both as a source of noise, signal-degrading speckle, and as a carrier of information, signal-carrying speckle. This signal information may be used to characterize the underlying nature of the imaged tissue [24].

Since most of the works in literature focus on signal-degrading speckle, the information

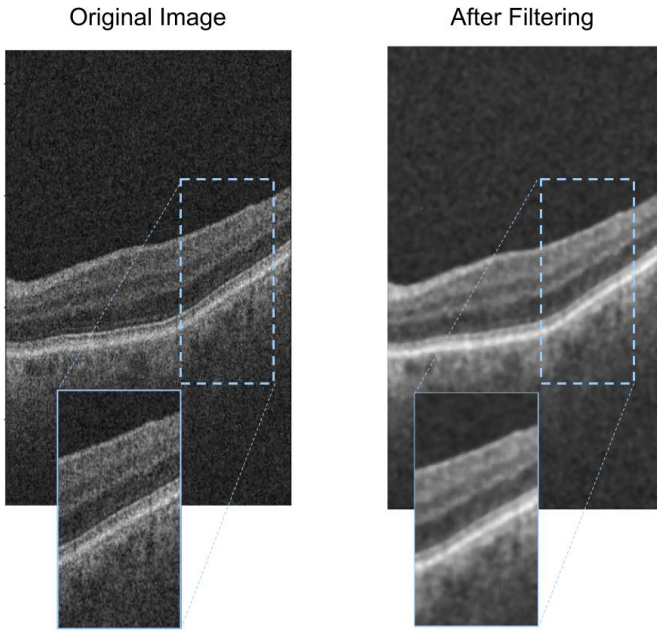

Figure 2.8: OCT image of the retina before (left) and after (right) median filtering.

regarding signal-carrying speckle analysis is more scarce and diffuse. Generally, two different types of speckle patterns analysis can be defined: static and dynamic. A static speckle pattern will remain unchanged over time, as all elements in the sample are static. However, if moving scatterers are present, e.g. blood cells in a blood vessel, they can cause a change of the pattern over time, resulting in a dynamic pattern. In this latter case, the temporal pattern changes contains information about the motion of the particles [9].

Although the interpretability of signal-carrying speckle remains a challenging task, a number of works have shown its feasibility for biomedical imaging related tasks including classification (e.g. healthy/pathological) [8, 10–27], segmentation (e.g. vessels) [76], motion quantification [6, 77–81], blood pulse pressure waveform estimation [82, 83], and cellular assessment in muscle tissue [84].

2.5 Glaucoma biomarkers in OCT

Due to its high resolution, OCT allows to have a more objective and clear evaluation of structures of the optic nerve and macular area, and can automatically provide parameters with high reproducibility [85]. These evaluations are especially important for early stages of glaucoma, as structural changes can happen before any functional damage is visible [86]. Nevertheless,

in advanced/severe cases of glaucoma, OCT can reach a "floor effect", a point where further functional progression does not lead to any alterations in the structural exams and can not be identified in OCT scans [56, 87].

OCT devices often provide automatic peripapillary RNFL thickness values, obtained from a ring outside the optic disc area. This is one of the most commonly used biomarkers in clinical practice. According to a study performed by Bowd *et al.* [88], the mean peripapillary RNFL thickness for normal eyes is around 85.8 μm, whereas in glaucomatous eyes the mean thickness is around 44.4μm, showing a significant thinning due to glaucoma. Optic nerve head parameters, like the neuroretinal rim area, the optic disc area, and the Bruch's membrane opening, are also frequently provided by OCT devices, and relevant for glaucoma diagnosis [89].

Macular parameters are also important, since the macula is the region with the highest density of ganglion cells. Macular parameters gained more popularity in the recent years, as recent technology advances in OCT devices have enabled more robust and automatic segmentation of the retinal layers. Even though more recently introduced into clinical practice, their importance has already been highlighted in several studies [51, 90–93]. Different OCT devices provide different macular parameters, but they usually consist in thickness measurements from the RNFL, or the ganglion cell layer (GCL)+inner plexiform layer (IPL) complex. Since they are calculated in the macular region, these parameters are not related with the anatomy of the optic nerve, which can be an advantage in patients who present other ophthalmic pathologies that affect the optic nerve, such as myopia. Contrarily, it is a disadvantage for patients with other pathologies that damage the retina, such as diabetic retinopathy. Macular parameters are the main focus of this project.

A combination of the previously defined OCT derived parameters (computed in both the peripapillary and macula regions) with visual field parameters and IOP values represent the standard biomarkers considered by clinicians for the diagnostic and progression monitoring of glaucoma patients.

The severity classification is mainly done using the visual field mean deviation (VF MD), as visual filed defects increase with disease severity [94]. However, visual field measurements require a great amount of concentration and cooperation from the patient, presenting a low repeatability and reproducibility, especially in advanced cases, where patients can present larger fluctuations in these exams. Although these factors hamper a correct monitoring of glaucoma progression [4], visual field testing remains the preferred exam type for this purpose. Moreover, the use of VF MD as a severity assessment metric implies that visual field loss has already

occurred, which would ideally be avoided.

2.6 State of the art

2.6.1 Speckle Analysis

Several methods have been proposed in the literature for retrieving information from speckle in OCT, for biomedical applications. A literature review was conducted in two databases on April 3rd, 2021. PubMed was chosen for being one of the largest databases in the medical field, and Scopus for combining articles from both a medical and a technical nature. The results of this review were submitted in the article titled "Signal-carrying speckle in Optical Coherence Tomography: a methodological review on biomedical applications" [28] (Section 1.5). In this section, some of the results of the review are summarized, namely the methods applied through the development of the project.

Features retrieved from OCT speckle can be successfully applied to different biomedical applications, such as classification [11, 12, 14] and segmentation [76] of a variety of tissues (e.g. cornea [19], skin [16], lung [25]). However, the results of the literature review show that speckle analysis is highly application-dependant and, hence, the best approach may vary between biomedical applications. Moreover, several of the reviewed publications are only performed either in a theoretical context or using phantom data, showing that signal-carrying speckle analysis in OCT imaging is still in its early stage, and further work is needed to validate its applicability and reproducibility in a clinical context. The specific formulation of the methods used in this project is detailed in Section 3.2.3.

Statistical Properties

Simple statistical properties of speckle intensities were used by Roy *et al.* [12], together with an estimate of optical attenuation and signal confidence measures to detect tissue's susceptibility to rupture using intravascular OCT, with the objective of assessing atherosclerosis. The retrieved features proved to have high performance in the identification of such tissues using a Random Forest predictive model with an area under the receiver operating characteristics curve (AUROC) of 0.967.

Statistical distributions

Applying different statistical distributions for speckle modeling is a common method to analyse speckle, having been proposed for various applications.

Two fundamental distributions have been proposed, the Rayleigh and K distribution. The Rayleigh distribution is valid when number of scatterers per coherence volume is high [95], and we have a fully developed speckle pattern. However, this argument is only partially applicable to biological tissues, as different tissues may have different natures, *i.e.* and some can be more heterogeneous with a lower number of scatterers [16]. On the other hand, the K distribution is said to be suitable for non-fully developed speckle patterns, which corresponds to the majority of real data.

However, the optimal probability density function (PDF) can differ depending on several other parameters, as the analysed tissue or the OCT device used. This led to various other distributions being tested for this purpose. Some authors [8, 76] have presented comparative studies among several PDFs to find the best fit for their application, specifically to model speckle from different skin layers, and from different groups in corneal data, respectively. Despite both works including Nagakami, Generalized Gamma (GG), and Lognormal distributions, their conclusions were vastly different. For skin layers [76], Nakagami distribution was found to be most adequate. However, for corneal data [8], the best fit was achieved by the GG distribution.

Other authors, such as Kirillin *et al.* [96] and Almasian *et al.* [97] studied speckle phenomenon statistics in OCT using theoretical modulations and simulations. Almasian *et al.* [97] also proved experimentally the goodness of fit of the Rayleigh PDF using controlled samples of silica microspheres suspended in water. They also claimed that OCT speckle amplitude distribution for homogeneous samples can be described by a Rayleigh distribution for images with low optical depth (coefficient of determination, $R^2 \approx 0.98$). Kirillin *et al.* [96] developed a Monte Carlo model for speckle statistic simulation of OCT data, and proved the validity of the model using a tissue phantom with known properties. Also, they demonstrated by visual inspection that the Gamma distribution was a good fit for both phantom and simulated data. They also showed that some features retrieved with this method depend on the scatterers' concentration, and that the analysis of fitting parameters may provide physical insight into the underlying tissue scattering characteristics.

The practical applicability of this approach to different tasks has also been highlighted, mainly to classify between different tissues. In the ophthalmology field, Niemczyk *et al.* [21] modeled speckle from corneal data using a Gamma distribution, in order to understand the relation of the speckle properties with IOP, and concluded that there are some IOP induced alterations in the optical scattering pattern, and therefore this method has the potential to be used as a non-invasive assessment of corneal properties and IOP measurement.

Jesus *et al.* [8, 14, 15] and Iskander *et al.* [19] also modeled speckle from the cornea, but using the GG distribution. Jesus *et al.* [8, 14] studied variations of the different parameters of the GG distribution between healthy subjects from three different age groups. A significant statistical difference, (p-*value*<0.05, Kruskal-Wallis test) was observed for all three parameters. In a later study [15], the parameters relation with micro-structural corneal properties were analysed. Significant correlation (p-*value*<0.001) was found between the statistical parameters and the intraocular pressure (IOP). The authors suggest that GG parameters can contribute to the correction of IOP measurements performed with non-contact tonometry. In the study performed by Iskander *et al.* [19], speckle information from the cornea was modelled with the GG distribution and used to differentiate between glaucoma suspects, glaucoma patients, and healthy controls. They were able to find significant differences between the parameters of the three study groups (p-*value*<0.0001, Fisher's Test). These studies highlight the potential of speckle statistics to infer micro-structure properties of corneal tissue, and their applicability to complement standard methods.

In a different application, Seevaratnam *et al.* [13] also used the GG distribution to investigate the effect of temperature variation in tissue phantoms, and found a correlation between GG parameters and the increase of the temperature of the tissue (p-*value*$=7.93 \times 10^{-6}$, Student's t-test), showing another possible application for speckle analysis.

Using OCT lymphangiography and neurography, Matveev *et al.* [18] and Demidov *et al.* [16] applied spatial speckle statistics to map lymphatic vessels. Their experiments, on normal skin and tumor tissues, showed that by using different regions of interest (ROI) in an image, Rayleigh distribution properties could be used as features for nerves and lymphatic vessels mapping.

Contrast ratio

Speckle contrast is a method that has been applied not only in OCT, but in several other imaging modalities, such as Laser Speckle Imaging [9]. In OCT, Hillman *et al.* [98] performed an experiment using controlled tissue phantoms to prove the existence of a correlation between local speckle contrast statistics and the scatterers density in the sample, suggesting that speckle contrast might be a promising method to analysis tissue properties. Duncan *et al.* [99] also computed information retrieved from the local contrast image of simulated synthetic speckle patterns to perform image segmentation. Experiments using chick embryo OCT images were performed and vessels and background were successfully segmented using differences in contrast parameters.

Spatial Gray-Level Dependence Matrices (SGLDM)

Kasaragod *et al.* [24] used the spatial gray-level dependence matrices (SGLDM) method to retrieve information from speckle OCT images. A Bayesian model was applied to the classification of tissue phantoms with different amount of scatterers, and to identify the invasion of melanoma cell into tissue engineered skin. Their results were satisfactory in classifying the number of scatterers in the tissue phantoms, shown visually by a receiver operating characteristic (ROC) plot. However, this approach provided limited results in the identification of the melanoma cells in the tissue. Finally, the results of these experiments showed that texture features vary with both size and concentration of scatterers in a sample.

Frequency Domain & SGLDM

Gossage *et al.* [25, 27] combined information retrieved from SGLDM and frequency domain methods for analysing and classifying texture of different tissues. The features were used to develop a model for differentiating mouse skin and fat, and normal versus abnormal mouse lung tissue, based on a minimum-error-rate Bayesian model. The results proved the efficiency of this method, resulting in a high accuracy classification of mouse skin and fat, of 98.5 and 97.3% respectively, and also resulting in a satisfactory performance for normal and abnormal mouse lung, tissues with a higher visual similarity, of 64.0 and 88.6% respectively.

Table A.1, included as an appendix, summarises the essential information to retain from all the articles aforementioned, including the application and aim of each study and OCT device details. The implementation of each of these methods are detailed in Section 3.2.3.

2.6.2 Disease Progression Modelling

Disease progression modelling (DPM) is a novel data-driven research technique dedicated to the understanding of the temporal progression of degenerative diseases, balancing clinical knowledge and machine learning. It has emerged from the need to further understand common neurodegenerative diseases, like Alzheimer's [100] and Parkinson's disease [101]. DPM attempts to model and understand the progression of different biomarkers for a specific disease, in order to further understand how and when the different biomarkers change, that is, to estimate and interpret the temporal sequence of biomarkers of a disease.

DPM techniques have to overcome several challenges due to the nature of the biomarkers and symptoms of degenerative diseases. Some of the most relevant include the disease heterogeneity between patients and the ability to accurately acknowledge the effect of genetics. Moreover, there

is usually a lack of information and available data about the silent phase of disease. Since data collection generally does not happen prior to the appearance of any symptoms, it is extremely hard to study the earliest pathological processes.

Several methods for DPM have been proposed, such as Disease Course mapping [102], Gaussian Process progression modeling [103] and Event-based modeling (EBM). Unlike Disease Course mapping and Gaussian Progression, which require the use of longitudinal datasets (datasets with information collected from the same subjects at multiple points in time), EBM is developed using cross-sectional data, data collected from a vast population with subjects in different stages of the disease at a specific point in time. This is a powerful advantage of EBM over other DPM methods, as cross-sectional datasets are generally more widely-available for research, compared to large longitudinal datasets, which are very time-consuming to collect and generally scarce. Also, it allows future patients to be diagnosed without need for personal patient history information.

EBM was proposed in 2011 by Fonteijn *et al.* [104] and has been used in multiple neurodegenerative diseases, such as Alzheimer's disease [105], Huntington's disease [106], posterior cortical atrophy [107], Parkinson's disease [101], and amyotrophic lateral sclerosis [108]. It estimates the temporal sequence of events in a certain disease, the order in which biomarkers become abnormal. A biomarker abnormality event is defined as the moment when a biomarker goes from a normal state to an abnormal state. EBM assumes monotonic increase or decrease of biomarker values with the increase of disease severity, which means: disease progression is irreversible. This assumption means that, once a symptom/biomarker becomes abnormal, it will not regress back to its normal state. The EBM uses this concept across multiple biomarkers to estimate the disease progression sequence.

Some of the assumptions made in the initial versions of EBM were too rigid to apply in real problems, and have been relaxed in more recent approaches. For example, in the original EBM proposed by Fonteijn *et al.* [105], the ordering of events was assumed to be common for all the subjects in a dataset, and healthy controls were assumed to present all biomarkers within the normal (healthy) ranges. However, degenerative diseases are notoriously heterogeneous in terms of pathology, as well as in their clinical presentation in patients, and healthy subjects can have presymptomatic disease. More recent version of EBM [109] do not make any of these assumptions.

DEBM model, introduced in 2017 by Venkatraghavan *et al.* [110], is a novel technique that estimates the orderings robustly in the presence of heterogeneity among subjects. The DEBM

also estimates the relative temporal distance between the events, as opposed to initial versions of EBM [104, 105], which performed a strict quantification of the position of each event.

Venkatraghavan *et al.* [111, 112] proved that DEBM is more accurate than previously presented EBM methods [105, 109, 113] for Alzheimer disease progression modeling, by performing extensive comparisons using both simulated and real data (1737 subjects). They also proposed a patient staging algorithm for classifying each patient based on the estimated disease progression timeline. This algorithm uses relative proximities between events, and has proved to correctly stage patients into severity disease levels. The results of their experiments were encouraging, and suggest that DEBM is a promising approach to disease progression modeling, having the potential to assist clinicians with computer aided diagnosis. Due to this results, DEBM was the technique selected for this project, and details on its implementation are presented in Section 3.4.

To the knowledge of the research team, neither EBM nor DEBM have been used for attempting glaucoma progression modeling.

Chapter 3

Methods

This chapter comprises the methodology behind the development of this project. First, details of the dataset used are given, including description of the data acquisition device (Section 3.1). Following, the techniques used for feature extraction from the OCT volumes are detailed (Section 3.2). Next, are presented the methods used for an initial statistical analysis of the features retrieved and consequent feature pre-selection for the model implementation (Section 3.3). Lastly, the details on DEBM's implementation, optimization and performance evaluation are shown (Section 3.4). Python was used for the implementation of all algorithms presented in this dissertation.

Figure 3.1 outlines the different steps taken in this project, and summarizes the structural organisation of this chapter. A more detailed version of this figure is included in the Appendix (Figure B.1), which can serve as a guide while reading this chapter.

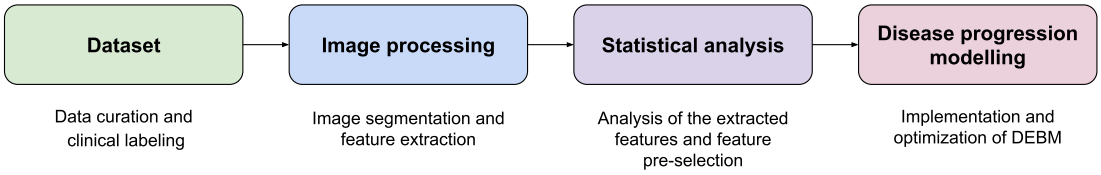

Figure 3.1: Outline of the project's steps.

3.1 Dataset

The dataset used in this project is from the Leuven Eye Study cohort [59], one of the largest clinical trials on glaucoma, with 280 patients. This cross-sectional database includes healthy

controls, glaucoma patients (including NTG, POAG and OHT), and glaucoma suspects.

There is one OCT volume of the macular region available for each patient, and a set of clinical relevant parameters, like age, gender, eye imaged (left or right), IOP and visual field mean deviation. Five patients were excluded due to the lack of quality of their OCT volume, as the surfaces between retinal layers were not clear, making the segmentation unrealiable.

The healthy controls were recruited from people accompanying the patients (while excluding blood relatives, those with a family history of glaucoma, an increased or asymmetrical cup/disc ratio or any other optic disc structural change or an IOP above 21 mmHg). Subjects in all groups with a history of ocular trauma or any other eye disease were excluded, as well as patients with diabetes mellitus, as the disease has known effects in the retina that can be considered confounders in this case [114,115]. The classification of each patient between glaucoma, glaucoma suspect and healthy was performed by a glaucoma specialist, based on an extensive clinical examination combining several clinical parameters and the clinician's expertise.

According to the European Glaucoma Society (EGS) guidelines [94], the glaucoma suspects (16 subjects) were required to have at least one abnormal or suspicious parameter among the following: visual field, parameters from the optic disc, or circumpapillary retinal nerve fiber layer (cpRNFL) thickness. Due to the uncertainty related with these cases, they were not included in this study. After excluding suspect cases and those with unreliable data, a final dataset of 259 patients was left.

Since the main goal was to study glaucoma disease progression, the group of glaucoma patients were clustered into three severity stages, according to the the criterion from the EGS guidelines [94] to classify glaucomatous eyes, where patients with a VF MD higher or equal to –6 dB were considered to have mild glaucoma, lower than –6 but higher or equal to –12 dB were considered moderate glaucoma, and those with a VF MD lower than –12 dB were considered severe glaucoma cases. The final dataset is presented in Table 3.1 for each of the four groups (healthy, mild, moderate, and severe glaucoma). Figure 3.2 presents the distribution of the VF MD values by severity groups. For an easier visual interpretation, a Kernel density estimate (KDE) was fitted to the distribution of each group, as a way to estimate their PDF. The *y-axis* on the left side is the number of patients, and in the right is the probability density, the probability per unit on the *x-axis*.

The Kruskal Wallis H test was used for assessing possible differences in gender, imaged eye, age and IOP between severity groups. This same test was applied later on to the speckle features, also to asses the level of differences between severity groups. Details on this test and reasoning

Table 3.1: Demographic information from the subjects in each group. *p-values* presented were obtained with the Kruskal Wallis H test and indicates *p-values* of overall differences between the diagnostic groups. n=number of subjects; M/F=Male/Female; L/R=Left/Right; SD=Standard deviation.

	n	**Gender (M/F)**	**Eye (L/R)**	**Age (Years) (Mean ± SD)**	**VF MD (dB) (Mean ± SD)**	**IOP (mmHg) (Mean ± SD)**
Healthy	57	32/25	28/29	62.9 ± 12.3	0.06 ± 1.02	15.1 ± 3.1
Mild	94	45/49	42/52	67.5 ± 10.3	-1.8 ± 2.3	14.0 ± 4.5
Moderate	48	27/21	19/29	65.4 ± 8.5	-8.5 ± 1.7	12.7 ± 4.1
Severe	60	35/25	32/28	68.8 ± 11.3	-18.1 ± 4.4	12.4 ± 4.4
p-value	-	0.56	0.51	0.02	-	0.0002
TOTAL	259	139/120	121/138	66.4 ± 10.9	-6.4 ± 7.6	13.6 ± 4.3

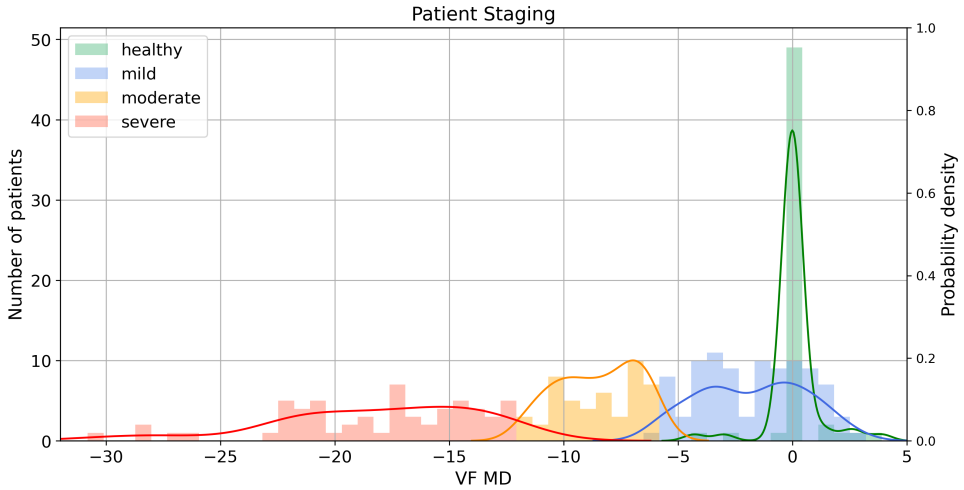

Figure 3.2: Distribution of the VF MD by severity groups, with a fitted KDE.

behind its use are presented later on in Section 3.3. No statistically significant differences were found in the gender or imaged eye (left/right). However, for the age, the differences can be considered statistically significant, if considering a confidence interval of 95%. Age is a known risk factor for developing glaucoma [116]. To account for this, and by suggestion of the clinical team, biomarkers were corrected for age and gender before being included in the DEBM model. Regarding the IOP, it is possible to see that the values are generally lower in more advanced cases of glaucoma, which could be considered unexpected, as glaucoma disease is often associated with high levels of IOP, as explained in Section 2.2. The reason for this incoherence is due to the topical medication administered to advanced glaucoma patients, that aims to lower the

IOP [117]. Thus, the slight decreasing trend seen in IOP values with the increase of disease severity is not an effect of the disease progression. This being said, the IOP information was not included in this study as a biomarker, in order to reduce biasing the prediction model with parameters dependent on medication.

The OCT device used for the acquisition of the 3D volumes of the macular retina in this dataset was the Cirrus 4000 HD OCT model (Carl Zeiss, Dublin, CA), a SD-OCT device with an axial resolution of 5 μm and a transverse resolution of 15 μm. Each 3D volume acquired has 128 B-scans, with 1024×512 pixels each, as a grayscale image (8-bit).

Although studying the optic disc may seem like the obvious choice for glaucoma diagnosis, as it is the region that shows to be more affected by glaucomatous damage, macular OCT scans have the advantage of being more common and easier to acquire for the patient, compared to the optic disc scans. Also, the effects of glaucomatous damage in the macular region have already been highlighted in several studies and proved to have a similar value for the glaucoma diagnosis [118–120].

3.2 Image Processing

This section details the methods used for the feature extraction from the OCT volumes.

3.2.1 Segmentation of retinal layers

The retina can be segmented in several different layers, with different structural properties (Section 2.1). For this reason, in order to perform a clinically meaningful analysis of the OCT volumes, the features had to be calculated individually for each layer, meaning the volumes had to be previously segmented. For this purpose, the Iowa Reference Algorithm (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, IA) [121–123], an automatic segmentation tool publicly available for research purposes, was used.

The segmentation algorithm was used to identify 11 surfaces (Figure 3.3):

1. internal limiting membrane (ILM)
2. RNFL - GCL
3. GCL - IPL
4. IPL - inner nuclear layer (INL)
5. INL - outer plexiform layer (OPL)
6. OPL - Henle fiber layer (HFL)

7. Boundary of myoidand ellipsoid of inner segments (BMEIS)
8. Inner-segment/Outer-segment junction (IS/OSJ)
9. Inner boundary of Outer Photoreceptor (IB_OPR)
10. Inner boundary of retinal pigment epithelium (IB_RPE)
11. Outer boundary of RPE (OB_RPE)

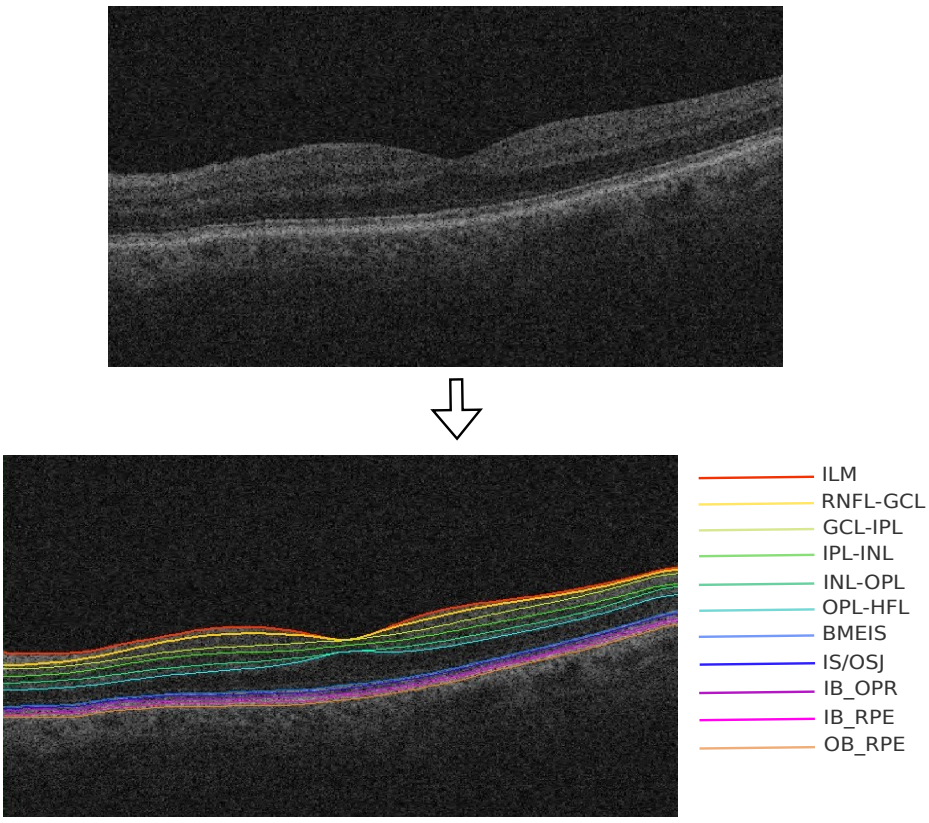

Figure 3.3: Example of a macula centred OCT B-scan from Leuven Eye Study [59]. The image was segmented using the Iowa Reference Algorithm [121–123].

The identified surfaces were then used to generate a mask labelled from 0 to 10. From 1 to 2: RNFL; From 2 to 3: GCL; From 3 to 4: IPL; From 4 to 5: INL; From 5 to 6: OPL; From 6 to 7: outer nuclear layer (ONL); From 7 to 8: inner-segment/outer-segment (IS/OS); From 8 to 9: outer segment layer (OSL); From 9 to 10: outer photoreceptor (OPR) ; From 10 to 11: retinal pigment epithelium (RPE). The segmented mask can be observed in Figure 3.4. All the methods for feature retrieval, presented in Sections 3.2.2 and 3.2.3, were applied separately to each of these 10 retinal layers.

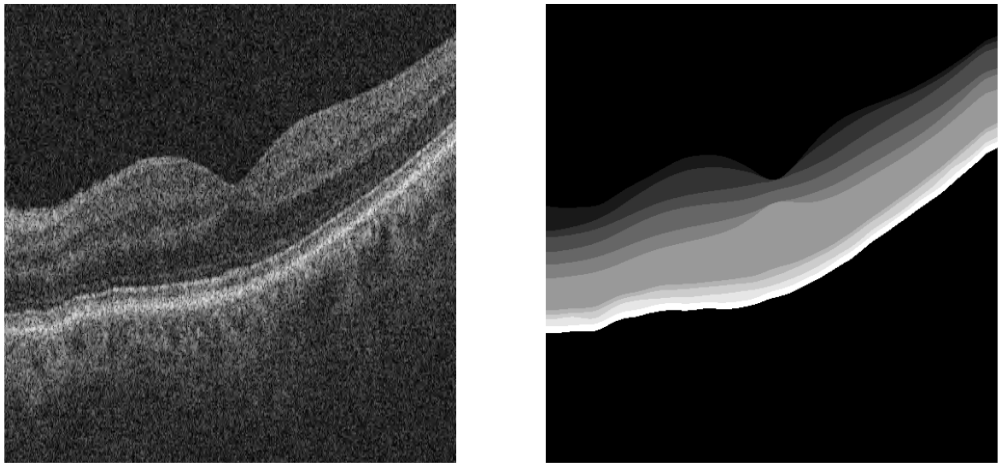

Figure 3.4: Visualization of the final segmentation mask (right) of a retinal B-scan (left) using the Iowa Reference Algorithm. The segmented layers, from the top (darker gray) to the bottom (lighter gray), are: RNFL; GCL; IPL; INL; OPL; ONL; IS/OS; OSL; OPR and RPE.

3.2.2 Thickness measurements

The mean thickness of the peripapillary RNFL is one of the main biomarkers used in glaucoma diagnosis, and several studies have shown that glaucomatous damage leads to a thinning of the RNFL in this area [88, 124].

Although RNFL measurements in the optic disc region are more commonly used, the thickness from the macular region has been proved to be affected by glaucoma progression as well [119, 120].

As opposed to speckle parameters, which provides information on the tissue microstructure, the thickness of each retinal layer provides information on the macrostructure of that layer. The thickness of each of the 10 layers was calculated for every individual A-scan in the volume, and then the mean value was taken. This resulted in 10 thickness features per patient.

3.2.3 Speckle features

One of the main goals of this project is to investigate whether speckle information from retinal OCT data is affected by glaucomatous damage, and consequently has the potential to be used as a discriminating feature for glaucoma. For this purpose, the methods presented in this section were implemented and tested. Every feature retrieved from speckle was calculated individually for each of the retinal layers.

Statistical properties

Local statistical properties of OCT speckle amplitude, presented by OCT pixel intensity, have been used for inferring speckle characteristics [12]. These properties include: mean (μ), standard deviation (σ), kurtosis (κ) and skewness (ν) of the pixel intensities.

$$\kappa = \frac{\mu_4}{\sigma^4} \tag{3.1}$$

where μ_4 is the fourth central moment and σ is the standard deviation.

$$\nu = \frac{m_3}{m_2^{3/2}} \tag{3.2}$$

where $m_i = \frac{1}{N}\sum_{n=1}^{N}(x[n] - \mu)^i$ and N is the size of the sample.

Statistical distributions

Speckle formation is a highly complex random phenomenon, and difficult to modulate. Nevertheless, its statistics within an OCT image can be described by fitting a PDF to the image pixel intensity histogram, corresponding to the speckle amplitude. Some authors have studied these statistical properties and several PDFs have been proposed for fitting OCT pixel intensity data from different applications. The parameters describing these distributions can change according to the light source properties and dimension/organization of the scatterers in the sample. Although there are different notations and formulations in literature, a coherence on the mathematical notation was maintained along this section in order to make the methods more comprehensive. This section is organized as follows: first, the two fundamental distributions used to represent fully developed and non-fully developed speckle patterns, Rayleigh and K-distribution, are presented; then, the Gamma and Generalized Gamma distributions are detailed. These two are the starting point for the distributions presented afterwards, which enclose Weibull, Nakagami, Rician, and Lognormal.

The optimal PDF for modelling OCT speckle data can be different depending on the analysed tissue, the application, and the OCT device. For this reason, in this project, 8 distributions were used to model speckle statistics of the pixels of each retinal layer, chosen according to the state of the art works found in the literature. The Python library *SciPy* [125] was used for this purpose, and the resultant parameters of the data fitting with each PDF were later studied as

potential discriminative features for the problem. Figure 3.5 presents a diagram showing the relations between the 8 presented distributions.

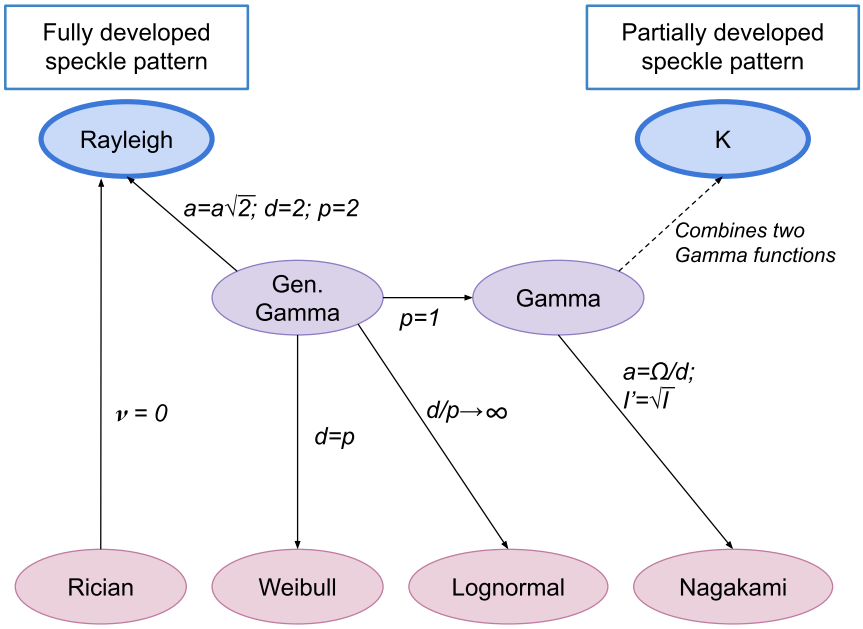

Figure 3.5: Diagram with the relation between the statistical distributions used in this project.

Fundamental distributions: The Rayleigh distribution is a 1-parameter distribution, used to model fully developed OCT speckle patterns. Its PDF is given by:

$$p_{RL}(I;a) = \frac{I}{a^2} e^{\left(-\frac{I^2}{2a^2}\right)}, \tag{3.3}$$

where a is the scale parameter, and I is the OCT pixel intensity. In most biological tissues, because of the heterogeneity in tissues natures, the resultant speckle pattern is not fully developed, leading to other distribution being proposed.

K-distribution's PDF is a 3-parameter distribution which combines two gamma functions and can be written as Equation 3.4. In speckle analysis, it is suitable for non-fully developed speckle patterns, which corresponds to the majority of real data. K-distribution has been used to model cases where a small number of scatterers are present in the sample [126].

$$p_K(x;\upsilon,\varphi,L) = \frac{2\xi^{(\beta+1)/2} x^{(\beta-1)/2}}{\Gamma(L)\Gamma(\varphi)} K_\alpha(2\sqrt{\xi x}) \tag{3.4}$$

$$K_n(z) = \frac{\pi}{2}\frac{I_{-n}(z) - I_n(z)}{sin(n\pi)} \tag{3.5}$$

$$I_n(z) = \frac{1}{2\pi i}\oint e^{(z/2)(t+1/t)}t^{-n-1}\,dt \tag{3.6}$$

$$\Gamma(n) = (n-1)!\ \forall\, n \in \mathbb{N}^+ \tag{3.7}$$

where $\alpha = \varphi - L$, $\beta = L + \varphi - 1$, $\xi = L\varphi/\mu$, K_α is a modified Bessel function of the second kind of order α and Γ is the Gamma Function (Equation (3.7)).

Gamma distributions: Besides the two previously presented distributions, several others, without a specific physical relation with speckle phenomenon, can be applied to achieve a better fit to data from multiple applications [21, 96]. One example is the Gamma distribution, a 2-parameters distribution (a and d) with a PDF defined as:

$$p_G(I; a, d) = \frac{I^{d-1}e^{-I/a}}{a^d\Gamma(d)} \text{ for } a, d > 0 \tag{3.8}$$

where d is the shape parameter, a is the scale parameter and Γ represents the Gamma Function (Equation (3.7)). Fitting a Gamma distribution to OCT data can provide a dependence of these two parameters to the optical properties of the samples, such as scatterers concentration [96].

Recent studies have also proposed the Generalized Gamma (GG) distribution as the best fit to represent speckle in OCT of tissue phantoms [13] and corneal data [8, 14, 15, 19]. It is a 3-parameter distribution that represents a generalization of the Gamma distribution, with a PDF given as:

$$p_{GG}(I; a, d, p) = \frac{pI^{d-1}}{a^d\Gamma(d/p)}e^{-(I/a)^p} \text{ for } p > 0, \tag{3.9}$$

where d and p are shape parameters and a the scale parameter. To obtain the Gamma PDF, p must be set to 1. Special cases of GG include the Rayleigh PDF (Equation (3.3)), by setting the parameters to $p_{GG}(I; a\sqrt{2}, 2, 2), a > 0$.

Gamma derived distributions: The 2-parameter Weibull distribution [8], can be obtained

from the GG distribution for d=p:

$$p_W(I;a,d) = \frac{dI^{d-1}}{a^d} e^{-(I/a)^d} \tag{3.10}$$

This distribution has been tested to model speckle corneal data [8].

The Nakagami distribution is a 2-parameter distribution, obtained from the Gamma distribution by setting $a = \Omega/d$ and taking the square root of the original random variable, $I' = \sqrt{I}$ [127]. Its PDF is represented in Equation (3.11):

$$p_{NK}(I;d,\Omega) = \frac{2d^d}{\Gamma(d)\Omega^d} I^{2d-1} e^{-\frac{d}{\Omega}I^2} \tag{3.11}$$

where d is a shape parameter and Ω is a spread parameter. This distribution has been tested for both corneal [8,14] and skin [76] speckle data against other distributions.

The Rician, or Rice distribution (Equation 3.12) is a 2-parameter generalization of the Rayleigh distribution, by the introduction of the noncentrality parameter.

$$p_{RI}(I;a,\nu) = \frac{I}{a^2} e^{-\frac{I^2+\nu^2}{2a^2}} I_0\left(\frac{I\nu}{a^2}\right) \tag{3.12}$$

$$I_0(z) = \sum_{k=0}^{\infty} \frac{(\frac{1}{4}z^2)^k}{(k!)^2} \tag{3.13}$$

where ν is the noncentrality parameter and I_0 is the zero order modified Bessel function of the first kind. To obtain the Rayleigh distribution (Equation (3.3)), the ν parameter is set to 0.

The Lognormal distribution is a 2-parameter distribution whose logarithm follows a normal distribution. Its PDF is given by Equation (3.14), and can be derived from GG distribution by setting $d/p \to \infty$.

$$p_L(I;\mu,\sigma) = \frac{1}{\sigma I\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(logI-\upsilon)^2}{2\sigma^2}} \ , \tag{3.14}$$

where υ is the logarithmic mean and σ is the logarithmic standard deviation of the random variable I. It has been tested to fit corneal [8] and skin speckle data [76].

Contrast ratio

The contrast ratio (C) of an OCT image can be defined as the ratio of the signal's standard deviation, (σ), to its mean, (μ) [98]:

$$C = \frac{\sigma}{\mu} \tag{3.15}$$

A correlation can be found between contrast statistics of OCT images, duee to their speckle component, and sample properties, namely scatterers concentration [98].

By calculating the local contrast, the contrast over a localized spatial region, of small neighbourhoods across an intensity image (e.g. 3×3 pixel neighbourhoods), a contrast map image can be created. Several statistical parameters can then be computed over this contrast image, the most commons being the mean and standard deviation. The spatial contrast equation, proposed by Briers and Webster [128] computes the contrast in square regions of lateral size n:

$$\mu_{i,j} = \frac{1}{n^2} \sum_{x=i-\frac{n-1}{2}}^{i+\frac{n-1}{2}} \sum_{y=j-\frac{n-1}{2}}^{j+\frac{n-1}{2}} I_{x,y} \ , \tag{3.16}$$

$$K_{i,j} = \frac{\sqrt{\frac{1}{n^2} \sum_{x=i-\frac{n-1}{2}}^{i+\frac{n-1}{2}} \sum_{y=j-\frac{n-1}{2}}^{j+\frac{n-1}{2}} (I_{x,y} - \mu_{i,j})^2}}{\mu_{i,j}} \ , \tag{3.17}$$

where n is the chosen size of the neighborhood, $I_{x,y}$ is the intensity value of the image at position (x,y) and $K_{i,j}$ is the contrast value of the element centered at (i, j). The size of the neighborhood element (n) must be an odd number larger than one, to ensure that the processing element is always associated with a single central pixel. This element is then moved through the image and its contrast is computed yielding a contrast map image. An adaptation of this algorithm for a 3D volume, using a $3 \times 3 \times 3$ is represented in Figure 3.6.

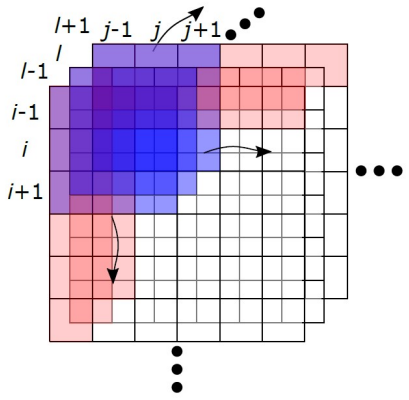

Figure 3.6: Visual representation of the 3D contrast map algorithm, using a neighbourhood of $3 \times 3 \times 3$. Image from [9].

To deal with the impossibility of evaluating the intensity at coordinates outside the image,

padding by duplication was used. In this approach, new borders are added to the image, obtained by duplicating the previous borders, and the new corners by taking the average of the two closest pixels. This padding approach was used so that the the resulting contrast map has the same size as the original image and the influence of the values in the padded borders would be minimum (as opposed to other padding approaches such as zero-padding).

The contrast maps of each OCT volume were calculated using the equation (3.17), adapted to a 3D volume, using a $3 \times 3 \times 3$ neighborhood element, in order to obtain a contrast map volume. After the contrast maps were computed, the mean and standard deviation of the voxels corresponding to each layer were calculated and used as features.

Spatial Gray-Level Dependence Matrices (SGLDM)

Texture is a result of local variations in brightness within one small region of an image. By looking at an image, different textures can easily be perceived or identified, that can be used to differentiate between different sets of images [129]. Texture analysis may help differentiate tissue without relying only on visible structures, and has been previously used to analyse ultrasound, magnetic resonance imaging (MRI), computed tomography (CT), fluorescence microscopy, light microscopy, and other digital images [25, 130, 131]. When speckle is present, one possible approach to OCT images texture analysis are spatial gray-level dependence matrices (SGLDM), also referred to as grey level co-ocorrence matrices [24–27].

SGLDM measure variations in the brightness of an image, and are determined by the estimation of the second-order joint-probability distribution of each combination of grey-level values that occur next to each other, averaged over directions of 0^{o}, 45^{o}, 90^{o}, and 135^{o}. Assuming the images are normalized with L grey scale levels, each $f(i,j \mid d,\theta)$ is the probability of a pixel with level i being at a distance d from a pixel with gray level of j in the θ direction. An $L \times L$ matrix can be created for each direction θ and for each chosen distance d, as is shown in Figure 3.7.

Several features can be extracted from these matrices, including energy (3.18), entropy (3.19), correlation (3.20), local homogeneity (3.21) and contrast (3.22):

$$Energy = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sqrt{[s_\theta(i, j \mid d)]^2} \tag{3.18}$$

$$Entropy = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} -s_\theta(i, j \mid d)\ log\big[s_\theta(i, j \mid d)\big] \tag{3.19}$$

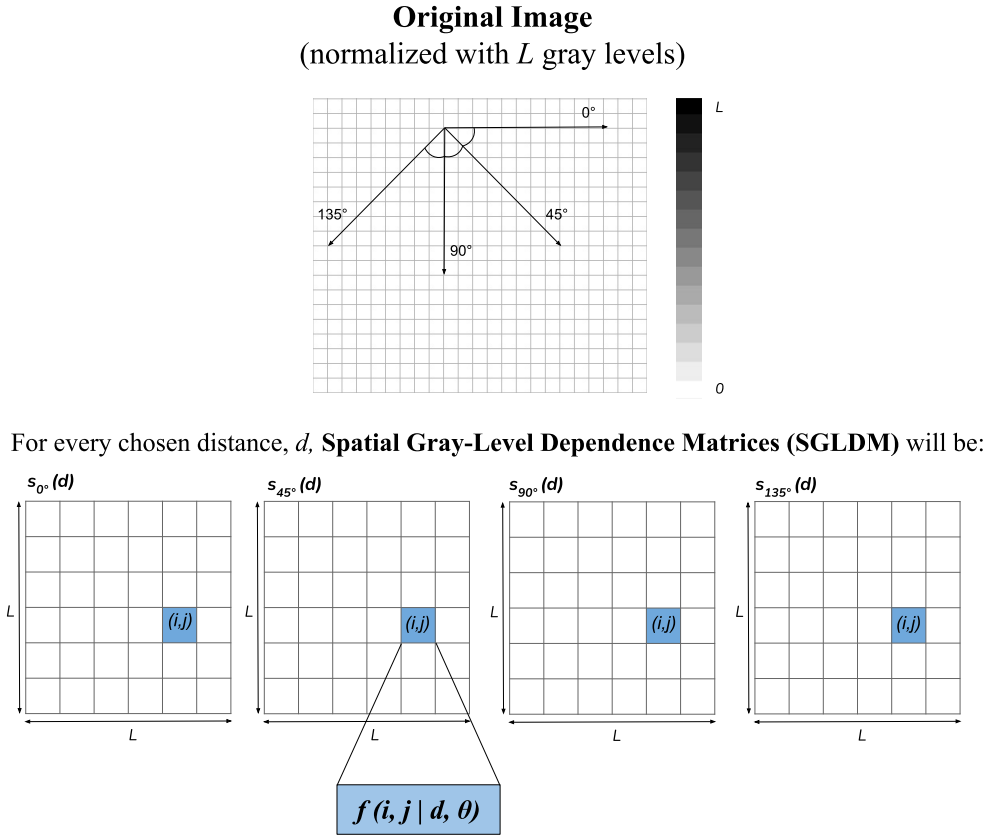

Figure 3.7: Diagram representing the final result of SGLDM. s_θ *(i,j | d)* is the SGLDM matrix for distance d and direction θ, and *f(i,j | d,θ)* represents the probability of a pixel with level i being at a distance d from a pixel with gray level j in the θ direction.

$$Correlation = \frac{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-\mu_x)(j-\mu_y)s_\theta(i,j \mid d)}{\sigma_x\sigma_y} \tag{3.20}$$

$$Local\ homogeneity = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{1}{1+(i-j)^2}s_\theta(i,j \mid d) \tag{3.21}$$

$$Contrast = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-j)^2 s_\theta(i,j \mid d) \tag{3.22}$$

where $s_\theta(i,j \mid d)$ is the (i,j) element of the SGLDM for distance d, and,

$$\mu_x = \sum_{i=0}^{L-1} i \sum_{j=0}^{L-1} s_\theta(i,j \mid d) \tag{3.23}$$

$$\mu_y = \sum_{i=0}^{L-1} j \sum_{j=0}^{L-1} s_\theta(i, j \mid d) \tag{3.24}$$

$$\sigma_x = \sum_{i=0}^{L-1} (i - \mu_x)^2 \sum_{j=0}^{L-1} s_\theta(i, j \mid d) \tag{3.25}$$

$$\sigma_y = \sum_{i=0}^{L-1} (j - \mu_y)^2 \sum_{j=0}^{L-1} s_\theta(i, j \mid d) \tag{3.26}$$

In this project, energy, entropy, correlation, local homogeneity and contrast were calculated from the SGLDM for each retinal layer, and later used as features. The angles used were 0^{o}, 45^{o}, 90^{o}, and 135^{o} and the distance 1 pixel. This resulted in 20 SGLDM features per layer. Since the calculation of a SGLDM requires an image, each B-scan was analysed individually, and then the mean of the property values in all the B-scans were calculated. The features were calculated for each layer, with the rest of the image set to 0 (background). Then, for the calculation of the SGLMD properties, the row and columns in the SGLDM corresponding to the intensity 0 were discarded, so that the background did not have any influence in any of the texture properties.

Frequency domain methods

The 2D Discrete Fourier transform (DFT) can also be used to retrieve information from OCT images. After calculating the 2D DFT, the image can be divided into a number of different regions, according to their frequency content, giving one different texture parameter for each region (Figure 3.8). The contribution of each different region to the total frequency magnitude in the image is calculated by summing all the values of the spatial frequencies in that region and dividing by the total frequency magnitude of the image. This percentage can be used as a feature, that represents the percentage of signal within a certain range of spatial frequencies, and be later used in tissue classification models.

Since the data used in this project are OCT volumes, a 3D DFT algorithm was used. In order to compute the features individually for each layer, the background was set to 0, leaving only the pixels corresponding to each layer. After, the 3D volume resulting from the DFT was divided into 4 regions, similarly to a 3D adaptation of Figure 3.8. Then, the contribution of each different region to the total frequency magnitude was calculated and later used as a feature.

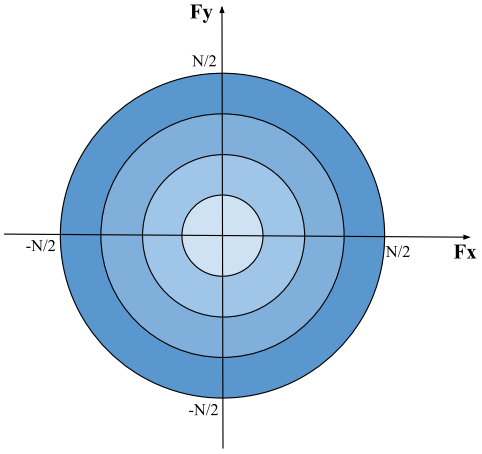

Figure 3.8: Diagram of the N × N 2D discrete Fourier transform space, divided into four regions.

3.3 Statistical analysis of extracted features

After the feature extraction and previously to the implementation of the disease progression model, a detailed statistical analysis of the retrieved features was conducted. This analysis is important not only to understand which features are the most discriminative for the problem, but also to attempt an interpretation and further understanding of OCT speckle.

Table 3.2 summarizes all the retrieved features from the speckle analysis methods (presented in Section 3.2.3), in a total of 470 speckle features (47 × 10 layers).

Additionally to the speckle features, the mean thickness was also calculated for each layer, resulting in a total of 480 features to study. The analysis of the speckle and thickness features were performed separately, for interpretation and comparison purposes.

For the implementation of all the statistical tests performed in this section, the Python library *Scipy* [125] was used. A level of significance of $\alpha = 0.01$ was applied throughout this section, unless stated otherwise. This indicates a 1% risk of concluding that a difference exists when there is no actual difference.

Shapiro-Wilk test

First, a Shapiro-Wilk test for normality was conducted, to assess whether the features followed a normal distribution.

The null hypothesis for this test is that the data is normally distributed [132]. From the 480 tested features (including both speckle and thickness features), 380 presented a *p-value* smaller than 0.01, meaning there is statistical evidence that they do not follow a normal distribution.

Table 3.2: Methods and number of speckle features per studied layer.

	Speckle Features		No. of features
Statistical properties	Mean Standard deviation Kurtosis Skewness		4
Statistical distributions	Rayleigh: 1 K: 3 Gamma: 2 Generalized Gamma: 3 Weibull: 2 Nakagami: 2 Rician: 2 Lognormal: 2		17
Contrast ratio	Mean Standard deviation		2
SGLDM	Energy Entropy Correlation Local homogeneity Contrast	distance = 1 $\theta = 0^{o}$, 45^{o}, 90^{o}, 135^{o}	20
Fourier Domain	Contribution of 4 regions to the total magnitude		4
TOTAL			47

For this reason, non-parametric statistical tests were used onward.

Kruskal-Wallis H test

Since the goal of this project is to study glaucoma disease progression, the Kruskal-Wallis test was used to evaluate which features presented the higher differences between the four severity groups (healthy controls, mild, moderate and severe glaucoma patients). This test was also used for the demographic analysis of the dataset, already presented in Section 3.1.

The Kruskal–Wallis H test, also known as one-way ANOVA on ranks, is a non-parametric statistical test for assessing whether two or more samples originate from the same distribution [132]. With a *p-value* lower than the significance level chosen, there is enough evidence to reject the null hypothesis and conclude that the group medians are not equal. However, it is important to note that the Kruskal-Wallis H test does not tell which specific groups are statistically significantly different from each other, only that at least two groups were different.

Wilcoxon rank-sum test

Glaucoma diagnosis and follow-up progression monitoring are two distinct tasks. As stated in Section 2.5, for the first, as there is no standard and determined tool, a combination of several clinical parameters is used, and the clinicians' interpretation of the parameters has a great influence in the diagnosis. For the latter, the visual field is the metric chosen for assessing severity and progression [94]. Even though OCT is already used for the assessment of some clinical parameters, speckle information has never been considered for neither diagnosis not progression study. Considering the groups in our dataset, a separate analysis was made for these two problems, first for assessing statistical significant differences between healthy and mild glaucoma patients, and between moderate *versus* severe glaucoma patients. For this purpose, the Wilcoxon rank-sum test was used.

The Wilcoxon rank-sum test, or Wilcoxon–Mann–Whitney test is a nonparametric statistical test used to verify whether two samples are likely to derive from the same population [132]. The aforementioned Kruskal-Wallis test is an extension of the Wilcoxon rank sum test for more than two samples. If the *p-value* of the Wilcoxon rank-sum test is less than the significance level ($\alpha = 0.01$), then the null hypothesis is rejected and there is statistical evidence that the two samples' distributions differ.

Feature pre-selection

Due to the high number of features retrieved from OCT speckle (470), it was necessary to perform some sort of feature pre-selection prior to the attempt of implementation of the model. For this purpose, the following two approaches were used. First, since there are 47 features retrieved from each layer, and several of them by very similar methods (*i.e.* 20 features per layer calculated from SGLDM), it is expected that they have a high correlation between them. For the implementation of the model, and considering computational and time optimization, it might not beneficial to include features that contain very similar information. For this reason, to lower model complexity without losing a significant amount of information, features were discarded based on their correlation with each other. If two features had a correlation higher than 0.99 (absolute value) between each other, only the one which presented a lower *p-value* of the correlation with VF MD was kept, while the other was discarded.

As stated in Section 2.6.2, the DEBM, the model used in this project for disease progression estimation, assumes monotonic increase or decrease of biomarker values with the increase of disease severity. The Kruskal Wallis and Wilcoxon tests only assess whether there are any

differences in the distributions of the different clinical groups, and do not verify whether the biomarker has a monotonic relation with disease progression, only analyses the groups medians. Since the main goal of this project is to study the progression of glaucoma, and the visual field mean deviation is the metric used by clinicians for clustering glaucoma patients into severity classes, the correlation between speckle features and VF MD was calculated and ultimately used as a method for feature pre-selection and "ranking method". It was used to assess which were the best features and rank them based on their *p-value*.

The feature pre-selection was performed combining the two aforementioned approaches, and establishing a threshold of 0.001 in the *p-value* of the correlation with the VF MD. This threshold was chosen to ensure that the number of features kept was not excessively high and unmanageable. However, a number of experiences with the model were performed to test and support all of these choices, which are explained in Section 3.4.1.

Spearman's correlation

The correlation metric used for the two previously mentioned approaches was the Spearman's correlation coefficient. This coefficient measures the strength and direction of association between the variables. It was chosen over Pearson's correlation, as the latter only evaluates for linear correlation. The Spearman's correlation also returns a *p-value*, which represents a measure of how probable it is that any observed correlation is due to chance [132]. It is also a non-parametric test, as it does not make any assumptions about the distribution of the data.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \ ; \tag{3.27}$$

where ρ is the Spearman's correlation coefficient, d_i is the difference between the ranks of corresponding variables and n is number of observations.

The results of this feature pre-selection phase are presented in Section 4.1.

3.4 Discriminative Event-Based Modelling

The discriminative event-based modeling (DEBM) [110, 111] is a method for estimating the central ordering of events in a degenerative disease. In this work, the model developed by Venkatraghavan et al [111] was used. The model consists of four main steps, summarized in Figure 3.9, and further explained afterwards.

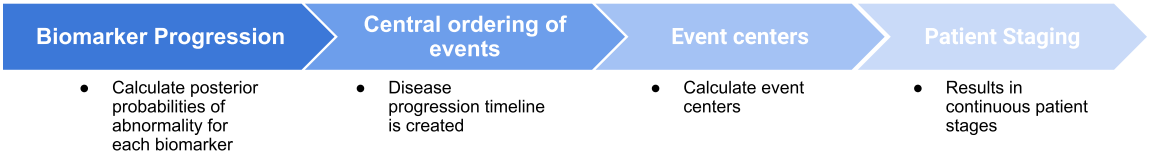

Figure 3.9: Overview of the four main steps in DEBM.

Biomarker Progression

In DEBM, an event, E_i, is characterized as one biomarker becoming abnormal, with i being a biomarker ($i \in [1..N]$), N being the number of biomarkers considered. In a cross-sectional dataset X of M subjects, X_j denotes a measurement of all biomarkers for subject j ($j \in [1..M]$), consisting of N biomarker values $x_{j,i}$. The posterior probability of a biomarker being abnormal, $p(E_i|x_{i,j})$, is assumed to represent the progression of a biomarker. This first step consists of a method to estimate this probabilities and later estimate the event order for each specific subject (s_j).

Age and gender were considered as confounding factors. Specially for the age, there are several studies that proved that age alone has an influence on the macular retinal thickness [133, 134]. For this reason the effects of these parameters were subsequently corrected for all biomarkers, before being used in the model.

The first step is to estimate the distributions of the pre-event and post-event classes for each biomarkers. For this purpose, the PDFs are assumed to be Gaussians. This estimation is not trivial and might be prone to bias, due to the fact that clinical labels might not always be correct, especially in diseases where the diagnostic is not always straightforward and is highly dependent on patient cooperation/clinicians opinion. A patient considered healthy might have early undetected disease, or the opposite case. Also, not all biomarkers will be normal for a healthy individual and not all biomarkers will be abnormal for a severe glaucoma patients, especially due to the high variability of neuropathologies like glaucoma.

An initial estimation of these PDFs is made using biomarkers from an "easily classifiable" subset of the data, and is later refined using the entire dataset. An "easily classifiable" subset is constructed as follows: first, a Bayesian classifier is trained for each biomarker, using only the two most extreme classes in the dataset (healthy and severe glaucoma); after, this classifier is applied to the training data (two extreme groups), and any missclassified cases are removed, resulting in two completely separable groups for each biomarker. The mean and standard deviation of the Gaussian distribution of both normal and abnormal classes is estimated for each biomarker based

on this subset. Naturally, these initial estimations are biased, especially the standard deviation, which is smaller than in reality. The PDFs initial estimation for each biomarker is then refined using a Gaussian mixture model (GMM) [135], a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters (each representing a label group), with an objective function for optimization, using the entire dataset. The objective optimization function used in the GMM for biomarker i is a summation of log-likelihoods, for all subjects:

$$C_i = \sum_{\forall j} log\big[\big(\theta_i \times p(x_{j,i}|E_{j,i})\big) + \big((1-\theta_i) \times p(x_{j,i}|\neg E_{j,i})\big)\big] \ , \tag{3.28}$$

where θ_i is the mixing parameter, which determines the proportion of abnormal biomarker data in the dataset and ranges between [0,1]. The mixing parameters (that control the proportion of abnormal biomarker data in the dataset) and the Gaussian PDF parameters (that control the mean and standard deviation of the distribution) are optimized alternately, until convergence.

The PDF obtained from the GMM is then used for classification of the biomarkers using a Bayesian classifier, where the final mixing parameters θ_i and $(1-\theta_i)$ are used as the prior probability of each class when estimating posterior probabilities for each biomarker. Since these posterior probabilities are assumed to be a measure of progression of a biomarker, sorting decreasingly the biomarkers based on their posterior probabilities results in an estimation of the event order for each specific subject, s_j.

Central order of events

After obtaining s_j independently for every subject, the heterogeneity in the disease is captured. The central event ordering S is the mean of the subject-specific estimates. To obtain this mean, a generalized Mallow's model, a distance-based probability model for distributions on permutations [136], is used.

The central ordering is calculated as the ordering that minimizes the sum of the distances to all subject-specific orders, s_j. The distance used between orderings was the a probabilistic version of Kendall's Tau distance. The Kendall's Tau distance between the central order, S and the subject specific order, s_j, is calculated as:

$$K(S, s_j) = \sum_{i=1}^{N-1} V_i(S, s_j) \ , \tag{3.29}$$

where $V_i(S, s_j)$ is the number of adjacent swaps needed so that event in position i is the same in

both orders. In the probabilistic version, $\hat{K}$, instead of giving each swap the same importance, each swap is penalized with a certain weight based on the proximity of the posterior probability of the two biomarkers being swapped. This way, it is accounted that not all the events are evenly distanced from each other. The optimum S is the one that minimizes the sum of all weighted Kendall distances of all s_j. This represents an estimation of the sequence in which biomarkers become abnormal during disease progression.

Figure 3.10 summarizes the steps described involved in this stage of DEBM, for calculating the central order of events from the biomakers' measurements of a cross-sectional dataset.

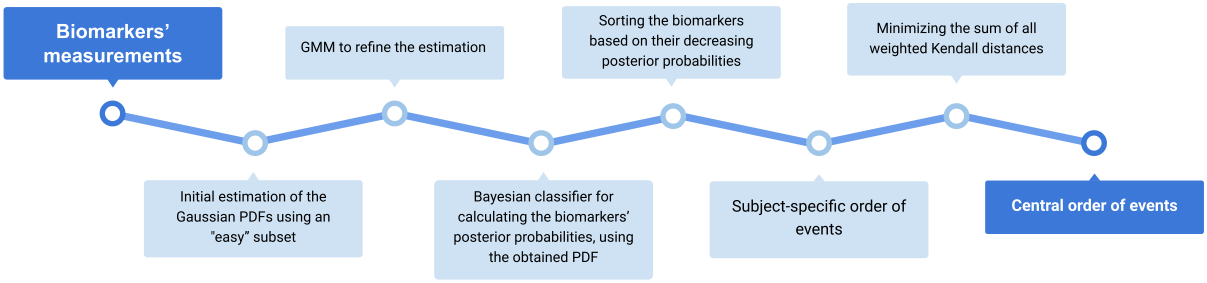

Figure 3.10: Detailed overview of the steps involved in the first stage of DEBM. Steps for obtaining central order of events from biomarkers' measurements.

Event centers

The previously calculated central order of events, S, is merely an order of biomarkers, and it does not represent the temporal distance between consecutive events. The distances between events are estimated by computing the cost of adjacent swaps in the event ordering, as the sum of probabilistic Kendall's Tau distance over all subjects:

$$\Gamma_{i+1,i} = \sum_{\forall j} \hat{K}(S_{i+1,i}, s_j) - \hat{K}(S, s_j) \ , \tag{3.30}$$

where $S_{i+1,i}, s_j$ is an event order identical to S except for events i and $i+1$ that are swapped and and $\Gamma_{i+1,i}$ is the cost of the swap. The closer two events are, the lower the cost of the swap is, so it is assumed that the cost is proportional to the proximity between two events.

Then, the event center (λ_k) of event k is calculated as follows:

$$\lambda_k = \sum_{i=0}^{k-1} \Gamma_{i+1,i} \tag{3.31}$$

The event centers are scaled to range from 0 to 1. We now have information about both the

order and the proximity between events.

Patient Staging

The final step of DEBM is to convert the estimated central order of events and events' centers of the disease into a staging measure for each patient. A patient stage (Υ_i) can be calculated as:

$$\Upsilon_i = \frac{\sum_{k=1}^{N} \lambda_k \; p(X_j|k,S) \; p(k,S)}{\sum_{k=1}^{N} p(X_j|k,S) \; p(k,S)} \tag{3.32}$$

where $p(k,S)$ is defined as:

$$p(k,S) = \frac{\prod_{i=1}^{k} \theta_{S(i)}^{E} \; \prod_{i=k+1}^{N} \theta_{S(i)}^{\neg E}}{Z} \tag{3.33}$$

with Z being a normalizing factor. This method results in continuous patient staging, from least to most severe. DEBM was implemented in this project using the Python toolbox *pyebm* [137].

3.4.1 Model's performance evaluation

The final output of the DEBM model is a set of continuous patient stages between 0 and 1 (0 being the least severe and 1 the most severe), as well as an ordering of biomarkers and respective event centers. Since the ground-truth event ordering is unknown, especially because speckle parameters were never considered in clinical practice yet, we resort to using the estimated patient staging to perform an indirect measure of the reliability of the event ordering.

As a way to measure the performance of the estimated order of events, following the methodology adopted by Venkatraghavan *et. al.* [111], the resultant patient staging was used for classification, and its resultant area under the receiver operating characteristics curve (AUC) was used as a performance metric, acting as an indirect way of evaluating the reliability of the calculated event ordering. The AUC is a performance metric used to evaluate classification models, calculated from a ROC, which shows the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds. In our case, for a set of continuous patient stages, the AUC can be interpreted as the probability of a randomly selected patient who has severe glaucoma having a higher predicted stage than a randomly selected patient who is healthy.

In a first stage, the AUC was computed for the classification of healthy *versus* severe cases of glaucoma (two extreme groups), for an assessment of how well the model could separate these two groups. The number of speckle features to include in the model was chosen to maximize this

metric. Then, for a more complete assessment of the model's performance and a more realistic idea of the DEBM's ability to model glaucoma disease progression, the AUC of all consecutive classes was also computed: healthy *versus* mild, mild *versus* moderate, and moderate *versus* severe. Other performance metrics were also computed at the optimal threshold: accuracy, precision, recall, and F1 score.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total} \tag{3.34}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{3.35}$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{3.36}$$

$$F1\text{-}score = 2 \times \frac{Precision + Recall}{Precision + Recall} \tag{3.37}$$

Additionally, since the correlation with the VF MD was the metric used to cluster glaucoma patients into disease severity groups (mild, moderate and severe), it acts as a measure of disease progression. For this reason, the Spearman's correlation between the patient staging performed by DEBM and the VF MD was also calculated, and used as a "global" performance metric. However, it is important to note that even though the VF MD was the parameter used to separate glaucoma patients into mild, moderate and severe groups, it was not the metric used to classify between healthy and glaucoma. This distinction, as stated in Section 3.1, was performed by a clinician based on a number of clinical factors. For this reason, healthy controls and mild glaucoma patients have overlapping VF MD distributions (Figure 3.2).

3.4.2 Validation of feature pre-selection and ranking techniques

Ideally, the more information given to the model, the better outcome should be achieved. However, this is not always the case, as adding less powerful or redundant features may decrease the performance, while increasing computational complexity and time.

In this section, two experiments were performed as a way to validate the decisions made in Section 3.3, on the speckle feature pre-selection stage, as well as to find the optimal subset of speckle features to include in the model. The performance metric used in this stage was the AUC of the classification of healthy *versus* severe glaucoma, the overall separation between the two

extreme groups. The decision to use this performance metric was based on the work performed by Venkatraghavan *et al.* [111]. At this stage, only speckle features were included in the model.

K-fold cross validation was used in the model evaluation. This resampling procedure splits the data in k groups arbitrarily, then leaves one out each time for testing, training the model with the other k-1 parts. A 5-fold cross validation was used, meaning the data was split into 5 parts and the model was trained five times, each time one of the folds was left out for testing. Cross-validation is used to return the performance metric as a mean and a standard deviation, offering a confidence interval to its true performance. With this approach, the model bias is reduced and a more robust estimation of how the model reacts to unseen data is obtained.

Effect of removing highly correlated features: To assess whether eliminating features with a high correlation between each other (more than 0.99) negatively affected the overall performance, the model performance was measured with an increasing number of biomarkers in two cases: with and without discarding features with high correlation between each other. The two plots and the maximum performance obtained were then compared.

Effect of pre-selecting features on the complete set: In the second step of the pre-selection phase, the entire dataset was used to perform a selection of the best features and rank them from most to least powerful, based on their correlation with the VF MD. To insure no high over-fitting occurred due to the fact that both the training and test sets were used in this stage, the performance was tested again with an increasing number of features, but this time performing the ranking of the features (with increasing *p-values* of the correlation with VF MD) using only the training set in each fold. The discarding of highly correlated features was performed in both cases. The two cases were then compared to assess whether there were any significant differences.

The logical and most accepted choice would be to perform the feature ranking and selection on the training set of each fold, to ensure that no overfitting or bias occurred by including the testing data in the feature selection stage. However, the main goal of DEBM is to provide an insight of the processes that occur during disease progression. By performing the feature selection individually in each fold, different sets of features would be used in each fold, hampering the model's interpretation and the choosing of the optimal subset of features to train the model with. This being said, using the same set of features in each fold is crucial to ensure proper interpretation, and to effectively perform the feature selection.

The results of these two experiences allow us to validate the feature pre-selection stage, and

also to evaluate the performance variance with the increase of the number of features used to train the model, and therefore choose the optimal number of features to keep.

3.4.3 Micro *versus* macrostructural parameters

In this project, two different types of features were retrieved from the retinal OCT volumes. The first, the features retrieved from the speckle analysis, due to the nature of speckle formation, provide information about the microstructure of the imaged tissue. The second, the measurements of the retinal layer thickness provide information on the macrostructure of the layers, their size. Although the thickness measures have been studied and used in clinical practice before, speckle information has not. In order to compare the power and potential of these two type of features, some comparisons were made.

After finding the optimal subset of speckle features, the next step was to assess whether speckle has any additional value to the macro features. For this purpose, a model was trained with the optimal number of speckle features chosen in the previous stage, and all the performance metrics detailed in Section 3.4.1 were computed. This model was then compared to a model using only macrostructural information. In order to assess potential added value of speckle features, the two types of parameters were combined. For this stage, a 5-fold cross validation was also used in every model trained.

3.4.4 Disease timeline variation analysis

DEBMs estimate the disease progression timeline as the order in which biomarkers become abnormal, and then use this ordering to infer patient staging. The purpose of these models is not only to allow the staging of patients into the disease severity, but also to provide a more insightful look and interpretation of events that occur during disease timeline. Because of this, the variation of the ordering computed by DEBM was studied using, bootstrapping [138]. Bootstrapping is a resampling technique with replacement, that creates a subset of the data by picking elements from the original dataset so that an element may appear multiple times in a specific sample. It allows to make estimations of the event centers with confidence levels, and to study the variation of the ordering of events over a high number of repetitions. Similarly to the previous analysis, the disease timeline was studied using only macro parameters, only micro parameters and the two combined.

Chapter 4

Results

The results presented in this section follow the same order and structure of the methods presented in Chapter 3. First, the statistical analysis of the extracted features is presented (Section 4.1), starting with the four severity groups, and followed by the tests comparing healthy *versus* mild and moderate *versus* severe cases, separately. A 99% confidence interval was used for all statistical tests presented. The results of the feature pre-selection are presented in this section as well. Next, results of the DEBM are shown (Section 4.2), starting with the two experiments meant to validate the feature pre-selection choices, and followed by the detailed results of the model's performance and the variance in the estimated disease timeline.

4.1 Statistical analysis of extracted features

4.1.1 Severity groups

The distributions of the four labeled groups were compared using the Kruskal-Wallis test. Of the 470 speckle features tested, 158 presented a *p-value* lower than 0.01. Figure 4.1 shows two speckle features (skewness of the RNFL and entropy of SGLDM calculated over angle 0° and distance 1 in the GCL), as an example of two of the features that presented the lowest p-values overall. The complete list of all the 158 features which presented statistical significant difference are included in the Appendix C, ordered by ascending *p-values*.

More important than the number of features which present statistical significant differences, is to understand which features are these, and how they were calculated. Table 4.1 shows the distribution of the features by layer and by method used.

The layers which present the features with the lowest *p-values* are RNFL and GCL. The most commonly selected speckle methods are SGLDM and statistical properties, but is highly

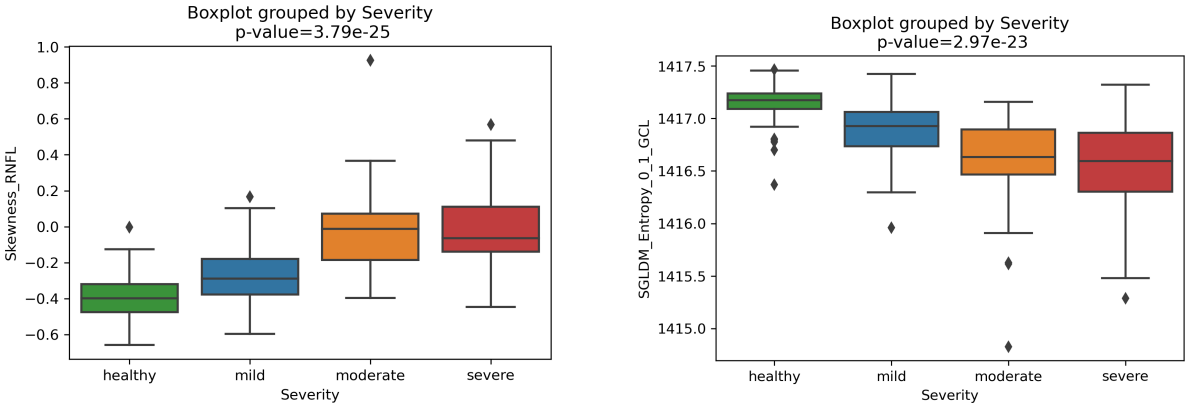

Figure 4.1: Box plot for the skewness of the RNFL (left) and entropy of SGLDM calculated over angle 0° and distance 1 in the GCL (right), with the correspondent *p-values* calculated with Kruskal Wallis H test.

variant between layers.

The same evaluation was also performed on the layers' thickness measurements. Out of the 10 analysed layers, the thickness of 3 of them presented statistically significant differences: RNFL, GCL, and IPL. Figure 4.2 shows the ordered *p-vales* of these layers, and it can be observed that all of them are well below the significance level $\alpha = 0.01$.

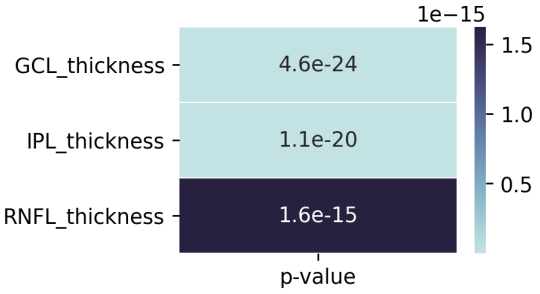

Figure 4.2: Layers' whose thickness measurements present statistically significant differences between groups, using the Kruskal-Wallis H test, ordered by ascending *p-values*.

4.1.2 Post-hoc analysis

As explained in Section 3.3, the Wilcoxon rank-sum test, was used to perform a pairwise comparison: healthy controls against mild glaucoma patients, and moderate against severe glaucoma patients.

Healthy *versus* Mild glaucoma

For the first case analysed, 50 features out of 470 presented a *p-value* less than 0.01. Figure 4.3 shows two examples of speckle features (entropy of SGLDM calculated over angle 0° and

Table 4.1: Distribution of the 158 features which presented a *p-value*<0.01 statistical significant differences between severity groups, using the Kruskal-Wallis test, by retinal layer and speckle analysis method. SP=statistical properties; SD=statistical distributions; CR=contrast ratio; SGLDM=spatial gray level dependence matrices; FD=frequency domain.

Retinal Layer	**No. of features *p-value*<0.01**	**Methods**	
RNFL	37	SP:4 SD:12 CR:2	SGLDM:16 FD:3
GCL	24	SP:2 SD:4	CR:2 SGLDM:16
IPL	30	SP:3 SD:7 CR:2	SGLDM:16 FD:2
INL	20	SP:1 SD:4 CR:1	SGLDM:12 FD:2
OPL	29	SP:3 SD:4 CR:2	SGLDM:18 FD:2
ONL	2	SGLDM:2	
IS/OS	5	SP:1 SD:2	FD:2
OSL	3	SGLDM:1	FD:2
OPR	3	SP:1	FD:2
RPE	5	SP:1 SD:1	SGLDM:1 FD:2

distance 1 in the GCL, and the mean of the contrast map in the RNFL) two of the features that presented the lowest *p-values* overall. The list of all the 50 features which presented statistical significant differences for this experiment are included in the Appendix C, ordered by ascending *p-values*.

Following the same approach as before, Table 4.2 shows the distribution of these features by layer and by method used. Features retrieved from the INL, OPL, and ONL did not present any statistically significant difference in any of the features studied.

Figure 4.4 shows the results for the layers' thickness performing the same analysis. The thickness of 3 layers presented statistically significant differences between healthy and mild glaucoma: RNFL, GCL and IPL.

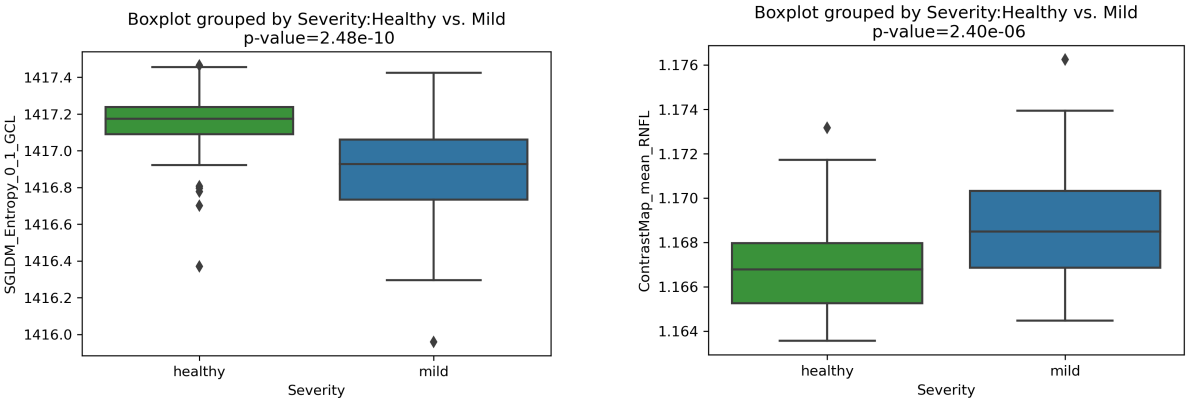

Figure 4.3: Box plot of entropy of SGLDM calculated over angle 0° and distance 1 in the GCL (left), and the mean of the contrast map in the RNFL (right), with correspondent *p-values* calculated with Wilcoxon rank-sum test, for healthy *versus* mild.

Table 4.2: Distribution of the 50 features which show statistical significant differences between the healthy controls and mild glaucoma patients, using the Wilcoxon rank-sum test, by retinal layer and speckle analysis method. SP=statistical properties; SD=statistical distributions; CR=contrast ratio; SGLDM=spatial gray level dependence matrices; FD=frequency domain.

Retinal Layer	**No. of features** ***p-value<0.01***	**Methods**
RNFL	25	SP:3 SD:8 CR:2 SGLDM:12
GCL	10	SP:1 SGLDM:9
IPL	8	SGLDM:8
IS/OS	3	SP:1 SD:2
OSL	1	SD:1
OPR	1	SD:1
RPE	2	SGLDM:1 FD:1

Moderate *versus* Severe glaucoma

When comparing the moderate against severe glaucoma patients, only 5 out of the 470 studied features presented a statistically significant difference between the two groups. Figure 4.5 shows the boxplot of two of these features (d-parameter of the Generalized Gamma distribution in INL and contrast of SGLDM calculated over distance 1 and angle 0°in IPL), the two features

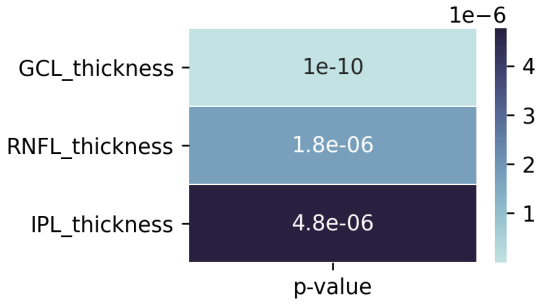

Figure 4.4: Layers whose thickness measurements presented a statistically significant difference between healthy controls and mild glaucoma patients, using the the Wilcoxon rank-sum test, ordered by ascending *p-values*.

with the lowest *p-values*. The list of all the 5 features with *p-values* lower than 0.01 are included in the Appendix C, ordered by ascending *p-values*. The distribution of features by layer and method is presented in Table 4.3. The layers not presented in the table (RNFL, GCL, ONL, IS/OS, OPR, RPE) did not present any statistically significant difference for any of the studied features.

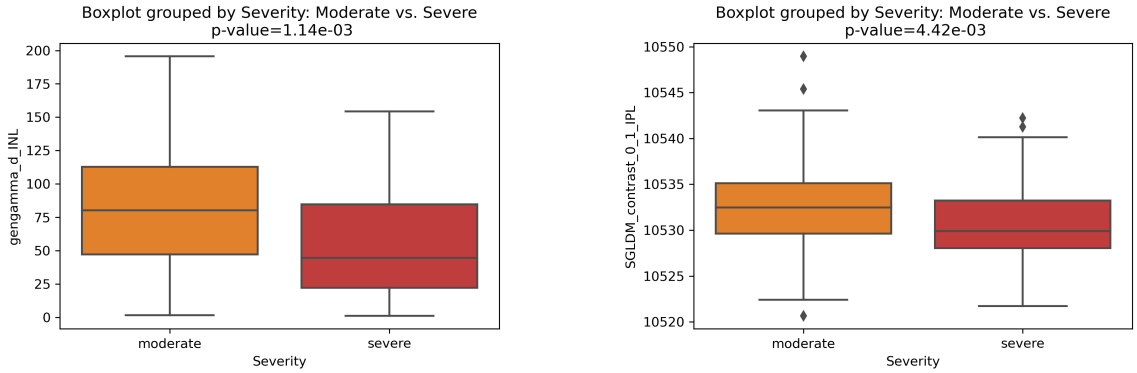

Figure 4.5: Box plot of *d*-parameter of the Generalized Gamma distribution in the INL (left) and contrast of SGLDM calculated over distance 1 and angle 0° in IPL (right), with correspondent *p-values* calculated with Wilcoxon rank-sum test, for moderate *versus* severe glaucoma.

Analysing the layers' thickness, 3 layers showed a statistically significant difference between moderate and severe patients: IPL, OSL and OPL.

The combined results of the three presented statistical tests already allow us to draw some important conclusions, and get an insight of the potential of speckle in glaucoma diagnosis and severity assessment. Interpretation of these results are presented in the Discussion chapter.

Table 4.3: Distribution of the 5 features which showed statistical significant differences between moderate and severe glaucoma patients, using the Wilcoxon rank-sum test, by retinal layer and speckle analysis method. SP=statistical properties; SD=statistical distributions; CR=contrast ratio; SGLDM=spatial gray level dependence matrices; FD=frequency domain.

Retinal Layer	**No. of features *p-value<0.01***	**Methods**
IPL	1	SGLDM:1
INL	2	SD:1 SGLDM:1
OPL	1	SGLDM:1
OSL	1	SGLDM:1

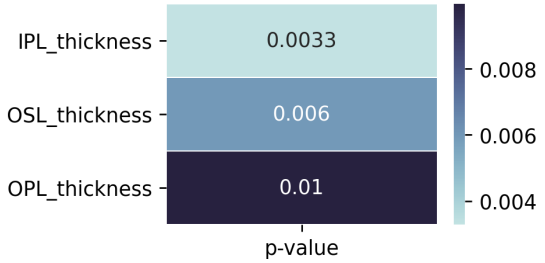

Figure 4.6: Layers whose thickness measurements present statistically significant differences between moderate and severe glaucoma patients, using the Kruskal-Wallis H test, ordered by ascending *p-values*.

4.1.3 Feature pre-selection and ranking

The feature pre-selection was performed combining two steps, as explained in Section 3.3. For the first step, the correlation between features was calculated. Figure 4.7 shows the heatmap matrix for the Spearman's correlation between speckle features from the RNFL, as an example. The heatmaps for the remaining layers are included in Appendix C. For easier visual understanding, the absolute values of the correlation were taken for the plotting the heatmaps.

By analysing the heatmaps, is possible to see that there are a large number of features with high correlation values (whether positive or negative) with each other. For example, features as the energy calculated over the SGLDM with different angles, have a high correlation between them (close to 1), but other combinations present very low correlations (close to 0). With this pre-selection approach, the total number of features was reduced from 470 to 418, without significant loss of information. In Section 4.2, the effect of this method in the model's performance is analysed.

In the second step, the correlation with visual field was used to rank the features from the

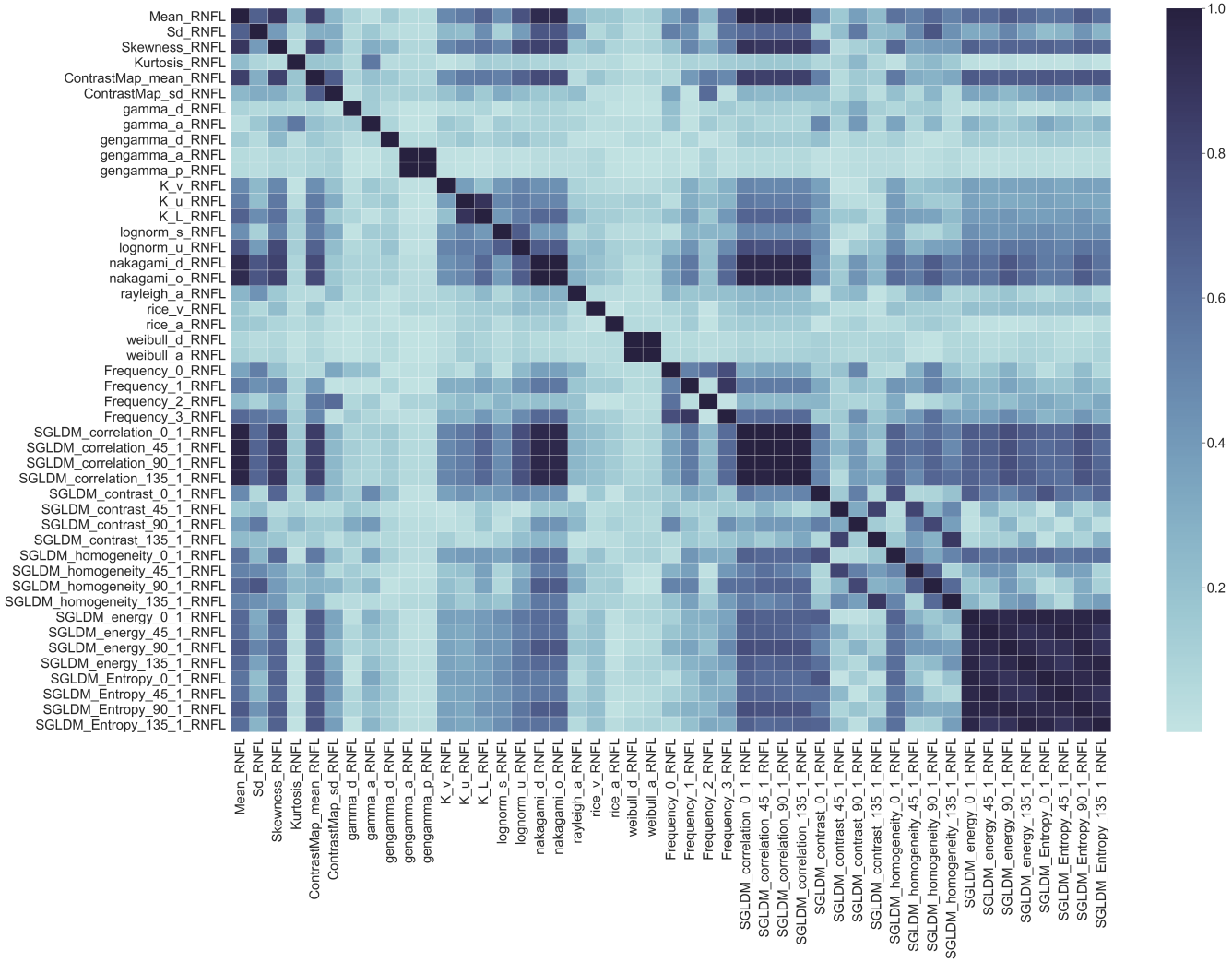

Figure 4.7: Heatmap of the absolute value of Spearman's correlation values between features for the RNFL.

most to the least correlated, based on their *p-values*. Features with a correlation with a *p-value* higher than 0.001 were discarded, to limit the number of features. After the application of this second step, a subset of 123 features was kept, presented in Figure 4.8.

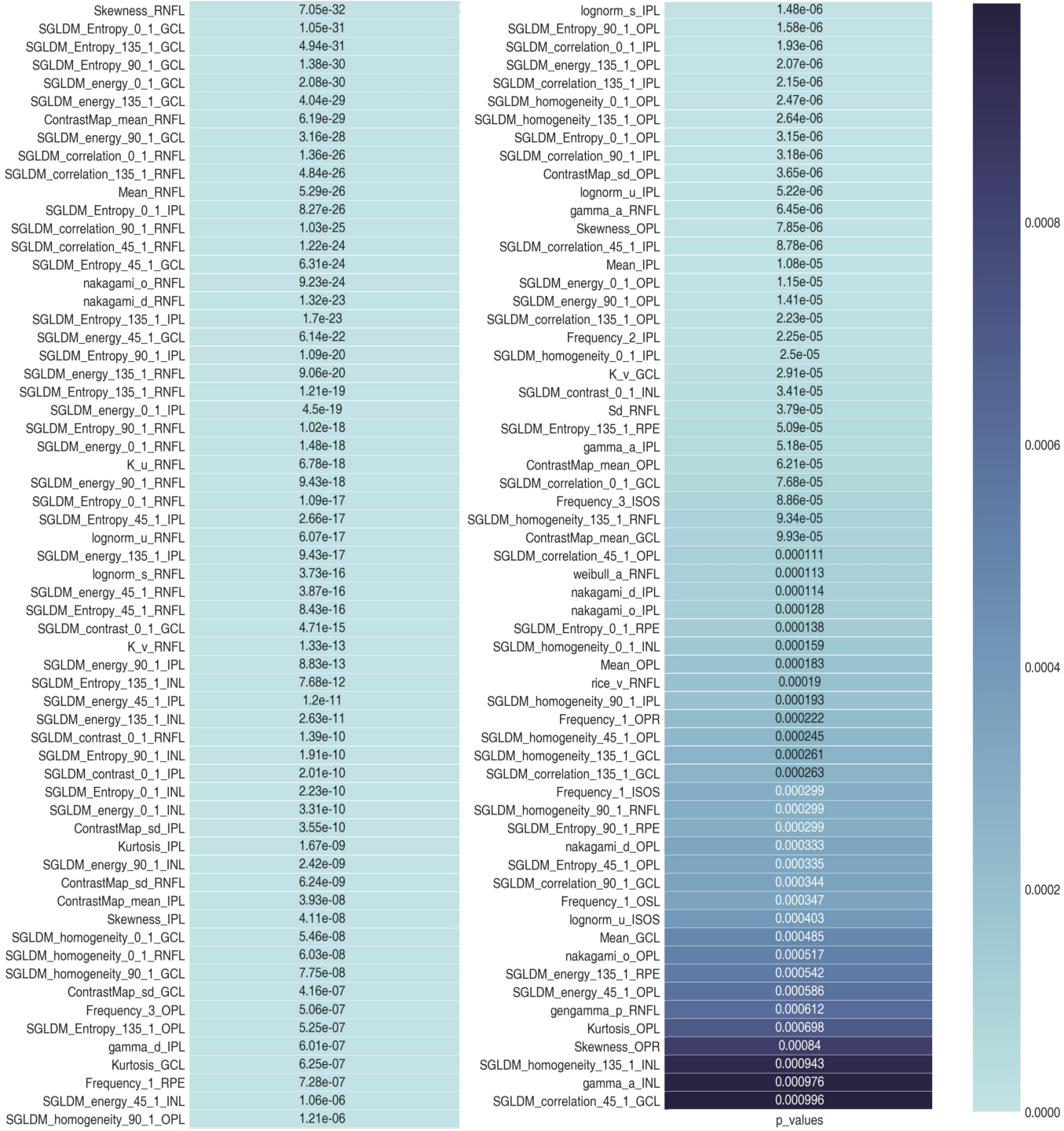

Figure 4.8: List of the 123 retained features after applying the feature selection. The features are ranked from most to less correlated with VF MD, with respective *p-values* for this correlation.

4.2 Discriminative Event-Based Modelling

4.2.1 Validation of feature pre-selection and ranking techniques

In the first experiments made with DEBM, only speckle features were included. To validate the effectiveness of the feature pre-selection approaches and to assess whether they had any effect on the model's performance, some tests were performed. The performance metric of the DEBM considered was the AUC of the classification of healthy *versus* severe glaucoma patients performed using the estimated event ordering. The order of the features to be included was the ranking based on increasing *p-values* of the correlation with the visual field (Figure 4.8).

Effect of removing highly correlated features

On the feature pre-selection stage, 123 features were kept (Figure 4.8). In order to study the effect of discarding highly correlated features, the model's performance was studied with and without following this approach (using a threshold of 0.001 on the correlation with the visual field on both cases). This leads to 123 and 138 features selected, respectively. The performance of the model with increasing number of features in these two cases is depicted in Figure 4.9.

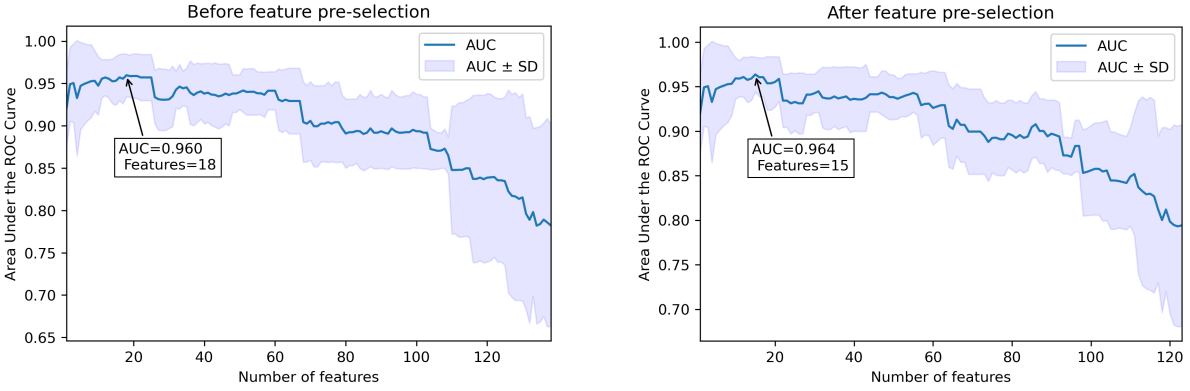

Figure 4.9: AUC as a function of the number of features used to train the DEBM before (left) and after (right) performing the feature selection based on the correlation step. A 5-fold cross validation was used to compute AUC mean and standard deviation.

Both plots are fairly similar in shape and values. Also, is is possible to observe that the maximum AUC is very similar for both plots, even slightly higher in the plot using the correlation pre-selection (0.960 versus 0.964), but with a lower number of features used (18 versus 15). These results align with what was expected, because even though 15 features in a total of 138 were eliminated, they can be considered redundant, as no significant information was lost. This experiment also allow us to find the number of features which optimizes the performance metric

chosen, which is 15. Moreover, after a certain point, around 100 features when not discarding highly correlated features and around 90 features when discarding, both plots show a clear downward trend. Because of this, we can assume that the chosen threshold for the *p-value* was an adequate choice, as choosing a higher *p-value* (and consequently including less powerful features in study) would not be beneficial in terms the model's performance.

Effect of pre-selecting features on the complete set

Performing the feature pre-selection only on the training subset of each fold would ensure that no bias is induced, as no information from the test sets is transferred to the model. However, using the entire dataset for this purpose yields a unique list of features for every fold, instead of one for each cross-validation fold, being more convenient for finding the optimal subset of features, and for the interpretation of the estimated disease timeline. To confirm that performing the feature ranking in the entire dataset was not highly biasing the model, a comparison between two cases was conducted: performing the feature ranking and selection on each training set, and using the feature ranking presented on Figure 4.8 (performed on the entire dataset). The two steps of the feature pre-selection were applied for both cases in this experiment.

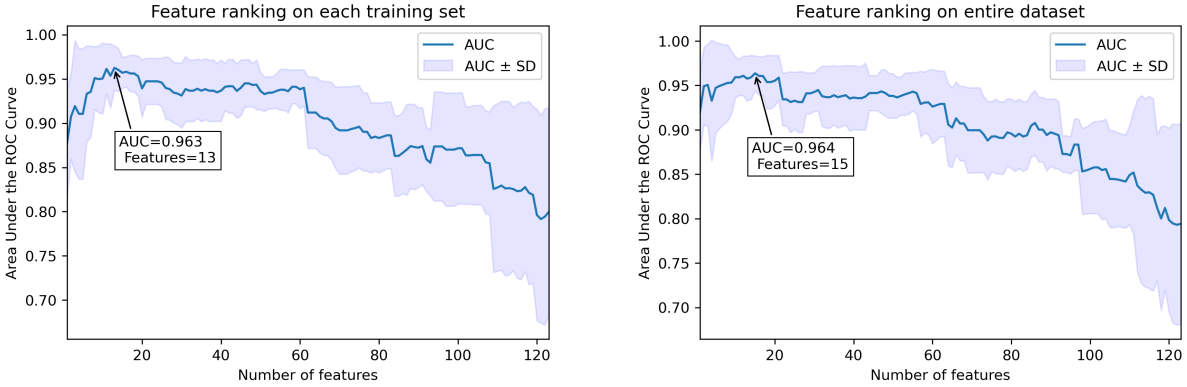

Figure 4.10: AUC as a function of the number of features used to train the DEBM, performing the feature pre-selection individually in each training set (left), and on the entire dataset (right). A 5-fold cross validation was used to compute AUC mean and standard deviation.

By visually analysing the results presented in Figure 4.10, we can conclude that the AUC plots are very similar, both in shape and values, indicating that performing the feature pre-selection before-hand using the entire dataset did not introduce a high bias in the model. The maximum performance is also similar (0.963 and 0.964) and is obtained from a comparable number of features (15/13), indicating that the variations are only due to normal fluctuations in choosing the features. The most significant difference between the two plots is that the standard

deviation is, generally, slightly higher when performing the feature selection only in the training set, especially when using higher number of features. This can be explained by the variability in the features used in the different folds.

4.2.2 Micro *versus* macrostructural parameters

The model's performance was evaluated using either micro (speckle features) or macrostructural parameters (thickness features), and also combining both set of parameters. In this stage, the AUC's of all consecutive disease progression classes were computed: healthy *versus* mild, mild *versus* moderate, and moderate *versus* severe, as well as the overall separation between healthy and severe. Additionally, the accuracy, recall, precision and F1-score of the classification at the best threshold are presented as well for a wider view of the model's performance.

Speckle features

Table 4.5 presents the performance metrics for DEBM trained with the first 15 speckle features present in Figure 4.8 (number chosen based on the previous experiments). All values are presented as mean ± SD, computed over 5-fold cross-validation.

Table 4.4: Performance metrics for DEBM using the 15 best speckle features. All metrics are presented as Mean ± SD.

	AUC	**Accuracy**	**Recall**	**Precision**	**F1-score**	**Correlation VF MD**
Healthy vs. Mild	0.715 ± 0.075	0.648 ± 0.092	0.493 ± 0.182	0.893 ± 0.066	0.614 ± 0.163	-0.662 ± 0.065
Mild vs. Moderate	0.829 ± 0.080	0.767 ± 0.050	0.791 ± 0.147	0.641 ± 0.075	0.693 ± 0.059	
Moderate vs. Severe	0.527 ± 0.139	0.593 ± 0.069	0.462 ± 0.293	0.564 ± 0.288	0.484 ± 0.276	
Healthy vs. Severe	0.964 ± 0.020	0.873 ± 0.019	0.856 ± 0.057	0.899 ± 0.087	0.871 ± 0.016	

Figure 4.11 shows the distribution of the computed patient staging output by each severity group, when training the DEBM with the entire dataset. A KDE was fitted to the distribution of each group, as a way to estimate their PDF and visualize how distinct they are from each other.

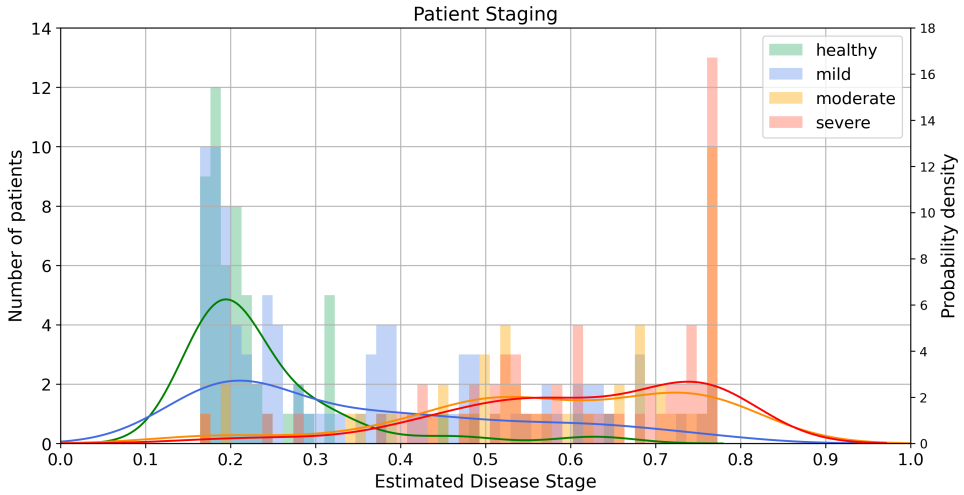

Figure 4.11: Distribution of patient stages with a KDE estimation for each group, when the 15 most significant speckle features were given as input to the DEBM, using the entire dataset for training and staging.

Thickness features

Next, the performance using only the macrostructure information was computed. For this, the thickness of the 3 layers that showed statistically significant differences between the four severity groups (Figure 4.2) were used: RNFL, GCL and IPL. The results of this model are presented in Table 4.6 and Figure 4.14.

Table 4.5: Performance metrics for DEBM using the thickness from RNFL, GCL and IPL. All metrics are presented as Mean ± SD.

	AUC	**Accuracy**	**Recall**	**Precision**	**F1-score**	**Correlation VF MD**
Healthy vs. Mild	0.821 ± 0.068	0.771 ± 0.055	0.730 ± 0.075	0.896 ± 0.088	0.799 ± 0.047	-0.689 ± 0.087
Mild vs. Moderate	0.778 ± 0.112	0.765 ± 0.094	0.674 ± 0.156	0.668 ± 0.176	0.661 ± 0.137	
Moderate vs. Severe	0.579 ± 0.130	0.624 ± 0.085	0.524 ± 0.170	0.762 ± 0.195	0.592 ± 0.129	
Healthy vs. Severe	0.981 ± 0.025	0.923 ± 0.034	0.863 ± 0.062	0.983 ± 0.033	0.917 ± 0.03	

From these results we can already assess that micro and macrostructural features have similar performances overall, considering the confidence intervals of most metrics. Nevertheless, the macrostructural model achieves a slightly higher correlation with the VF MD and overall

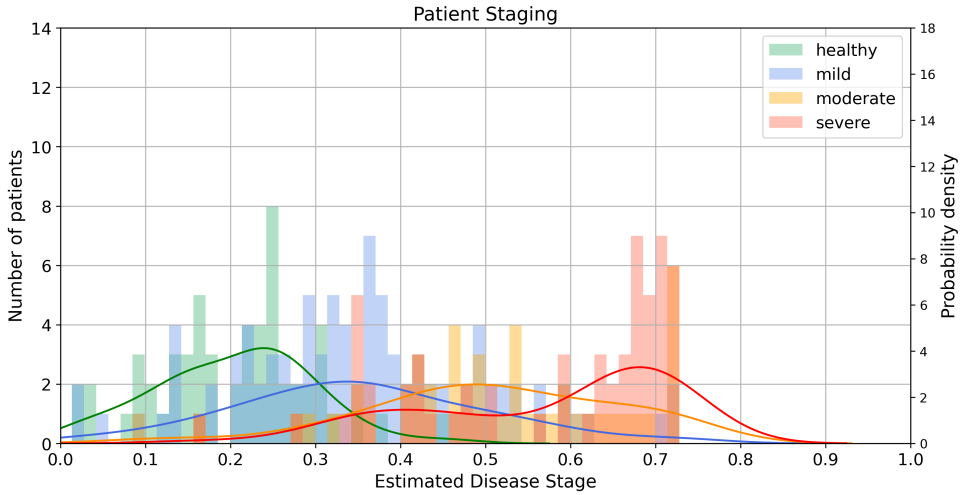

Figure 4.12: Distribution of patient stages with a KDE estimation for each group, when the thickness of RNFL, GCL and IPL were given as input to DEBM, using the entire dataset for training and staging.

separation between healthy and severe glaucoma.

Combining speckle and thickness features

Since the thickness features seem to perform slightly better, the next experiment made was to assess whether adding speckle features to the macrostructural features would bring an additional value to the model. The performance of the model was studied with an increasing number of speckle features added to the three significant thickness features. Figure 4.13 shows the AUC of the healthy *versus* severe classification with an increasing number of speckle features added (using the same order presented in Figure 4.8). The plot starts at 0 (only the thickness features) and shows an increasing number of features, up to 70, as the previous experiments showed that increasing this number further only results in a performance decrease.

As it can be seen from Figure 4.13, the overall higher AUC was obtained for only one speckle feature added, the *skewness of the RNFL*. However, it is important to note that the performance has very little variations until around 20 speckle features added. Table 4.7 presents the remaining performance metrics for this case, also using a 5-fold cross validation to compute mean and standard deviation for each metric. Figure 4.14 shows the patient staging for the entire population in study.

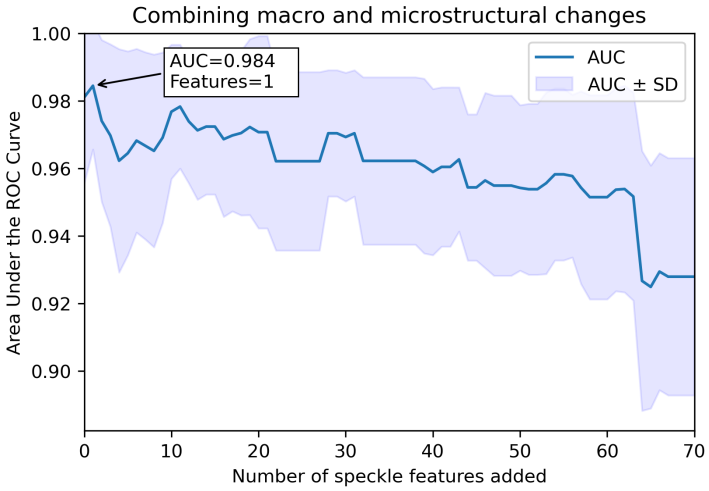

Figure 4.13: AUC as a function of the number of speckle features added to the thickness features to train the DEBM.

Table 4.6: Performance metrics for DEBM using the thickness feature combined with the Skewness of RNFL. All metrics are presented as Mean ± SD.

	AUC	**Accuracy**	**Recall**	**Precision**	**F1-score**	**Correlation VF MD**
Healthy vs. Mild	0.829 ± 0.075	0.771 ± 0.055	0.721 ± 0.077	0.906 ± 0.068	0.798 ± 0.042	-0.723 ± 0.066
Mild vs. Moderate	0.821 ± 0.112	0.772 ± 0.070	0.731 ± 0.196	0.657 ± 0.094	0.676 ± 0.111	
Moderate vs. Severe	0.580 ± 0.097	0.599 ± 0.064	0.503 ± 0.152	0.704 ± 0.117	0.567 ± 0.117	
Healthy vs. Severe	0.984 ± 0.018	0.932 ± 0.020	0.897 ± 0.032	0.965 ± 0.043	0.929 ± 0.021	

Moderate *versus* severe optimization

The overall results from the combined features are slightly better than the previous results, indicating that the speckle information from the RNFL might be beneficial when combined with the macrostructural parameters. However, event though the model presented an overall good performance, the differentiation between the two most advanced stages (moderate and severe glaucoma) presented a lower performance than the remaining cases. The results shown in Section 4.1, depicted in Figures 4.6 and C.3 show that there are still some features that show statistically significant differences between the two groups, computed in the outer retinal layers. Nevertheless, these features are not among the most significant to the overall problem

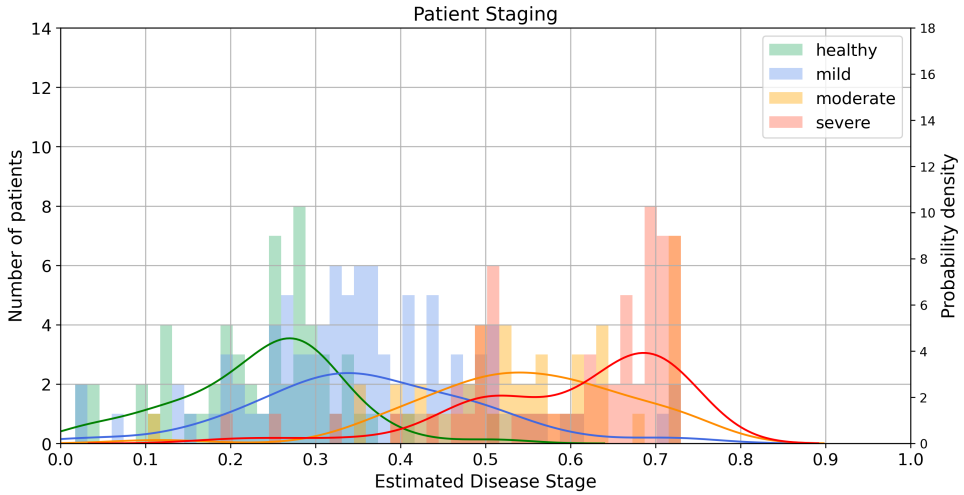

Figure 4.14: Distribution of patient stages with a KDE estimation for each group, when the thickness of the RNFL, GCL, and IPL, and the *skewness of the RNFL* were given as input to the DEBM, using the entire dataset for training and staging.

(to distinguish among the four classes) and they present a lower correlation with the VF MD than most features from the inner retinal layers. The only exception are the features computed on the IPL, that proved to remain significant for all disease severity stages, even in the more advanced stages.

The best feature overall (considering thickness and speckle features) was the d parameter of the Generalized Gamma distribution for the INL. To verify if this feature could represent an added value for the staging estimation, it was added to the previous model. The performance of this new model is presented in Table 4.7, and the distribution of the patient stages by label is depicted in Figure 4.15.

The performance for the differentiation between moderate and severe glaucoma increased slightly, but at the cost of a worse performance on the other cases (where the chosen feature was not as significant). Moreover, the improvement in the moderate *versus* severe case by adding this parameter was very small, as the variance intervals for all the performance metrics overlap. Also, it can be seen in Figure 4.15 that the predicted stages for moderate and severe patients have a strong overlap, indicating that the model is still not capable of successfully differentiating between these two classes.

Table 4.7: Performance metrics for DEBM using the optimal model (thickness features + Skewness of RNFL) and adding the d parameter of the generalized gamma of INL. All metrics are presented as Mean ± SD.

	AUC	Accuracy	Recall	Precision	F1-score	Correlation VF MD
Healthy vs. Mild	0.757 ± 0.081	0.734 ± 0.052	0.650 ± 0.137	0.892 ± 0.047	0.742 ± 0.094	-0.630 ± 0.074
Mild vs. Moderate	0.727 ± 0.134	0.709 ± 0.059	0.678 ± 0.264	0.576 ± 0.066	0.579 ± 0.155	
Moderate vs. Severe	0.595 ± 0.143	0.617 ± 0.089	0.539 ± 0.223	0.698 ± 0.071	0.583 ± 0.147	
Healthy vs. Severe	0.938 ± 0.045	0.882 ± 0.047	0.831 ± 0.100	0.941 ± 0.089	0.875 ± 0.054	

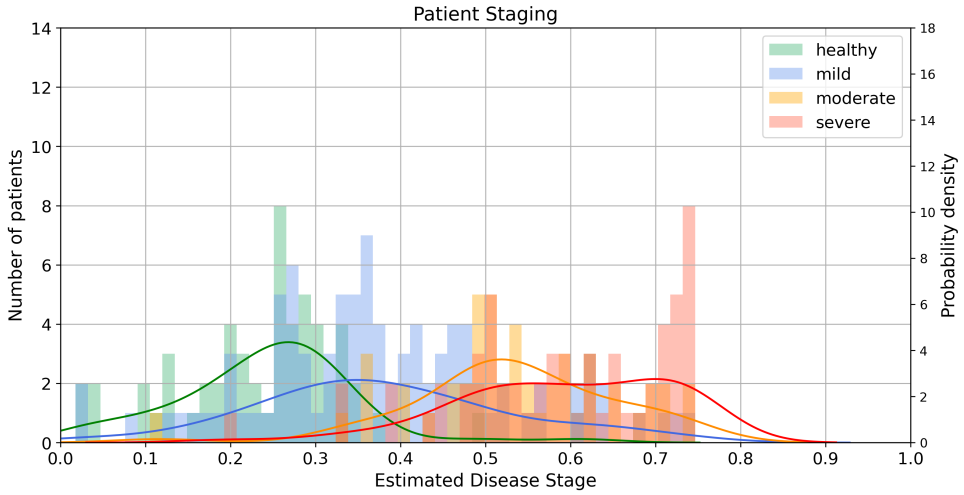

Figure 4.15: Distribution of patient stages with a KDE estimation for each group, when the thickness of the RNFL, GCL, and IPL, the *skewness of the RNFL*, and the *d parameter of the generalized gamma distribution from INL*, were given as input to the DEBM, using the entire dataset for training and staging.

4.2.3 Disease timeline variation analysis

The main goal of disease progression modelling is not only to stage patients into a disease progression group, but to successfully be able to model the disease timeline, and understand which biomarkers are more important in each stage of the disease. For this purpose, and to assess if this timeline remains constant while running the DEBM multiple times, a bootstrapping technique was used.

Thickness features

First, the timeline was studied considering only the thickness features which presented statistically significant differences between groups (Figure 4.2). The positional variance of the obtained disease timeline was studied over 100 bootstraps, in order to evaluate the uncertainty of the central event ordering estimation.

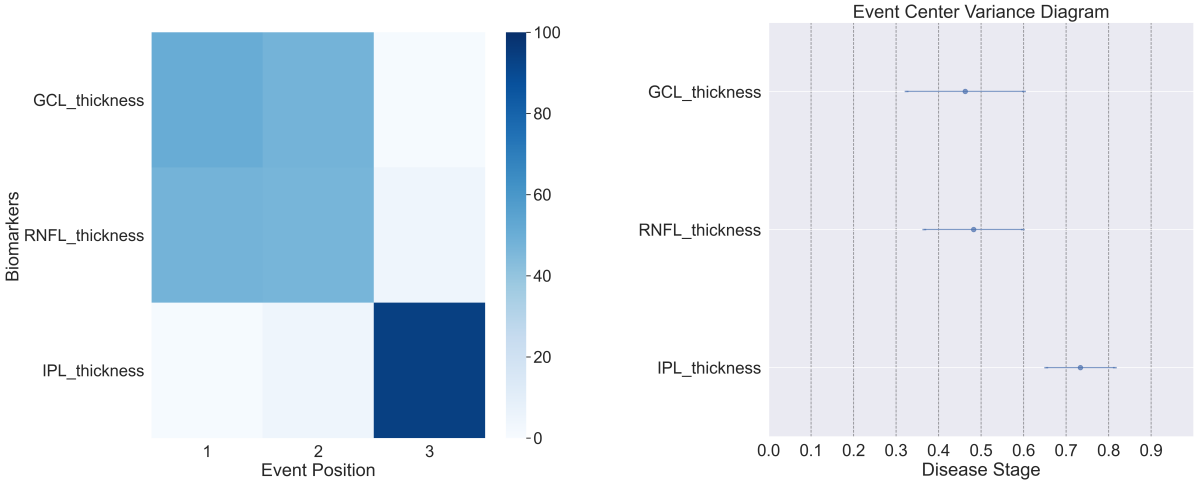

Figure 4.16: Positional variance diagram (left) shows the uncertainty in estimating the central event ordering and event center variance diagram (right) shows the standard error of estimated event-centers, for DEBM trained with thickness from RNFL, GCL and IPL. These variations were measured by 100 repetitions of bootstrapping.

Speckle features

Next, the same experiment was repeated using only the speckle features, with the optimal chosen number: 15. The results are presented in Figures 4.16 and 4.17.

By analysing both the positional variance diagram and the event center variance diagram of the DEBM trained with thickness features, is possible to see that, during the 100 bootstrap iterations, the thickness of the IPL was consistently placed in the last position, in nearly all the iterations. The RNFL and GCL, on the other hand, presented a larger variation in the estimation of the ordering, appearing almost the same number of times in the first and second position. Regarding the model trained with speckle features, the results are similar. Features from the GCL and RNFL appear in the first positions, with close event centers, and the two features from the IPL appear in the last two positions. However, the variance in the estimation of the event centers is much higher than in the model trained with the thickness.

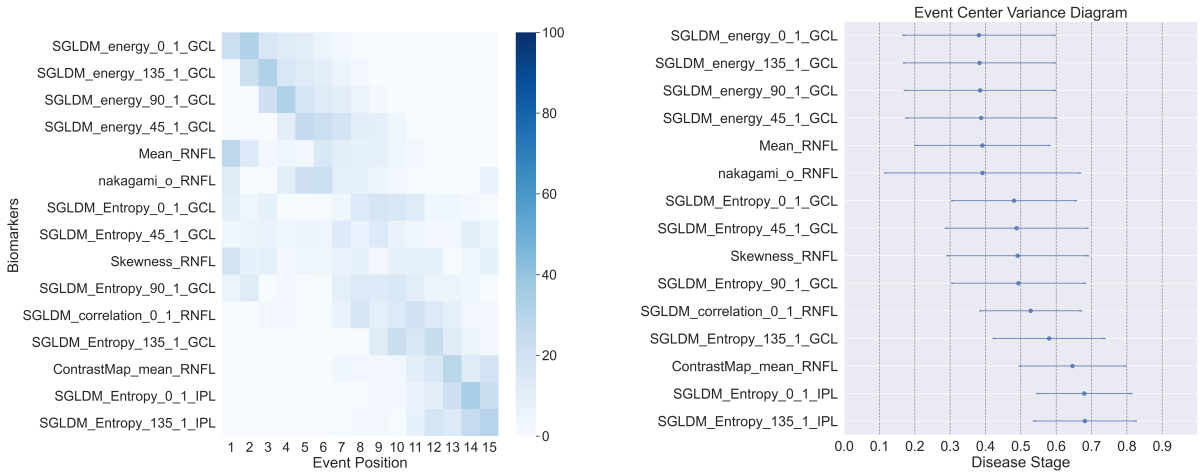

Figure 4.17: Positional variance diagram (left) that shows the uncertainty in estimating the central event ordering, and event center variance diagram (right) that shows the standard error of estimated event-centers, for DEBM trained with the 15 best speckle features. These variations were measured by 100 repetitions of bootstrapping.

Combining speckle and thickness features

The next experiment done was trying to combine both speckle and thickness features. Although the best performing model was obtained for only one speckle features added, from Figure 4.13 the performance of the model remained fairly constant between 0 and 20 speckle features added to the thickness features. For this reason, the same 15 speckle features used in the previous model were kept, for an easier interpretation and comparison between models.

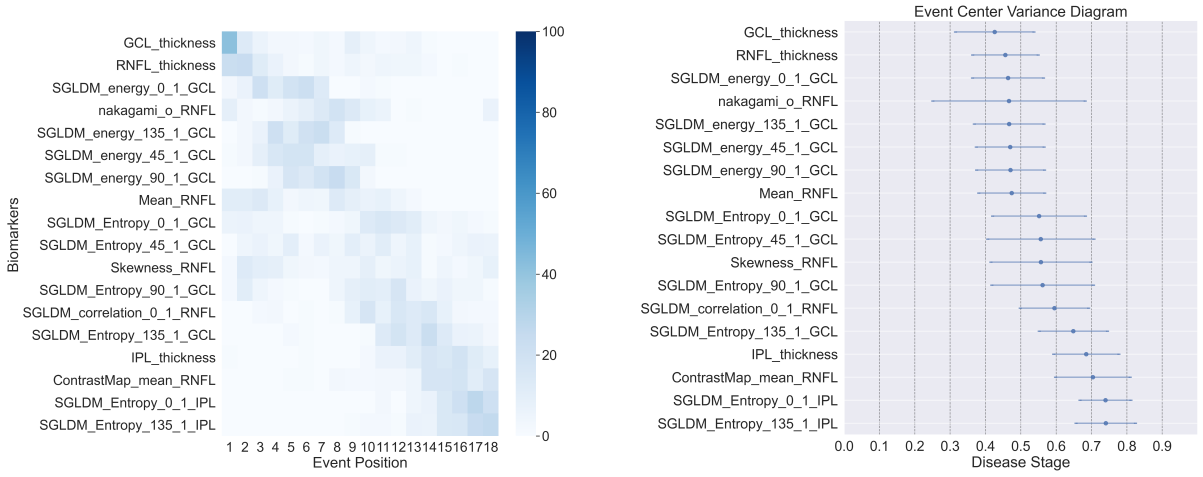

Figure 4.18: Positional variance diagram (left) shows the uncertainty in estimating the central event ordering and event center variance diagram (right) shows the standard error of estimated event-centers, for DEBM trained with thickness from RNFL, GCL and IPL, combined with the 15 best speckle features. These variations were measured by 100 repetitions of bootstrapping.

Chapter 5

Discussion

This chapter constitutes of a critical analysis of the main findings of this project, as well as their biomedical interpretation. When possible, results are related to the current state of the art.

This chapter is organized in two sections. In the first section, the results of the statistical analysis of both speckle and thickness features are discussed and interpreted. The second section, comprises the results obtained for the DEBM and respective analysis, as well as the reasoning behind the estimated disease timeline.

5.1 Statistical analysis of extracted features

In this project, several methods for the analysis of OCT speckle were applied to data from the retina. Since analysis of signal-carrying speckle is a relatively new and unexplored research topic in OCT imaging, and the physical meaning of speckle information is not yet fully understood [139], a set of different statistical tests were applied. These tests were meant to assess whether speckle information has the potential to be included into glaucoma research, and eventually into clinical practice. Additionally, they provide an insight of the role of each retinal layer in different disease stages, as well as the effectiveness of each speckle analysis method applied.

The overall results for the statistical tests on speckle features, presented in section 4.1, show that they were capable of assessing differences between the four stages of glaucoma, confirming the existence of a signal-carrying component in OCT speckle. Since speckle in OCT can be linked to microstructural properties of the imaged tissue [97], we can further conclude that glaucomatous damage has an influence on the microstructures of the retina, and their study and clinical interpretation may lead to a deeper understanding of the development of the disease.

Glaucoma severity groups

When comparing the four severity levels, several speckle parameters have shown to be statistically different (4.1). The RNFL, GCL and IPL, the three inner retinal layers that form the ganglion cell complex (GCL), seem to be more affected by glaucomatous damage than the outer layers, which is in accordance with the literature [140, 141]. These three layers present a higher number of features with statistically significant differences between groups when compared with the remaining layers (Figure C.1), as well as the lowest *p-values* overall. Nevertheless, all retinal layers seem to be affected by glaucoma, as every layer presents statistically significant features, even if not equally relevant or significant.

All speckle analysis methods applied were able to capture signal information from the imaged speckle, resulting in statistically significant features for the problem. Nevertheless, some methods presented overall better results, e.g. the parameters of the SGLDM. Contrast map parameters and statistical properties like skewness, mean, and kurtosis, also performed well in several layers. From all the statistical distributions tested, the Nakagami, Lognormal, K distribution, and Gamma seem to be the most adequate for the analysed datasets, as their parameters better capture differences between severity groups. From all methods, the frequency domain approach had a lower performance for inner layers (RNFL, GCL, IPL), although its results were still statistically significant between groups. Surprisingly, this method was able to assess significant differences in outer layers (IS/OS, OSL, OPR, RPE), when other methods that performed better in the inner layers did not. This further confirms the hypothesis stated in the literature review [28] that speckle analysis methods are highly application dependent, and the optimal method to study the signal-carrying component can vary depending on the imaged tissue.

Regarding the thickness of the retinal layers, the results were, in general, consistent with those obtained for the speckle features, as the three innermost layers (RNFL, GCL and IPL) were the ones that presented a statistically significant difference between groups.

Glaucoma diagnosis and follow-up progression monitoring can be seen as two distinct problems, as glaucoma has different symptoms and structural characteristics depending on its severity. For this reason, two additional tests were performed: one to account for the potential role of speckle in the diagnosis of early stage glaucoma, and the second, for differentiating between advanced stages.

Healthy *versus* mild glaucoma

The thicknesses for the three inner retinal layers, RNFL, GCL, and IPL, have showed sta-

tistically significant differences between healthy and mild glaucoma. Regarding the speckle information, the same three layers also presented the highest number of discriminating features. With this, it can be concluded that the innermost layers are the most affected in early stage glaucoma, which is also in accordance with the literature.

The cpRNFL thickness (measured in the optic disc area) is one of the most popular OCT parameters used in clinical practice for the glaucoma diagnosis [118,124,142–144]. Even though in this project OCT data from the macula is used, and not from the optic disc, the role and importance of the RNFL remains. In fact, macular SD-OCT imaging measures an area encompassing half of the RGCs in the eye [145]. This realization has been acknowledged by Kim *et al.* [146], where extensive comparisons were made between the thickness of the three layers of macular ganglion cell complex (GCC), as well as the cpRNFL, for healthy and glaucomatous eyes. The comparisons showed that all layers were significantly thinner in subjects with glaucoma. Furthermore, the cpRNFL, macular retinal nerve fiber layer (mRNFL) and GCL showed similar performance in differentiating between normal and glaucoma, contrary to the IPL, which was slightly lower. Our statistical tests provide similar results, with the difference that the thickness of IPL presents a very similar *p-value* to the RNFL, but both slightly lower than GCL (Figure 4.4). The role of the macular GCL and IPL, often considered together as the ganglion cell-Inner plexiform layer (GCIPL), or together as the GCC (RNFL+GCL+IPL) has also been the focus of interest of other works, which claim it can achieve similar results to the values of cpRNFL alone, or even provide some added value to the diagnostic [145, 147–151].

Besides the three layers of GCC, the IS/OS, OSL, OPR, and RPE presented at least one significantly different feature when comparing healthy and mild glaucoma, even though with *p-values* generally higher than the GCC layers (Figure C.2). Nevertheless, it is still an interesting result, as structural changes in these outer layers are far less explored than the layers in GCC. A few studies have suggested that glaucoma may cause abnormalities of the outer retina [152–154]. However, this affirmation seems to be controversial, as other works indicate the exact opposite [155,156]. Even though it is not a consensual affirmation, histological experiments have suggested that glaucoma can cause loss and injury of photoreceptors [157], present in OPR, OSL. This might explain the results obtained for the outer retinal layers, as well as the results which are discussed next, for more advanced glaucoma, where these layers seem to gain more importance.

Moderate *versus* severe glaucoma

Comparing moderate to severe glaucomatous eyes, the results obtained are fairly different

than those obtained previously. First of all, the number of speckle features which are statistically significantly different is considerably lower, only 5 compared to 50 in the healthy *versus* mild test. Also, the layers from which they were extracted differ as well, as the RNFL and GCL no longer show any differences, but the outer layers do. The most discriminative features come from the INL, followed by IPL, OSL, and OPL. Looking at the layers' thickness information, the IPL, OSL, and OPL show statistically significant differences, which is in accordance with the information from speckle, with the exception of the INL. Also, looking at the *p-values* of the Wilcoxon rank sum test for both sets of features, although they are significant, they have generally higher *p-values* than most discriminant features in the healthy *versus* mild glaucoma problem. This indicates that differentiation between moderate and severe cases of glaucoma based on OCT macula imaging is a "harder" problem.

Several works in the literature have explored the "floor effect' phenomenon (Section 2.5), the designation for a point in advanced stages of glaucoma where functional defects (visual field) no longer seem to cause any significant or detectable structural changes, and disease progression is therefore indistinguishable in OCT [87]. Although the reasons behind this effect are not completely known, the presence of non-neural tissue in the retinal layers, such as blood vessels or retinal glial cells, is thought to contribute to this effect [87]. Even if all ganglion cells die due to glaucomatous damage, non-neuronal elements do not degenerate and the layers' thickness will never decrease to 0, reaching a certain threshold above which it remains constant. The "floor effect" in OCT generally refers to the RNFL peripapillary thickness measurements (measured in the optic disc area), but our results indicate that it is also valid for its thickness measured in the macular region. Also, the same conclusion can be drawn for the thickness of the GCL, which also seems to reach a floor point in advanced stages. The speckle parameters from these two layers also show similar behaviours, showing that further changes at a cellular level in these layers are also undetectable after a certain point in disease progression. However, micro and macrostructural parameters from the IPL still show statistically significant differences between moderate and severe glaucoma, even though less significant than in the previous problem. This can explain why the thickness of IPL presented a higher discriminative power than the RNFL thickness when considering the four severity groups, but lower for only healthy and mild glaucoma: even though the RNFL changes more in early disease stages, it reaches the "floor effect" sooner, while IPL's importance remains even for advanced progression.

In the INL, although its thickness (macrostructure) was not discriminant to the problem, the speckle information (microstructure) presented a statistically significant difference between

groups. Even with *p-values* generally lower than the best features for the healthy *versus* mild test, and close to the level of significance, it is crucial to highlight the significance of these results. In advanced stages, when the RNFL and GCL reach the floor point, the outer layers seem to gain more importance. In a study performed by Hasegawa *et al.* [158], subjects with glaucoma showed a significantly higher chance of developing Microcystic INL changes than healthy subjects. This can lead focal and progressive damage in glaucoma, and potentially explaining the microstuctural differences in the INL found in our results. Kim *et al.* [154] stated that INL thickness is significantly associated with VF MD, and suggested that reactive responses of neuronal or glial cells in the INL could be the reason. In our results, the INL thickness did not show any statistically significant difference in any of the tests performed (all severity groups; healthy *versus* mild; moderate *versus* severe). However, the speckle information did, which can be explained by arguing that the microstructural changes occur prior to the macrostructural being detectable.

Changes in OSL also have been acknowledged in the literature. Matsuura *et al.* [159] compared the prediction performance of the GCC to the performance of a model combining GCC, OSL and RPE information. The latter showed a higher performance, indicating that outer layers like OSL and RPE might have some added value to the prediction of glaucoma. Nevertheless, these changes have not been claimed to gain importance, or improve distinction, on later stages of the disease. On the other hand, changes in OPL, to the teams' knowledge have not yet been stated. In fact, the opposite has been claimed [160], and OPL is though not to be affected by glaucoma, which differs from our results. For advanced stages, its thickness and one speckle feature showed statistically significant differences, as well as 29 out of 47 speckle features when considering overall differences between groups. Since these are pioneer findings and the role of the outer retinal layers in glaucoma is still fairly unexplored, further studies with different and larger datasets are needed to validate these results.

General observations

In general, the results presented in this section can be related with the literature. However, due to the varying subject inclusion criteria, the different OCT devices, parameters extracted, and the analysis algorithms used, detailed and quantitative comparisons are difficult to achieve.

In this project, both for the thickness and the speckle parameters, average measurements over the entire volume of the macular region were used. It can be argued that by using global average measures, it is not possible to detect early localized loss [145]. As an alternative, some

OCT devices measure local parameters, by dividing the retina into several regions. For example, the Early Treatment Diabetic Retinopathy Study (ETDRS) grid [161] divides the retina into 9 regions, defined by three rings: a central foveal ring with 1 mm diameter, an inner macular ring with 3mm diameter, and an outer macular ring with 6mm diameter, with the inner and outer rings being further divided into nasal, temporal, superior, and inferior quadrants. Thickness maps created from this grid can capture more local information and variations, and assess the thickness variation with respect to each anatomically distinct location [162, 163]. This analysis, although being more complex and requiring more advanced segmentation algorithms, can prevent local positive or negative deviations on certain regions to be canceled by averaging, allowing more sensitive measurements.

Overall, the results of the statistical tests confirm the presence of a signal-carrying component in OCT speckle, that can be used for the early diagnosis and follow-up of advanced stages of glaucoma. Additionally, speckle analysis may help improve the comprehension of the information retrieved from OCT imaging, namely its physical and physiological meaning. It has been proven that speckle provides information on the size and distribution of the scatterers [24, 98, 164], and that it can be linked to microstructural properties of the imaged tissue [97]. Since the RNFL thinning that occurs in glaucoma is known to be due to ganglion cell loss, it was expected that these changes at the cellular level would result in changes in the speckle pattern, which is supported by these results. The question that still arises is whether speckle information has any added value for the diagnosis of glaucoma when compared with the macrostructural information, a simpler and more interpretable parameter.

Possible drawbacks and challenges of speckle

Even though micro and macrostructural parameters seem to present similar *p-values* for the inner retinal layers, there are cases where speckle features present lower *p-values* than the thickness of the corresponding layers. For the outer layers, even though the thickness of OSL and OPL shows statistically significant differences between moderate and severe glaucoma, only one out of 47 speckle features for each layer is able to assess these differences (the entropy of SGLDM calculated over 135° and distance 1 of OSL, and the contrast of SGLDM calculated over 0° and distance 1 of OPL). This might indicate that if microstructural changes occur in the outer layers, they are harder to capture compared to the inner layers. Considering the physical nature of speckle might help explaining this phenomenon. The tissue is scanned by an optical beam, and the inner retinal layers are the first layers by which the light source passes

when acquiring an image. Since most of the light is scattered, as depicted in Figure 2.6, the incident light in the outer layers will logically suffer both multiple backscatterering and forward scatterering in multiples points along the retina before reaching the image plane, in comparison with the speckle from the inner layers, where the length the light travels is inferior. For this reason, it is expected that speckle patterns from the outer layers have a higher signal-degrading speckle component (commonly referred as speckle noise) comparing to speckle from the inner layers, and is therefore a less powerful and less reliable source of information.

On another note, speckle intensity depends highly on the imaging parameters and image quality. For example, any type of pre-processing of the image, such as filtering or averaging, can affect speckle. Also, OCT image quality may depend on the user experience, the system settings, or the presence of artifacts, which may also be seen as a drawback of this kind of analysis. Thus, the existence of raw unprocessed images and knowledge of imaging parameters is crucial for the development of a complete model, and to ensure proper reproducibility of these measurements between devices. This can be seen as a disadvantage of speckle information that needs to be overcame, contrary to thickness measurements, which are already proved to have high reproducibility between devices [165].

Also, the speckle size is dependent of the wavelength of the light used [9, 166]. If the size of each individual speckle is smaller than the pixel size, spatial averaging can occur. Therefore, the minimum resolvable speckle size will have an influence on the size of the microstructures that can be imaged. For this reason, further studies are in urge to understand the level of influence that changing the OCT device wavelength might have on the microstructural changes that can be captured in glaucomatous eyes.

Lastly, we can conclude that the best speckle features are, in general, similarly discriminant as the thickness features. Speckle information interpretability still remains a challenging task, mostly due to the vast quantity of speckle analysis methods proposed. Speckle information is intrinsically related to the spatial arrangement and biomechanical properties of the scatterers in the sample. Scatterers, in our problem are assumed to be cells or fibers in the imaged tissue, may these be ganglion cells, glial cells, photoreceptors, among others. The optimal method for speckle analysis has shown to be highly dependent on the layer and problem considered. This hampers the interpretability of speckle in a clinical context, as comparing information retrieved from different methods is difficult, and by choosing only one method, we might not achieve the full potential of speckle information. Further research in interpretability is important for applicability of OCT speckle quantification in a clinical context.

5.2 Disease Progression Modelling

The real disease progression timeline for glaucoma is unknown, especially due to the fact that speckle features and EBM have not been used before for this purpose. Because of this, a few indirect ways to measure the performance of DEBM were explored. Two "global" metrics were used: the overall separation between the two most extreme groups (healthy and severe glaucoma) and the correlation of the estimated patient staging with VF MD. The first metric was chosen following the approach proposed by [111] for the evaluation of the DEBM's performance. The second metric was chosen due to the VF MD being the parameter used for classifying glaucoma patients between severity groups, and being considered a measure of disease severity. Also, contrary to the remaining metrics used, it does not depend on the clustering of the patients between groups. However, it is important to remember that the VF MD was not the metric used for classifying between healthy and glaucoma (details on this classification are explained in Section 3.1), so the VF MD does not have a total correlation with the disease stage (the correlation of VF MD with the severity labels is 0.89). Nevertheless, it is still a useful measure.

Overall, some very promising results were achieved, with the highest performing model reaching an AUC of 0.984 ± 0.018 for the separation between healthy controls and severe glaucoma and a correlation of 0.72 between the estimated disease stage and the VF MD.

Micro *versus* macrostructural parameters

The overall performance of the models using only speckle or thickness features separately were similar, with most of the performance metrics presenting overlapping confidence intervals.

The model trained with thickness features presented better results for the differentiation between healthy *versus* mild and moderate *versus* severe, while the model trained with speckle features presented a better performance differentiating mild *versus* moderate glaucoma. Both in the overall separation between healthy and severe and the correlation with the VF MD (the two "global" metrics used), the thickness model obtained a slightly higher performance, although the confidence intervals overlap. When combining both types of features, a higher performance was achieved for almost all metrics. However, these results need to be analysed thoughtfully, as the improvements are very small for most cases and the confidence intervals also overlap. Also, even though the highest performing model was obtained for the thickness features combined with only one speckle feature, the performance remained fairly similar until around 20 speckle features were added. Nevertheless, it is unclear whether speckle information has a significant added value for the DEBM, and if it could have any clinical advantage.

In general, for all presented models, the DEBM's performance in differentiating moderate and severe glaucoma is lower than in the other adjacent severity groups. This was an expected result, as from the statistical tests it was clear that this was a "harder" differentiating problem, with less significant features, due to the previously discussed OCT "floor effect". The differentiation between healthy *versus* mild and mild *versus* moderate presented generally a higher performance, reaching an AUC of 0.829 ± 0.075 and 0.821 ± 0.112, respectively. The fact that the two intermediate groups (mild and moderate glaucoma) were well separated further indicates that DEBM was successful in understanding and modelling glaucoma disease progression, as the model was not explicitly trained to separate the intermediate groups (DEBM is only explicitly trained to separate the two most extreme severity groups, and it expects biomarkers to follow a monotonic behaviour with the increase of disease severity).

Disease timeline variation analysis

DEBM aims to provide a new insight to the processes behind the progression of a disease. Three positional variance diagrams were presented (Figures 4.16, 4.17 and 4.18), and some phenomenon are consistently observed in all of them. The changes on the GCL appear always in the first positions, although both features from the GCL and RNFL appear to become abnormal in the initial stages of the disease, with event centers very close to each other. On the other hand, structural changes in the IPL appear consistently in the later positions, with event centers in more advanced stages. The role of RNFL has been well established in the literature. The role of GCL and IPL, although not so popular and not so widely used, has also been highlighted before [145, 167]. However, these two layers are normally analysed together, as the GCIPL [4, 168–172]. From our results, GCL and IPL present different behaviours, with event centers at different disease stages. The GCL seems to be more important at an early stage, while IPL seems to gain importance in more advanced stages, even after GCL and RNFL reach the "floor effect". This being said, using the information from the two layers together, might mitigate some of their local changes and hamper their correct interpretation. Analysing them separately might be beneficial for both the differentiating performance and disease interpretation. In a work performed by Ustaoglu *et al.* [169], the GCIPL thickness showed a higher performance discriminating moderate-to-severe from early glaucoma patients, when compared to the cpRNFL thickness, while in differentiating healthy subject from glaucoma suspects, the cpRNFL presented a higher performance. This is in accordance with the hypothesis that IPL gains importance in later disease stages, although the results from our project and Ustaoglu *et*

al.'s [169] work are hard to compare quantitatively, as they did not attempt a differentiation between moderate and severe glaucoma patients (only moderate/severe *versus* early glaucoma) and did not consider information from the IPL alone. Other works [173–175] have also claimed that the thickness of the GCIPL has a stronger relationship with VF MD severity than RNFL thickness for advanced disease stages, which further supports our hypothesis.

Recently, more research has been performed using GCL and IPL separately [146,176], mainly due to technological advancements in OCT devices enabling the development of more reliable and reproducible macular retinal segmentation algorithms. However, to the teams knowledge, no work has reported the importance of IPL for more advanced stages of the disease (moderate *versus* severe). From a clinical perspective, Aydin *et al.* [176] claimed that the structural alterations that occur in the IPL can be explained by RGC's dendrites (which are arborized in the IPL) atrophy. The number of synapses established between RGC's dendrites and retinal bipolar and amacrine cells decreases as neurodegeneration progresses, leading to axon degeneration. This is expected to cause thinning of the IPL and consequent visual function worsening.

Micro *versus* macrostructural parameters

Comparing the disease timeline obtained by micro and macrostructural parameters separately, the results regarding the positioning of parameters from the RNFL, GCL and IPL are in accordance. However, in the disease timeline computed with speckle features, the event centers present much higher variance intervals compared to the model trained with thickness features. This can be explained by the higher number of features in study and, most importantly, due to the fact that we are considering several features computed over the same layer. Even though during the feature pre-selection stage, features with a correlation higher than 0.99 with other features were discarded, there are still features with high correlations (same layers and similar methods) that were included. This can hamper the staging algorithm, as features that provide very similar information are expected to change at very similar times, making it harder to estimate the order in which they become abnormal. For example, the events in the first four positions in Figure 4.17 present nearly similar event centers and intervals of variance, which is understandable considering they are all computed with the same method and in the same layer, just at different angles (Energy of SGLDM). Also, it is interesting to note that the variation of the events in the middle are much higher than the events which appear in the beginning and in the end of the mean central ordering. This might also be a result of the high correlation between features, but it can also indicate that they become abnormal at similar times.

When combining features computed from speckle with thickness features (Figure 4.18), the ordering of the changes in the three layers remains the same. Nonetheless, it is also interesting to compare the positioning of the speckle and thickness features from the same layers. As speckle is though to be linked to tissue microstructure, and macrostructural changes in the inner retinal layers are known to be due to RGC loss, it could be hypothesized that the death of RGC would lead to microstructural changes simultaneously (or even before) to any macrostructural changes being detectable. Even though this was not verified by the current results, as thickness features were always placed before the speckle features from the corresponding layers, this hypothesis should not be completely discarded. Speckle parameters are not a direct representation of any microstructural processes, but rather surrogate measures that can be linked to them. Since speckle has a signal-degrading component, speckle biomarkers can have a high noise component in comparison with thickness biomarkers, making it harder for small alterations to be noticed. If macro and microstructural changes happen simultaneously, it can happen that changes in speckle are only noticeable later on, due to the higher noise component. Finally, the OCT resolution might not be high enough to capture subtle microstructural changes, as even though speckle is able to capture structures smaller than the OCT resolution, it is still highly dependent of imaging parameters. Further studies with multiple OCT devices are needed to confirm or refute this hypothesis.

General observations

From the positional variance diagrams (Figures 4.16, 4.17 and 4.18), is possible to observe that events present variant distances between each other, sometimes even with overlapping variance intervals. For example, in Figure 4.16, the RNFL and GCL thickness have very close event centers, while IPL's event center is located in a later stage of the disease. The relative distance between events centers adds a new insight to the interpretation of the disease timeline, and is a great advantage of the DEBM over earlier versions of EBM.

Another important remark about the DEBM results, is the importance of good biomarkers. In the work performed by Venkatraghavan *et al.* [111], it was pointed that the quality of biomarkers plays a huge role in the performance of EBMs [112], confirmed by experiments on simulated data. From our results, this is also clear. In Section 4.2.1, when DEBM's performance was studied using an increasing number of speckle features, ordered by most to least powerful, it is clear that using a high number of features (therefore including less powerful features) has a negative impact on the model's performance. Also, in Section 4.2.2, the best parameter for the

moderate *versus* severe case was added to the best performing model in an attempt to increase the performance in differentiating these two groups, even though it was not a "strong" biomarker overall. Adding this biomarker had a negative impact in the overall performance of DEBM, further confirming that good biomarkers are crucial to achieve a good DEBM performance. This being said, the highest DEBM performance was obtained using only features from the three first inner layers: RNFL, GCL and IPL. Although features from more outer layers showed to be statistically significant to the problem, they did not improve the model's performance. For this reason, parameters from the outer layers were not included in the disease timeline variation study, as a low performing model would make the timeline unreliable. However, these results do not imply that the study of the outer retinal layers is not relevant. Although not improving the overall performance, adding a feature from the INL improved differentiation between moderate and severe glaucoma. It could be argued that a "perfect" model for glaucoma prediction should be a multi-model approach, where information from different layers would be used for different stages of the disease. For earlier stages, information from the RNFL and GCL would be preferred, while in more advanced stages, information from the outer layers would be prioritized.

The common conclusions drawn from both the statistical tests and from the disease timeline variation study were that earlier changes occur in RNFL and GCL, followed by IPL. Also, it was seen that information from other outer layers became important in later disease stages (INL, OSL, OPL). From a physical point of view, it could be hypothesized that these changes are linked to damage caused by increased IOP, one of the most important risk factors for developing glaucoma [46,49], and the main causes of damage in the retina, especially in patients with POAG (Figure 2.3). As the increase of the IOP causes an increase of the mechanical stress and strain on the posterior structures of the eye, it is expected that the layers closer to the vitreous body were the first to be affected by this pressure, contrary to the outer layers that would be affected in later stages of the disease, when the IOP is higher.

Chapter 6

Conclusions

6.1 General conclusions

Speckle has historically been considered a source of noise in coherent light imaging and, even though some works have proved otherwise, it is still a research field with a lot of uncharted territory. The overall results presented in this project indicate that speckle can be used for evaluating the microstructural properties of the retina. However, from the obtained results it is still not clear whether it provides an added value for glaucoma diagnosis when compared to the commonly used biomarkers. Further studies in this field are in urge to achieve a deeper understanding of the physical processes that originate changes in speckle patterns.

Measurements from the macular thickness of RNFL and GCIPL have been the aim of recent research on glaucoma. However, separate measures from the GCL, IPL, or other outer retinal layers are not usually considered. Our results indicate that the outer retinal layers can contribute to a deeper comprehension of the processes behind glaucoma progression and aid a correct severity assessment, especially in more advanced cases.

The DEBM was able to successfully model glaucoma disease progression, and to provide an interpretable overview of the timeline in which processes occur during glaucoma progression. The results highlight the importance of structural changes in the IPL at later disease stages, which are deeply unexplored and barely used in clinical practice. Information from the IPL might aid glaucoma monitoring, by providing robust and reproducible measures to assess disease severity, without having to perform additional tests on the patient.

6.2 Future work

Further research on larger cohorts is in urge in order to validate or refute this work's findings and the formulated hypothesis. The use of longitudinal data could allow the implementation of different disease progression models, and eventually provide a new and essential overview of the disease. Also, the existence of raw unprocessed data, as well as knowledge of OCT device specifications is key.

The difficulty in interpreting speckle information is still one of its main drawbacks when compared with traditional, interpretable quantifications, such as layer's thickness. To take a step further in this topic, and allow speckle information to be included in a wider range of biomedical applications, more studies using tissue phantoms or artificially simulated speckle patterns that can accurately model real biological tissues could be beneficial. Also, further histological studies might help to understand the origin of the microstructural changes that happen in the retina.

While it was shown that EBM is a promising approach to give a profound insight into the processes that occur during a neuropathology like glaucoma, Section 4.2.3 showed that the use of multiple speckle features computed over the same layer and with similar methods did not benefit the interpretation of the estimated disease timeline, as it introduced a lot of variance in this estimation. This indicates that disease progression models might benefit from the use of a wide range of biomarkers, with different origins (different imaging modalities, different regions of interest) and that provide distinct information. Additionally, this would provide clinicians a wider overview of the disease timeline, and would help to understand which type of parameters are better for each disease stage. For example, parameters from the optic disc area, as the mean RNFL thickness, the Bruch's membrane opening, the neuroretinal rim area, the average cup-to-disc area, and measurements from the lamina cribrosa [177, 178] could be included in future models, and potentially increase their predictive performance. Speckle information from the cornea have also been claimed to have an influence on the IOP measurements [15]. Specifically, Iskander *et al.* [19] were able to find differences in corneal speckle from glaucoma patients, glaucoma suspects and healthy controls. The inclusion of parameters obtained using multiple imaging modalities in the DEBM is also an option for future research. For example, optical coherence tomography angiography (OCTA) has been recently introduced and gained popularity in glaucoma research. This novel imaging approach is capable of analyzing retinal vascular flow that have been reported to be abnormal in glaucomatous eyes, especially in more advanced

stages [56]. Other imaging modalities that have been used to analyze glaucoma and, thus, could be included in the model, are fundus photography, confocal scanning laser ophthalmoscopy, scanning laser polarimetry, or adaptive optics imaging.

A challenge that can emerge when using any type of imaging modality is ensuring proper repeatability of the measures between distinct devices. This is a bottleneck that has to be tackled in order to ensure comparable results between studies and to be able to implement them in real applications. This is specially relevant in speckle features, since OCT devices exist in a wide range of parameters and configurations, and the speckle is dependent on them. Therefore, a deeper knowledge of the imaging parameters' influence in the speckle patterns is needed.

In Table 3.1 it is possible to see the effect of topical medication prescribed in order to lower the IOP, one of the most common used treatments for glaucoma. Even though glaucoma disease results in higher levels of IOP, due to topical medication, the IOP is maintained within the normal range even in advanced stages. For this reason, it was decided to not include IOP measurements in the DEBM in order to not bias the model with medication effects. However, it is impossible to guarantee that the model is completely free of bias introduced by the medication, as it might influence the microstructural properties of the tissue, as the increase of IOP causes mechanical stress to the tissue, with possible effect on its microstructure arrangement. This possible effect should also be further explored.

Bibliography

[1] Y.-C. Tham, X. Li, T. Y. Wong, H. A. Quigley, T. Aung, and C.-Y. Cheng, "Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis," *Ophthalmology*, vol. 121, no. 11, pp. 2081–2090, 2014.

[2] E. D. P. R. Group *et al.*, "Prevalence of open-angle glaucoma among adults in the united states," *Archives of ophthalmology*, vol. 122, no. 4, p. 532, 2004.

[3] P. G. Spry and C. A. Johnson, "Identification of progressive glaucomatous visual field loss," *Survey of ophthalmology*, vol. 47, no. 2, pp. 158–173, 2002.

[4] J. W. Shin, J. Lee, J. Kwon, J. Choi, and M. S. Kook, "Regional vascular density–visual field sensitivity relationship in glaucoma according to disease severity," *British Journal of Ophthalmology*, vol. 101, no. 12, pp. 1666–1672, 2017.

[5] C. Stephen and L.-M. Benjamin, "The east london glaucoma prediction score: web-based validation of glaucoma risk screening tool," *International journal of ophthalmology*, vol. 6, no. 1, p. 95, 2013.

[6] L. R. De Pretto, G. E. Nogueira, and A. Z. Freitas, "New speckle analysis method for optical coherence tomography signal based on autocorrelation," in *Biophotonics South America*, vol. 9531, p. 95313P, International Society for Optics and Photonics, 2015.

[7] J. M. Schmitt, S. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *Journal of biomedical optics*, vol. 4, no. 1, pp. 95–105, 1999.

[8] D. A. Jesus and D. R. Iskander, "Assessment of corneal properties based on statistical modeling of oct speckle," *Biomedical Optics Express*, vol. 8, no. 1, pp. 162–176, 2017.

[9] P. G. Vaz, A. Humeau-Heurtier, E. Figueiras, C. Correia, and J. Cardoso, "Laser speckle imaging to monitor microvascular blood flow: a review," *IEEE reviews in biomedical engineering*, vol. 9, pp. 106–120, 2016.

[10] S. Wang, C.-H. Liu, V. P. Zakharov, A. J. Lazar, R. E. Pollock, and K. V. Larin, "Three-dimensional computational analysis of optical coherence tomography images for the detection of soft tissue sarcomas," *Journal of biomedical optics*, vol. 19, no. 2, p. 021102, 2013.

[11] P. Ossowski, A. Raiter, A. Szkulmowska, and M. Wojtkowski, "Detection of small biological objects by phase-sensitive optical coherence tomography," in *Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XIX*, vol. 9312, p. 93122A, International Society for Optics and Photonics, 2015.

[12] A. G. Roy, S. Conjeti, S. G. Carlier, A. König, A. Kastrati, P. K. Dutta, A. F. Laine, N. Navab, D. Sheet, and A. Katouzian, "Bag of forests for modelling of tissue energy interaction in optical coherence tomography for atherosclerotic plaque susceptibility assessment," in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 428–431, IEEE, 2015.

[13] S. Seevaratnam, A. Bains, M. Farid, G. Farhat, M. Kolios, and B. A. Standish, "Quantifying temperature changes in tissue-mimicking fluid phantoms using optical coherence tomography and envelope statistics," in *Optical Fibers and Sensors for Medical Diagnostics and Treatment Applications XIV*, vol. 8938, p. 89380R, International Society for Optics and Photonics, 2014.

[14] D. A. Jesus and D. R. Iskander, "Age-related changes of the corneal speckle by optical coherence tomography," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5659–5662, IEEE, 2015.

[15] D. A. Jesus, M. Majewska, P. Krzyżanowska-Berkowska, and D. R. Iskander, "Influence of eye biometrics and corneal micro-structure on noncontact tonometry," *PloS one*, vol. 12, no. 5, p. e0177180, 2017.

[16] V. Demidov, L. A. Matveev, O. Demidova, A. L. Matveyev, V. Y. Zaitsev, C. Flueraru, and I. A. Vitkin, "Analysis of low-scattering regions in optical coherence tomography: applications to neurography and lymphangiography," *Biomedical optics express*, vol. 10, no. 8, pp. 4207–4219, 2019.

[17] L. A. Matveev, A. L. Matveyev, V. Demidov, A. A. Sovetsky, G. V. Gelikonov, V. Y. Zaitsev, and I. A. Vitkin, "Assessment of optical coherence tomography speckle patterns in low-scatterer-concentration regions: simulations for lymphatic vessels mapping," in *European Conference on Biomedical Optics*, p. 11075_2, Optical Society of America, 2019.

[18] L. A. Matveev, V. Demidov, M. A. Sirotkina, D. A. Karashtin, A. A. Moiseev, I. Popov, A. A. Sovetsky, A. L. Matveyev, O. Demidova, G. V. Gelikonov, *et al.*, "Oct lymphangiography based on speckle statistics evaluation," in *Saratov Fall Meeting 2018: Optical and Nano-Technologies for Biology and Medicine*, vol. 11065, p. 1106502, International Society for Optics and Photonics, 2019.

[19] D. R. Iskander, M. A. Kostyszak, D. A. Jesus, M. Majewska, M. E. Danielewska, and P. Krzyzanowska-Berkowska, "Assessing corneal speckle in optical coherence tomography: A new look at glaucomatous eyes," *Optometry and Vision Science*, vol. 97, no. 2, pp. 62–67, 2020.

[20] K. H. Cheng, A. Mariampillai, K. K. Lee, B. Vuong, T. W. Luk, J. Ramjist, M. A. Curtis, H. Jakubovic, P. Kertes, M. Letarte, *et al.*, "Histogram flow mapping with optical coherence tomography for in vivo skin angiography of hereditary hemorrhagic telangiectasia," *Journal of Biomedical Optics*, vol. 19, no. 8, p. 086015, 2014.

[21] M. Niemczyk, M. E. Danielewska, M. A. Kostyszak, D. Lewandowski, and D. R. Iskander, "The effect of intraocular pressure elevation and related ocular biometry changes on corneal oct speckle distribution in porcine eyes," *Plos one*, vol. 16, no. 3, p. e0249213, 2021.

[22] C. Photiou, E. Bousi, I. Zouvani, and C. Pitris, "Measuring tissue dispersion using optical coherence tomography speckle," in *European Conference on Biomedical Optics*, p. 104160H, Optical Society of America, 2017.

[23] C. Photiou, E. Bousi, I. Zouvani, and C. Pitris, "Using speckle to measure tissue dispersion in optical coherence tomography," *Biomedical optics express*, vol. 8, no. 5, pp. 2528–2535, 2017.

[24] D. K. Kasaragod, Z. Lu, L. E. Smith, and S. J. Matcher, "Speckle texture analysis of optical coherence tomography images," in *Speckle 2010: Optical Metrology*, vol. 7387, p. 73871V, International Society for Optics and Photonics, 2010.

[25] K. W. Gossage, T. S. Tkaczyk, J. J. Rodriguez, and J. K. Barton, "Texture analysis of optical coherence tomography images: feasibility for tissue classification," *Journal of biomedical optics*, vol. 8, no. 3, pp. 570–576, 2003.

[26] K. W. Gossage, T. S. Tkaczyk, J. J. Rodriquez, and J. K. Barton, "Texture analysis for tissue classification of optical coherence tomography images," in *Advanced Biomedical and Clinical Diagnostic Systems*, vol. 4958, pp. 109–117, International Society for Optics and Photonics, 2003.

[27] K. W. Gossage, C. M. Smith, E. M. Kanter, L. P. Hariri, A. L. Stone, J. J. Rodriguez, S. K. Williams, and J. K. Barton, "Texture analysis of speckle in optical coherence tomography images of tissue phantoms," *Physics in Medicine & Biology*, vol. 51, no. 6, p. 1563, 2006.

[28] V. B. Silva, D. A. D. Jesus, S. Klein, T. van Walsum, J. Cardoso, L. S. Brea, and P. G. Vaz, "Signal-carrying speckle in optical coherence tomography: a methodological review on biomedical applications," 2021.

[29] J. Zhu, E. Zhang, and K. Del Rio-Tsonis, "Eye anatomy," *eLS*, 2012.

[30] H. Kolb, "Gross anatomy of the eye," *Webvision: The Organization of the Retina and Visual System [Internet]*, 2007.

[31] Jmarchn, "Diagram of human eye without labels." `https://commons.wikimedia.org/wiki/File:Diagram_of_human_eye_without_labels.svg`, 2014.

[32] D. A. Jesus, R. Kedzia, and D. R. Iskander, "Precise measurement of scleral radius using anterior eye profilometry," *Contact Lens and Anterior Eye*, vol. 40, no. 1, pp. 47–52, 2017.

[33] D. A. Atchison, G. Smith, and G. Smith, *Optics of the human eye*, vol. 2. Butterworth-Heinemann Oxford, 2000.

[34] A. Malhotra, F. J. Minja, A. Crum, and D. Burrowes, "Ocular anatomy and cross-sectional imaging of the eye," in *Seminars in Ultrasound, CT and MRI*, vol. 32, pp. 2–13, Elsevier, 2011.

[35] D. C. Hood, A. S. Raza, C. G. V. de Moraes, J. M. Liebmann, and R. Ritch, "Glaucomatous damage of the macula," *Progress in retinal and eye research*, vol. 32, pp. 1–21, 2013.

[36] M. Häggström, "Medical gallery of Mikael Häggström 2014." DOI:`10.15347/wjm/2014.008`, 2014.

[37] D. A. Jesus and D. R. Iskander, "Simplifying numerical ray tracing for two-dimensional non circularly symmetric models of the human eye," *Applied optics*, vol. 54, no. 34, pp. 10123–10127, 2015.

[38] K. Irsch and D. L. Guyton, "Anatomy of eyes.," *Encyclopedia of Biometrics*, vol. 1, 2009.

[39] J. F. Koretz and G. H. Handelman, "How the human eye focuses," *Scientific American*, vol. 259, no. 1, pp. 92–99, 1988.

[40] P. Albrecht, M. Ringelstein, A. Müller, N. Keser, T. Dietlein, A. Lappas, A. Foerster, H. Hartung, O. Aktas, and A. Methner, "Degeneration of retinal layers in multiple sclerosis subtypes quantified by optical coherence tomography," *Multiple Sclerosis Journal*, vol. 18, no. 10, pp. 1422–1429, 2012.

[41] S. Lemmens, T. Van Craenendonck, J. Van Eijgen, L. De Groef, R. Bruffaerts, D. A. de Jesus, W. Charle, M. Jayapala, G. Sunaric-Mégevand, A. Standaert, *et al.*, "Combination of snapshot hyperspectral retinal imaging and optical coherence tomography to identify alzheimer's disease patients," *Alzheimer's research & therapy*, vol. 12, no. 1, pp. 1–13, 2020.

[42] C. S. Lee and R. S. Apte, "Retinal biomarkers of alzheimer disease," *American journal of ophthalmology*, vol. 218, pp. 337–341, 2020.

[43] E. M. Lad, D. Mukherjee, S. S. Stinnett, S. W. Cousins, G. G. Potter, J. R. Burke, S. Farsiu, and H. E. Whitson, "Evaluation of inner retinal layers as biomarkers in mild cognitive impairment to moderate alzheimer's disease," *PloS one*, vol. 13, no. 2, p. e0192646, 2018.

[44] N. N. Samani, F. A. Proudlock, V. Siram, C. Suraweera, C. Hutchinson, C. P. Nelson, M. Al-Uzri, and I. Gottlob, "Retinal layer abnormalities as biomarkers of schizophrenia," *Schizophrenia bulletin*, vol. 44, no. 4, pp. 876–885, 2018.

[45] D. G. Sevim, M. Unlu, M. Gultekin, and C. Karaca, "Retinal single-layer analysis with optical coherence tomography shows inner retinal layer thinning in huntington's disease as a potential biomarker," *International ophthalmology*, vol. 39, no. 3, pp. 611–621, 2019.

[46] R. N. Weinreb, T. Aung, and F. A. Medeiros, "The pathophysiology and treatment of glaucoma: a review," *Jama*, vol. 311, no. 18, pp. 1901–1911, 2014.

[47] H. A. Quigley, "Glaucoma," *Lancet*, vol. 377, pp. 1367–1377, 2011.

[48] J. Barbosa-Breda, L. Abegão-Pinto, K. Van Keer, D. A. Jesus, S. Lemmens, E. Vandewalle, A. Rocha-Sousa, and I. Stalmans, "Heterogeneity in arterial hypertension and ocular perfusion pressure definitions: Towards a consensus on blood pressure-related parameters for glaucoma studies," *Acta ophthalmologica*, vol. 97, no. 4, pp. e487–e492, 2019.

[49] B. J. Fan, D. Y. Wang, D. S. C. Lam, and C. P. Pang, "Gene mapping for primary open angle glaucoma," *Clinical biochemistry*, vol. 39, no. 3, pp. 249–258, 2006.

[50] J. Lusthaus and I. Goldberg, "Current management of glaucoma," *Medical Journal of Australia*, vol. 210, no. 4, pp. 180–187, 2019.

[51] W.-W. Su, S.-T. Cheng, W.-J. Ho, P.-K. Tsay, S.-C. Wu, and S. H. Chang, "Glaucoma is associated with peripheral vascular endothelial dysfunction," *Ophthalmology*, vol. 115, no. 7, pp. 1173–1178, 2008.

[52] B. S. Shastry, "Genetic susceptibility to primary angle closure glaucoma (pacg)," *Discovery medicine*, vol. 15, no. 80, pp. 17–22, 2013.

[53] R. N. Weinreb, C. K. Leung, J. G. Crowston, F. A. Medeiros, D. S. Friedman, J. L. Wiggs, and K. R. Martin, "Primary open-angle glaucoma," *Nature Reviews Disease Primers*, vol. 2, no. 1, pp. 1–19, 2016.

[54] A. Greco, M. I. Rizzo, A. De Virgilio, A. Gallo, M. Fusconi, and M. De Vincentiis, "Emerging concepts in glaucoma and review of the literature," *The American journal of medicine*, vol. 129, no. 9, pp. 1000–e7, 2016.

[55] J. Mallick, L. Devi, P. K. Malik, and J. Mallick, "Update on normal tension glaucoma," *Journal of ophthalmic & vision research*, vol. 11, no. 2, p. 204, 2016.

[56] D. A. De Jesus, L. S. Brea, J. B. Breda, E. Fokkinga, V. Ederveen, N. Borren, A. Bekkers, M. Pircher, I. Stalmans, S. Klein, *et al.*, "Octa multilayer and multisector peripapillary microvascular modeling for diagnosing and staging of glaucoma," *Translational Vision Science & Technology*, vol. 9, no. 2, pp. 58–58, 2020.

[57] D. A. Jesus, J. B. Breda, K. Van Keer, A. R. Sousa, L. A. Pinto, and I. Stalmans, "Quantitative automated circumpapillary microvascular density measurements: a new angiooct-based methodology," *Eye*, vol. 33, no. 2, pp. 320–326, 2019.

[58] G. Beykin, A. M. Norcia, V. J. Srinivasan, A. Dubra, and J. L. Goldberg, "Discovery and clinical translation of novel glaucoma biomarkers," *Progress in Retinal and Eye Research*, p. 100875, 2020.

[59] L. Abegão Pinto, K. Willekens, K. Van Keer, A. Shibesh, G. Molenberghs, E. Vandewalle, and I. Stalmans, "Ocular blood flow in glaucoma–the leuven eye study," *Acta ophthalmologica*, vol. 94, no. 6, pp. 592–598, 2016.

[60] A. C. Shetgar and M. B. Mulimani, "The central corneal thickness in normal tension glaucoma, primary open angle glaucoma and ocular hypertension," *Journal of clinical and diagnostic research: JCDR*, vol. 7, no. 6, p. 1063, 2013.

[61] J. Barbosa-Breda, K. Van Keer, L. Abegão-Pinto, V. Nassiri, G. Molenberghs, K. Willekens, E. Vandewalle, A. Rocha-Sousa, and I. Stalmans, "Improved discrimination between normal-tension and primary open-angle glaucoma with advanced vascular examinations–the leuven eye study," *Acta ophthalmologica*, vol. 97, no. 1, pp. e50–e56, 2019.

[62] J. Flammer, K. Konieczka, and A. J. Flammer, "The primary vascular dysregulation syndrome: implications for eye diseases," *EPMA Journal*, vol. 4, no. 1, pp. 1–33, 2013.

[63] A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser, "Optical coherence tomography-principles and applications," *Reports on progress in physics*, vol. 66, no. 2, p. 239, 2003.

[64] J. S. Schuman, M. R. Hee, A. V. Arya, T. Pedut-Kloizman, C. A. Puliafito, J. G. Fujimoto, and E. A. Swanson, "Optical coherence tomography: a new tool for glaucoma diagnosis.," *Current opinion in ophthalmology*, vol. 6, no. 2, pp. 89–95, 1995.

[65] I. I. Bussel, G. Wollstein, and J. S. Schuman, "Oct for glaucoma diagnosis, screening and detection of glaucoma progression," *British Journal of Ophthalmology*, vol. 98, no. Suppl 2, pp. ii15–ii19, 2014.

[66] A. P. Yow, R. Srivastava, J. Cheng, A. Li, J. Liu, L. Schmetterer, H. L. Tey, and D. W. Wong, "Techniques and applications in skin oct analysis," *Deep Learning in Medical Image Analysis*, pp. 149–163, 2020.

[67] T.-H. Tsai, C. L. Leggett, A. J. Trindade, A. Sethi, A.-F. Swager, V. Joshi, J. J. Bergman, H. Mashimo, N. S. Nishioka, and E. Namati, "Optical coherence tomography in gastroenterology: a review and future outlook," *Journal of biomedical optics*, vol. 22, no. 12, p. 121716, 2017.

[68] J. G. Fujimoto, S. A. Boppart, G. Tearney, B. E. Bouma, C. Pitris, and M. E. Brezinski, "High resolution in vivo intra-arterial imaging with optical coherence tomography," *Heart*, vol. 82, no. 2, pp. 128–133, 1999.

[69] P. H. Tomlins and R. K. Wang, "Theory, developments and applications of optical coherence tomography," *Journal of Physics D: Applied Physics*, vol. 38, no. 15, p. 2519, 2005.

[70] A. G. Podoleanu, J. A. Rogers, D. A. Jackson, and S. Dunne, "Three dimensional oct images from retina and skin," *Optics Express*, vol. 7, no. 9, pp. 292–298, 2000.

[71] Q. Wang, "Discussion on the fully developed speckle field," *Optik-International Journal for Light and Electron Optics*, vol. 124, no. 17, pp. 2948–2950, 2013.

[72] D. D. Duncan and S. J. Kirkpatrick, "Algorithms for simulation of speckle (laser and otherwise)," in *Complex Dynamics and Fluctuations in Biomedical Photonics V*, vol. 6855, p. 685505, International Society for Optics and Photonics, 2008.

[73] A. Ozcan, A. Bilenca, A. E. Desjardins, B. E. Bouma, and G. J. Tearney, "Speckle reduction in optical coherence tomography images using digital filtering," *JOSA A*, vol. 24, no. 7, pp. 1901–1910, 2007.

[74] M. Szkulmowski, I. Gorczynska, D. Szlag, M. Sylwestrzak, A. Kowalczyk, and M. Wojtkowski, "Efficient reduction of speckle noise in optical coherence tomography," *Optics express*, vol. 20, no. 2, pp. 1337–1359, 2012.

[75] S. Xiang, L. Zhou, and J. M. Schmitt, "Speckle noise reduction for optical coherence tomography," in *Optical and Imaging Techniques for Biomonitoring III*, vol. 3196, pp. 79–88, International Society for Optics and Photonics, 1998.

[76] A. Mcheik, C. Tauber, H. Batatia, J. George, and J.-M. Lagarde, "Speckle modelization in oct images for skin layers segmentation.," in *VISAPP (1)*, pp. 347–350, 2008.

[77] G. Farhat, G. J. Czarnota, M. C. Kolios, V. X. Yang, and A. Miriampillai, "Detecting apoptosis using dynamic light scattering with optical coherence tomography," *Journal of biomedical optics*, vol. 16, no. 7, p. 070505, 2011.

[78] X. Liu, J. C. Ramella-Roman, Y. Huang, Y. Guo, and J. U. Kang, "Robust spectral-domain optical coherence tomography speckle model and its cross-correlation coefficient analysis," *JOSA A*, vol. 30, no. 1, pp. 51–59, 2013.

[79] N. Uribe-Patarroyo, M. Villiger, and B. E. Bouma, "Quantitative technique for robust and noise-tolerant speed measurements based on speckle decorrelation in optical coherence tomography," *Optics express*, vol. 22, no. 20, pp. 24411–24429, 2014.

[80] N. Uribe-Patarroyo and B. E. Bouma, "Rotational distortion correction in endoscopic optical coherence tomography based on speckle decorrelation," *Optics letters*, vol. 40, no. 23, pp. 5518–5521, 2015.

[81] N. G. Ferris, T. M. Cannon, M. Villiger, B. E. Bouma, and N. Uribe-Patarroyo, "Forward multiple scattering dominates speckle decorrelation in whole-blood flowmetry using optical coherence tomography," *Biomedical optics express*, vol. 11, no. 4, pp. 1947–1966, 2020.

[82] P. Vaz, P. Santos, E. Figueiras, C. Correia, A. Humeau-Heurtier, and J. Cardoso, "Laser speckle contrast analysis for pulse waveform extraction," in *European Conference on Biomedical Optics*, p. 954007, Optical Society of America, 2015.

[83] Y. Beiderman, I. Horovitz, N. Burshtein, M. Teicher, J. Garcia, V. Micó, and Z. Zalevsky, "Remote estimation of blood pulse pressure via temporal tracking of reflected secondary speckles pattern," *Journal of biomedical optics*, vol. 15, no. 6, p. 061707, 2010.

[84] O. P. Maksymenko, L. I. Muravsky, and M. I. Berezyuk, "Application of biospeckles for assessment of structural and cellular changes in muscle tissue," *Journal of biomedical optics*, vol. 20, no. 9, p. 095006, 2015.

[85] M. Sehi and D. S. Greenfield, "Assessment of retinal nerve fiber layer using optical coherence tomography and scanning laser polarimetry in progressive glaucomatous optic neuropathy," *American journal of ophthalmology*, vol. 142, no. 6, pp. 1056–1059, 2006.

[86] T. M. Kuang, C. Zhang, L. M. Zangwill, R. N. Weinreb, and F. A. Medeiros, "Estimating lead time gained by optical coherence tomography in detecting glaucoma before development of visual field defects," *Ophthalmology*, vol. 122, no. 10, pp. 2002–2009, 2015.

[87] S. Asrani, L. Essaid, B. D. Alder, and C. Santiago-Turla, "Artifacts in spectral-domain optical coherence tomography measurements in glaucoma," *JAMA ophthalmology*, vol. 132, no. 4, pp. 396–402, 2014.

[88] C. Bowd, R. N. Weinreb, J. M. Williams, and L. M. Zangwill, "The retinal nerve fiber layer thickness in ocular hypertensive, normal, and glaucomatous eyes with optical coherence tomography," *Archives of ophthalmology*, vol. 118, no. 1, pp. 22–26, 2000.

[89] J. M. Gmeiner, W. A. Schrems, C. Y. Mardin, R. Laemmer, F. E. Kruse, and L. M. Schrems-Hoesl, "Comparison of bruch's membrane opening minimum rim width and peripapillary retinal nerve fiber layer thickness in early glaucoma assessment," *Investigative Ophthalmology & visual science*, vol. 57, no. 9, pp. OCT575–OCT584, 2016.

[90] O. Tan, G. Li, A. T.-H. Lu, R. Varma, D. Huang, A. I. for Glaucoma Study Group, *et al.*, "Mapping of macular substructures with optical coherence tomography for glaucoma diagnosis," *Ophthalmology*, vol. 115, no. 6, pp. 949–956, 2008.

[91] D. E. Lederer, J. S. Schuman, E. Hertzmark, J. Heltzer, L. J. Velazques, J. G. Fujimoto, and C. Mattox, "Analysis of macular volume in normal and glaucomatous eyes using optical coherence tomography," *American journal of ophthalmology*, vol. 135, no. 6, pp. 838–843, 2003.

[92] Y. H. Hwang, Y. C. Jeong, H. K. Kim, and Y. H. Sohn, "Macular ganglion cell analysis for early detection of glaucoma," *Ophthalmology*, vol. 121, no. 8, pp. 1508–1515, 2014.

[93] V. Mohammadzadeh, N. Fatehi, A. Yarmohammadi, J. W. Lee, F. Sharifipour, R. Daneshvar, J. Caprioli, and K. Nouri-Mahdavi, "Macular imaging with optical coherence tomography in glaucoma," *Survey of ophthalmology*, 2020.

[94] E. G. Society, *Terminology and Guidelines for Glaucoma. 5th ed.* Savona: Dogma, 2020.

[95] P. Lee, W. Gao, and X. Zhang, "Speckle properties of the logarithmically transformed signal in optical coherence tomography," *JOSA A*, vol. 28, no. 4, pp. 517–522, 2011.

[96] M. Y. Kirillin, G. Farhat, E. A. Sergeeva, M. C. Kolios, and A. Vitkin, "Speckle statistics in oct images: Monte carlo simulations and experimental studies," *Optics letters*, vol. 39, no. 12, pp. 3472–3475, 2014.

[97] M. Almasian, T. G. van Leeuwen, and D. J. Faber, "Oct amplitude and speckle statistics of discrete random media," *Scientific reports*, vol. 7, no. 1, pp. 1–11, 2017.

[98] T. R. Hillman, S. G. Adie, V. Seemann, J. J. Armstrong, S. L. Jacques, and D. D. Sampson, "Correlation of static speckle with sample properties in optical coherence tomography," *Optics Letters*, vol. 31, no. 2, pp. 190–192, 2006.

[99] D. D. Duncan, S. J. Kirkpatrick, and R. K. Wang, "Statistics of local speckle contrast," *JOSA A*, vol. 25, no. 1, pp. 9–15, 2008.

[100] S. H. Cho, S. Woo, C. Kim, H. J. Kim, H. Jang, B. C. Kim, S. E. Kim, S. J. Kim, J. P. Kim, Y. H. Jung, *et al.*, "Disease progression modelling from preclinical alzheimer's disease (ad) to ad dementia," *Scientific reports*, vol. 11, no. 1, pp. 1–10, 2021.

[101] N. P. Oxtoby, L.-A. Leyland, L. M. Aksman, G. E. Thomas, E. L. Bunting, P. A. Wijeratne, A. L. Young, A. Zarkali, M. M. Tan, F. D. Bremner, *et al.*, "Sequence of clinical and neurodegeneration events in parkinson's disease progression," *Brain*, vol. 144, no. 3, pp. 975–988, 2021.

[102] I. Koval, A. Bône, M. Louis, T. Lartigue, S. Bottani, A. Marcoux, J. Samper-Gonzalez, N. Burgos, B. Charlier, A. Bertrand, *et al.*, "Ad course map charts alzheimer's disease progression," *Scientific Reports*, vol. 11, no. 1, pp. 1–16, 2021.

[103] C. Abi Nader, N. Ayache, P. Robert, M. Lorenzi, A. D. N. Initiative, *et al.*, "Monotonic gaussian process for spatio-temporal disease progression modeling in brain imaging data," *Neuroimage*, vol. 205, p. 116266, 2020.

[104] H. M. Fonteijn, M. J. Clarkson, M. Modat, J. Barnes, M. Lehmann, S. Ourselin, N. C. Fox, and D. C. Alexander, "An event-based disease progression model and its application to familial alzheimer's disease," in *Biennial International Conference on Information Processing in Medical Imaging*, pp. 748–759, Springer, 2011.

[105] H. M. Fonteijn, M. Modat, M. J. Clarkson, J. Barnes, M. Lehmann, N. Z. Hobbs, R. I. Scahill, S. J. Tabrizi, S. Ourselin, N. C. Fox, *et al.*, "An event-based model for disease progression and its application in familial alzheimer's disease and huntington's disease," *NeuroImage*, vol. 60, no. 3, pp. 1880–1889, 2012.

[106] P. A. Wijeratne, A. L. Young, N. P. Oxtoby, R. V. Marinescu, N. C. Firth, E. B. Johnson, A. Mohan, C. Sampaio, R. I. Scahill, S. J. Tabrizi, *et al.*, "An image-based model of brain volume biomarker changes in huntington's disease," *Annals of clinical and translational neurology*, vol. 5, no. 5, pp. 570–582, 2018.

[107] N. C. Firth, S. Primativo, E. Brotherhood, A. L. Young, K. X. Yong, S. J. Crutch, D. C. Alexander, and N. P. Oxtoby, "Sequences of cognitive decline in typical alzheimer's disease and posterior cortical atrophy estimated using a novel event-based model of disease progression," *Alzheimer's & Dementia*, vol. 16, no. 7, pp. 965–973, 2020.

[108] M. C. Gabel, R. J. Broad, A. L. Young, S. Abrahams, M. E. Bastin, R. A. Menke, A. Al-Chalabi, L. H. Goldstein, S. Tsermentseli, D. C. Alexander, *et al.*, "Evolution of white matter damage in amyotrophic lateral sclerosis," *Annals of clinical and translational neurology*, vol. 7, no. 5, pp. 722–732, 2020.

[109] A. L. Young, N. P. Oxtoby, P. Daga, D. M. Cash, N. C. Fox, S. Ourselin, J. M. Schott, and D. C. Alexander, "A data-driven model of biomarker changes in sporadic alzheimer's disease," *Brain*, vol. 137, no. 9, pp. 2564–2577, 2014.

[110] V. Venkatraghavan, E. E. Bron, W. J. Niessen, and S. Klein, "A discriminative event based model for alzheimer's disease progression modeling," in *International Conference on Information Processing in Medical Imaging*, pp. 121–133, Springer, 2017.

[111] V. Venkatraghavan, E. E. Bron, W. J. Niessen, S. Klein, A. D. N. Initiative, *et al.*, "Disease progression timeline estimation for alzheimer's disease using discriminative event based modeling," *NeuroImage*, vol. 186, pp. 518–532, 2019.

[112] V. Venkatraghavan, *The symphony of cacophony: understanding the order in neurodegenerative diseases*. PhD thesis, E, June 2021.

[113] J. Huang and D. Alexander, "Probabilistic event cascades for alzheimer's disease," in *Advances in neural information processing systems*, pp. 3095–3103, Citeseer, 2012.

[114] M. Lorenzi and C. Gerhardinger, "Early cellular and molecular changes induced by diabetes in the retina," *Diabetologia*, vol. 44, no. 7, pp. 791–804, 2001.

[115] A. J. Barber, E. Lieth, S. A. Khin, D. A. Antonetti, A. G. Buchanan, T. W. Gardner, *et al.*, "Neural apoptosis in the retina during experimental and human diabetes. early onset and effect of insulin.," *The Journal of clinical investigation*, vol. 102, no. 4, pp. 783–791, 1998.

[116] G. Guedes, J. C Tsai, and N. A Loewen, "Glaucoma and aging," *Current aging science*, vol. 4, no. 2, pp. 110–117, 2011.

[117] J. J. Armstrong, T. Wasiuta, E. Kiatos, M. Malvankar-Mehta, and C. M. Hutnik, "The effects of phacoemulsification on intraocular pressure and topical medication use in patients with glaucoma: a systematic review and meta-analysis of 3-year data," *Journal of glaucoma*, vol. 26, no. 6, pp. 511–522, 2017.

[118] Y. Nakatani, T. Higashide, S. Ohkubo, H. Takeda, and K. Sugiyama, "Evaluation of macular thickness and peripapillary retinal nerve fiber layer thickness for detection of early glaucoma using spectral domain optical coherence tomography," *Journal of glaucoma*, vol. 20, no. 4, pp. 252–259, 2011.

[119] A. Giovannini, G. Amato, and C. Mariotti, "The macular thickness and volume in glaucoma: an analysis in normal and glaucomatous eyes using oct," *Acta Ophthalmologica Scandinavica*, vol. 80, pp. 34–36, 2002.

[120] F. A. Medeiros, L. M. Zangwill, C. Bowd, R. M. Vessani, R. Susanna Jr, and R. N. Weinreb, "Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography," *American journal of ophthalmology*, vol. 139, no. 1, pp. 44–55, 2005.

[121] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE reviews in biomedical engineering*, vol. 3, pp. 169–208, 2010.

[122] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 1, pp. 119–134, 2005.

[123] M. K. Garvin, M. D. Abramoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," *IEEE transactions on medical imaging*, vol. 28, no. 9, pp. 1436–1447, 2009.

[124] C. K. Leung, N. Choi, R. N. Weinreb, S. Liu, C. Ye, L. Liu, G. W. Lai, J. Lau, and D. S. Lam, "Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: pattern of rnfl defects in glaucoma," *Ophthalmology*, vol. 117, no. 12, pp. 2337–2344, 2010.

[125] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, *et al.*, "Scipy 1.0: fundamental algorithms for scientific computing in python," *Nature methods*, vol. 17, no. 3, pp. 261–272, 2020.

[126] G. R. Ge, J. P. Rolland, and K. J. Parker, "Speckle statistics of biological tissues in optical coherence tomography," *arXiv preprint arXiv:2101.12265*, 2021.

[127] L.-F. Huang, "The nakagami and its related distributions," *WSEAS Trans. Math*, vol. 15, no. 44, pp. 477–485, 2016.

[128] J. D. Briers and S. Webster, "Laser speckle contrast analysis (lasca): a nonscanning, full-field technique for monitoring capillary blood flow," *Journal of biomedical optics*, vol. 1, no. 2, pp. 174–179, 1996.

[129] D. A. 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.

[130] A. S. Gaudêncio, P. G. Vaz, M. Hilal, G. Mahé, M. Lederlin, A. Humeau-Heurtier, and J. M. Cardoso, "Evaluation of covid-19 chest computed tomography: A texture analysis based on three-dimensional entropy," *Biomedical Signal Processing and Control*, vol. 68, p. 102582, 2021.

[131] A. S. F. Gaudêncio, P. G. Vaz, M. Hilal, J. M. Cardoso, G. Mahé, M. Lederlin, and A. Humeau-Heurtier, "Three-dimensional multiscale fuzzy entropy: validation and application to idiopathic pulmonary fibrosis," *IEEE journal of biomedical and health informatics*, vol. 25, no. 1, pp. 100–107, 2020.

[132] G. K. Kanji, *100 statistical tests*. Sage, 2006.

[133] K. R. Sung, G. Wollstein, R. A. Bilonick, K. A. Townsend, H. Ishikawa, L. Kagemann, R. J. Noecker, J. G. Fujimoto, and J. S. Schuman, "Effects of age on optical coherence tomography measurements of healthy retinal nerve fiber layer, macula, and optic nerve head," *Ophthalmology*, vol. 116, no. 6, pp. 1119–1124, 2009.

[134] C. Liu, W. Wang, Z. Li, Y. Jiang, X. Han, J. Ha, W. Meng, and M. He, "Biological age estimated from retinal imaging: A novel biomarker of aging," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 138–146, Springer, 2019.

[135] D. A. Reynolds, "Gaussian mixture models.," *Encyclopedia of biometrics*, vol. 741, pp. 659–663, 2009.

[136] J. Ceberio, E. Irurozki, A. Mendiburu, and J. A. Lozano, "A review of distances for the mallows and generalized mallows estimation of distribution algorithms," *Computational Optimization and Applications*, vol. 62, no. 2, pp. 545–564, 2015.

[137] V. Venkatraghavan, "pyebm - a toolbox for event based models." `https://github.com/88vikram/pyebm`, 2017.

[138] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.

[139] M. Niemczyk and D. R. Iskander, "Statistical modeling of corneal oct speckle. a distributional model-free approach," *arXiv preprint arXiv:2107.07256*, 2021.

[140] H. Yamada, M. Hangai, N. Nakano, K. Takayama, Y. Kimura, M. Miyake, T. Akagi, H. O. Ikeda, H. Noma, and N. Yoshimura, "Asymmetry analysis of macular inner retinal layers for glaucoma diagnosis," *American journal of ophthalmology*, vol. 158, no. 6, pp. 1318–1329, 2014.

[141] J.-C. Mwanza, J. D. Oakley, D. L. Budenz, R. T. Chang, J. K. O'Rese, and W. J. Feuer, "Macular ganglion cell–inner plexiform layer: automated detection and thickness reproducibility with spectral domain–optical coherence tomography in glaucoma," *Investigative ophthalmology & visual science*, vol. 52, no. 11, pp. 8323–8329, 2011.

[142] T. Ojima, T. Tanabe, M. Hangai, S. Yu, S. Morishita, and N. Yoshimura, "Measurement of retinal nerve fiber layer thickness and macular volume for glaucoma detection using optical coherence tomography," *Japanese journal of ophthalmology*, vol. 51, no. 3, pp. 197–203, 2007.

[143] C. K.-s. Leung, C. Y.-l. Cheung, R. N. Weinreb, Q. Qiu, S. Liu, H. Li, G. Xu, N. Fan, L. Huang, C.-P. Pang, *et al.*, "Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study," *Ophthalmology*, vol. 116, no. 7, pp. 1257–1263, 2009.

[144] H. Wu, J. F. De Boer, and T. C. Chen, "Diagnostic capability of spectral-domain optical coherence tomography for glaucoma," *American journal of ophthalmology*, vol. 153, no. 5, pp. 815–826, 2012.

[145] K. Nouri-Mahdavi, S. Nowroozizadeh, N. Nassiri, N. Cirineo, S. Knipping, J. Giaconi, and J. Caprioli, "Macular ganglion cell/inner plexiform layer measurements by spectral domain optical coherence tomography for detection of early glaucoma and comparison to retinal nerve fiber layer measurements," *American journal of ophthalmology*, vol. 156, no. 6, pp. 1297–1307, 2013.

[146] H. J. Kim, S.-Y. Lee, K. H. Park, D. M. Kim, and J. W. Jeoung, "Glaucoma diagnostic ability of layer-by-layer segmented ganglion cell complex by spectral-domain optical coherence tomography," *Investigative ophthalmology & visual science*, vol. 57, no. 11, pp. 4799–4805, 2016.

[147] J.-C. Mwanza, D. L. Budenz, D. G. Godfrey, A. Neelakantan, F. E. Sayyad, R. T. Chang, and R. K. Lee, "Diagnostic performance of optical coherence tomography ganglion cell–inner plexiform layer thickness measurements in early glaucoma," *Ophthalmology*, vol. 121, no. 4, pp. 849–854, 2014.

[148] Y. Nakatani, T. Higashide, S. Ohkubo, and K. Sugiyama, "Influences of the inner retinal sublayers and analytical areas in macular scans by spectral-domain oct on the diagnostic ability of early glaucoma," *Investigative ophthalmology & visual science*, vol. 55, no. 11, pp. 7479–7485, 2014.

[149] M. Pazos, A. A. Dyrda, M. Biarnés, A. Gómez, C. Martín, C. Mora, G. Fatti, and A. Antón, "Diagnostic accuracy of spectralis sd oct automated macular layers segmentation to discriminate normal from early glaucomatous eyes," *Ophthalmology*, vol. 124, no. 8, pp. 1218–1228, 2017.

[150] N. R. Kim, E. S. Lee, G. J. Seong, J. H. Kim, H. G. An, and C. Y. Kim, "Structure–function relationship and diagnostic value of macular ganglion cell complex measurement using fourier-domain oct in glaucoma," *Investigative ophthalmology & visual science*, vol. 51, no. 9, pp. 4646–4651, 2010.

[151] M.-S. Sung, J.-H. Yoon, and S.-W. Park, "Diagnostic validity of macular ganglion cell-inner plexiform layer thickness deviation map algorithm using cirrus hd-oct in preperimetric and early glaucoma," *Journal of glaucoma*, vol. 23, no. 8, pp. e144–e151, 2014.

[152] S. Panda and J. B. Jonas, "Decreased photoreceptor count in human eyes with secondary angle-closure glaucoma.," *Investigative ophthalmology & visual science*, vol. 33, no. 8, pp. 2532–2536, 1992.

[153] S. S. Choi, R. J. Zawadzki, M. C. Lim, J. D. Brandt, J. L. Keltner, N. Doble, and J. S. Werner, "Evidence of outer retinal changes in glaucoma patients as revealed by ultrahigh-resolution in vivo retinal imaging," *British journal of ophthalmology*, vol. 95, no. 1, pp. 131–141, 2011.

[154] E. K. Kim, H.-Y. L. Park, and C. K. Park, "Relationship between retinal inner nuclear layer thickness and severity of visual field loss in glaucoma," *Scientific reports*, vol. 7, no. 1, pp. 1–7, 2017.

[155] K. R. Kendell, H. A. Quigley, L. A. Kerrigan, M. E. Pease, and E. N. Quigley, "Primary open-angle glaucoma is not associated with photoreceptor loss.," *Investigative ophthalmology & visual science*, vol. 36, no. 1, pp. 200–205, 1995.

[156] M. Wang, D. C. Hood, J.-S. Cho, Q. Ghadiali, G. V. De Moraes, X. Zhang, R. Ritch, and J. M. Liebmann, "Measurement of local retinal ganglion cell layer thickness in patients with glaucoma using frequency-domain optical coherence tomography," *Archives of ophthalmology*, vol. 127, no. 7, pp. 875–881, 2009.

[157] T. M. Nork, J. N. Ver Hoeve, G. L. Poulsen, R. W. Nickells, M. D. Davis, A. J. Weber, S. H. Sarks, H. L. Lemley, L. L. Millecchia, *et al.*, "Swelling and loss of photoreceptors in chronic human and experimental glaucomas," *Archives of ophthalmology*, vol. 118, no. 2, pp. 235–245, 2000.

[158] T. Hasegawa, T. Akagi, M. Yoshikawa, K. Suda, H. Yamada, Y. Kimura, H. Nakanishi, M. Miyake, N. Unoki, H. O. Ikeda, *et al.*, "Microcystic inner nuclear layer changes and retinal nerve fiber layer defects in eyes with glaucoma," *PloS one*, vol. 10, no. 6, p. e0130175, 2015.

[159] M. Matsuura, Y. Fujino, T. Kanamoto, H. Murata, M. Yanagisawa, K. Hirasawa, T. Inoue, N. Shoji, K. Inoue, J. Yamagami, *et al.*, "Improving the structure-function relationship in

glaucomatous and normative eyes by incorporating photoreceptor layer thickness," *Scientific reports*, vol. 8, no. 1, pp. 1–9, 2018.

[160] H. Ishikawa, D. M. Stein, G. Wollstein, S. Beaton, J. G. Fujimoto, and J. S. Schuman, "Macular segmentation with optical coherence tomography," *Investigative ophthalmology & visual science*, vol. 46, no. 6, pp. 2012–2017, 2005.

[161] E. T. D. R. S. R. Group *et al.*, "Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified airlie house classification: Etdrs report number 10," *Ophthalmology*, vol. 98, no. 5, pp. 786–806, 1991.

[162] M. N. Menke, S. Dabov, P. Knecht, and V. Sturm, "Reproducibility of retinal thickness measurements in healthy subjects using spectralis optical coherence tomography," *American journal of ophthalmology*, vol. 147, no. 3, pp. 467–472, 2009.

[163] M. Röhlig, R. K. Prakasam, J. Stüwe, C. Schmidt, O. Stachs, and H. Schumann, "Enhanced grid-based visual analysis of retinal layer thickness with optical coherence tomography," *Information*, vol. 10, no. 9, p. 266, 2019.

[164] S. J. Kirkpatrick, D. D. Duncan, R. K. Wang, and M. T. Hinds, "Quantitative temporal speckle contrast imaging for tissue mechanics," *JOSA A*, vol. 24, no. 12, pp. 3728–3734, 2007.

[165] H. Hirasawa, M. Araie, A. Tomidokoro, H. Saito, A. Iwase, S. Ohkubo, K. Sugiyama, T. Ootani, S. Kishi, K. Matsushita, *et al.*, "Reproducibility of thickness measurements of macular inner retinal layers using sd-oct with or without correction of ocular rotation," *Investigative ophthalmology & visual science*, vol. 54, no. 4, pp. 2562–2570, 2013.

[166] A. Ennos, "Speckle interferometry," in *Laser speckle and related phenomena*, pp. 203–253, Springer, 1975.

[167] A. S. Raza, J. Cho, C. G. de Moraes, M. Wang, X. Zhang, R. H. Kardon, J. M. Liebmann, R. Ritch, and D. C. Hood, "Retinal ganglion cell layer thickness and local visual field sensitivity in glaucoma," *Archives of ophthalmology*, vol. 129, no. 12, pp. 1529–1536, 2011.

[168] C. Bowd, L. M. Zangwill, R. N. Weinreb, F. A. Medeiros, and A. Belghith, "Estimating optical coherence tomography structural measurement floors to improve detection of progression in advanced glaucoma," *American journal of ophthalmology*, vol. 175, pp. 37–44, 2017.

[169] M. Ustaoglu, N. Solmaz, and F. Onder, "Discriminating performance of macular ganglion cell-inner plexiform layer thicknesses at different stages of glaucoma," *International journal of ophthalmology*, vol. 12, no. 3, p. 464, 2019.

[170] A. J. Prager, D. C. Hood, J. M. Liebmann, C. G. De Moraes, L. A. Al-Aswad, Q. Yu, G. A. Cioffi, and D. M. Blumberg, "Association of glaucoma-related, optical coherence tomography–measured macular damage with vision-related quality of life," *JAMA ophthalmology*, vol. 135, no. 7, pp. 783–788, 2017.

[171] W. J. Lee, S. U. Baek, Y. K. Kim, K. H. Park, and J. W. Jeoung, "Rates of ganglion cell-inner plexiform layer thinning in normal, open-angle glaucoma and pseudoexfoliation glaucoma eyes: a trend-based analysis," *Investigative ophthalmology & visual science*, vol. 60, no. 2, pp. 599–604, 2019.

[172] J.-W. Lee, E. Morales, F. Sharifipour, N. Amini, F. Yu, A. A. Afifi, A. L. Coleman, J. Caprioli, and K. Nouri-Mahdavi, "The relationship between central visual field sensitivity and macular ganglion cell/inner plexiform layer thickness in glaucoma," *British Journal of Ophthalmology*, vol. 101, no. 8, pp. 1052–1058, 2017.

[173] E. Ghahari, C. Bowd, L. M. Zangwill, J. Proudfoot, K. A. Hasenstab, H. Hou, R. C. Penteado, P. I. C. Manalastas, S. Moghimi, T. Shoji, *et al.*, "Association of macular and circumpapillary microvasculature with visual field sensitivity in advanced glaucoma," *American journal of ophthalmology*, vol. 204, pp. 51–61, 2019.

[174] A. Belghith, F. A. Medeiros, C. Bowd, J. M. Liebmann, C. A. Girkin, R. N. Weinreb, and L. M. Zangwill, "Structural change can be detected in advanced-glaucoma eyes," *Investigative ophthalmology & visual science*, vol. 57, no. 9, pp. OCT511–OCT518, 2016.

[175] M.-S. Sung, B.-W. Kang, H.-G. Kim, H. Heo, and S.-W. Park, "Clinical validity of macular ganglion cell complex by spectral domain-optical coherence tomography in advanced glaucoma," *Journal of glaucoma*, vol. 23, no. 6, pp. 341–346, 2014.

[176] R. Aydın, M. Barış, C. Durmaz-Engin, L. A. Al-Aswad, D. M. Blumberg, G. A. Cioffi, J. M. Liebmann, T. H. Tezel, and G. Tezel, "Early localized alterations of the retinal inner plexiform layer in association with visual field worsening in glaucoma patients," *Plos one*, vol. 16, no. 2, p. e0247401, 2021.

[177] J. C. Downs and C. A. Girkin, "Lamina cribrosa in glaucoma," *Current opinion in ophthalmology*, vol. 28, no. 2, p. 113, 2017.

[178] J. Morgan-Davies, N. Taylor, A. Hill, P. Aspinall, C. O'Brien, and A. Azuara-Blanco, "Three dimensional analysis of the lamina cribrosa in glaucoma," *British journal of ophthalmology*, vol. 88, no. 10, pp. 1299–1304, 2004.

Appendix A

Speckle analysis: state of the art

Table A.1: Characteristics of the speckle analysis in OCT state of the art studies.

Authors & Publication year	**Method**	**Aim**	**Application / Data used**	**OCT technique (brand)**	**Light wavelength (nm)**
Roy et al. [12] (2015)	Statistical properties	Classification	Coronary artery	SD-OCT (CV-M2)	1320
Schmitt et al. [7] (1999)	Statistical distributions	Theoretical modeling	-	-	-
Mcheik et al. [76] (2008)	Statistical distributions	Segmentation	Skin	SD-OCT (SkinDex 300)	1300
Kirillin et al. [96] (2014)	Statistical distributions	Theoretical modeling	Tissue phantoms (polystyrene microspheres)	SS-OCT (custom made)	1310
Seevaratnam et al. [13] (2014)	Statistical distributions	Classification	Tissue phantoms (polystyrene microspheres)	SS-OCT (Biophotonics and Bioengineering Laboratory's)	1310
Jesus et al. [14] (2015)	Statistical distributions	Classification	Cornea	SD-OCT (Copernicus HR)	850
Almasian et al. [97] (2017)	Statistical distributions	Theoretical modeling	Tissue phantoms (silica microspheres)	SS-OCT (Santec IVS 2000)	1309
Jesus et al. [8] (2017)	Statistical distributions	Classification	Cornea	SD-OCT (IOLMaster 700)	850
Jesus et al. [15] (2017)	Statistical distributions	Classification	Cornea	SD-OCT (Copernicus HR)	851
Demidov et al. [16] (2019)	Statistical distributions	Classification	Mice (skin)	SS-OCT (custom made)	1320
Matveev et al. [18] (2019)	Statistical distributions	Classification	Mice (skin)	SS-OCT (custom made)	1320
Iskander et al. [19] (2020)	Statistical distributions	Classification	Cornea	SD-OCT (HRT 3)	850
Niemczyk et al. [21] (2021)	Statistical distributions	Classification	Cornea (porcine eyes)	SD-OCT (Copernicus REVO)	830
Hillman et al. [98] (2006)	Contrast ratio	Theoretical modeling	Tissue phantoms (polystyrene microspheres)	SD-OCT (custom made)	1330

Authors & Publication year	Method	Aim	Application / Data used	OCT technique (brand)	Light wavelength (nm)
Duncan et al. [99] (2008)	Contrast ratio	Theoretical modeling / Segmentation	Embryonic chick heart	-	-
Kasaragod et al. [24] (2010)	SGLDM	Classification	Tissue phantoms (agar intralipid solution) / Tissue engineered (skin)	SS-OCT (custom made)	1315
Gossage et al. [25] (2003)	SGLDM/Frequency Domain methods	Classification	Mouse lung	SS-OCT (custom made)	1300
Gossage et al. [26] (2003)	SGLDM/Frequency Domain methods	Classification	Mouse lung / Bovine tissues	SS-OCT (custom made)	1300
Gossage et al. [27] (2006)	SGLDM/Frequency Domain methods	Classification	Tissue phantoms (silica microspheres) / Bovine aorta endothelial cells	SS-OCT (custom made)	1300

SGLDM = Spatial Gray Level Dependence Matrices;

SD = Spectral Domain;

SS = Swept Source;

IOLMaster 700 = IOLMaster 700 (Carl Zeiss Meditec AG, Germany);

CV-M2= CV-M2, LightLab Imaging Inc. (Westford, MA, USA);

Copernicus HR = Copernicus HR (Optopol, Zawiercie, Poland);

HRT 3 = HRT 3, Heidelberg Engineering GmbH (Heidelberg, Germany);

Copernicus REVO = Copernicus REVO, (Optopol, Zawiercie, Poland).

Appendix B

Project Outline

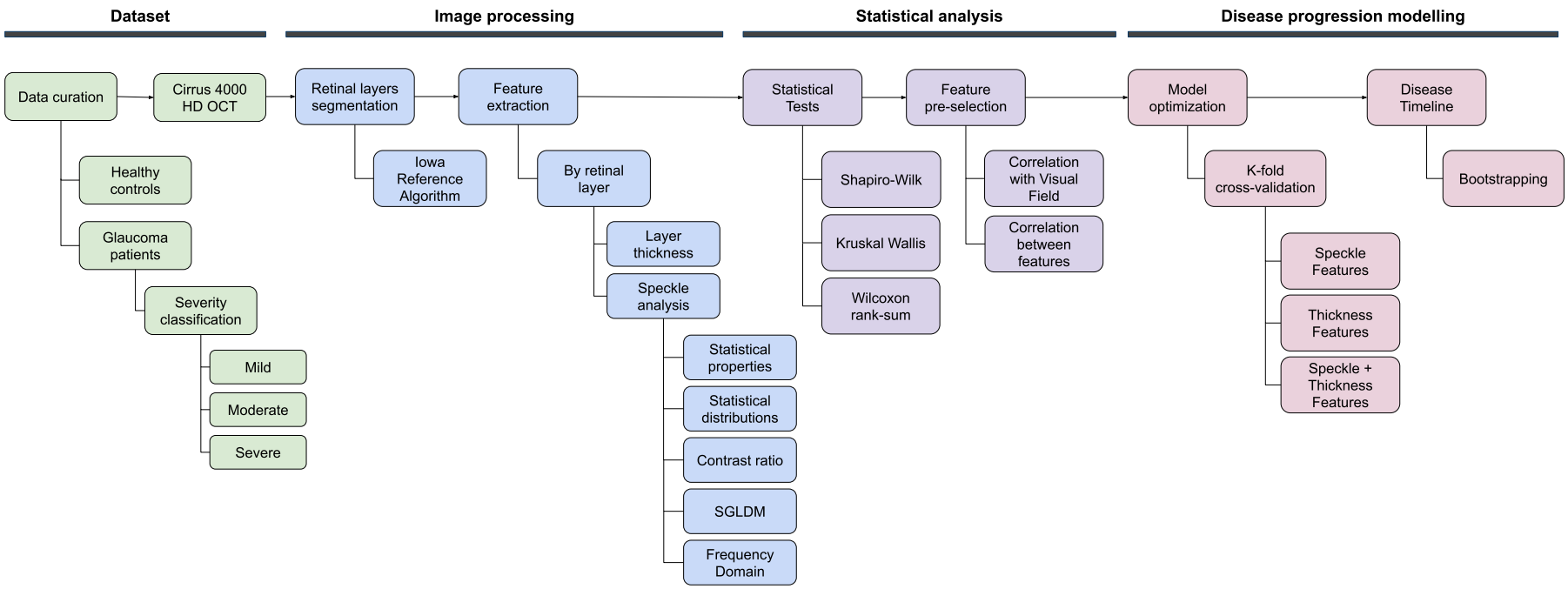

Figure B.1: Detailed outline of the different steps taken in the development of the project.

Appendix C

Statistical tests' results

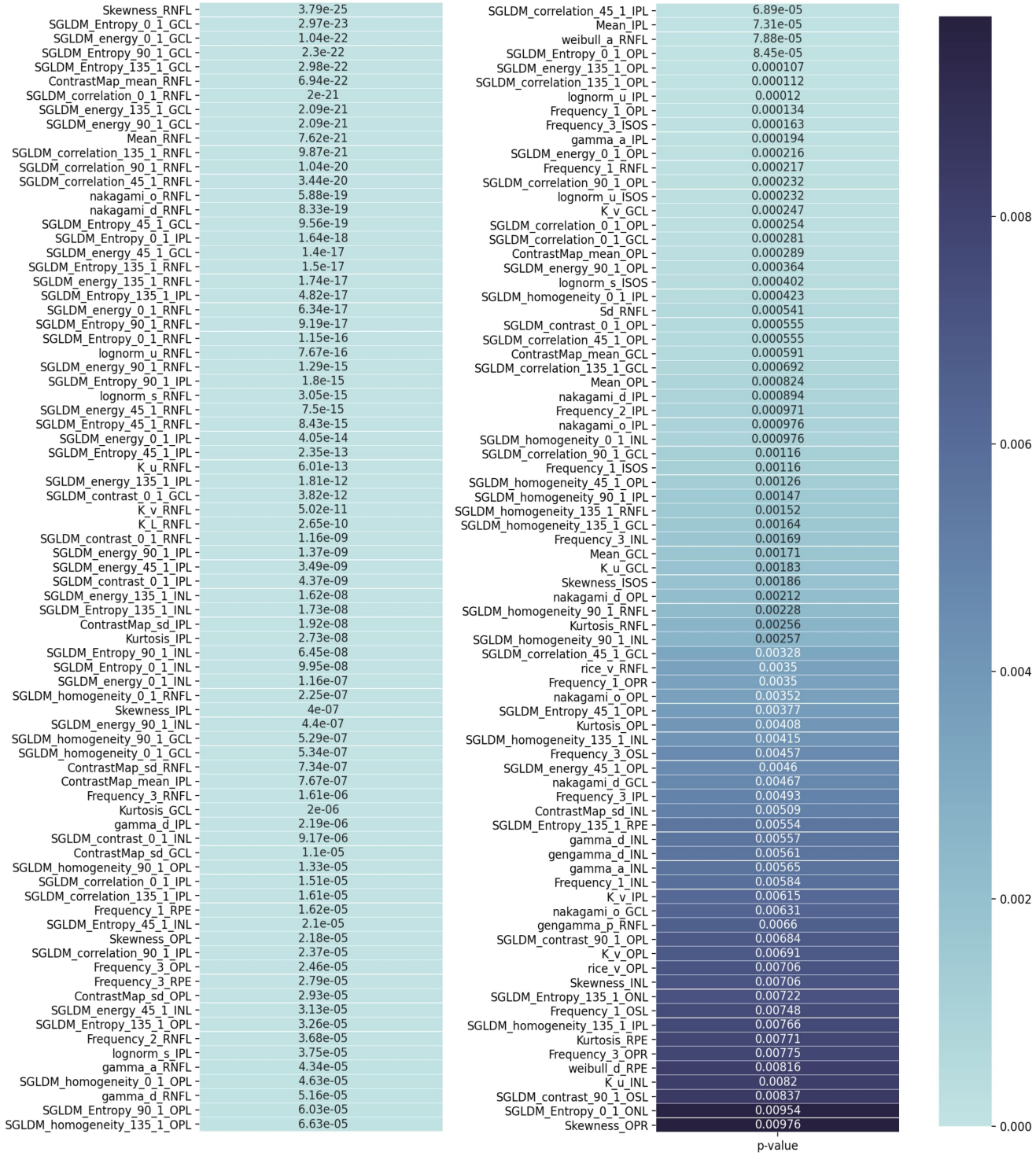

Figure C.1: List of the 158 features that presented statistical differences ($\alpha = 0.01$) between healthy and glaucoma patients, using the Kruskal-Wallis H test.

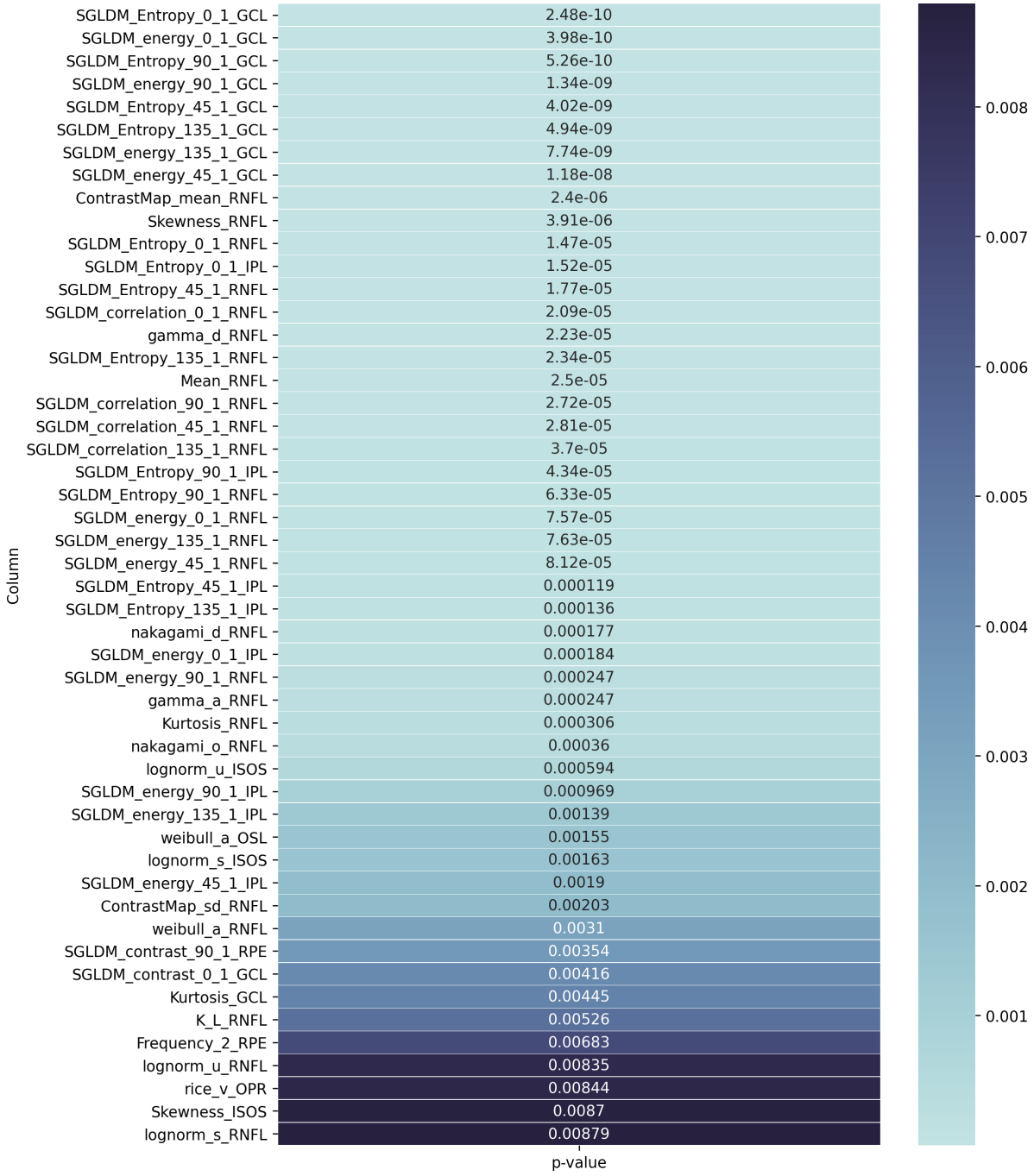

Figure C.2: List of the 50 features that presented statistical differences ($\alpha = 0.01$) between healthy controls and mild glaucoma patients, using the Wilcoxon rank-sum test.

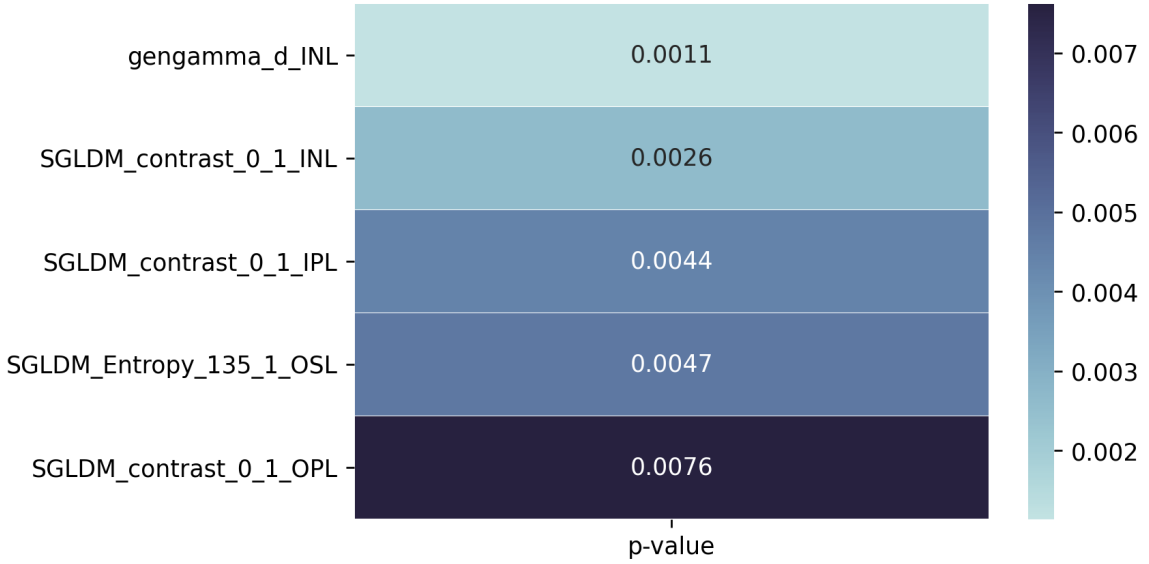

Figure C.3: List of the 5 features that presented statistical differences ($\alpha = 0.01$) between moderate and severe glaucoma patients, using the Wilcoxon rank-sum test.

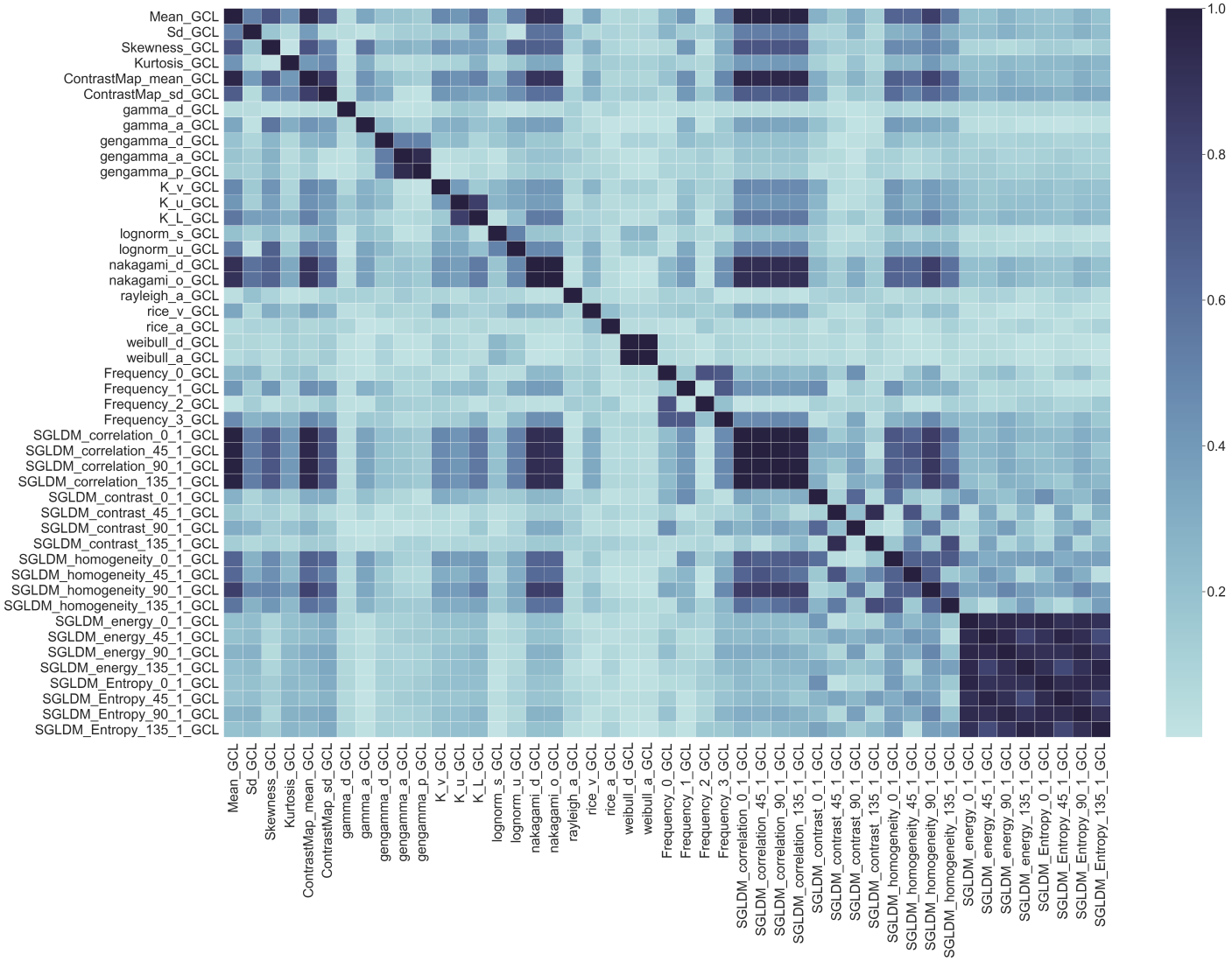

Figure C.4: Heatmap of the absolute value of Spearman's correlation values between features, for GCL.

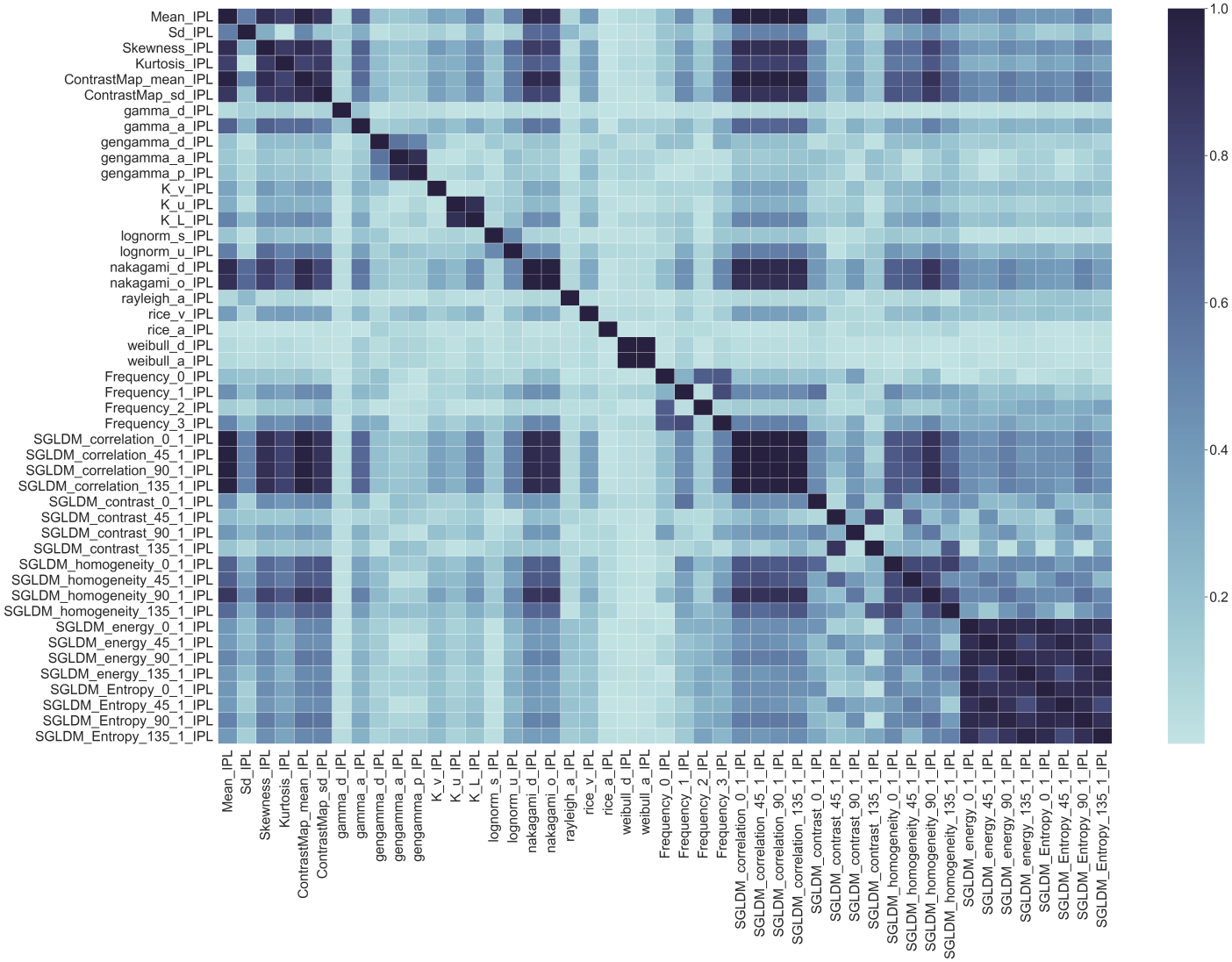

Figure C.5: Heatmap of the absolute value of Spearman's correlation values between features, for IPL.

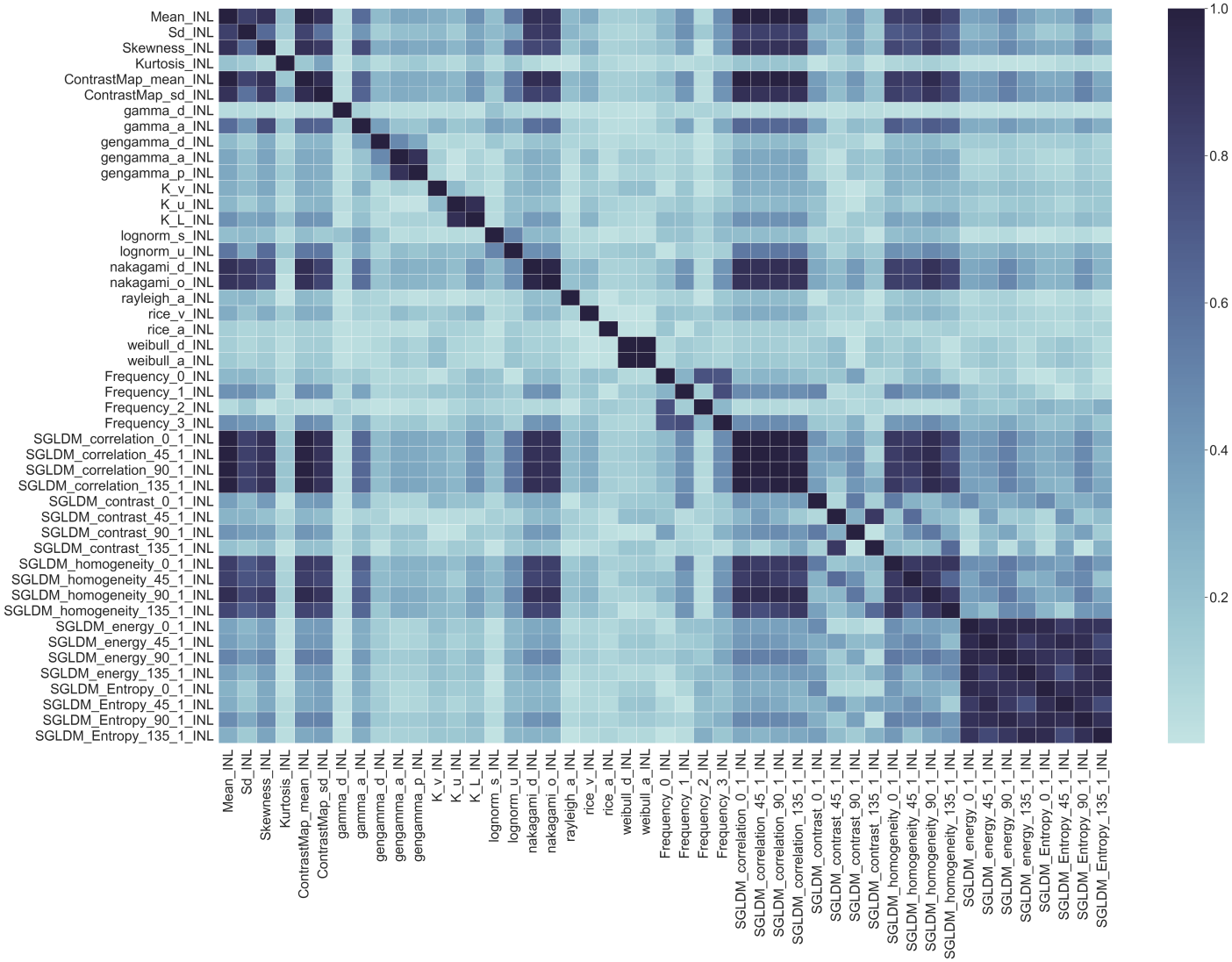

Figure C.6: Heatmap of the absolute value of Spearman's correlation values between features, for INL.

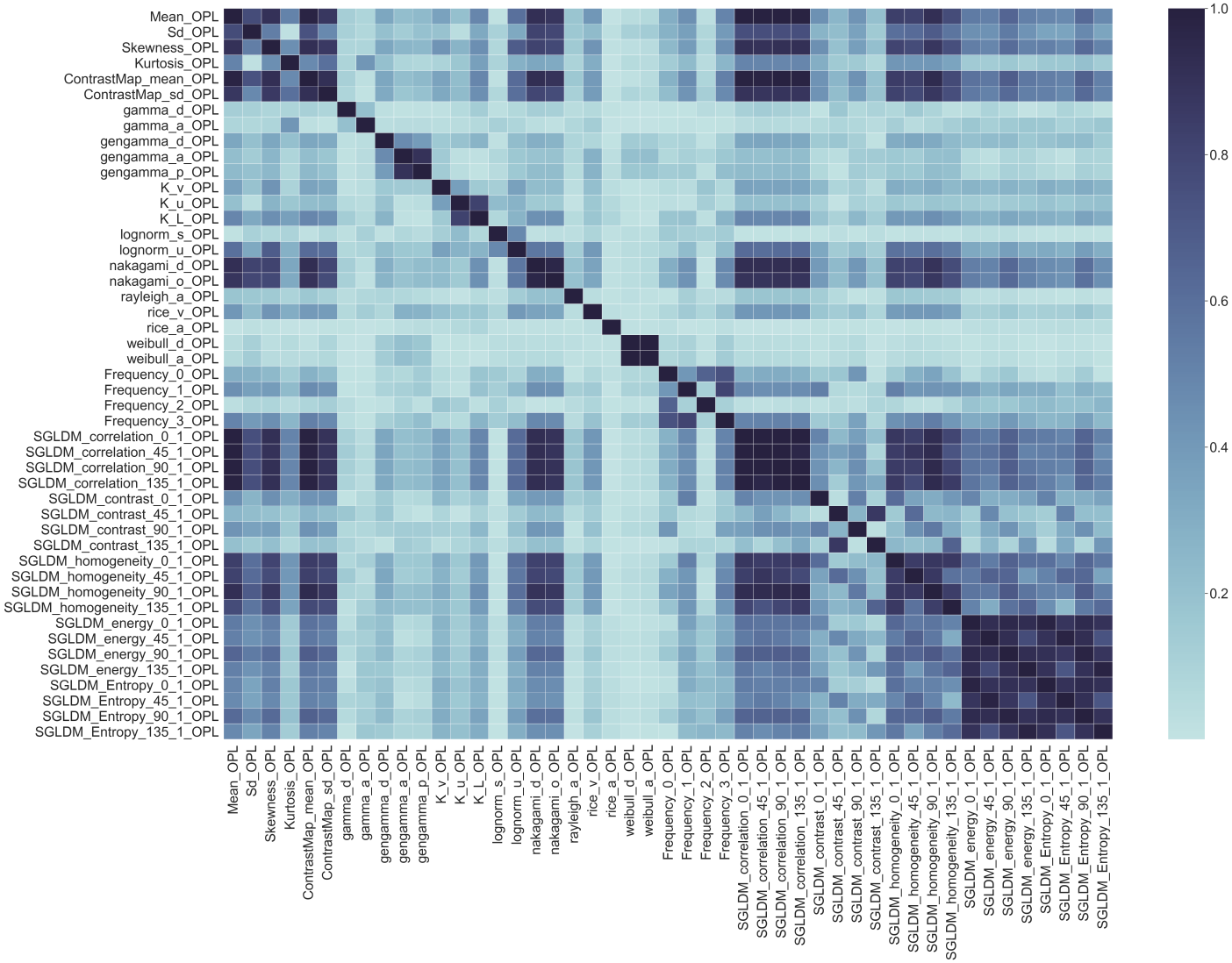

Figure C.7: Heatmap of the absolute value of Spearman's correlation values between features, for OPL.

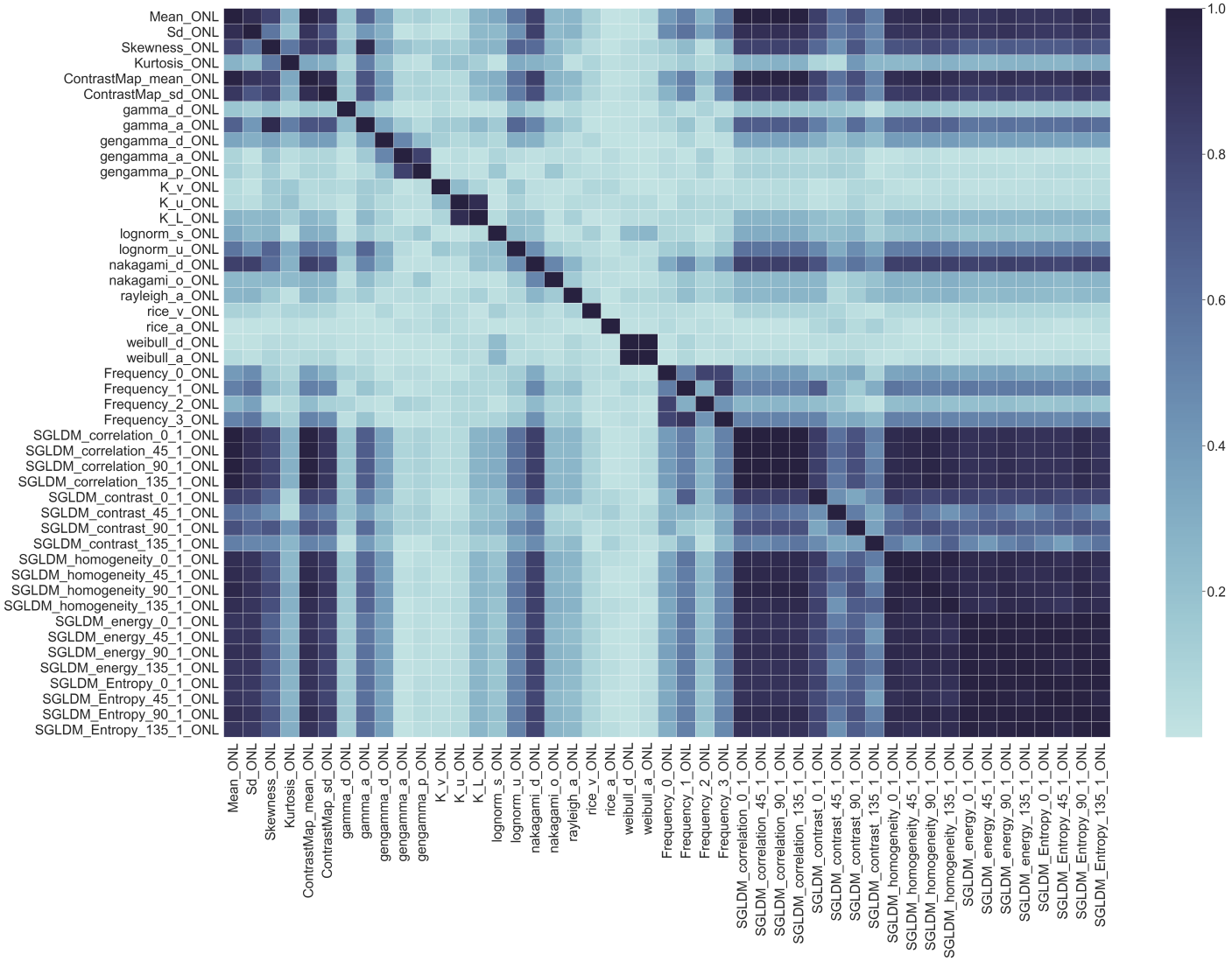

Figure C.8: Heatmap of the absolute value of Spearman's correlation values between features, for ONL.

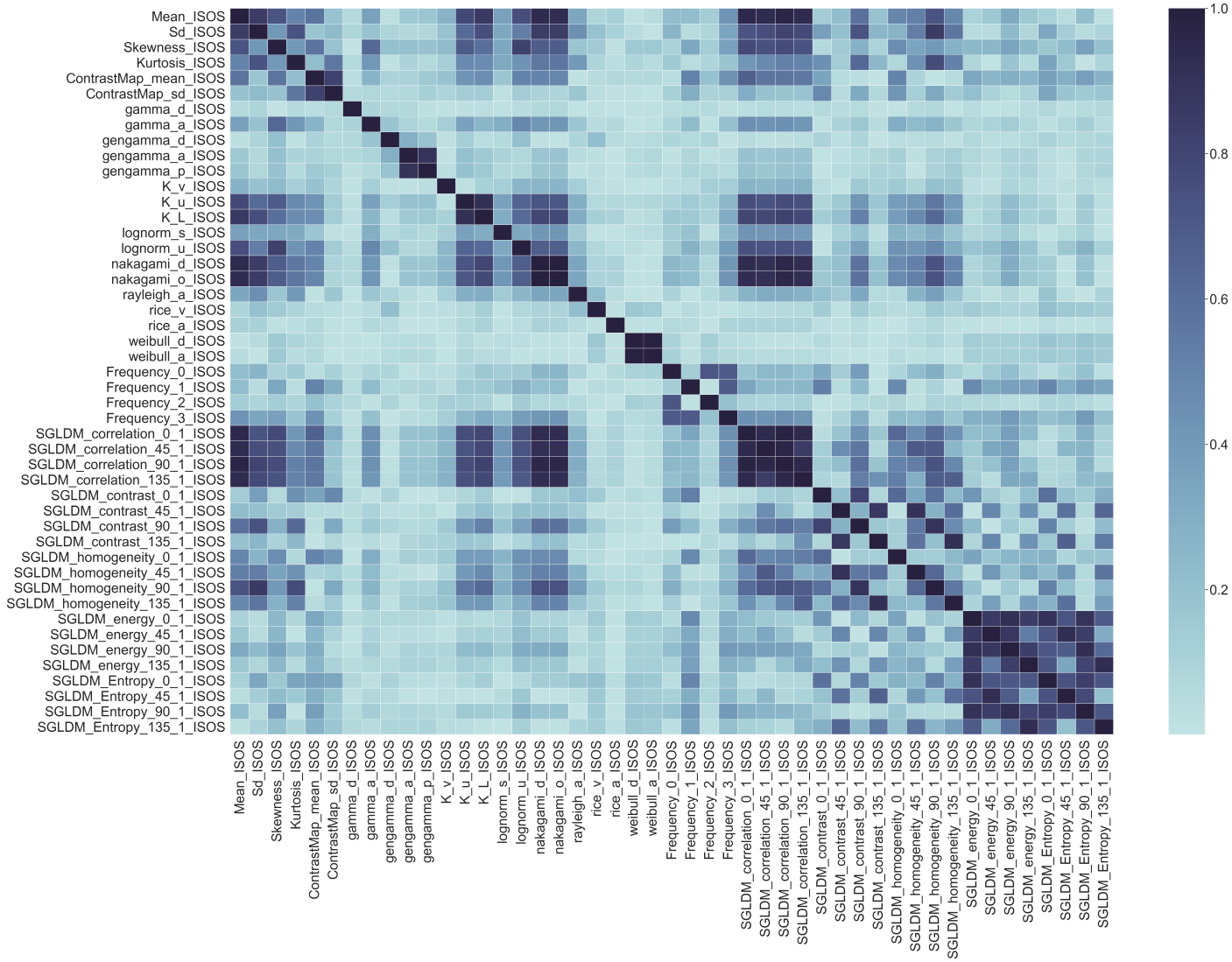

Figure C.9: Heatmap of the absolute value of Spearman's correlation values between features, for IS/OS.

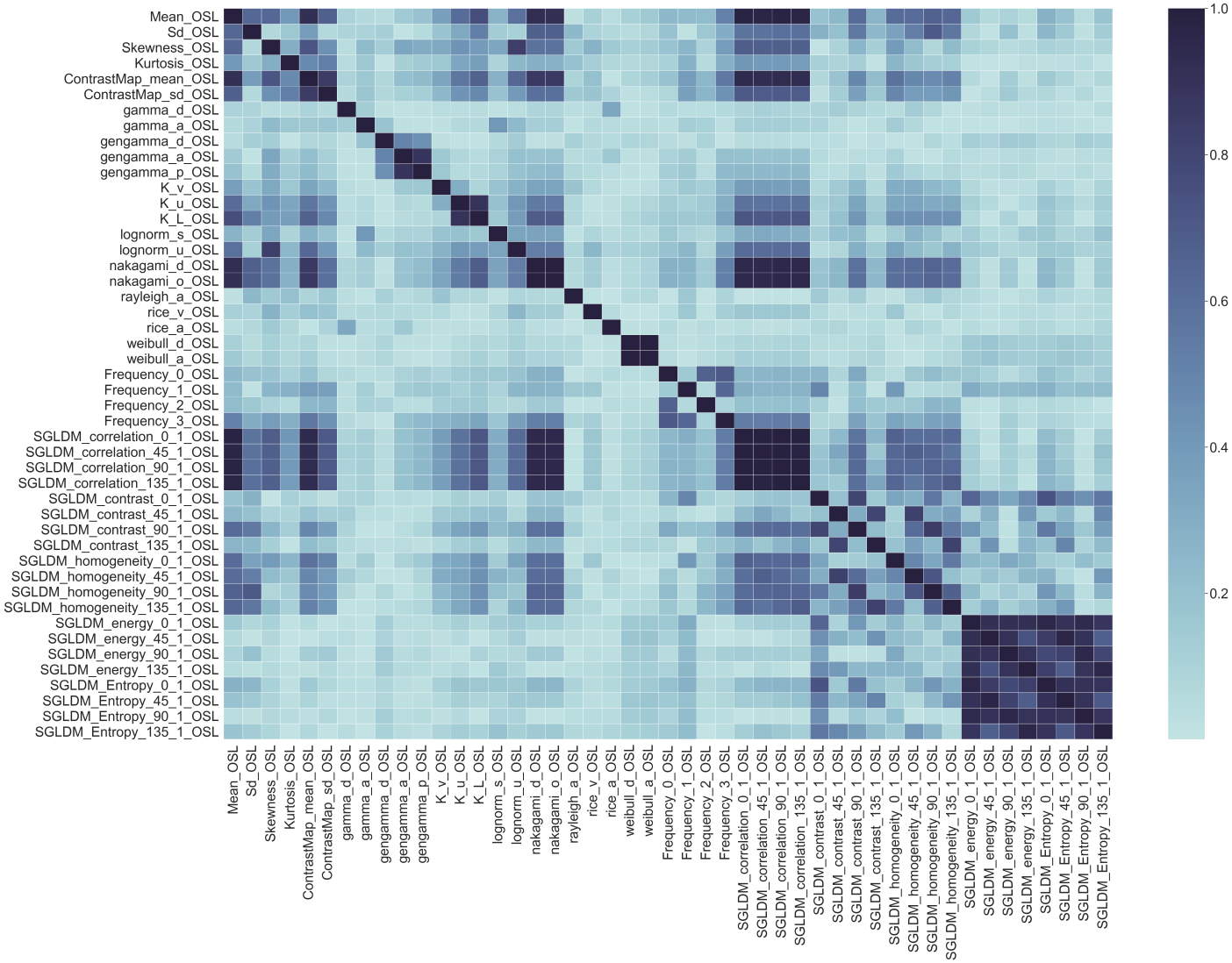

Figure C.10: Heatmap of the absolute value of Spearman's correlation values between features, for OSL.

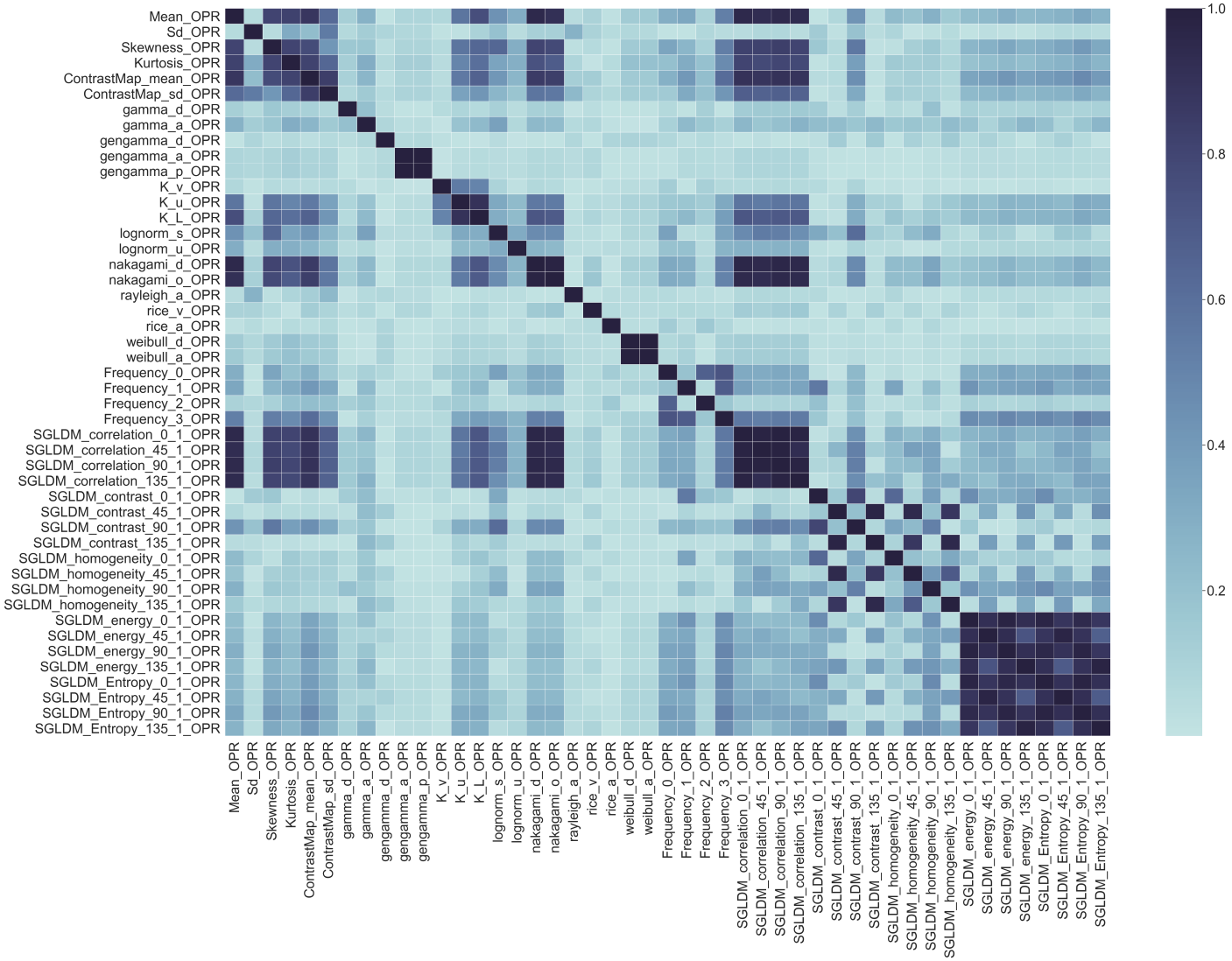

Figure C.11: Heatmap of the absolute value of Spearman's correlation values between features, for OPR.

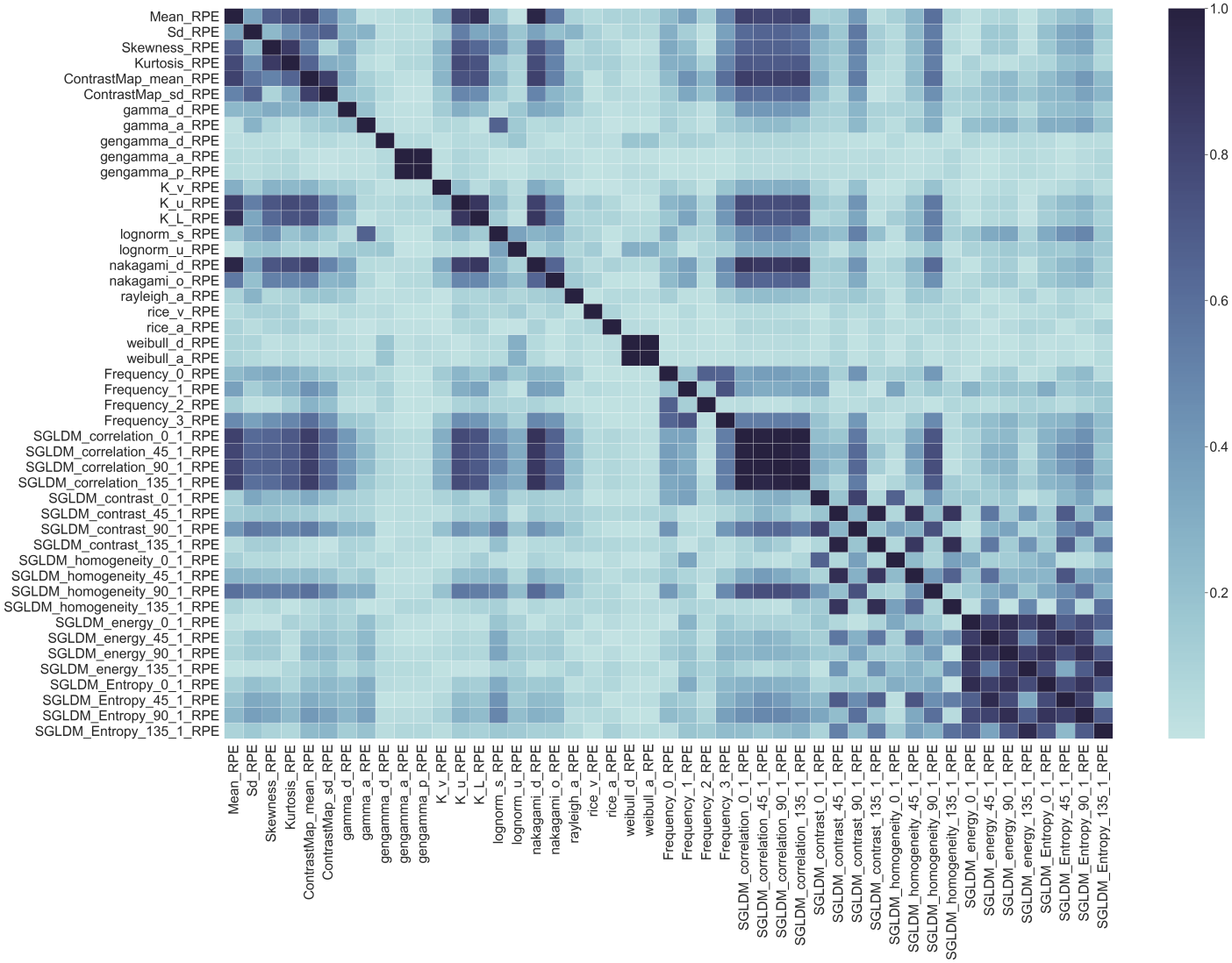

Figure C.12: Heatmap of the absolute value of Spearman's correlation values between features, for RPE.