

AMERICAN UNIVERSITY OF BEIRUT

AUTOMATED DETECTION AND
MEASUREMENT OF CORNEAL HAZE
USING OPTICAL COHERENCE
TOMOGRAPHY IMAGES

by

MANAL MOUIN CHOKR

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science
to the Department of Computer Science
of the Faculty of Arts and Sciences
at the American University of Beirut

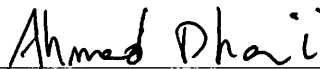
Beirut, Lebanon
May 2017

AMERICAN UNIVERSITY OF BEIRUT

AUTOMATED DETECTION AND
MEASUREMENT OF CORNEAL HAZE
USING OPTICAL COHERENCE
TOMOGRAPHY IMAGES

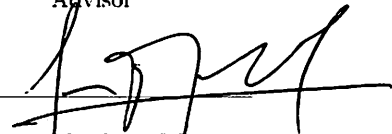
by
MANAL MOUIN CHOKR

Approved by:



Dr. Ahmad R. Dhaini, Assistant Professor
Computer Science

Advisor



Dr. George Turkiyyah, Professor
Computer Science

Member of Committee

Dr. Shady Elbassuoni, Assistant Professor
Computer Science

Member of Committee

Date of thesis defense: May 4, 2017

AMERICAN UNIVERSITY OF BEIRUT

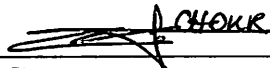
THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Chokr Manal Mouin
Last First Middle

Master's Thesis Master's Project Doctoral Dissertation

I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes after: **One ___ year from the date of submission of my thesis, dissertation or project.**
Two ___ years from the date of submission of my thesis, dissertation or project.
Three ___ years from the date of submission of my thesis, dissertation or project.

 08/05/2017
Signature Date

This form is signed when submitting the thesis, dissertation, or project to the University Libraries

Acknowledgements

I would like to thank all the people who contributed in some way to the work described in this thesis.

First and foremost, I would like to thank Professor Ahmad Dhaini, my thesis advisor. I am very grateful for his patience and support in overcoming numerous obstacles I have been facing through my research.

I would like to thank Dr. Shady Awwad and Dr. Maamoun Abdul Fatah for their cooperation which made this thesis possible with high quality results.

I would like to thank my committee members, Professor George Turkiyyah and Professor Shady Elbassuoni, I am gratefully indebted for their very valuable comments on this thesis.

I would like to thank Professor Abbas El Hakim for offering help in some material I did not have a background on.

I would like to thank my fellow graduate students for their feedback, cooperation and of course friendship.

I would like to thank my friends for accepting nothing less than excellence from me.

I must express my very profound gratitude to my beloved parents and brother and sister for supporting me spiritually throughout writing this thesis and my life in general.

Last but not least, I would like to thank, from the bottom of my heart, my closest friend Mohamad Hmede, who died last October, for his endless support and continuous encouragement in the last five years.

An Abstract of the Thesis of

Manal Mouin Chokr for Master of Science
Major: Computer Science

Title: Automated Detection and Measurement of Corneal Haze
using Optical Coherence Tomography Images

The cornea typically constitutes two-thirds of the eye's optical power; a healthy one has a dome shape. However, a vision disorder so-called Keratoconus, which is especially prevalent in the Middle East and Gulf area, may result in a cone shape of the cornea, leading to significant loss of visual acuity. As a healing procedure, ophthalmologists use a technique so-called corneal cross-linking to halt the progression of Keratoconus. One indicator of corneal cross-linking success is arguably the presence and depth of the stromal demarcation line. In addition, corneal haze beyond the demarcation line can be an ominous sign of loss of corneal transparency after cross-linking, which is a much-dreaded side-effect of the procedure. To date, to observe corneal haze, ophthalmologists use slit lamps and/or they observe Optical Coherence Tomography (OCT) micrometer resolution images of the corneal tissue. These techniques are subjective, time consuming and error-prone. In this thesis, we propose a novel technology to automatically detect and measure corneal haze and demarcation line in OCT images. To achieve so, we propose new image analysis algorithms and make use of existing libraries such as OpenCV; additionally, we propose a customized machine learning approach used for OCT image classification. The new automated tool provides the user with haze statistics as well as visual annotation reflecting the shape and location of the haze and demarcation line in the cornea; it also detects subtle changes in the cornea over time. Our experimental results and analysis demonstrate the efficacy and effectiveness of our new tool in accurately detecting and measuring the demarcation line depth, in comparison to manual measurements, in a much faster manner. The Intraclass correlation coefficients (ICC) and Pearson correlation coefficients (PCC) between the automated detection software and two human operators are measured as 0.945 and 0.951, and 0.910 and 0.918,

respectively. The ICC for inter-operators reproducibility is 0.882, and for intra-operator repeatability is 0.917. The ICC for inter-software repeatability is 1, and the average time spent by the operator per OCT examination is 31 seconds, as compared to 1 second by the software. We expect our new technology to help ophthalmologists have a better understanding of corneal haze, which will help in clinical decision making. Moreover, the software can be used as a standardized tool for corneal haze measurements or in longitudinal studies.

Contents

Acknowledgements	v
Abstract	vi
1 Introduction and Motivation	1
1.1 Goals of the Thesis	6
1.2 Organization of the Thesis	7
2 Related Work	8
2.1 Biomechanical Properties Measurement	8
2.2 Central Corneal Thickness Measurement	11
2.3 Automated Detection and Classification of Corneal Haze	12
3 Proposed Solution	15
3.1 Elimination of Invalid Cuts	17
3.1.1 Data Preparation	17
3.1.2 Feature Selection and Data Compression	18
3.1.3 Bias Variance Analysis	19
3.1.4 Experiments	21
3.1.5 Chosen Model and Results	24
3.2 Input Image	25
3.2.1 Finding the Corneal Region	26
3.2.2 Bright Region Detection	30
3.3 Computation of $S(c)$	33
3.3.1 Identifying Candidate Demarcation Line(s)	35
3.4 Minimum Score Value	37
4 Results and Analysis	41
5 Conclusions and Future Work	65
5.1 Conclusions	65
5.2 Future Work	66
5.2.1 Detection of Demarcation Line over Time	66
5.2.2 Processing 8mm/9mm OCT Images	67

5.2.3 3D Reconstruction of OCT Cuts	67
A Glossary	68
B Learning Tables	69

List of Figures

1.1	Anatomy of the human eye with a cross section fo the cornea. . .	2
1.2	On the left, picture of an eye with corneal haze. On the right, picture of a healthy eye.	3
1.3	Representative slitlamp biomicroscopy image of clinical demarcation line after CXL.	4
1.4	HD-OCT Anterior Segment Scan Review (Cube 512x128).	5
1.5	Corneal Image taken using Cirrus HD-OCT Scanner.	6
2.1	Positioning the Ruler for the Central Corneal Thickness Measurement.	11
2.2	(a) Original OCT Image; (b) Annotated Image.	12
2.3	(a) Original OCT Image; (b) Boundaries Detected.	13
3.1	An instance (i.e., OCT image of Cut 13) of the video taken using the OCT scanner. The image to the left one shows the location and the cut number the image to the right shows the transversal version of the cut.	16
3.2	Four examples of damaged cuts occurring in OCT images.	17
3.3	Illustration of bias and variance and the corresponding models typically used.	20
3.4	Learning curves of high bias and high variance datasets.	21
3.5	Learning curve of the training data of OCT images using Logistic Regression.	22
3.6	Region used for computing the Thresholding parameter θ	27
3.7	Image after applying the <i>Thresholding</i> technique.	28
3.8	Image after applying the <i>Contouring</i> method.	29
3.9	Image after blurring.	31
3.10	Image after applying the <i>Sobel</i> derivative.	31
3.11	Suggested demarcation line highlighted in color green, and candidate demarcation line in color red.	32
3.12	One cut of an OCT image showing the different regions used in the score computation.	34
3.13	Histogram of candidate demarcation line thickness.	37

3.14	Distribution of the Candidates and Suggested Demarcation Lines' Scores.	38
3.15	Probability Density Function of the Distribution of the Candidates and Suggested Demarcation Lines' Scores.	40
4.1	Depth of <i>suggested</i> and <i>candidate</i> demarcation line by the software and the measurements of the ophthalmologists on two occasions expressed in microns for all patients at 1-month and 3-months post-operation periods.	42
4.2	Differences of depths of the demarcation lines as measured by ophthalmologists and software in microns for all patients at 1-month and 3-months post-operation periods.	43
4.3	Raw and annotated cuts with depth in microns of demarcation line based on human and software operators for the patients with relatively large differences in measurements. The software suggested demarcation line is highlighted yellow, the candidate demarcation line is highlighted in orange when available. The demarcation lines measured by the first and second ophthalmologist are highlighted in green and in red, respectively. The annotation is made on the operator's chosen cut of the OCT image.	57
4.4	Depth of detected demarcation lines for different patients as measured by the software (in yellow), the first operator (in green) and the second operator (in red). The absence of a highlighted line indicates that a demarcation line was not detected by software / operator.	61
4.5	Raw and annotated cuts of detected demarcation lines for different patients as measured by the software (<i>suggested</i> , in yellow), the first operator (in green) and the second operator (in red) based on the 16 th percentile minimum score value. The absence of a highlighted line indicates that a demarcation line was not detected by software / operator.	62
4.6	Depth of suggested and candidate demarcation lines for different patients in microns and in percentage. The suggested demarcation line is highlighted in green and the candidates in red.	63
4.7	The progress in the depth of <i>suggested</i> demarcation line for one sample patient over time.	64
4.8	Intraclass correlation coefficients (ICC) of the demarcation lines' depths for all patients.	64

List of Tables

2.1	Summary of measurement techniques applied to determine the biomechanical properties of corneal tissue.	9
3.1	Measurements of the best performances of the selected models. . .	25
3.2	Results and hyperparameters of the Support Vector Machine model on testing data.	25
B.1	Logistic Regression	69
B.2	SMO with a Polynomial Kernel of degree 1 and Complexity Constant 'c' = 1	70
B.3	SMO with a Polynomial Kernel of degree 1 and Complexity Constant 'c' = 10	71
B.4	SMO with a Polynomial Kernel of degree 1 and Complexity Constant 'c' = 100	72
B.5	SMO with a Polynomial Kernel of degree 2 and Complexity Constant 'c' = 1	73
B.6	SMO with a Polynomial Kernel of degree 2 and Complexity Constant 'c' = 10	74
B.7	SMO with a Polynomial Kernel of degree 2 and Complexity Constant 'c' = 100	75
B.8	Multilayer Perceptron with 'a' Hidden Layers	76
B.9	Multilayer Perceptron with 'i' Hidden Layers	77
B.10	Multilayer Perceptron with 't' Hidden Layers	78
B.11	Multilayer Perceptron with 'o' Hidden Layers	79

Chapter 1

Introduction and Motivation

The cornea is the transparent front part of the eye that covers the iris, pupil, and anterior chamber; this transparent layer, along with the anterior chamber and lens, refracts light. The cornea accounts for approximately two-thirds of the eye's total optical power [1]. As illustrated in Figure 1.1, it has five layers:

- a. Corneal Epithelium: this is an exceedingly thin tissue of fast-growing and easily regenerated cells. This layer is approximately 50 microns (i.e., μm) thick.
- b. Bowman's Membrane: this is a tough layer composed of collagen that protects the corneal stroma. This layer is 8 to 14 microns thick.
- c. Corneal Stroma: this is a thick, transparent middle layer, consisting of regularly arranged collagen fibers. Up to 90% of the corneal thickness is composed of stroma. This layer is around 500 microns thick.
- d. Descemet's Membrane: this is a thin acellular layer that serves as the modified basement membrane of the corneal Endothelium, from which the cells are derived. This layer is 10 to 12 microns thick.
- e. Corneal Endothelium: this is a simple monolayer, of approximately 5 μm thick. Unlike the corneal epithelium, the cells of the endothelium do not regenerate. Instead, they stretch to compensate for dead cells, which reduces the overall cell density of the endothelium. If the latter can no longer maintain a proper balance, loss of transparency will occur, thereby impairing the image formed.

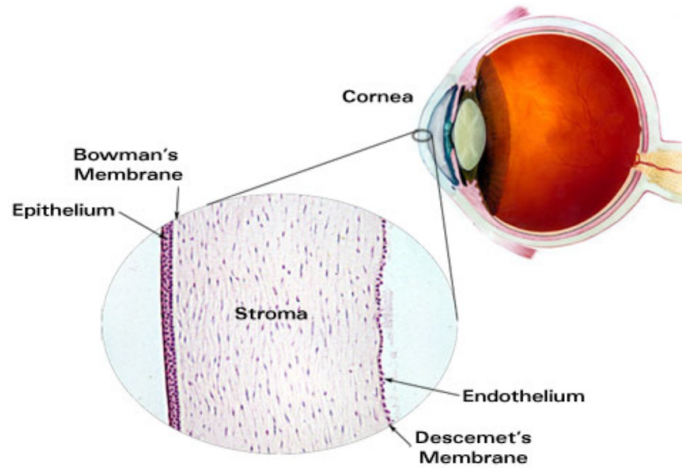


Figure 1.1: Anatomy of the human eye with a cross section for the cornea.

The cornea is one of the most sensitive tissues of the body. In [2], Kling et al. affirm that the corneal shape is a determinant of ocular refraction, which is determined by its biomechanical properties. Furthermore, Elsheikh et al. used human donor eyes to show that the contribution of the Epithelium to corneal stiffness is much lower to that of the stroma; therefore, it can most likely be neglected [3]. Hence, the stroma representing the largest part of the cornea is the layer mainly defining the biomechanical properties of the cornea.

Keratoconus is an ectatic disorder that affects the cornea; it is characterized by progressive thinning and steepening of the cornea, which leads to apical scarring and eventual loss of visual acuity [4, 5, 6]. This eye condition affects approximately 1 in every 2000 persons. More importantly, the prevalence of this phenomenon in Lebanon has been estimated to be around 3.3% when studied on a medical students sample population [7].

Corneal collagen cross-linking (CXL), which was first described by Wollensak et al. [8], is a surgical procedure that aims at halting or slowing the progression of Keratoconus. It induces the formation of additional covalent bonds within collagen fibrils using photopolymerization. As a result, it increases the biomechanical stiffening and the biochemical resistance of the cornea [9]. CXL is thought to be the most promising treatment modality in the progression of Keratoconus [10].

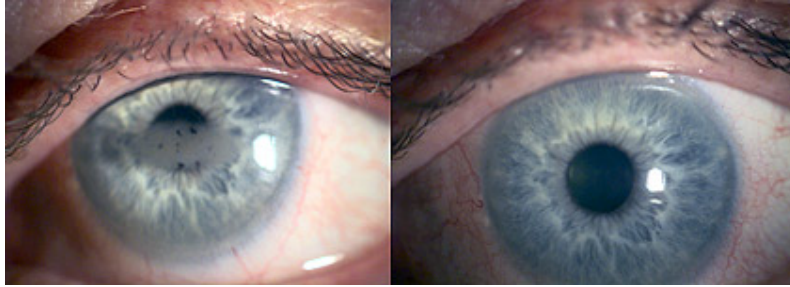


Figure 1.2: On the left, picture of an eye with corneal haze. On the right, picture of a healthy eye.

Typical corneal haze has generally been observed on clinical examination as a main effect of CXL, as reported by Greenstein et al. [11]. As depicted in Figure 1.2, haze can be defined as a cloudy or opaque appearance of the cornea, as well as a stromal demarcation line, shown in Figure 1.3. It is most often caused by inflammatory cells and other debris activated during trauma, infection, or eye procedures and surgeries such as corneal cross-linking for keratoconus eyes and kerato-refractive surgeries like photo-refractive keratectomy (PRK). The cornea is normally clear, and corneal haze can occur in any part of the cornea; thus it can greatly impair vision. The demarcation line marks the transition from cross-linked to non-cross-linked tissue; it can be observed using a slit lamp (see Figure 1.3). The stromal demarcation line appears as an anterior stromal zone with higher reflectivity up to a depth of approximately $300 \mu\text{m}$.

The backscattered light from a normal (clear) cornea is typically low; nonetheless, corneal haze, after corneal dystrophies or after injury, can be associated with pathology that often indicates the corneal structures responsible for poor vision. Therefore, changes in haze can be used to track progression of disease or response to treatment such as cross-linking. More importantly, the same processes that produce haze are often responsible for forward scatter that degrades vision; therefore, observation of haze has been very valuable to ophthalmologists.

To date, clinical instruments for measuring haze include slit-lamp biomicroscopy, clinical confocal microscopes, and Scheimpflug cameras. The spatial resolution needed to identify the source of the backscattered light varies among

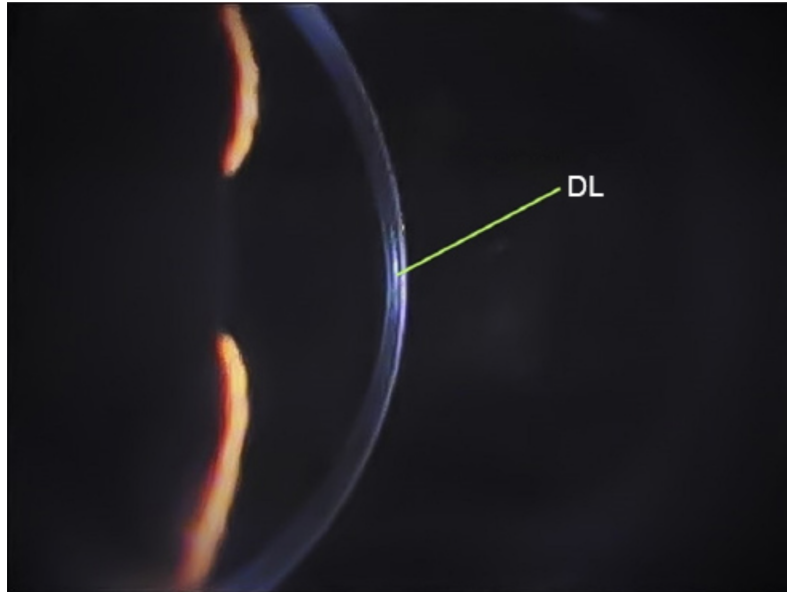


Figure 1.3: Representative slitlamp biomicroscopy image of clinical demarcation line after CXL.

instruments. Some slit-illumination instruments can only resolve approximately two thirds of the full corneal thickness. For instance, the Cirrus HD-OCT is a spectral-domain Optical Coherence Tomography (OCT) scanner that is primarily used for imaging and measuring the structures in the posterior segment of the eye. However, by changing the focus of the OCT beam, it can also be used to measure the structures in the anterior segment of the eye such as the cornea. This scanner operates in-vivo and employs axial cross-sectional imaging [12]. An OCT scanner is often likened to an Magnetic Resonance Imaging (MRI) or X-ray of the eye. This procedure is currently the only one that shows in-depth images of the eyes internal structures.

By scanning the anterior segment of the eye, HD-OCT generates a volume of data through a 4 millimeter square grid, by acquiring a series of 128 horizontal scan lines each composed of 512 amplitude-scans. As illustrated in Figure 1.4, the analysis screen for the segment cube of dimensions 512x128 shows the scan area, the 4mm square covered, and scan navigators superimposed. In the upper OCT image on the right, the X slice is shown; and the Y slice is shown below it. As shown in Figure 1.5, the corneal epithelium, Bowman's membrane and stroma

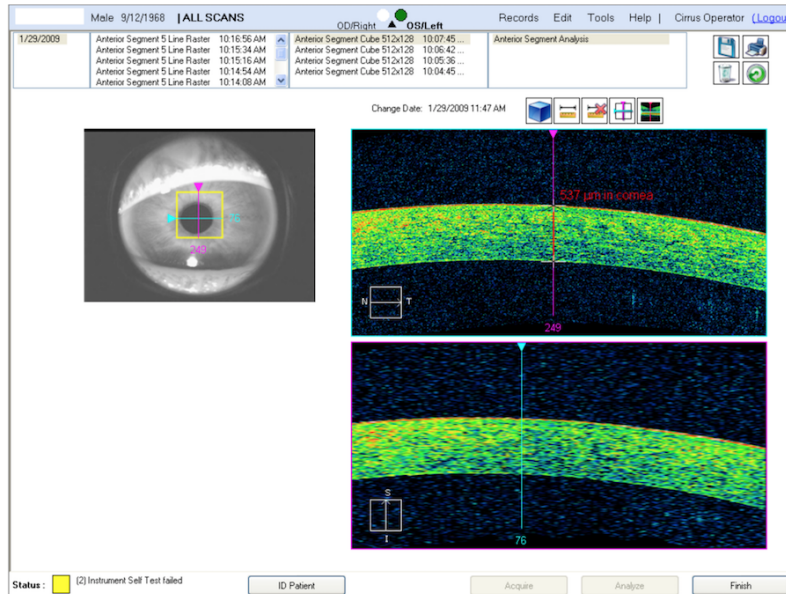


Figure 1.4: HD-OCT Anterior Segment Scan Review (Cube 512x128).

are generally visible in the tomograms generated by the HD-OCT scanners.

The presence and depth of the demarcation line is an easy and useful tool for clinicians for assessing the depth of the cross-linking effect, and in comparing the efficacy of the many variants of the procedure, which have been practiced in recent years [13, 14, 15, 16, 17]. A demarcation line can be accurately detected and measured as a hyper-reflective line within the corneal stroma by corneal and anterior segment OCT [18, 19]. However, the problem with evaluating the presence and depth of the demarcation line on OCT is that it is tedious and time consuming, as the clinician or an experienced operator has to go through many OCT sections of the cornea to determine the demarcation line among other reactive and non-specific hazy lines that can mimic the actual demarcation line. More importantly, the whole evaluation is in part subjective and operator-dependent, with intra-observer repeatability and inter-observer reproducibility still to be investigated. Therefore, quantification of corneal haze by means of an objective method that can adequately assess the overall amount of opacification would lead to better understanding of the phenomenon.

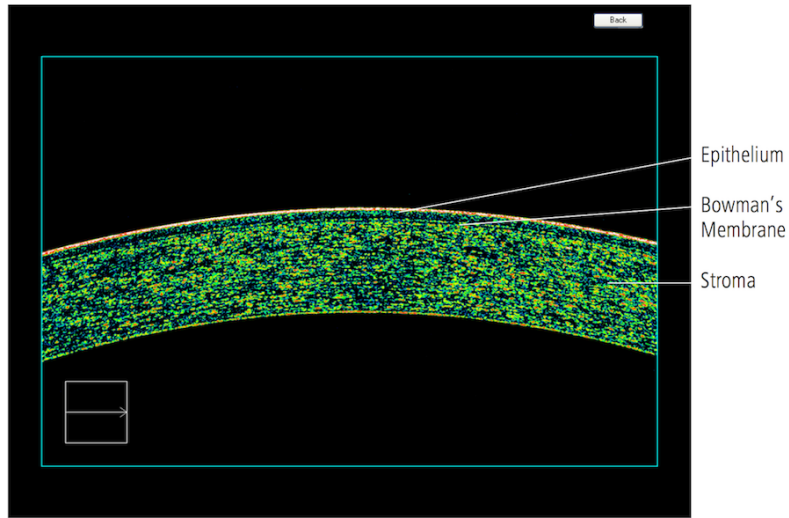


Figure 1.5: Corneal Image taken using Cirrus HD-OCT Scanner.

1.1 Goals of the Thesis

The main goal of this thesis is to design, implement and experiment with the first technology that offers the measurements needed for corneal haze and demarcation line examination, in a fast, accurate, automated and objective manner. Using image analysis and machine learning libraries such as OpenCV [20], Weka [21], new image analysis techniques, and a customized machine learning approach, our technology shall take as input an OCT scan of the anterior segment of the cornea and provide as output an annotated image exhibiting the detection and measurement of the corneal demarcation line, as well as haze statistics.

The new technology has the promise to be employed as standardized method of care for stromal haze measurement of individual patients or in aggregate data for the purpose of longitudinal studies. The novel technology will enable ophthalmologists to obtain fast objective detection and measurement of corneal haze, which may improve clinical decision making after corneal surgeries such as cross-linking. Using our tool, automated detection and measurement of the demarcation line depth can become "the standard of care" in cross-linking surgery in assessing treatment success. By the same token, software analysis of stromal haze could potentially become instrumental in objectively assessing cross-linking side effects, notably corneal haze and loss of translucency, with impact vision.

1.2 Organization of the Thesis

The remainder of this thesis is organized as follows. Chapter 2 surveys related work on corneal haze detection and measurement. Chapter 3 presents a detailed description of the novel technology. First, we discuss the input validation and processing using Machine Learning. Second, we explain in details the image analysis to detect and measure the demarcation line. Third, we discuss the statistical analysis. Lastly, we discuss the demarcation line validation. We discuss in Chapter 4 our experimental results and analysis. In Chapter 5, we conclude our work and discuss future extensions. In Appendix A, we provide a glossary of the abbreviations used in the text, sorted alphabetically. In Appendix B, we provide an extensive list of the machine learning results.

Chapter 2

Related Work

After corneal cross-linking treatment, the largest post-surgical change in the biomechanical properties of the cornea is observed to occur within one week post surgery [22]. Multiple imaging techniques have been so far used to assess the presence, and measure corneal haze.

2.1 Biomechanical Properties Measurement

Patient-specific modeling has been a necessity after observing that corneal stiffness and thickness vary between individuals [2]. Multiple techniques for the measurement of biomechanical properties have been used clinically. Table 2.1 shows the primarily used techniques along with some of their characteristics. Studies that are in-vivo are those, where the experiments are tested on living organisms (animals or humans), whereas studies that are ex-vivo are those, where the measurements are done in or on a tissue from an organism in an external environment. Details about these techniques are provided as follows:

a. **Strip extensometry:**

Stress-strained extensometry is the gold standard method in engineering to measure the macroscopy mechanical properties. In this technique, a strip of corneal tissue with a constant width is dissected and attached to the grips of a slow rate tension machine while monitoring its behaviour [23]. Furthermore, the length of the specimen, being part of a spherical surface, along its longitudinal centreline is longer than along the sides. This variation in

Table 2.1: Summary of measurement techniques applied to determine the biomechanical properties of corneal tissue.

	Static	Dynamic	Invasive
Strip extensometry	Yes	Yes	Yes
Brillouin microscopy		Yes	
Air-puff systems		Yes	Minimally
Elastography		Yes	Minimally
Enzymatic Digestion	Yes		Yes

Ex vivo	In vivo	Macroscopic	Microscopic
Yes		Yes	
Yes	Yes		Yes
Yes	Yes	Yes	
Yes	Yes		Yes
Yes		Yes	

specimen length inevitably leads to a non-uniform stress distribution across the width of the specimen. As a result, the corneal deformation can not be separated from scleral deformation; in addition, only rough estimation of the average macroscopic corneal stiffness can be obtained.

b. Brillouin microscopy:

Brillouin microscopy consists of a confocal microscope with numerical aperture (NA) of 0.3 and a high-resolution optical spectrometer [24, 25]. Brillouin light scattering arises from the interaction between photons and acoustic phonons in a sample; in other words, it is based on the propagation of thermodynamic fluctuations. The feasibility of Brillouin microscopy for mapping corneal modulus in three dimensions with a high spatial resolution has been demonstrated [26], offering a spatially resolved map of corneal stiffness. This technique does not suffer from any major disadvantage allowing accurate measurement of corneal biomechanical properties.

c. Air-puff tonometry and related systems:

In air-puff tonometry, a central area of the cornea is deformed by an air-pulse of linearly increasing force. The instant of applanation is determined by a monitoring system that senses the light reflected from the corneal

surface [27]. The required air pressure to applanate the corneal curvature is considered to be equivalent to the intraocular pressor. However, the measured properties are only geometrical and not directly related to actual biomechanical parameters. Hence, only minor to no changes can be detected. Consequently, more detailed analysis of the recorded deformation profile is required. Recently, air-puff tonometry was combined with high-speed Scheimpflug and OCT imaging, allowing for the first time to capture both the complete temporal and spatial deformation profile of the cornea during the air-puff event.

d. **Elastography:**

In elastography, ultrasonic waves are typically used to induce microscopic strains in corneal tissue. The resulting shear wave is then directly recorded by ultrasonic imaging; thus, corneal elasticity is measured [28]. However, in this method, contact with a coupling liquid is required; thereby, making it uncomfortable for the patient. Recently, non-contact elastography based on OCT imaging has been developed [29]. However, since light has less impact on the corneal tissue than ultrasound, low-amplitude elastics waves need to be induced using a different source.

e. **Enzymatic digestion:**

The speed of digestion of certain enzymes can be used to infer the original stage of crosslinking and estimate corneal stiffness [30]. Using selective enzymatic digestion would allow studying the impact of certain components on the resulting biomechanical properties. Nevertheless, enzymatic digestion is not a direct measurement of corneal stiffness.

In summary, we can observe that good biomechanical properties measurement techniques already exist; however, most of the techniques that are either macroscopic or invasive are now combined with OCT imaging. As a result, Optical Coherence Tomography has been primarily employed to examine the ocular fundus and anterior segment in a noncontact and noninvasive way [31].

2.2 Central Corneal Thickness Measurement

One of the functionalities implemented in the HD-OCT scanner is the ability to measure the Central Corneal Thickness (CCT). An example of such measurement is shown in Figure 2.1.

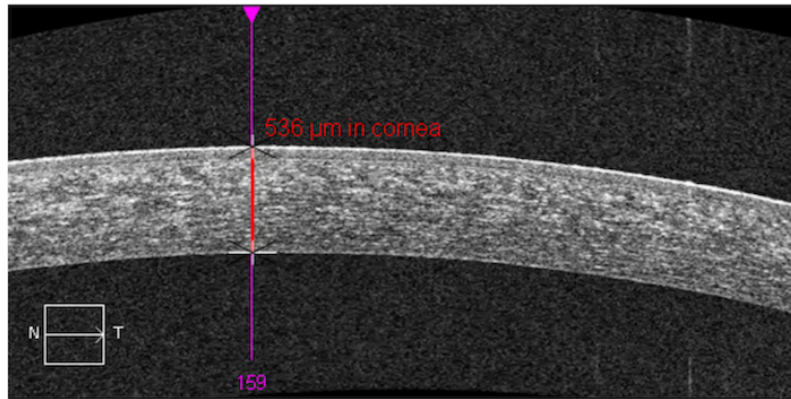


Figure 2.1: Positioning the Ruler for the Central Corneal Thickness Measurement.

The repeatability and reproducibility of the Cirrus HD-OCT instrument measurements of CCT were studied in two phases. In the first phase, 28 subjects were enrolled; each subject was imaged 3 times by one operator. The second phase enrolled 22 subjects; during this phase, each subject was imaged three times by three different operators. In these phases, different subjects were examined. The Cirrus HD-OCT repeatability is expressed as $4.08 \mu\text{m}$ standard deviation; whereas, the Cirrus scanner reproducibility is expressed as $4.23 \mu\text{m}$. The mean thickness of the cornea as measured in the first phase is $544.25 \mu\text{m}$, in the second phase, the mean thickness is $532.25 \mu\text{m}$; and the overall thickness is $538.25 \mu\text{m}$ [12]. The mean difference in CCT measurements between Cirrus HD-OCT and Ultrasound Pachymetry was computed [32]. This study enrolled a total of 50 subjects. The measurements were made by a single operator for each device. The results show a mean difference equal to $-9.06 \mu\text{m}$ with a standard deviation equal to 5.63 . A negative difference indicates that the Cirrus CCT measurement is thinner than the ultrasound CCT measurement.

OCT devices in general measure thinner than ultrasound pachymetry. In the

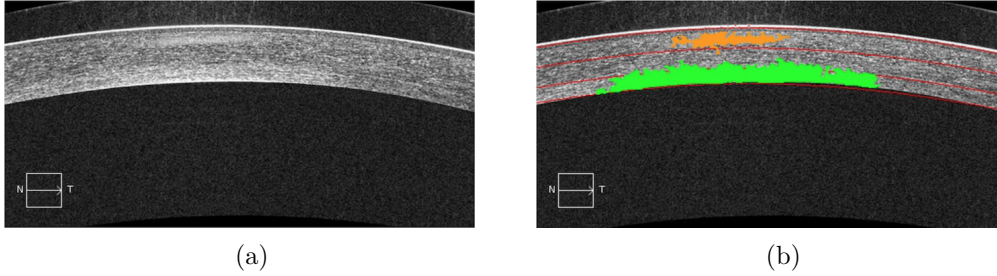


Figure 2.2: (a) Original OCT Image; (b) Annotated Image.

literature, the reported absolute differences between OCT and ultrasound pachymetry range from 11.64 to 49.4 microns (i.e. μm) [33, 34].

The results of the CCT measurement are considered to be relatively good. Nonetheless, this process is manual; therefore, it is error-prone and subjective. Hence, an automated objective mechanism to accurately measure the central corneal thickness can be of great help for the ophthalmologists whilst examining patients.

2.3 Automated Detection and Classification of Corneal Haze

Oud et al. offered the first automated objective detection and classification of corneal haze [35], where the cloudy opaque corneal haze appears in the cornea as dense regions of bright pixel intensity. The proposed novel software solution takes as input an OCT image of the anterior segment of the cornea (see Figure 2.2a), and provides as output the OCT image with visual annotations that reflect the detection and classification of the haze (see Figure 2.2b).

The detection and classification of corneal haze in an OCT image is achieved by the proposed software via the following steps:

- a. Detecting corneal boundaries (excluding the Epithelium's bright surface) (see Figure 2.3)
- b. Computing corneal properties such as area, intensity, and average brightness

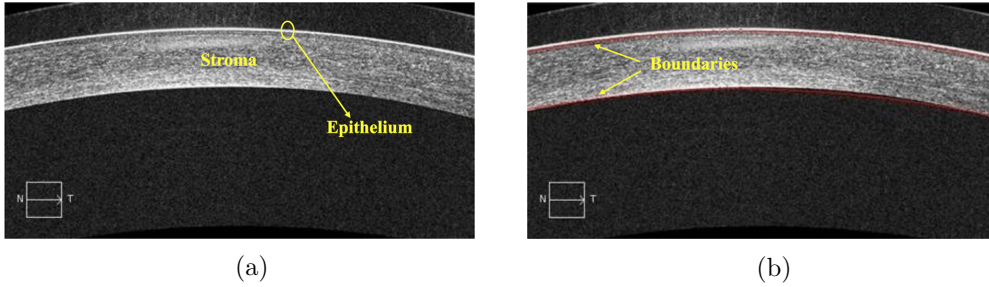


Figure 2.3: (a) Original OCT Image; (b) Boundaries Detected.

c. Splitting the cornea into three regions:

- (a) Anterior Stroma
- (b) Mid Stroma
- (c) Posterior Stroma

d. Detecting haze regions

e. Aggregating close haze regions

f. Classifying haze into corresponding region(s)

A study was conducted to experiment on the new technology proposed. In this study, 44 patients were enrolled whom underwent crosslinking for keratoconus. OCT images were collected at pre-operation and then at 2-weeks, 3-months, 6-months and 12-months post-operation. The average brightness of different corneal regions for all patients, collected at different periods, shows that the most significant changes occurs in the anterior stroma region 2-weeks post-operation. This phenomenon is actually typical after corneal collagen cross-linking: the greatest difference appears within two weeks after the surgery and the anterior region is the region affected the most by corneal haze.

The work that has been done by Oud et al. was the first to tackle the problem of automating the procedure of detection and classification of corneal haze. However, it did not address the problem of demarcation line detection and depth measurement in a similar manner: automated and objective. This is of utmost importance as the presence and depth of the demarcation line is an easy and useful tool for clinicians for assessing the depth of the cross-linking effect, and

in comparing the efficacy of the many variants of the procedure, as practiced in recent years [13, 14, 15, 16, 17].

Hence, in this work, we propose a novel automated objective detection of demarcation line. Our proposed solution make use of some of the methods proposed by Oud et al. Thus, our solution can be integrated with the solution proposed in [35] to be used as a unified tool for detecting and measuring corneal haze and demarcation line in OCT images.

Chapter 3

Proposed Solution

The demarcation does not normally appear in corneal hazy images. However, when it does, it indicates that the eye is healing as expected. Thus, automatically detecting it (when available) and measuring its corneal depth has important clinical indications.

Using the OCT scanner, multiple OCT images are typically captured by the experienced operator for the same eye (in the same period) in the form of a video of 128 cuts as illustrated in Figure 3.1. For analyzing corneal haze, only the middle cut (i.e., cut 64) is needed to detect and classify haze, since it does not vary across cuts. However, detecting and measuring the depth of the demarcation line requires the cut in which the demarcation line appears the "clearest". To achieve so, we split each given video into 128 images and apply a scoring scheme on each image. Here, a demarcation line is considered a candidate if it consistently appears across multiple cuts (e.g., in at least 15% of the cuts).

To account for the fact that the demarcation line may not typically exist at the same depth across different cuts, we consider a demarcation line as being the same one across cuts if it appears at a fluctuating depth of a certain range, which is obtained statistically from the available images. Consequently, the candidate lines that satisfy the property of continuation are left as candidate demarcation lines, and the one that, after running our algorithm, gets the highest score, is picked as suggested demarcation line; subsequently, its corresponding cut is taken as the best cut.

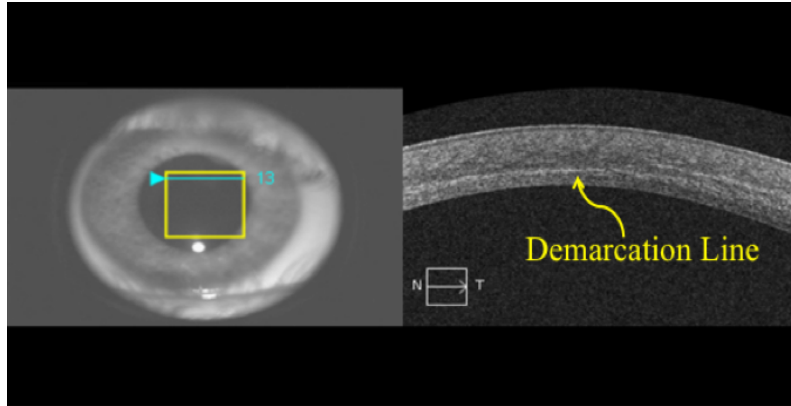


Figure 3.1: An instance (i.e., OCT image of Cut 13) of the video taken using the OCT scanner. The image to the left shows the location and the cut number, the image to the right shows the transversal version of the cut.

In summary, the detection and measurement of demarcation line in OCT image is achieved by the new software via the following steps:

- a. Loading the OCT video into the new software and dividing into 128 image cuts
- b. Filtering out invalid cuts
- c. Detecting the boundaries of the cornea
- d. Dynamically detecting and measuring the thickness of the Epithelium
- e. Computing a score for every candidate line across all valid cuts
- f. Determining the lines that should be taken as candidate demarcation line(s)
- g. Selecting the top two candidate lines
- h. Reporting the line with the highest score as "suggested" and second best as "candidate"

Details about the steps applied by the new solution to detect and measure the demarcation line depth, are provided in the following sub-sections.

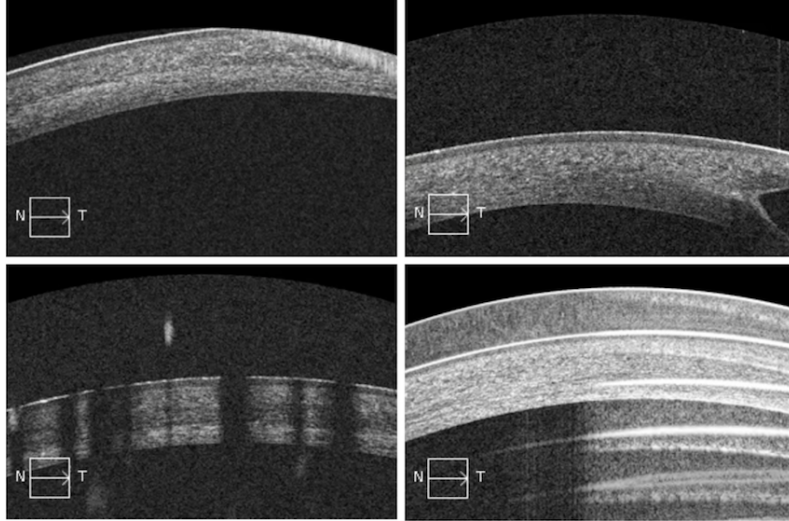


Figure 3.2: Four examples of damaged cuts occurring in OCT images.

3.1 Elimination of Invalid Cuts

As illustrated in Figure 3.2, some OCT images/videos may have (out of the 128 cuts) some damaged or scrambled cuts. However, this should not disqualify the whole image, since other cuts of the same image may be valid. Nevertheless, damaged cuts must be filtered out; otherwise, they would falsify the output of the software. To address this problem, we use Machine Learning to classify each cut in an OCT image as valid or invalid for demarcation line detection.

3.1.1 Data Preparation

Our data is composed of 8,064 cuts taken from 63 OCT images. Each cut is represented by 36,450 features such that each cut is of size 270x135 pixels, and where each pixel is depicted by its RGB value. Note that the width and height of each image have been reduced to half so as the machine training time is reduced, as this may take weeks to finish.

Each cut is annotated by class "0" as invalid, and class "1" as valid. This annotation was done manually and carefully in order to produce accurate learning. Note that with our data, we observed 14.7% of the instances classified as invalid.

3.1.2 Feature Selection and Data Compression

We use OpenCV to extract raw features from the images. Due to the large feature space (36,450 features per instance), feature reduction is employed to guarantee good generalization of the learnt models and reduce training time. To achieve the latter, the Principal Component Analysis (PCA) method [36] is employed. PCA is one of the most commonly used methods for dimensionality reduction in machine learning. We reduce the number of features by combining them; that is, we project the high-dimensional data into a low dimensional space without impairing the variance in the data. PCA requires defining a set of principal components as follows: the 1st component e_1 represents the direction of the greatest variance in the data, and the 2nd component e_2 represents the second greatest direction of the variance, perpendicular to e_1 , and so on until generating d principal components (where d is defined based on the analyzed data). The first m principal components become the new m dimensions, such that m is picked based on the required variance.

To perform PCA, we first split our dataset into two subsets: a *training* set consisting of 80% of the images and a *testing* set consisting of the remaining 20% of images. PCA is then applied on the training or validation set only and the first m eigen vectors that preserve 95% of the data variance are retrieved. Since PCA is cubic in the number of original dimensions, it did not scale well. To address this problem, we first perform feature selection by using the Information Gain (InfoGain) criteria for ranking the features with respect to the accuracy of the different prediction tasks. For each feature x_i , we compute its gain value $IG(x_i)$ using decision trees as follows:

$$IG(x_i) = E(t) - \sum_{j=1}^{m_i} \frac{N_j}{N} E(v_j) \quad (3.1)$$

where $E(t)$ and $E(v_j)$ are the entropy at nodes t and v_j , respectively; N_j and N are the number of instances at nodes t and v_j , respectively. The goal is to maximize the info gain; equivalent to minimizing the average impurity, also known as entropy, defined as follows:

$$E(t) = - \sum_{k=1}^K p(k|t) \log_2 p(k|t) \quad (3.2)$$

where $p(k|t)$ is the fraction of instances that belong to class k at node t .

That is, after applying InfoGain, we get list of features ranked according to their influence on the prediction accuracy [37].

Due to the large dataset size, it was not possible to test on different cut-off points. However, the top 1/5 of the features (i.e., 8,192 features per instance) as ranked by InfoGain, performs well when later combined with PCA. Therefore, the top 5th is used as our cut-off.

Hence, we use these 8,192 features to represent an image instead of the original 36,450 features.

Once each image is represented using these features, PCA is applied to further dimensionality reduction, ending up with 172 features to represent each image. Note that both the feature selection and PCA in this case were also applied on the training set only. The learnt 172 Eigen vectors along with InfoGain ranking resulting from applying this two subsequent steps are then used to transform each test image into the new 172-dimensional space.

3.1.3 Bias Variance Analysis

To improve data fitting and train highly accurate models, it is necessary to understand the different sources of error leading to bias and variance [38].

We denote the class we are trying to predict as y and our features as vector X . We assume that there is a function f such that $y = f(X)$. Hence our goal is to estimate a model $\hat{f}(X)$ that emulates $f(X)$. For a feature vector X , the expected out-of-sample squared prediction error is:

$$Err(X) = E[(y - \hat{f}(X))^2] \tag{3.3}$$

This error is also decomposed into bias and variance terms:

$$Err(X) = (E[\hat{f}(X)] - f(X))^2 + E[(\hat{f}(X) - E[\hat{f}(X)])^2] \tag{3.4}$$

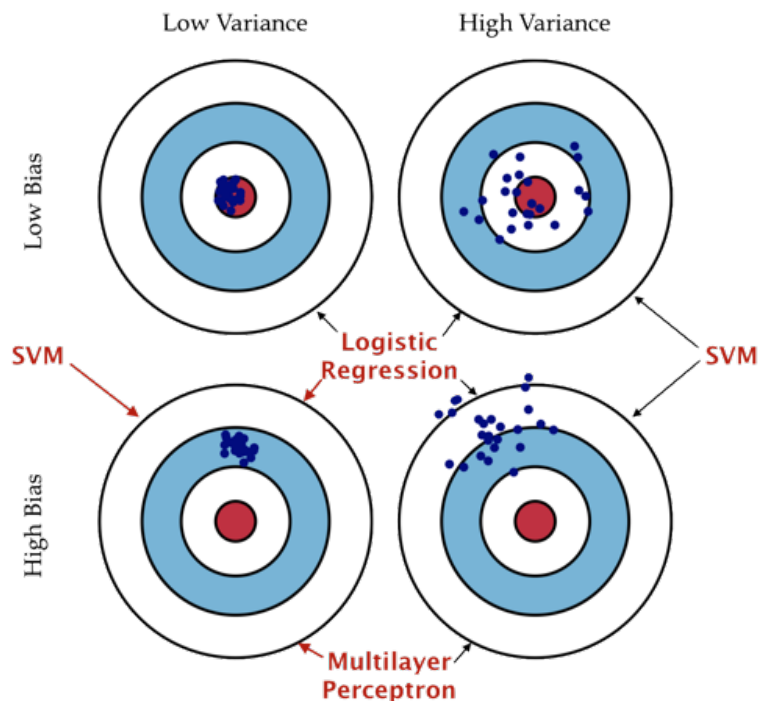


Figure 3.3: Illustration of bias and variance and the corresponding models typically used.

$$Err(X) = bias^2 + variance \quad (3.5)$$

Our purpose is to reduce both the bias and the variance. The error due to bias is taken as the difference between the expected (or average) prediction of our model and the correct value that we are trying to predict; whereas the error due to variance is taken as the variability of model prediction for a given data point. The variance is how much the predictions for a given point vary among different realizations of the model.

In order to identify whether the data suffers from high bias or high variance (illustrated in Figure 3.3), which affects the model employed for the training phase, we plot the learning curve. In order to do so, we start by creating samples of the original dataset: 10% 20% ... 90% and 100% of the training data. We then pick a very simple learning model, say linear regression or logistic regres-

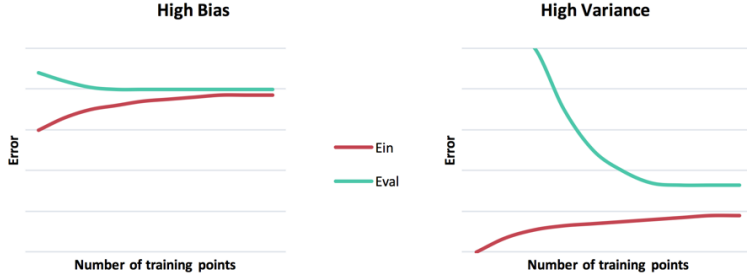


Figure 3.4: Learning curves of high bias and high variance datasets.

sion, and train it. Consequently, we measure the error on the validation dataset. We plot the curve of E_{in} , the measured error in the training data, and of E_{val} , the measured error when using the validation data, evaluated as the number of misclassified instances, with respect to the different dataset sizes.

As illustrated in Figure 3.4, the learning curve of high bias data is categorized by a small gap between E_{in} and E_{val} , and high values for both E_{in} and E_{val} , whereas the learning curve of high variance data is categorized by a constant decrease of E_{val} , and a value of E_{in} that is much lower than E_{val} .

Since validation sets have no longer been used as standard procedure in the literature, we do not consider them. Instead, we use 10-folds cross validation [39]. As seen in Figure 3.5, our data suffers from high bias; thus, logistic regression, SVM and multilayer perceptron are valid models that can provide accurate classification.

3.1.4 Experiments

The experiments were done on the training data only, using the machine learning library Weka [21]. Applying the 10-folds cross validation technique, we tune the hyperparameters of each of the considered models. As for the model evaluation, and since our goal is to classify a cut on whether it is a valid cut or an invalid cut, accuracy and f-measure can be computed using the following formulas:

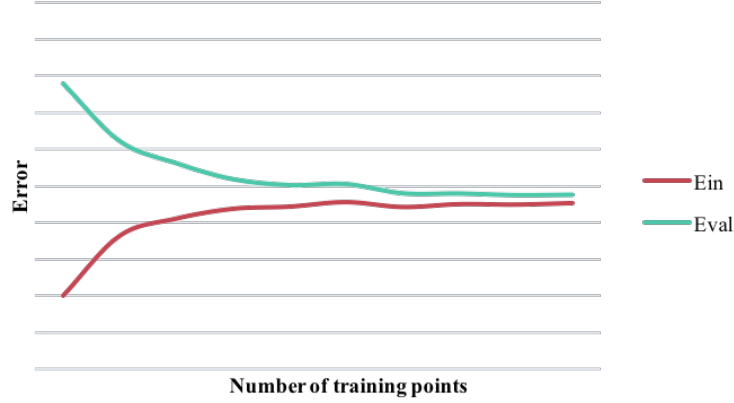


Figure 3.5: Learning curve of the training data of OCT images using Logistic Regression.

$$\begin{aligned}
 \text{Accuracy} &= \frac{TP + TN}{TP + FP + FN + TN} \\
 \text{Precision} &= \frac{TP}{TP + FP} \\
 \text{Recall} &= \frac{TP}{TP + FN} \\
 \text{F-Measure} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
 \end{aligned} \tag{3.6}$$

where, if considering a valid cut to be positive, and a damaged cut to be negative:

- TP is the number of true positive instances, i.e. the number of cuts that are annotated as valid and the model classified them as valid.
- FP is the number of false positive instances, i.e. the number of cuts that are annotated as damaged and the model classified them as valid.
- FN is the number of false negative instances, i.e. the number of cuts that are annotated as valid and the model classified them as damaged.
- TN is the number of true negative instances, i.e. the number of cuts that are annotated as damaged and the model classified them as damaged.

Our interest is in minimizing the number of damaged cuts classified as valid cuts, rather than to minimize the number of incorrectly classified instances. In other words, fewer valid cuts will not affect the overall results of an OCT image, whereas one damaged cut would detect imprecise demarcation line(s), and thus provide wrong output. Consequently, our goal is to minimize the number of false positive instances. Thus, we evaluate the models by the largest precision value.

In the following, we describe the models examined for the problem of classification of valid and invalid cuts in OCT images.

Logistic Regression

Logistic regression takes as input the feature vector X and output a value y [40]. The key characteristic of this model is that the output value is binary: 0 or 1. Input values are combined using weights β , also called coefficient values, in a linear manner. The coefficients of the logistic regression model are estimated using the training data. This is done using the maximum-likelihood estimation fashion. Therefore, the best coefficients, of the logistic regression model, result in a model that would predict a value extremely close to 1 for the positive class and a value extremely close to 0 for the negative class.

The Logistic Regression model is considered to be a simple supervised learning model; this model is exploited for the learning curve plotting due to its simple binary classification nature that will be exploited in the model estimation as well.

Support Vector Machines

Support Vector Machine (SVM) model is a supervised machine learning algorithm that can be used for both classification and regression problems (it is mainly used in classification) [41]. In this model, each instance is plotted in a n -dimensional space, such that n is the number of features per instance. Afterwards, the classification is performed by finding lines or hyperplanes that segregate the different classes by drawing the best frontier.

Sequential minimal optimization (SMO) is an algorithm for solving the quadratic programming problem that arises during the training of support vector machines [42]. It is applied by the Weka machine learning library instead of SVM. Using some heuristics, SMO partitions the training problem into smaller problems that can be solved analytically. Advantageously, it can significantly decrease the training time.

The SVM model is mostly used for binary classification. The bias-variance analysis of this model shows that, based on the tuning of the hyperparameters of the model, it can be used for both high variance, high bias, or high bias and high variance datasets [38].

Multilayer Perceptron

The Multilayer Perceptron model, also known as the feedforward artificial neural network, is a supervised learning algorithm [43]. Given an input composed of a set of features of size n , it can learn a non-linear function and then provides an output of m dimensions. This model works for both regression and classification. The difference between logistic regression and the multilayer perceptron models is that the latter can employ multiple non-linear function layers between the input layer and the output layer.

The major cure for high bias data is the use of non-linear model for the learning of classification. Hence, the Multilayer Perceptron can be used in the case where the learning curve of the training dataset presents the typical characteristics of high bias data [44].

3.1.5 Chosen Model and Results

Table 3.1 summarizes the best performance of each chosen model. As observed, all of the models are able to attain good performance after several rounds of tuning the hyperparameters. However, since our objective is to minimize the number of false positives, we adopted the SVM model, as it was able to provide only 14 false positives (FP) among all cuts of all images.

Table 3.1: Measurements of the best performances of the selected models.

	Accuracy	F-Measure	Precision
Logistic Regression	0.8620	0.9236	0.9835
Support Vector Machines	0.8628	0.9250	0.9974
Multilayer Perceptron	0.8504	0.9151	0.9508

Table 3.2: Results and hyperparameters of the Support Vector Machine model on testing data.

	<i>FP</i>	Accuracy	F-Measure	Precision
C: 1 Kernel: Polynomial Degree: 1 Epsilon: 1.00E-10 Tolerance: 0.01	3	0.8828	0.9366	0.9978

Note: Detailed results of the three models with the different hyperparameters can be found in Appendix B.

Table 3.2 shows the SVM model (i.e., its hyperparameters) chosen to show best testing results. As noticed, only 3 *FP* were achieved among all cuts, which is considered to be excellent, since the impact of these on determining the consistency of the demarcation line would be negligible.

Subsequently, the model is saved and is embedded in the software as part of the OCT image filtering process, which ensures that only valid cuts are used for analysis.

3.2 Input Image

The input image is a 2D digital image extracted from the OCT scanner in *jpeg* format. It has a width w of 540 pixels and height h of 270 pixels. The OCT scanner typically generates images of length $l = 4000\mu\text{m}$ [12]; thus, the *pixel-to- μm* conversion ratio can be obtained by l/w .

The OCT image is modeled in our program using a matrix \mathbf{M} with dimensions $w \times h$. As shown in Eq. (3.7), the matrix consists of pixel intensity values; each pixel intensity value is denoted as $p_{i,j}$ (also referred to as $[i,j]$; example, $\mathbf{M}[1,4] = p_{1,4}$), where $i = 0, \dots, h - 1$ and $j = 0, \dots, w - 1$. For a 8-bit grey-scale image, $0 \leq p_{i,j} \leq 255$.

$$\mathbf{M} = \begin{bmatrix} p_{0,0} & \cdots & p_{0,w-1} \\ \vdots & \ddots & \vdots \\ p_{h-1,0} & \cdots & p_{h-1,w-1} \end{bmatrix} \quad (3.7)$$

Note that the image is seen by the program with a coordinate system that has the origin (0,0) at the upper left corner of the image.

3.2.1 Finding the Corneal Region

Finding the relevant corneal region in the image, where the demarcation line occurs, requires finding the coordinate vectors of the top and bottom boundaries. In this subsection, we expose the methodology used by Oud et al. to find the boundaries of the cornea [35]. The same steps are used in our approach.

Since the Epithelium typically has a very bright surface that may greatly alter the average brightness calculated, the top boundary must be shifted down so that the Epithelium's surface is not considered as part of the analyzed region.

Finding the Top Boundary

To find the top boundary, the image is first processed using the *Thresholding* technique [45, 46]. Using this technique, the pixels in the corneal region are emphasized as opposed to the remainder of the image (i.e., the black region - with the legend - outside the cornea as shown in Figure 3.6). In the *Thresholding* technique, a brightness parameter θ is defined such that pixels with values smaller than θ are referred to as *background* pixels, and those with values larger than θ are referred to as *foreground* pixels.

In [35], θ is computed as the average brightness of the area below the cornea.

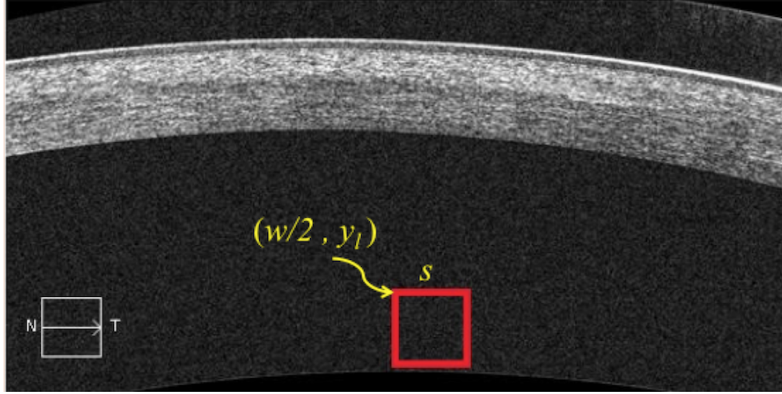


Figure 3.6: Region used for computing the Thresholding parameter θ .

A sample collection of pixels in that area is extracted with a square of side s (as shown in Figure 3.6), and then computing their average pixel value.

Thus, the brightness parameter θ is computed as follows:

$$\theta = \frac{1}{s^2} \times \left(\sum_{x=\frac{w}{2}}^{\frac{w}{2}+s} \sum_{y=y_l}^{y_l+s} \mathbf{M}[x, y] \right) \quad (3.8)$$

where y_l is the y-coordinate value of the legend in the images.

As a result, a binary version of the matrix \mathbf{M} is generated, denoted \mathbf{M}_B , with:

$$p'_{i,j} = \begin{cases} 1, & p_{i,j} \geq \theta \\ 0, & \text{otherwise} \end{cases} \quad (3.9)$$

The new matrix \mathbf{M}_B is illustrated in Figure 3.7. This new matrix is however used only to find the top boundary. In fact, the high intensity of the Epithelium's bright surface implies a relatively uniform intensity of the top boundary, whereas the lowest section does not always have this uniformity; thus, it may not have a continuous curvature.

The new image enables to extract the *contours* (i.e., borders) of the cornea; a contour is defined as the set of points that enclose a certain region. To achieve so, Suzuki's algorithm [47], implemented in OpenCV, is applied for obtaining



Figure 3.7: Image after applying the *Thresholding* technique.

external contours. However, the contouring function requires a "connected" pixel region enclosed in another "connected" region (in an OCT cut, it's the "black+legend" region).

Since this is not the case in Figure 3.7, the image is augmented with black borders to the left and right of \mathbf{M}_B with arbitrary thickness. Using the contouring function provided by OpenCV, the image shown in Figure 3.8 is obtained. Consequently, by searching for the first pixel in the y -axis that has a non-zero value, the top boundary $top(x)$ vector can be extracted. The vector $top(x)$ is in fact defined as a function that for a given x -coordinate, it returns the y -coordinate falling on the top boundary. Additionally, to ignore the Epithelium's surface, every y -coordinate in the vector $top(x)$ is increased by the thickness of the Epithelium's surface e , which we measure dynamically. Typically, the average value of e is about 3 pixels $\approx 22\mu m$.

Finding the Bottom Boundary

Since the bottom boundary does not have a "clear" curvature (see Figure 3.8), the technique used to detect the top boundary will not achieve the desired results. Therefore, a different technique using the original image matrix \mathbf{M} is used to capture it. Based on the type of OCT images extracted, to find the bottom boundary, the following assumptions are made:

- The bottom boundary of the cornea has the same "curvature" as the top boundary (which is normally the case in almost all images)

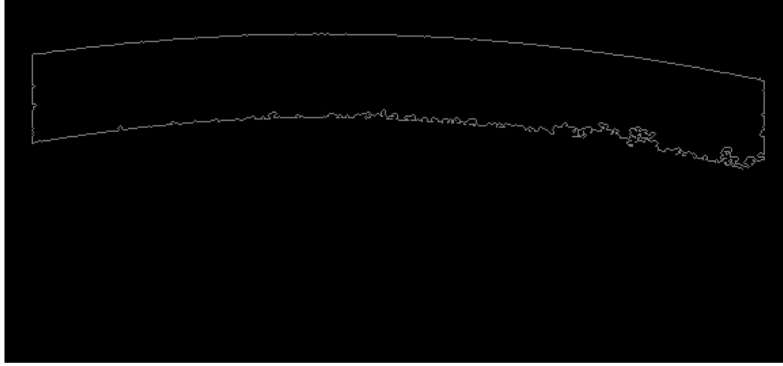


Figure 3.8: Image after applying the *Contouring* method.

- The change in pixel intensity (i.e., brightness) around the top and bottom boundaries (as seen in Figure 3.7 for example) is the highest among all other regions

Based on these two assumptions, a new technique is employed, which performs the following steps:

- Starting from $top(x)$ downwards, for every pixel curve $c(x)$, compute the change in image intensity at a distance y_d above and below $c(x)$. The change in image intensity is computed as the ratio \bar{I}_r of intensities above and below $c(x)$, denoted \bar{I}_a and \bar{I}_b respectively, as follows:

$$\begin{aligned}
 \bar{I}_a &= \sum_{x=0}^w \sum_{y=c(x)-y_d}^{c(x)-1} \mathbf{M}[x, y] \\
 \bar{I}_b &= \sum_{x=0}^w \sum_{y=c(x)+1}^{c(x)+y_d} \mathbf{M}[x, y] \\
 \bar{I}_r &= \frac{\bar{I}_a}{\bar{I}_b}
 \end{aligned} \tag{3.10}$$

- Find the minimum ratio \bar{I}_r (thus, the curve with the highest pixel intensity value) and mark its y-position as y_{shift} .
- The bottom boundary curve is thus represented by the following function:

$$bottom(x) = top(x) + y_{shift} \tag{3.11}$$

3.2.2 Bright Region Detection

To detect the demarcation line in the OCT image (see Figure 3.1) and measure its depth, we make the following two assumptions:

- a. The demarcation line has the same "curvature" as the top boundary (which is normally the case in almost all images, especially for 4mm-wide cross sections of the cornea)
- b. The total intensity value and continuity of bright pixels of the demarcation line (except the Epithelium's bright surface) is the highest among all other regions.

As before, the first assumption implies the pixel vector of the demarcation line can be found by shifting $top(x)$ down, such that the appropriate position is found based on the second assumption. This approach is valid since the demarcation line comprises a dense curve of high-intensity pixels with few pixels in thickness (unlike haze, which is usually thicker and does not necessarily span the width of the cornea).

Our proposed mechanism applies a formula that computes a "score" (based on the intensity value and continuity of bright pixels) for every curve starting from the top boundary downwards; and then based on a well-defined criterion, we detect a set of "candidate" demarcation lines. Subsequently, we select/suggest the "best" candidate (i.e., the curve with the highest score) to be the actual demarcation line.

To identify the foregoing dense region of high-intensity pixels, we first blur the image so as to eliminate any sharp intensity fluctuation, thereby obtaining the image in Figure 3.9. Next, using an OpenCV function, we apply the Sobel Derivative [48] in the y-direction (i.e., vertical derivative), which multiplies each pixel value by a kernel matrix \mathfrak{S} , so that its gets assigned a value representing the vertical change across the pixel.

The Sobel operator is a discrete differentiation operator. It computes an approximation of the gradient of the image intensity function [49]. The gradient of the image intensity at each point is calculated, giving the direction of the largest possible increase from light to dark and the rate of change in that direction. As

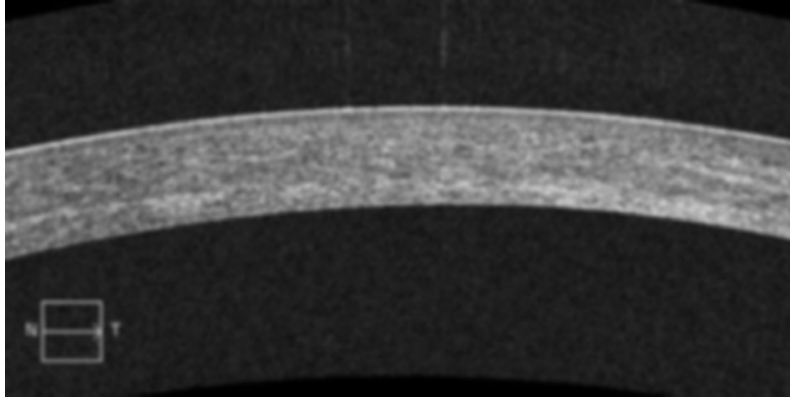


Figure 3.9: Image after blurring.

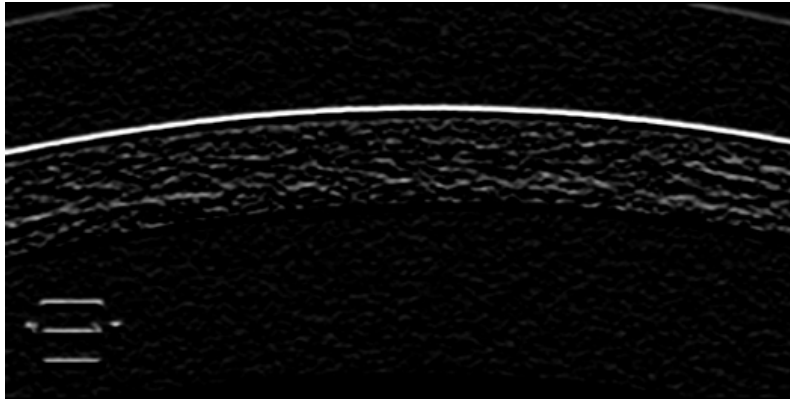


Figure 3.10: Image after applying the *Sobel* derivative.

a result, the computed value shows how "abruptly" or "smoothly" the image changes at that point and therefore how likely it is on an edge [50].

The Sobel kernel matrix is defined as follows:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

The resulting image is shown in Figure 3.10.

Finally, using the blurred and "Sobelled" image, we apply the following methodology to extract the demarcation line:

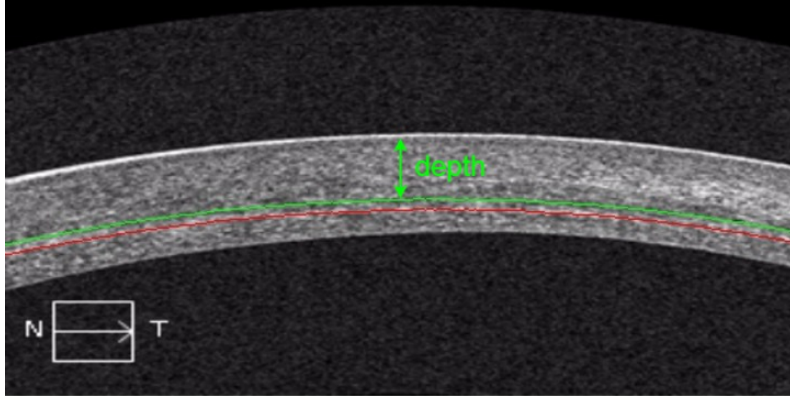


Figure 3.11: Suggested demarcation line highlighted in color green, and candidate demarcation line in color red.

- a. For each curve $c(x)$ from $top(x)$ to $bottom(x)$, calculate a score $S(c)$ per Eq. 3.16.
- b. Identify and mark the candidates demarcation lines; these will be the *outliers* in the data set collected from each image (as described later).
- c. Scan through the candidate demarcation lines, and mark the y -position, which will be the position of the demarcation line, as y_d of the highest value of $S(c)$.
- d. The suggested demarcation is thus represented by the following function:

$$dem(x) = top(x) + y_d \quad (3.12)$$

As illustrated in Figure 3.11, by applying the foregoing method, we get the suggested demarcation line $dem(x)$ highlighted in color green, and the candidate demarcation line is highlighted in color red on the original image. The figure also shows how the demarcation line depth is measured.

3.3 Computation of $S(c)$

The computation of score $S(c)$ for each curve at depth d is obtained by scanning every pixel $p_{i=1..n} \in c(x)$, where n the width of the curve, from left to right and computing the total pixel intensity value of $c(x)$. We need to account for both continuity and brightness in the computation of $S(c)$. However, continuity should have a higher weight such that if we have, for example, a line at depth d_1 that is continuous throughout whole image cut but at the same time has a low overall level of brightness, and a line at depth d_2 that has a high intensity value but only appears in a small section of the curve, then the line at depth d_1 would have a higher score since it is more probable to be a potential demarcation line. In other words, to account for continuity of bright pixels, which is a key feature for identifying a demarcation line, we use a scoring technique that is designed to distinguish between any two curves that might have equal total intensity values, but one is continuous and the other is not, in which case, the applied technique will give a higher score to the continuous line.

Given a curve $c(x)$ at depth d of pixels values p_1, p_2, \dots, p_n , such that $0 \leq p_i \leq 255 \forall i \in \{1, \dots, n\}$, the computation of its score $S(c)$ is performed via the following four steps:

- a. Compute the mean position, μ , of the non-zero-pixel values. For example, for $n = 540$, if at depth d , we have a curve of pixel values:

$$\begin{aligned} p_3 &= 18 \\ p_{29} &= 103 \\ p_{413} &= 107 \\ \text{and } p_i &= 0 \forall i \in \{1, \dots, n\} \setminus \{3, 29, 413\} \end{aligned}$$

Then, the mean position is $\mu = 29$.

- b. As illustrated in Figure 3.12, we divide the cut into three regions and classify the mean value according to the following via a coefficient, \mathbf{c}_μ , as follows:

$$\mathbf{c}_\mu = \begin{cases} 3, & \text{if } \mu \in \text{Region III} : \{\frac{2n}{5}, \frac{3n}{5}\} \\ 2, & \text{if } \mu \in \text{Region II} : \{\frac{n}{5}, \frac{2n}{5} | \frac{3n}{5}, \frac{4n}{5}\} \\ 1, & \text{if } \mu \in \text{Region I} : \{0, \frac{n}{5} | \frac{4n}{5}, n\} \end{cases} \quad (3.13)$$

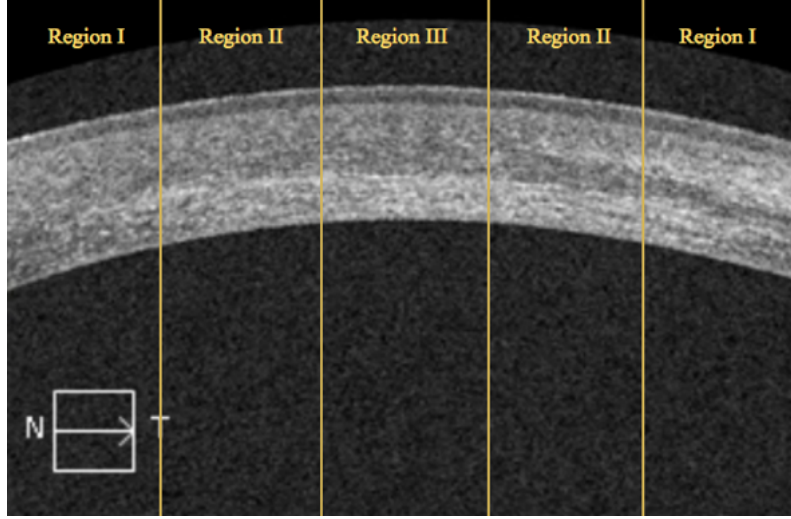


Figure 3.12: One cut of an OCT image showing the different regions used in the score computation.

By setting the parameter \mathbf{c}_μ , we give more weight to the curves having a μ value in the center of the cornea, which can be explained as a homogeneous distribution of the bright pixels across the curve, and hence a more continuous potential demarcation line.

- c. Compute the standard deviation, Σ , of the position of the non-zero-pixel brightness values. It is quantified using the below formula:

$$\Sigma^2 = \frac{\mathcal{S}}{n} \quad (3.14)$$

where,

$$\mathcal{S} = \sum_{i=1}^n (i - n)^2 \quad \text{if } p_i > 0 \quad (3.15)$$

In other words, \mathcal{S} is the summation of the squared difference between the mean, μ , and the different bright pixels; Σ^2 is the average difference between the bright pixels and the mean μ .

- d. Compute $S(c)$ as follows:

$$S(c) = \mathbf{c}_\mu \times \Sigma \times \sum_{i=1}^n p_i \quad (3.16)$$

For example, assume the following two curves, at depths d_1 and d_2 , with their corresponding pixel values for $n = 10$:

$$c_1(x) = \{25, 35, 20, 20, 0, 0, 0, 0, 0, 0\}$$

$$c_2(x) = \{25, 0, 0, 35, 0, 0, 20, 0, 0, 20\}$$

Their scores will then be computed as follows:

- For $c_1(x)$:

$$\mu = 2$$

The mean appears in the first fifth of the total width of the line,

hence $\mathbf{c}_\mu = 1$

$$S = (1 - 2)^2 + (2 - 2)^2 + (3 - 2)^2 + (4 - 2)^2 = 6$$

Therefore, $\Sigma^2 = \frac{6}{10} = 0.6$ and $\Sigma = 0.77$

Concequently, $S(c_1) = 1 \times 0.77 \times (25 + 35 + 20 + 20) = 77$

- For $c_2(x)$:

$$\mu = 4$$

The mean appears in the first fifth of the total width of the line,

hence $\mathbf{c}_\mu = 2$

$$S = (1 - 4)^2 + (4 - 4)^2 + (7 - 4)^2 + (10 - 4)^2 = 54$$

Therefore, $\Sigma^2 = \frac{54}{10} = 5.4$ and $\Sigma = 2.32$

Concequently, $S(c_2) = 2 \times 2.32 \times (25 + 35 + 20 + 20) = 464$

Note that in case a candidate is detected near or at the bottom boundary (i.e., in the last 35 microns of the cornea), it is considered the suggested line if and only if it is the only candidate detected. Otherwise, it is considered as candidate even if it has the highest score (the second highest score is considered as the suggested). This is due to the fact that, as confirmed by manual observation, a high intensity (i.e., high $S(c)$ score) candidate falling in the last 35 microns (i.e., approximately last 5 pixels of the image) will most likely be a reflection in the cornea and not a demarcation line, except if it was the only candidate.

3.3.1 Identifying Candidate Demarcation Line(s)

To identify the scores that are considered outliers (and thus taken as candidate demarcation lines), we apply a statistical approach [51], where the set of scores is divided into four quartiles with three delimiters $Q1$, $Q2$ and $Q3$. The delimiter

$Q1$ represents the lower 25th percentile score $S(c)$; $Q2$ stands for the mid-50th percentile score, and $Q3$ represents the 75th percentile score. Typically, in the foregoing statistical approach, upper and lower *fences* are employed (i.e., the cut-off scores that separate two quartiles) and are usually set at fixed distance from the interquartile range (i.e., $Q3 - Q1$). These are set by default as 1.5 times the interquartile range. Subsequently, any score that falls outside these fences is considered a potential *outlier*. Even when the scores are not normally distributed, this approach can be used safely because it depends on the median and not the average of the scores.

In our scheme, only upper fence is used as the threshold for considering a curve as candidate demarcation line. This threshold \mathfrak{T} is thus computed as follows:

$$\mathfrak{T} = Q3 + (Q3 - Q1) \times 1.5 \quad (3.17)$$

Hence, a curve $c(x)$ is considered a candidate demarcation line if its score $S(c) > \mathfrak{T}$.

In a hazy image (i.e., an image with a lot of detected haze), the scores of the different curves that are to be taken as candidate demarcation lines would be high, which makes these curves belong to $Q3$ subject to a "higher" upper fence value. This ensures that these curves are not considered as part of a haze region. Thus, our mechanism is able to distinguish/separate demarcation line candidates from curves that belong to haze.

Finally, to ensure that our scheme does not consider multiple curves as multiple candidates, when they should be taken as one demarcation line, we have evaluated the typical thickness of a demarcation line that we need consider as one candidate. To achieve so, we have first confirmed from manual observations by ophthalmologists that a demarcation line is considered a candidate if it is thicker than 1 pixel, which is approximately equivalent to 7.4 microns. To verify this observation, we generate and plot in Figure 3.13 the histogram of average thicknesses (and its polynomial curve fitting) of all extracted candidates for all images. As noticed, the average thickness of a demarcation line is approximately 4 pixels.

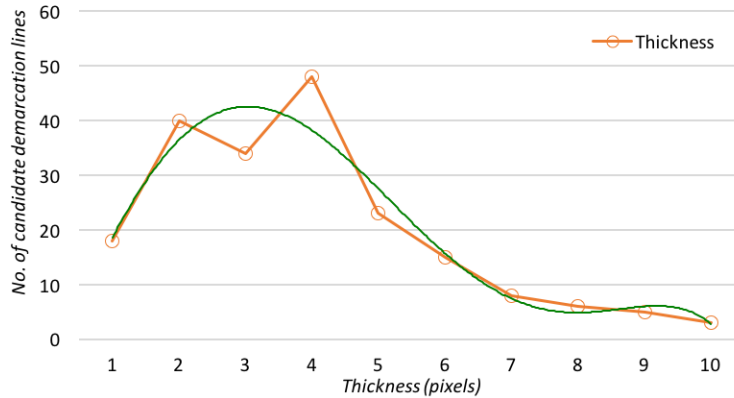


Figure 3.13: Histogram of candidate demarcation line thickness.

Thus in our scheme, when scanning the curves from the upper boundary to the bottom boundary, whenever a curve is considered a candidate (i.e., the score greater than the upper fence threshold), the scores of the next three curves are "tagged", and then the one with the highest score among these four "tagged" curves, is picked as candidate; the rest are discarded.

3.4 Minimum Score Value

In some cuts, a demarcation line must not be detected, as it is barely visible. However, in the same cut, a small bright region in the stroma may have a "high-enough" score and can exist in multiple cuts (i.e., is consistent). In this case, our technology must not report it as a demarcation line; therefore, a minimum score must be set, below which a curve is not considered a candidate demarcation line.

To statistically set the minimum score, denoted Υ , we perform the following:

- a. We first calculate the scores of all the candidate and suggested demarcation lines for all available OCT images. The total number of entries in our dataset is obtained as 5,228.
- b. To reduce the set space, we aggregate the scores such that, starting with the minimum score \mathbf{m} (of the available values), all scores belonging to $\mathbf{m} + \sigma$, are grouped (e.g., $\sigma = 2000$). Subsequently, we keep updating \mathbf{m} by σ score

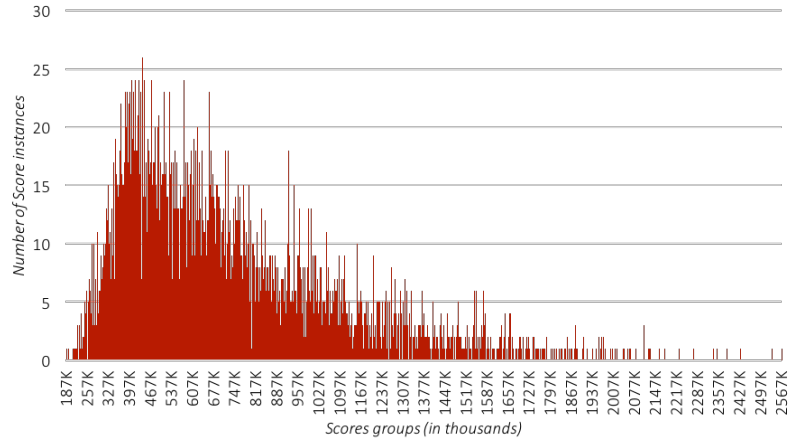


Figure 3.14: Distribution of the Candidates and Suggested Demarcation Lines' Scores.

(i.e., $\mathbf{m} = \mathbf{m} + \mathbf{o}$) until we reach the maximum score (which is obtained from the available values).

- c. We plot the histogram of the distribution of the grouped scores.
- d. We identify the type of distribution obtained [52].
- e. We compute the 16th percentile minimum risk, which is the commonly used percentile in the literature [53]; this would be our minimum score.

Figure 3.14 shows the distribution of scores with respect to the aggregated score groups. This positively skewed distribution can be a:

- a. Weibull distribution
- b. Gamma distribution
- c. Lognormal distribution
- d. Power lognormal distribution

After *studying* the obtained distribution, we associate our distribution with the Gamma family [54, 55], for which we compute the shape and scale parameters α and β , respectively.

These two values were computed based on the mean $\alpha \times \beta$ and the standard deviation of the distribution $\alpha \times \beta^2$.

We define the Gamma function, for any positive real number x to be:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad (3.18)$$

The cumulative distribution (CDF) function F of the gamma distribution is:

$$F(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^x t^{\alpha-1} e^{-\frac{t}{\beta}} dt \quad (3.19)$$

Figure 3.15 shows the probability density function of our distribution with:

$$\alpha = 4.128$$

$$\beta = 175718.39$$

In order to find the 16th percentile of the distribution, we need to find x such that:

$$F(x; 4.128, 175718.39) = 0.16$$

Hence,

$$\Upsilon = x = 384489.6$$

Hence, a suggested or candidate demarcation line is considered if and only if its score $S(c) > \Upsilon$; otherwise it is ignored.

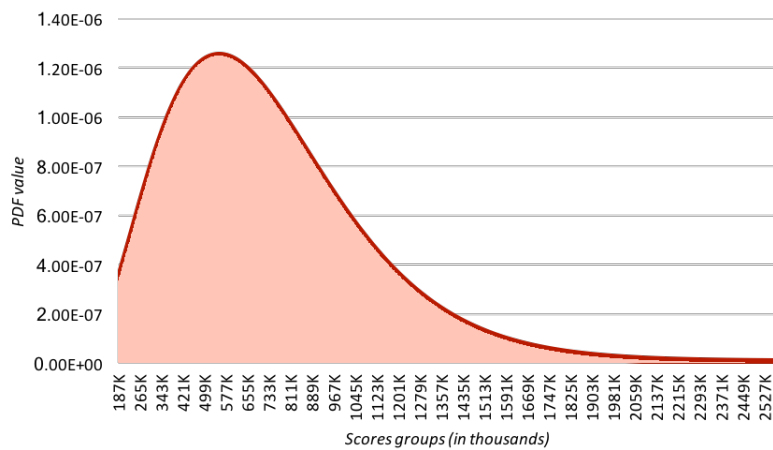


Figure 3.15: Probability Density Function of the Distribution of the Candidates and Suggested Demarcation Lines' Scores.

Chapter 4

Results and Analysis

After applying our algorithm on all videos, we detect and measure the depth of the suggested and candidate demarcation lines in microns, such that the depth is measured from the beginning of the Epithelium's bright surface to $dem(x)$ as shown in Figure 3.11.

Two independent operators and the automated detection software examined corneal OCTs of 40 eyes post corneal cross-linking, at 1 and 3 months postoperatively. Operators evaluated the presence of the demarcation line and measured its depth by looking at OCT images (128 cuts) on two separate occasions one week apart. The automated software examined all 128 cuts of each OCT measurement. The operators were blinded to patients' names, the results of other examiners (human or software), and their own previous results. Intra-observer measurements repeatability and intra-observer reproducibility were calculated and compared.

Figure 4.1 shows the depths of the demarcation lines in microns for all considered patients. These measurements are taken by two experienced human operators (i.e., Dr. Shady Awwad and Dr. Maamoun Abdul Fattah from AUBMC) and the novel technology. The absence of a depth value indicates that the operator (software/human) did not detect any (candidate or suggested) demarcation line (e.g., Patient 6 at 1-month post-operation).

As noticed in the figure, some of the measurements of the demarcation line depth of the novel software and the two human operators are approximately equal (e.g., patient 24 at 1-month post-operation), or the margin of difference is very

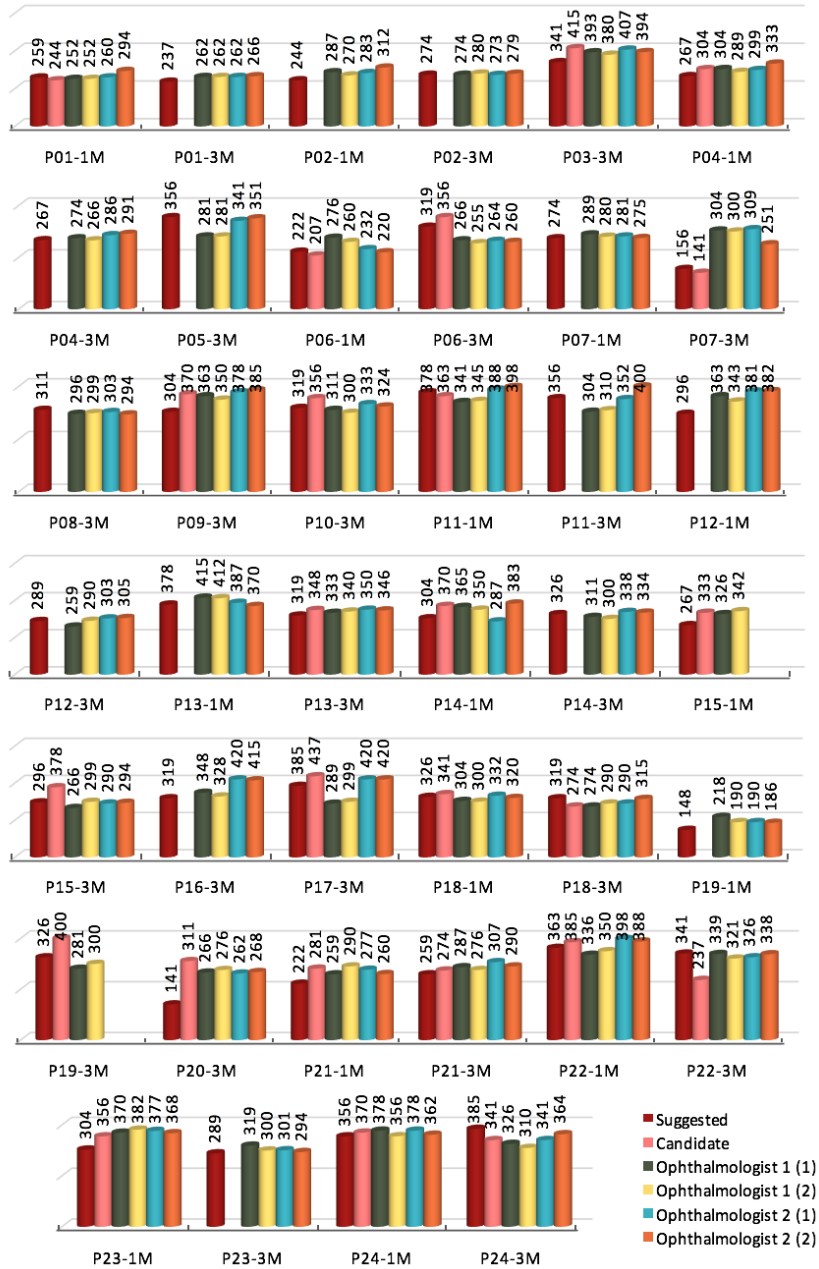
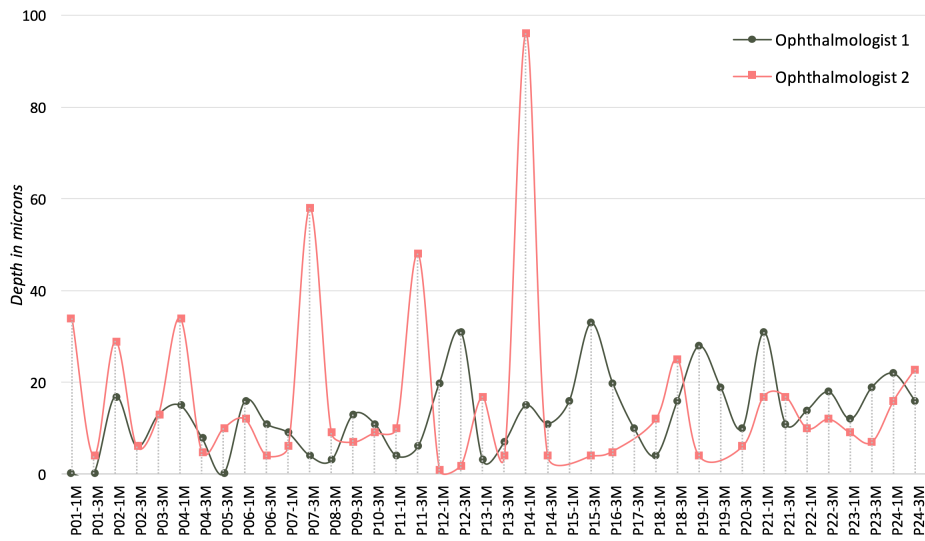
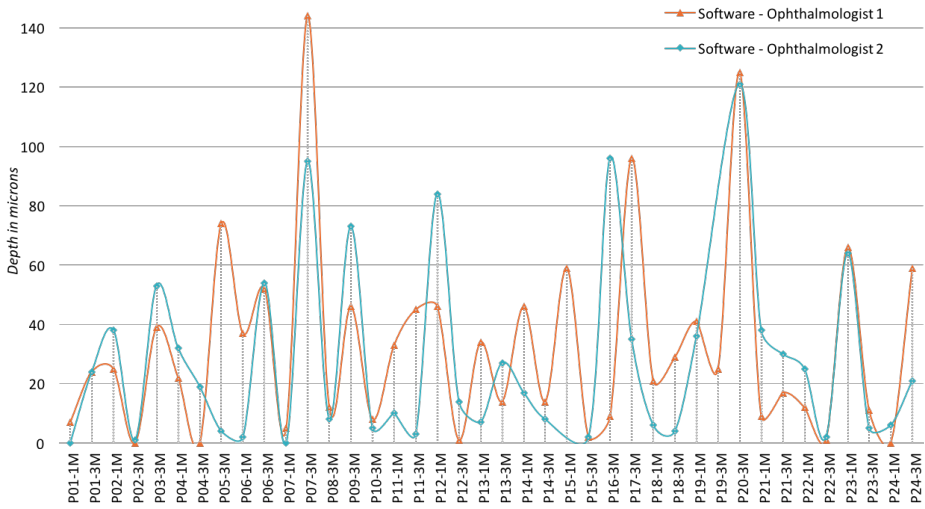


Figure 4.1: Depth of *suggested* and *candidate* demarcation line by the software and the measurements of the ophthalmologists on two occasions expressed in microns for all patients at 1-month and 3-months post-operation periods.



(a) Differences of depths for all patients among ophthalmologists on two occasions.



(b) Differences of depths for all patients between software and the average measurement of ophthalmologists.

Figure 4.2: Differences of depths of the demarcation lines as measured by ophthalmologists and software in microns for all patients at 1-month and 3-months post-operation periods.

small. Here, we note that every pixel is equal to 7.4 microns, and it is very natural that a human operator "drifts" by ± 5 pixels when doing manual measurements, which is equivalent to a drift of ~ 40 microns.

We also observe in some measurements outstanding cases where the software provides a measurement close to one of the human operators' measurements only (e.g., patient 13 at 1-month post-operation) or may provide a measurement that is far from the ones provided by both human operators (e.g., patient 7 at 3-months post-operation). Interestingly, since the software is designed to detect and report not only the best candidate line (i.e., the suggested one), but also the second best candidate (if applicable), those two candidate(s) can be fairly used to assess the performance of the software vs. the different measurements of human operators. Here, a candidate demarcation line (when it exists) might be the line chosen by ophthalmologists as "*the*" demarcation line. However, with our technology, for a line to be chosen as suggested and not candidate, this means that this line that has the **highest** score among all **consistent** potential demarcation lines (i.e., exists in 15% of the valid cuts).

The difference of measurements in microns, between the different operators, is quantified in Figure 4.2. As observed in Figure 4.2a, the measurements provided by the human operators differ between the two measurements of the same operator, and across operators. These results highlight the human error and inconsistency factor in obtaining manual subjective measurements, which motivates the need for an automated tool and software. Similarly, Figure 4.2b shows the differences in measurement between the software and the average measurements provided by the two human operators. As observed, although blinded to the human operator's measurements, the software is able to provide very close (and sometimes almost identical) results taking into account the aforementioned human drift in measurement.

This reflects and proves the accuracy and robustness of our technology in being the main tool used for corneal haze measurement, with the promise of replacing manual measurements. However, the same figure shows some discrepancy of measurement for some patients between the automated and manual operators. This is due to the fact that unlike the human operators, the novel software takes into account multiple factors (i.e., valid cuts, consistency among cuts, continuity,

etc.) for considering a demarcation line as a "legal" one, thereby having the human operators selecting a cut and measurement that shouldn't be considered.

To shed the light on this interesting observation, we plot in Figure 4.3 the measurements of the cases where discrepancy is observed (i.e., P6-3M, P7-3M, P9-3M, P12-1M, P15-1M, P20-3M and P23-1M) as obtained by the different operators (human and software) on the different cuts that were selected to be the "best" by the different operators. Since the difference between the first and second measurement of one human operator is observed to be minor, for simplicity, we only show one of the human measurements. Discussion and analysis of these results are presented as follows:

- (a) As exhibited in Figure 4.3a, for the case of P6-3M, our software picked cut 14 in which the demarcation line appears the clearest (i.e., has highest score among all consistent demarcation lines), and its depth is measured to be $319 \mu\text{m}$. Conversely, the two human operators selected cuts 45 and 46, respectively, in which they obtained depths of 266 and $264 \mu\text{m}$, respectively. Interestingly, it can be noticed that the thickness of the cornea has varied between the different cuts. More specifically, the thickness in cuts 45 and 46 remained almost the same; whereas at cut 14, it appears to have expanded. To verify this observation, we measured the relative depth of the line, and we obtained a depth of $\approx 62\%$ for the software, and a depth of $\approx 58\%$ for the two human operators measurements. Therefore, we can claim that the demarcation line chosen by the ophthalmologists is most likely to be the same line as the one chosen by the software.
- (b) As exhibited in Figure 4.3b, for the case of P7-3M, after examining the scores of the lines (suggested/candidate) picked by the operators, we observed that the lines chosen by the ophthalmologists have obtained scores that are *high enough* to be potential demarcation lines. However, they did not rank 1^{st} or 2^{nd} (i.e., suggested and candidate, respectively) to be reported by the software. In fact, a line with a depth near the line chosen by the human operators, was detected by the novel technology; however, it was ranked 3^{rd} .
- (c) As exhibited in Figure 4.3c, for the case of P9-3M, the suggested demarcation line as measured by the software is captured at cut 121, in which it

appears the most continuous across the curvature of the cornea, and consistent among cuts. The line with the highest score appears at depth 304 μm , whereas the candidate demarcation line is at depth 370 μm . Interestingly, the two human operators picked a demarcation line at approximately the same depth as the candidate line (i.e., at depth 350 and 378 μm , respectively). However, in cut 121, as observed in the figure, the thickness of the demarcation line is observed to be relatively high (i.e., greater than 4 pixels, which is computed statistically as the average thickness of a demarcation line based on the available data set). Thus, it is most likely that the two lines pertain to the same demarcation line.

- (d) As exhibited in Figure 4.3d, for the case of P12-1M, the cut picked by the software exhibits the presence of a "clear" demarcation line at a depth around 290 μm . Nevertheless, by comparing the cuts chosen by the different operators, we can affirm that this line appears the clearest in cut 122. In general, ophthalmologists tend to look mainly for potential lines in central cuts (in this case, cuts 51 and 82) as they assume that "most probably" a potential line appears the clearest in middle cuts, which is not the case for P12-1M.
- (e) As exhibited in Figure 4.3e, for the case of P15-1M, the second operator's observation indicates that there is no demarcation line. Nonetheless, the first human operator found a demarcation line at depth 326 μm . As for the software, a suggested and a candidate demarcation lines were observed at depths 267 μm and 333 μm , respectively, which adheres with the first operator's measurement. Discussion of the cases with no reported demarcation lines is presented later in the text.
- (f) As exhibited in Figure 4.3f, the case of P20-3M is a similar to the case of P7-3M. Here, the line chosen by the human operators was indeed captured by the software, but it did not have the highest score. The candidate's score is *slightly* smaller than that of the suggested one; hence, it was ranked 2nd by the novel technology.
- (g) Lastly, as exhibited in Figure 4.3g, for the case of P23-1M, after examining the OCT image, we observed that the demarcation line appearing for this patient's eye fluctuates across multiple cuts. However, the software

considers two lines to be the same if the difference in depth is less than or equal to 4 pixels (following the assumption that the demarcation line is the same if it is within 4 pixels of depth range). Interestingly, the fluctuation in this case appears to be slightly larger than 4 pixels. However, the software was able to capture this very high fluctuation by considering the fluctuation as two lines. Nevertheless, the measurements of the human operators were very close to the obtained candidate demarcation line (the difference is less than $24 \mu\text{m}$).

As noticed, the selection of the cut for all the cases has a big impact on the demarcation line measurement. This highlights the advantage of the automated tool in selecting the "best cut" statistically based on a well-defined criterion (which incorporates the score of the line and the consistency of the line among multiple cuts) as opposed the error-prone human approximation. In addition, as observed, the automated tool is able to accurately detect and measure the depth of the demarcation line in the selected cut.

In Figure 4.4, we emphasize this note by showing examples of the four scenarios that might occur in the demarcation line depth measurements: one of the operators does not find any potential demarcation line, (as observed in Figure 4.4a); the three operators, human and software, agree on the same demarcation line (as observed in Figure 4.4b); the gap between the demarcation line depth as measured by the novel software and the human operators is relatively large (about 100-140 microns, as observed in Figure 4.4c); and lastly, as observed in Figure 4.4d, the difference between the human operators is large, yet one of the human operators measurements is close to the software measurement.

Although a minimum score value was set, we notice that the software was still able to detect a suggested (and a candidate) demarcation line in all OCT images, even though one of the human operators did not detect a demarcation line in exactly two images: P15-1M and P19-3M. We examined the scores of the demarcation lines of the latter two images, and we noticed that by making the software more conservative (in this case, by picking 30th percentile as minimum score range as opposed to 16th percentile), the software would not detect a demarcation line in these two images; however, a demarcation line would also not be detected in two other images. In Figure 4.5, we show the cuts of the affected four

images. Interestingly, in these OCT images, the demarcation line chosen by the software and one (or two) human operator(s) is "faded" and with relatively the same brightness intensity. Hence, the software was able to give a well-computed score that accurately assesses the brightness and continuity of a stromal line; whereas, a human operator might give two different examinations for two lines with the same characteristics. Nevertheless, this highlights the flexibility of the tool to be customized as either *conservative* or *loose* by the user. Alternatively, in addition to the scoring technique, the tool may also employ machine learning to "learn" from the user about what should be taken as demarcation line, which is a feature we plan to implement in our future work.

Figure 4.6 exhibits some example of OCT images of patients across time: at 1-month and 3-months post-operative periods. The automation of the process of detection and measurement of demarcation line across time simplifies the comparison of the depth of the latter at different periods after the CXL treatment. Hence, to examine the progress of the treatment, the operator can see and compare the most recent OCT image vis-a-vis the OCT images of the patient from previous periods. Hence, the novel technology can detect subtle changes in the depth of demarcation line effortlessly.

In Figure 4.7, we show a graph exhibiting the progress of the demarcation line at different periods: pre-operation, 2-weeks, 1-month, 2-months and 3-months post-operation. We need to mention here that OCT images of these periods were not all available for the different patients. However, the software can process any OCT image from any period; and the more images are provided, a more precise correspondence of demarcation lines' depth can be made across time.

To assess the overall performance of the software, we take the different measurements of the operators and we compute the Pearson Correlation Coefficients (PCC) and the Intraclass Correlation Coefficients (ICC) [56, 57]. In the following, we explain the characteristics of these coefficients:

The **Pearson Correlation Coefficient** (PCC) is defined as:

$$\rho(x, y) = \frac{E(x, y)}{\sigma_x \sigma_y} \quad (4.1)$$

where,

- x and y are two real valued variables.
- $E(x, y)$ the cross-correlation between x and y .
- And, σ_x, σ_y the variances of x and y , respectively.

This coefficient is a measure of the linear correlation between the two variables, x and y , that has a value between -1 and 1 inclusively such that:

- If $\rho(x, y) < 0$, then x and y are negatively correlated.
- If $\rho(x, y) = 0$, then x and y are not correlated.
- If $\rho(x, y) > 0$, then x and y are positively correlated.

Therefore, the closer $\rho(x, y)$ is to the value 1, or -1, the stronger is the correlation between x and y [56, 57].

The **Intra-class Correlation Coefficients** (ICC) describes how strongly the values of two variables x and y resemble each other [58]. ICC operates on data as groups. We define $Y_{i,j}$, the i^{th} observation in the j^{th} group, as the following:

$$Y_{i,j} = \mu + \alpha_j + \epsilon_{i,j} \quad (4.2)$$

where,

- μ an unobserved overall mean.
- α_j - an unobserved random effect shared by the values in group j .
- $\epsilon_{i,j}$ an unobserved noise term.

The population ICC is hence computed using the following formula:

$$\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\epsilon^2} \quad (4.3)$$

where,

- σ_α^2 the variance of α_j .
- σ_ϵ^2 the variance of $\epsilon_{i,j}$

The value of ICC is strictly positive and can reach at most the value 1. For a value of the ICC less than 0.363, we consider the groups to have a poor agreement; on the other hand, when the value of ICC is greater than 0.75, we consider the data to have excellent agreement.

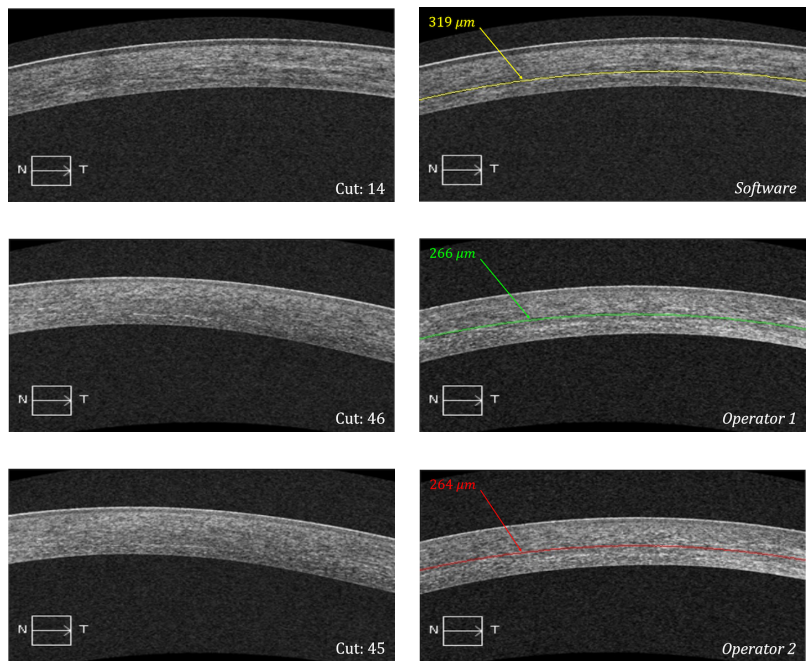
The main difference between PCC and ICC is that, in the first, each variable is centered and scaled by its own mean and standard deviation; whereas, in the second one, the data are scaled by a pool mean.

In Figure 4.8, we report the values computed for 40 Keratoconus patients considered in our experiments. As noticed, the human-software (for both human operators) is very high (i.e., > 0.94).

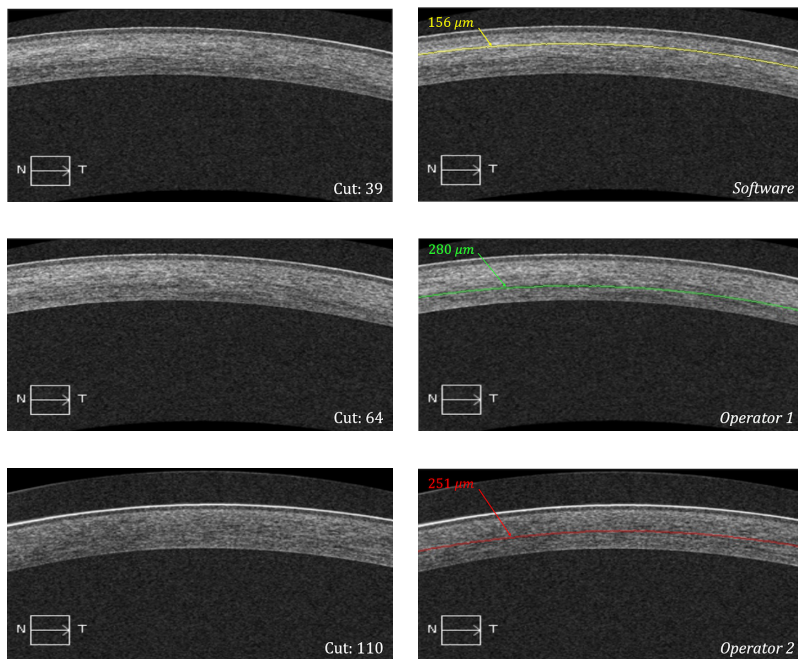
Furthermore, our experimental results report that:

- The mean corneal demarcation line depth was 295.9 ± 59.8 microns as computed by the automated technique.
- The mean corneal demarcation line depth was 314.5 ± 48.4 microns as computed by the human operators.
- The Pearson correlation coefficient between the software and the first operator is 0.918.
- The Pearson correlation coefficient between the software and the second operator is 0.910.
- The average time per OCT examination by the human operator was 31 seconds.
- The average time per OCT examination by the software scored less than a second.

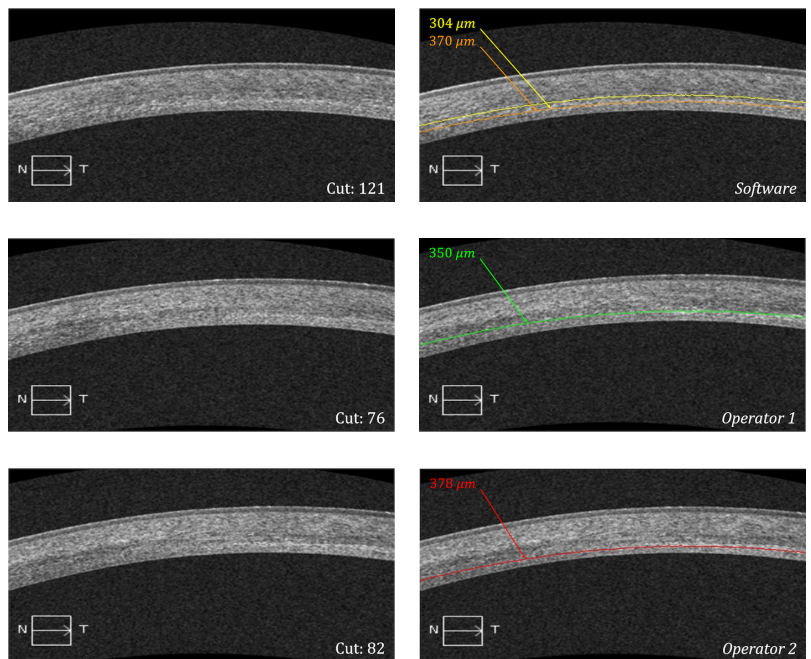
All these results validate the effectiveness and efficiency of the automated tool, and its promise in replacing manual human measurements.



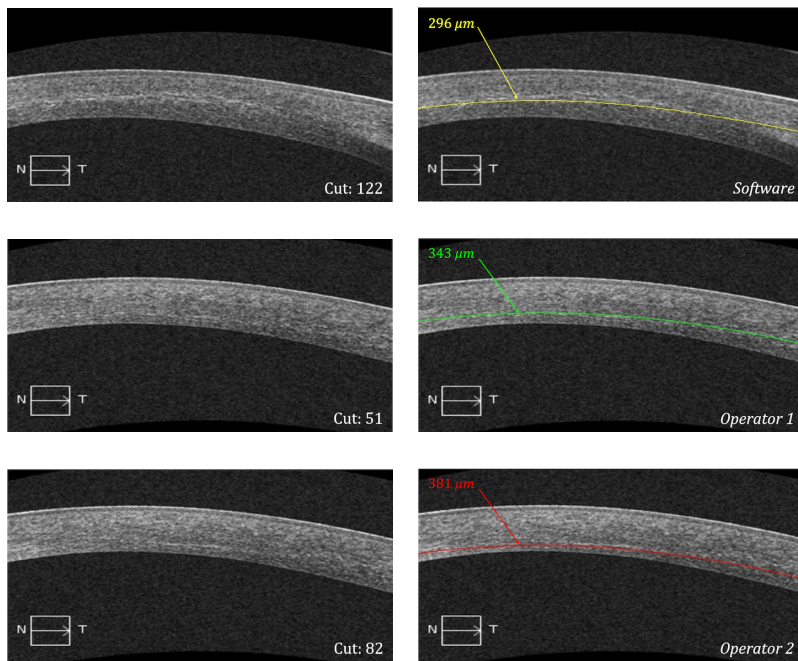
(a) P6: 3-months post-op



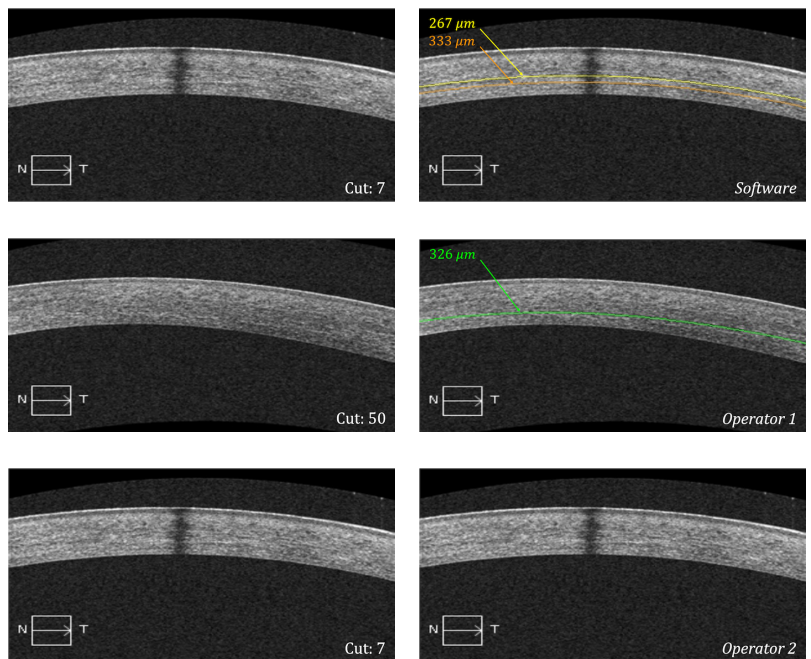
(b) P7: 3-months post-op



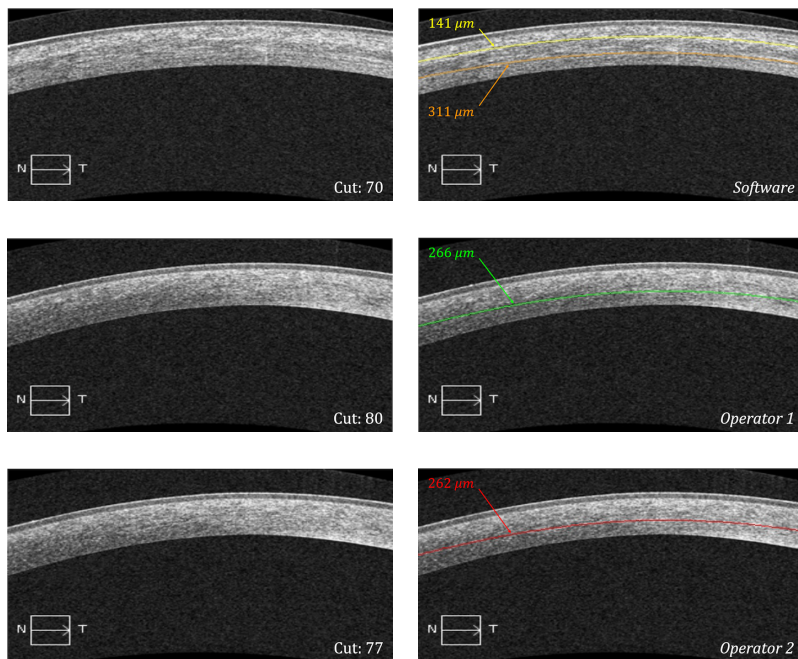
(c) P9: 3-months post-op



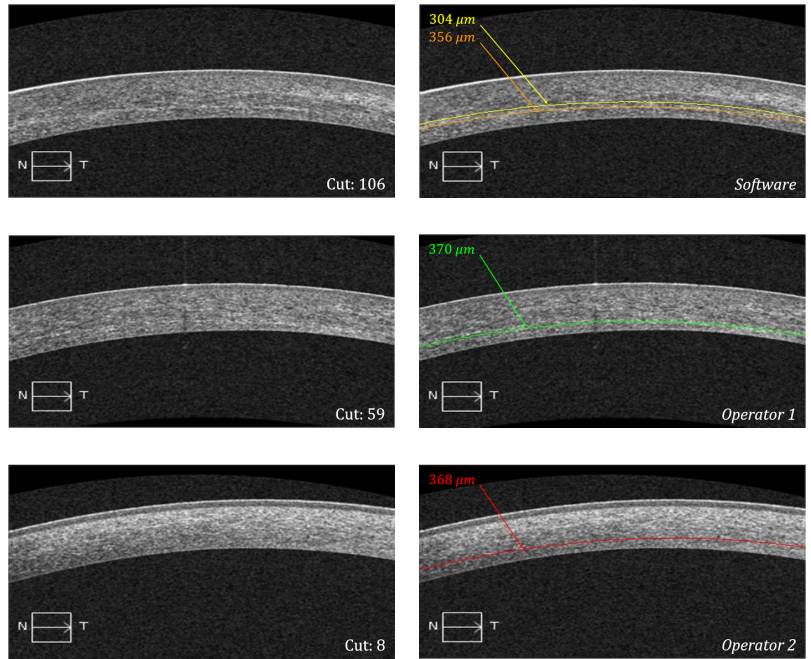
(d) P12: 1-month post-op



(e) P15: 1-month post-op

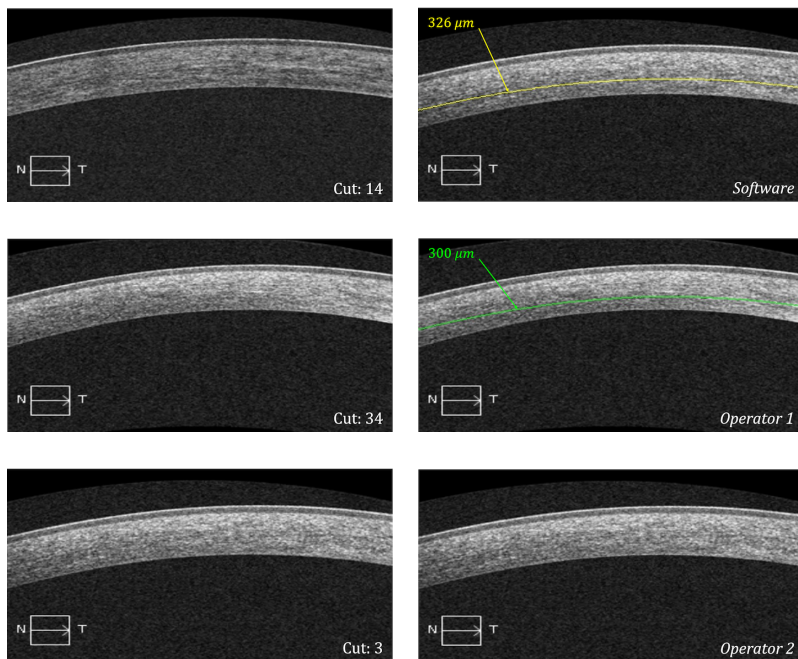


(f) P20: 3-months post-op

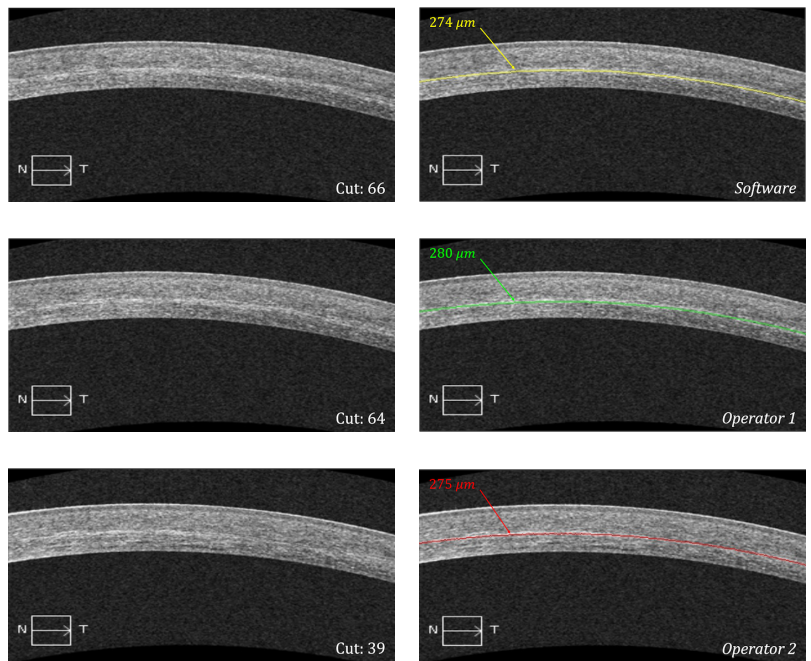


(g) P23: 1-month post-op

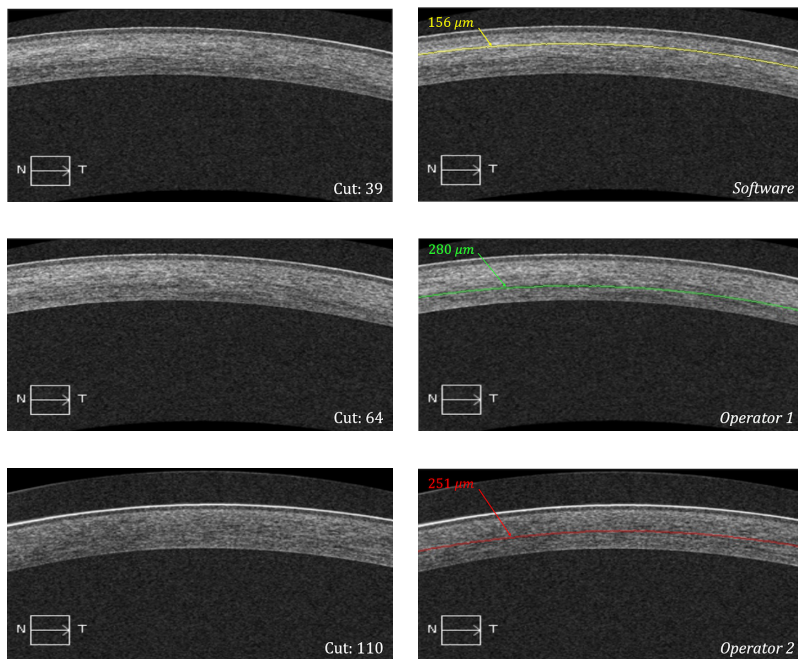
Figure 4.3: Raw and annotated cuts with depth in microns of demarcation line based on human and software operators for the patients with relatively large differences in measurements. The software suggested demarcation line is highlighted yellow, the candidate demarcation line is highlighted in orange when available. The demarcation lines measured by the first and second ophthalmologist are highlighted in green and in red, respectively. The annotation is made on the operator's chosen cut of the OCT image.



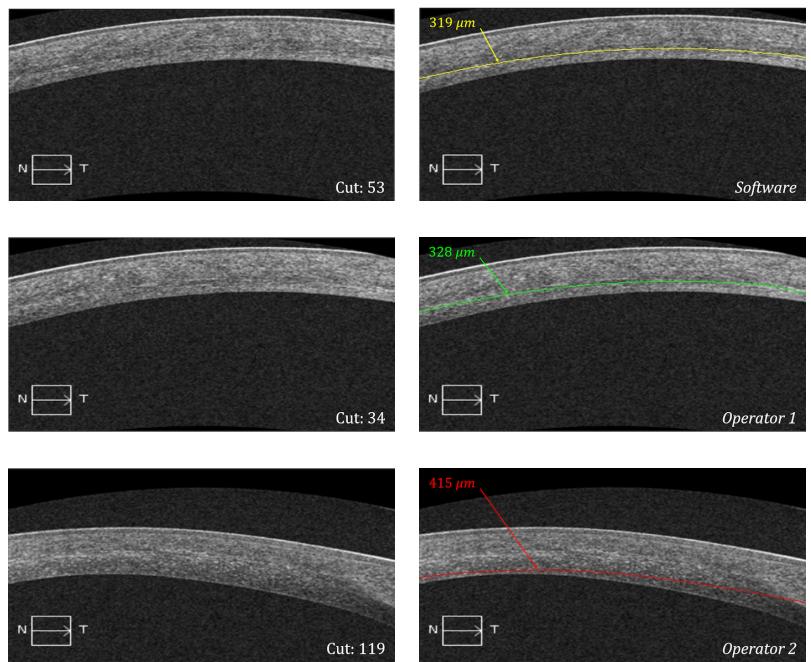
(a) P19: 3-months post-op



(b) P7: 1-month post-op

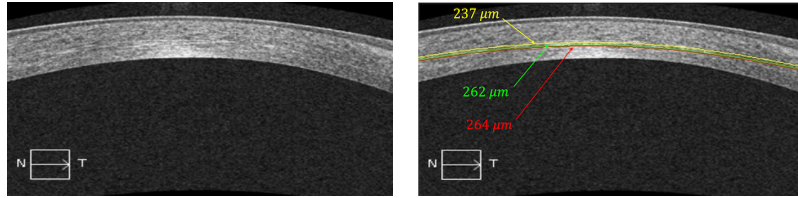


(c) P7: 3-months post-op

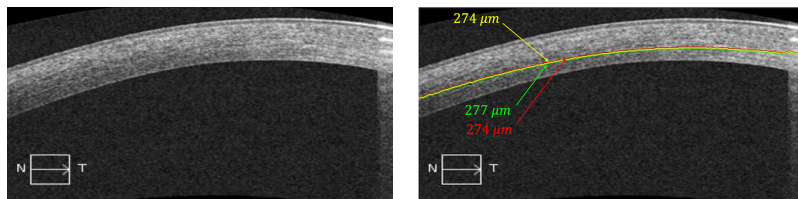


(d) P16: 3-months post-op

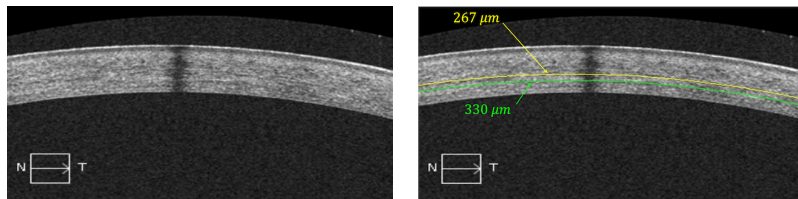
Figure 4.4: Depth of detected demarcation lines for different patients as measured by the software (in yellow), the first operator (in green) and the second operator (in red). The absence of a highlighted line indicates that a demarcation line was not detected by software / operator.



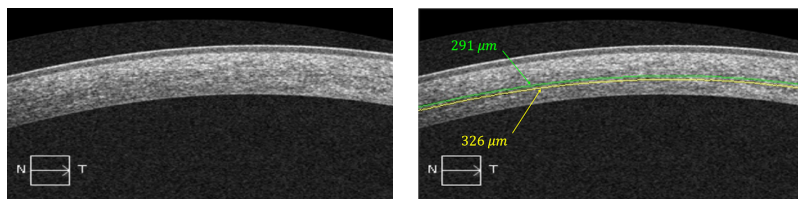
(a) P1: 3-months post-op



(b) P2: 3-months post-op



(c) P15: 1-month post-op



(d) P19: 3-months post-op

Figure 4.5: Raw and annotated cuts of detected demarcation lines for different patients as measured by the software (*suggested*, in yellow), the first operator (in green) and the second operator (in red) based on the 16th percentile minimum score value. The absence of a highlighted line indicates that a demarcation line was not detected by software / operator.

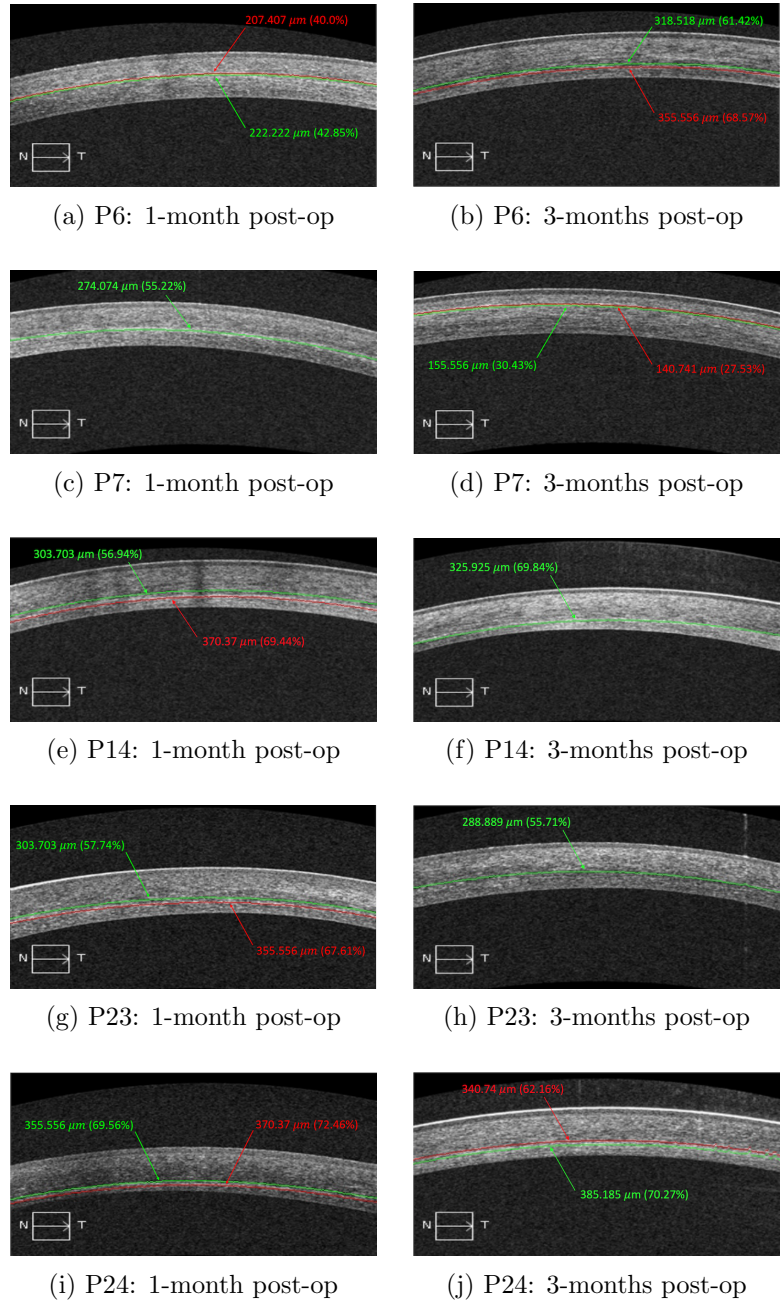


Figure 4.6: Depth of suggested and candidate demarcation lines for different patients in microns and in percentage. The suggested demarcation line is highlighted in green and the candidates in red.

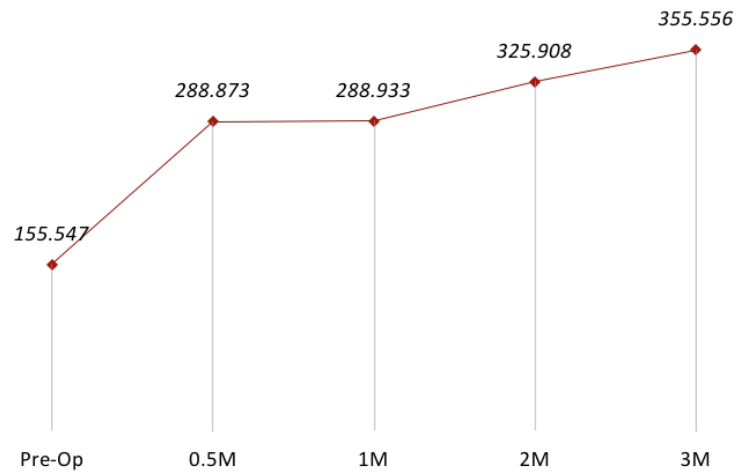


Figure 4.7: The progress in the depth of *suggested* demarcation line for one sample patient over time.

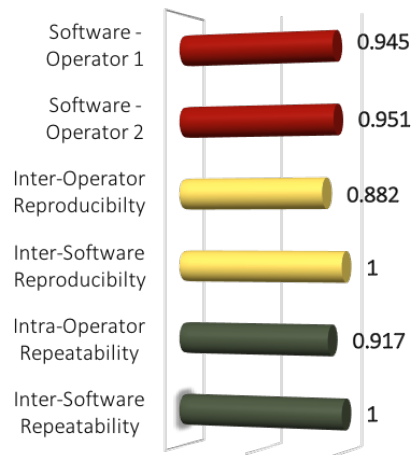


Figure 4.8: Intra-class correlation coefficients (ICC) of the demarcation lines' depths for all patients.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

In this thesis, we designed, implemented and experimented with the first solution for offering automated detection and measurement of demarcation line in OCT images. The new technology makes use of image analysis libraries such as OpenCV; it also includes new image processing techniques, and a customized machine learning approach. The tool takes as input an OCT scan (in the form of a video of 128 transversal cuts) of the anterior segment of the cornea and provides as output an annotated image exhibiting the detection and measurement of the corneal demarcation line, and related haze statistics.

Some of the image cuts taken by the OCT machine can be damaged or distorted. These cuts affect the measurements and therefore should be filtered out. Thus, as a first step in the process of demarcation line detection, we use a supervised machine learning approach to validate the different cuts of an OCT image. Hence, only valid cuts (i.e., showing the full structure of the cornea without any distortion) are processed by the software in the subsequent steps.

Once the OCT image is validated, the stromal region of each valid cut is explored to detect potential demarcation lines. We give a score to each line in that region; this score takes into account the brightness level and the continuity of bright pixels. The scoring technique is complemented with two statistical approaches in order to validate and rank potential demarcation lines. Subsequently, our novel technology chooses the cut in which the demarcation line gets the

highest score and *enough* consistency (i.e., present in more than 15% of the total number of valid cuts of one OCT image).

Experiments were done on a set of 40 OCT images (provided by our research partners at AUBMC). We compared between the values reported by ophthalmologists and different values computed using our automated approach. We interpreted the results of the comparison using the following statistical indicators: ICC and PCC. The results show an ICC and a PCC between the automated detection software and two human operators equal to 0.945 and 0.951, and 0.910 and 0.918, respectively. The ICC for inter-operators reproducibility is 0.882, and for intra-operator repeatability is 0.917. The ICC for inter-software repeatability is 1, and the average time spent by the operator per OCT examination is 31 seconds, as compared to 1 second by the software. Note that when the value of PCC is near 1 and when the value of ICC is greater than 0.75, the data is considered to have **excellent positive** agreement.

These results highlight the effectiveness and efficiency of the automated tool. The novel technology will enable ophthalmologists to obtain **fast**, **automated**, **accurate** and **objective** detection and measurement of corneal haze. The novel technology also has the promise of replacing the manual human measurements in the future; thus, it can become the standard of care in cross-linking surgeries for Keratoconus patients or any other corneal thinning processes.

5.2 Future Work

As future work, the following sections describe potential extensions and improvements to our solution:

5.2.1 Detection of Demarcation Line over Time

In the case where we have the OCT images of the same eye for different periods, the choice of suggested/candidate demarcation lines is no longer based purely on one video/image (i.e., at one period). A demarcation line's depth may vary across time, but the variation is relatively low. Thus, an algorithm, that takes

into consideration n sets of suggested/candidate demarcation lines at n different post-op periods, and that would then pick the most probable combination of demarcation lines depths, may be proposed.

5.2.2 Processing 8mm/9mm OCT Images

Most widely-used OCT scanners generate images that include a 4mm wide shot of the cornea, which has approximately an 11mm of diameter. However, most recent OCT machines generate a single OCT scan, that has 8mm/9mm shots of the cornea. The algorithm proposed for 4mm images can hence be extended so that it is accurately applied on 8mm/9mm images.

5.2.3 3D Reconstruction of OCT Cuts

The detection and depth measurement of the demarcation line can be done via modeling and reconstructing the OCT cuts in 3D (where the third dimension is the cut number). This can provide a holistic view of the demarcation line, which can help measure its depth more accurately. Subsequently, the software can pick the "clearest" slice and report it to the ophthalmologists.

Appendix A

Glossary

CCT	Central Corneal Thickness
CXL	Collagen Cross-Linking
FN	False Negative
FP	False Positive
ICC	Intraclass Correlation Coefficient
InfoGain	Information Gain
MRI	Magnetic Resonance Imaging
OCT	Optical Coherence Tomography
PCA	Principal Components Analysis
PCC	Pearson Correlation Coefficient
PRK	Photo-Refractive Keratectomy
SMO	Sequential Minimal Optimization
SVM	Support Vector Machine
TN	True Negative
TP	True Positive

Appendix B

Learning Tables

Table B.1: Logistic Regression

Ridge	TP	FP	TN	FN	Precision	Recall	Accuracy	F-measure
1.0E-8	5386	90	175	800	0.9835	0.8706	0.8620	0.9236

Table B.2: SMO with a Polynomial Kernel of degree 1 and Complexity
Constant 'c' = 1

Epsilon	Tolerance	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
1.00E-12	0.001	5456	20	105	870	0.9963	0.8625	0.862	0.9246
	0.01	5460	16	103	872	0.9971	0.8623	0.8623	0.9248
	0.1	5457	19	108	867	0.9965	0.8629	0.8627	0.9249
1.00E-10	0.001	5460	16	106	869	0.9971	0.8627	0.8628	0.925
	0.01	5462	14	104	871	0.9974	0.8625	0.8628	0.9251
	0.1	5457	19	103	872	0.9965	0.8622	0.8619	0.9245
1.00E-08	0.001	5434	42	143	832	0.9923	0.8672	0.8645	0.9256
	0.01	5427	49	144	831	0.9911	0.8672	0.8636	0.925
	0.1	5443	33	139	836	0.994	0.8669	0.8653	0.9261
1.00E-06	0.001	4843	633	332	643	0.8844	0.8828	0.8022	0.8836
	0.01	5388	88	153	822	0.9839	0.8676	0.8589	0.9221
	0.1	5104	372	236	739	0.9321	0.8735	0.8278	0.9018

Table B.3: SMO with a Polynomial Kernel of degree 1 and Complexity
Constant 'c' = 10

Epsilon	Tolerance	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
1.00E-12	0.001	5454	22	125	850	0.996	0.8652	0.8648	0.926
	0.01	5454	22	122	853	0.996	0.8648	0.8644	0.9257
	0.1	5453	23	117	858	0.9958	0.864	0.8634	0.9253
1.00E-10	0.001	5451	25	119	856	0.9954	0.8643	0.8634	0.9252
	0.01	5451	25	119	856	0.9954	0.8643	0.8634	0.9252
	0.1	5441	35	124	851	0.9936	0.8647	0.8627	0.9247
1.00E-08	0.001	4504	972	379	596	0.8225	0.8831	0.7569	0.8517
	0.01	5188	288	225	750	0.9474	0.8737	0.8391	0.9091
	0.1	4549	927	369	606	0.8307	0.8824	0.7624	0.8558
1.00E-06	0.001	4634	842	301	674	0.8462	0.873	0.765	0.8594
	0.01	4678	798	310	665	0.8543	0.8755	0.7732	0.8648
	0.1	4634	842	335	640	0.8462	0.8786	0.7703	0.8621

Table B.4: SMO with a Polynomial Kernel of degree 1 and Complexity
Constant 'c' = 100

Epsilon	Tolerance	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
1.00E-12	0.001	5455	21	124	851	0.9962	0.865	0.8648	0.926
	0.01	5448	28	122	853	0.9949	0.8646	0.8634	0.9252
	0.1	5452	24	124	851	0.9956	0.865	0.8644	0.9257
1.00E-10	0.001	5454	22	119	856	0.996	0.8643	0.8639	0.9255
	0.01	5447	29	124	851	0.9947	0.8649	0.8636	0.9253
	0.1	5448	28	125	850	0.9949	0.865	0.8639	0.9254
1.00E-08	0.001	4977	499	240	735	0.9089	0.8713	0.8087	0.8897
	0.01	5087	389	199	776	0.929	0.8676	0.8194	0.8973
	0.1	5208	268	213	762	0.9511	0.8724	0.8403	0.91
1.00E-06	0.001	3725	1751	425	550	0.6802	0.8713	0.6433	0.764
	0.01	3532	1944	489	486	0.645	0.879	0.6233	0.744
	0.1	4418	1058	336	639	0.8068	0.8736	0.7369	0.8389

Table B.5: SMO with a Polynomial Kernel of degree 2 and Complexity
Constant 'c' = 1

Epsilon	Tolerance	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
1.00E-12	0.001	5428	48	256	719	0.9912	0.883	0.8811	0.934
	0.01	5422	54	266	709	0.9901	0.8844	0.8817	0.9343
	0.1	5426	50	259	716	0.9909	0.8834	0.8813	0.9341
1.00E-10	0.001	5422	54	265	710	0.9901	0.8842	0.8816	0.9342
	0.01	5429	47	255	720	0.9914	0.8829	0.8811	0.934
	0.1	5439	37	260	715	0.9932	0.8838	0.8834	0.9353
1.00E-08	0.001	4530	946	394	581	0.8272	0.8863	0.7633	0.8558
	0.01	4359	1117	417	558	0.796	0.8865	0.7404	0.8388
	0.1	4442	1034	411	564	0.8112	0.8873	0.7523	0.8475
1.00E-06	0.001	4641	835	354	621	0.8475	0.882	0.7743	0.8644
	0.01	4683	793	346	629	0.8552	0.8816	0.7796	0.8682
	0.1	4334	1142	454	521	0.7915	0.8927	0.7422	0.839

Table B.6: SMO with a Polynomial Kernel of degree 2 and Complexity
Constant 'c' = 10

Epsilon	Tolerance	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
1.00E-12	0.001	5247	229	345	630	0.9582	0.8928	0.8668	0.9243
	0.01	5248	228	331	644	0.9584	0.8907	0.8648	0.9233
	0.1	5242	234	316	659	0.9573	0.8883	0.8616	0.9215
1.00E-10	0.001	5239	237	331	644	0.9567	0.8905	0.8634	0.9224
	0.01	5250	226	328	647	0.9587	0.8903	0.8647	0.9232
	0.1	5235	241	331	644	0.956	0.8905	0.8628	0.9221
1.00E-08	0.001	4330	1146	427	548	0.7907	0.8877	0.7374	0.8364
	0.01	4606	870	341	634	0.8411	0.879	0.7669	0.8596
	0.1	4337	1139	411	564	0.792	0.8849	0.736	0.8359
1.00E-06	0.001	3840	1636	555	420	0.7012	0.9014	0.6813	0.7888
	0.01	4658	818	370	605	0.8506	0.885	0.7794	0.8675
	0.1	4486	990	392	583	0.8192	0.885	0.7562	0.8508

Table B.7: SMO with a Polynomial Kernel of degree 2 and Complexity
Constant 'c' = 100

Epsilon	Tolerance	TP	FP	TN	FN	Precision	Recall	Accuracy	F-measure
1.00E-12	0.001	4960	516	411	564	0.9058	0.8979	0.8326	0.9018
	0.01	4949	527	405	570	0.9038	0.8967	0.8299	0.9002
	0.1	4961	515	408	567	0.906	0.8974	0.8323	0.9017
1.00E-10	0.001	4964	512	388	587	0.9065	0.8943	0.8296	0.9003
	0.01	4958	518	406	569	0.9054	0.8971	0.8315	0.9012
	0.1	4967	509	398	577	0.907	0.8959	0.8317	0.9015
1.00E-08	0.001	4505	971	384	591	0.8227	0.884	0.7579	0.8523
	0.01	4167	1309	424	551	0.761	0.8832	0.7117	0.8175
	0.1	4751	725	336	639	0.8676	0.8814	0.7886	0.8745
1.00E-06	0.001	4867	609	326	649	0.8888	0.8823	0.805	0.8856
	0.01	3579	1897	568	407	0.6536	0.8979	0.6428	0.7565
	0.1	4509	967	375	600	0.8234	0.8826	0.7571	0.852

Table B.8: Multilayer Perceptron with 'a' Hidden Layers

Learning Rate	Training Time	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
0.1	500	5124	352	394	581	0.9357	0.8982	0.8554	0.9166
	700	5103	373	400	575	0.9319	0.8987	0.853	0.915
	900	5093	383	409	566	0.9301	0.9	0.8529	0.9148
0.2	500	5011	465	395	580	0.9151	0.8963	0.838	0.9056
	700	5045	431	390	585	0.9213	0.8961	0.8425	0.9085
	900	5089	387	410	565	0.9293	0.9001	0.8524	0.9145
0.3	500	5010	466	445	530	0.9149	0.9043	0.8456	0.9096
	700	5061	415	388	587	0.9242	0.8961	0.8447	0.9099
	900	5042	434	428	547	0.9207	0.9021	0.8479	0.9113
0.4	500	5125	351	392	583	0.9359	0.8979	0.8552	0.9165
	700	5010	466	418	557	0.9149	0.8999	0.8414	0.9074
	900	5091	385	397	578	0.9297	0.898	0.8507	0.9136
0.5	500	5071	405	408	567	0.926	0.8994	0.8493	0.9125
	700	5064	405	408	567	0.9259	0.8993	0.8492	0.9124
	900	5066	410	436	539	0.9251	0.9038	0.8529	0.9144
0.6	500	5067	409	434	541	0.9253	0.9035	0.8527	0.9143
	700	5109	367	389	586	0.933	0.8971	0.8523	0.9147
	900	5098	378	406	569	0.931	0.8996	0.8532	0.915
0.7	500	5117	359	426	549	0.9344	0.9031	0.8592	0.9185
	700	5061	415	428	547	0.9242	0.9025	0.8509	0.9132
	900	5070	406	418	557	0.9259	0.901	0.8507	0.9133
0.8	500	5104	372	404	571	0.9321	0.8994	0.8538	0.9154
	700	5106	370	416	559	0.9324	0.9013	0.856	0.9166
	900	5119	357	380	595	0.9348	0.8959	0.8524	0.9149
0.9	500	5086	390	435	540	0.9288	0.904	0.8558	0.9162
	700	5137	339	430	545	0.9381	0.9041	0.863	0.9208
	900	5076	400	423	552	0.927	0.9019	0.8524	0.9143

Table B.9: Multilayer Perceptron with 'i' Hidden Layers

Learning Rate	Training Time	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
0.1	500	5075	401	405	570	0.9268	0.899	0.8495	0.9127
	700	5067	409	402	573	0.9253	0.8984	0.8478	0.9117
	900	5006	470	426	549	0.9142	0.9012	0.842	0.9076
0.2	500	4978	498	433	542	0.9091	0.9018	0.8388	0.9054
	700	5023	453	407	568	0.9173	0.8984	0.8417	0.9077
	900	5079	397	382	593	0.9275	0.8955	0.8465	0.9112
0.3	500	5078	398	407	568	0.9273	0.8994	0.8503	0.9131
	700	5014	462	407	568	0.9156	0.8982	0.8403	0.9069
	900	5115	361	378	597	0.9341	0.8955	0.8515	0.9144
0.4	500	5039	437	422	553	0.9202	0.9011	0.8465	0.9106
	700	5052	424	420	555	0.9226	0.901	0.8482	0.9117
	900	5072	404	397	578	0.9262	0.8977	0.8478	0.9117
0.5	500	5117	359	394	581	0.9344	0.898	0.8543	0.9159
	700	5075	401	416	559	0.9268	0.9008	0.8512	0.9136
	900	5116	360	403	572	0.9343	0.8994	0.8555	0.9165
0.6	500	5123	353	418	557	0.9355	0.9019	0.8589	0.9184
	700	5124	352	396	579	0.9357	0.8985	0.8557	0.9167
	900	5058	418	419	556	0.9237	0.901	0.849	0.9122
0.7	500	5138	338	406	569	0.9383	0.9003	0.8594	0.9189
	700	5134	342	421	554	0.9375	0.9026	0.8611	0.9197
	900	5068	408	450	525	0.9255	0.9061	0.8554	0.9157
0.8	500	5048	428	448	527	0.9218	0.9055	0.852	0.9136
	700	5074	402	435	540	0.9266	0.9038	0.854	0.9151
	900	5128	348	393	582	0.9364	0.8981	0.8558	0.9169
0.9	500	5125	351	426	549	0.9359	0.9032	0.8605	0.9193
	700	5079	397	440	535	0.9275	0.9047	0.8555	0.916
	900	5138	338	413	562	0.9383	0.9014	0.8605	0.9195

Table B.10: Multilayer Perceptron with 't' Hidden Layers

Learning Rate	Training Time	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
0.1	500	5114	362	396	579	0.9339	0.8983	0.8541	0.9157
	700	4962	514	448	527	0.9061	0.904	0.8386	0.9051
	900	5079	397	403	572	0.9275	0.8988	0.8498	0.9129
0.2	500	5052	424	388	587	0.9226	0.8959	0.8433	0.909
	700	5058	418	415	560	0.9237	0.9003	0.8484	0.9118
	900	5083	393	412	563	0.9282	0.9003	0.8518	0.914
0.3	500	5046	430	405	570	0.9215	0.8985	0.845	0.9098
	700	5033	443	447	528	0.9191	0.9051	0.8495	0.912
	900	5073	403	411	564	0.9264	0.8999	0.8501	0.913
0.4	500	5020	456	428	547	0.9167	0.9017	0.8445	0.9092
	700	5049	427	399	576	0.922	0.8976	0.8445	0.9096
	900	5052	424	408	567	0.9226	0.8991	0.8464	0.9107
0.5	500	5072	404	417	558	0.9262	0.9009	0.8509	0.9134
	700	5129	347	405	570	0.9366	0.9	0.8579	0.9179
	900	5107	369	404	571	0.9326	0.8994	0.8543	0.9157
0.6	500	5140	336	381	594	0.9386	0.8964	0.8558	0.917
	700	5146	330	395	580	0.9397	0.8987	0.8589	0.9188
	900	5127	349	383	592	0.9363	0.8965	0.8541	0.9159
0.7	500	5130	346	404	571	0.9368	0.8998	0.8579	0.918
	700	5111	365	424	551	0.9333	0.9027	0.858	0.9178
	900	5094	382	401	574	0.9302	0.8987	0.8518	0.9142
0.8	500	5175	301	390	585	0.945	0.8984	0.8627	0.9211
	700	5117	359	417	558	0.9344	0.9017	0.8579	0.9178
	900	5086	390	406	569	0.9288	0.8994	0.8513	0.9138
0.9	500	5124	352	398	577	0.9357	0.8988	0.856	0.9169
	700	5061	415	430	545	0.9242	0.9028	0.8512	0.9134
	900	5047	429	434	541	0.9217	0.9032	0.8496	0.9123

Table B.11: Multilayer Perceptron with 'o' Hidden Layers

Learning Rate	Training Time	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	Precision	Recall	Accuracy	F-measure
0.1	500	5145	331	290	685	0.9396	0.8825	0.8425	0.9101
	700	5165	311	299	676	0.9432	0.8843	0.847	0.9128
	900	5158	318	311	664	0.9419	0.8859	0.8478	0.9131
0.2	500	5183	293	290	685	0.9465	0.8833	0.8484	0.9138
	700	5090	386	313	662	0.9295	0.8849	0.8375	0.9067
	900	5104	372	310	665	0.9321	0.8847	0.8392	0.9078
0.3	500	5051	425	311	664	0.9224	0.8838	0.8312	0.9027
	700	5068	408	307	668	0.9255	0.8835	0.8332	0.904
	900	5137	339	283	692	0.9381	0.8813	0.8402	0.9088
0.4	500	5135	340	295	680	0.9379	0.8831	0.8419	0.9097
	700	5047	429	320	655	0.9217	0.8851	0.832	0.903
	900	5076	400	320	655	0.927	0.8857	0.8365	0.9059
0.5	500	5148	328	277	698	0.9401	0.8806	0.841	0.9094
	700	5135	341	294	681	0.9377	0.8829	0.8416	0.9095
	900	5066	410	305	670	0.9251	0.8832	0.8326	0.9037
0.6	500	5165	311	274	701	0.9432	0.8805	0.8431	0.9108
	700	5159	317	274	701	0.9421	0.8804	0.8422	0.9102
	900	5044	432	303	672	0.9211	0.8824	0.8289	0.9014
0.7	500	5100	376	308	667	0.9313	0.8843	0.8383	0.9072
	700	5165	311	282	693	0.9432	0.8817	0.8444	0.9114
	900	5114	362	300	675	0.9339	0.8834	0.8392	0.9079
0.8	500	5137	339	322	653	0.9381	0.8872	0.8462	0.9119
	700	5101	375	301	674	0.9315	0.8833	0.8374	0.9068
	900	5046	430	312	663	0.9215	0.8839	0.8306	0.9023
0.9	500	5031	445	308	667	0.9187	0.8829	0.8276	0.9005
	700	5207	269	279	696	0.9509	0.8821	0.8504	0.9152
	900	5092	384	309	666	0.9299	0.8843	0.8372	0.9065

Bibliography

- [1] B. Cassin, S. Solomon, and M. L. Rubin, *Dictionary of eye terminology*. Triad Pub. Co., 1984.
- [2] S. Kling and F. Hafezi, “Corneal biomechanics—a review,” *Ophthalmic and Physiological Optics*, 2017.
- [3] A. Elsheikh, D. Alhasso, and P. Rama, “Assessment of the epithelium’s contribution to corneal biomechanics,” *Experimental eye research*, vol. 86, no. 2, pp. 445–451, 2008.
- [4] T. Kurt, E. B. Ozgurhan, Y. Yildirim, B. I. S. Akcay, M. G. Cosar, E. Bozkurt, and M. Taskapili, “Accelerated (18 mw/cm²) corneal cross-linking for progressive keratoconus: 18-month results,” *Journal of Ocular Pharmacology and Therapeutics*, vol. 32, no. 4, pp. 186–191, 2016.
- [5] G. Wollensak, “Crosslinking treatment of progressive keratoconus: new hope,” *Current opinion in ophthalmology*, vol. 17, no. 4, pp. 356–360, 2006.
- [6] Y. S. Rabinowitz, “Keratoconus,” *Survey of ophthalmology*, vol. 42, no. 4, pp. 297–319, 1998.
- [7] N. Waked, A. Fayad, A. Fadlallah, and H. El Rami, “Keratoconus screening in a lebanese students’ population,” *Journal francais d’ophtalmologie*, vol. 35, no. 1, pp. 23–29, 2012.
- [8] G. Wollensak, E. Spoerl, and T. Seiler, “Riboflavin/ultraviolet-a-induced collagen crosslinking for the treatment of keratoconus,” *American journal of ophthalmology*, vol. 135, no. 5, pp. 620–627, 2003.
- [9] E. B. Ozgurhan, B. I. S. Akcay, T. Kurt, Y. Yildirim, and A. Demirok, “Accelerated corneal collagen cross-linking in thin keratoconic corneas,” *Journal of Refractive Surgery*, vol. 31, no. 6, pp. 386–390, 2015.
- [10] E. B. Ozgurhan, N. Kara, K. I. Cankaya, T. Kurt, and A. Demirok, “Accelerated corneal cross-linking in pediatric patients with keratoconus: 24-month outcomes,” *Journal of Refractive Surgery*, vol. 30, no. 12, pp. 843–849, 2014.

- [11] S. A. Greenstein, K. L. Fry, J. Bhatt, and P. S. Hersh, “Natural history of corneal haze after collagen crosslinking for keratoconus and corneal ectasia: Scheimpflug and biomicroscopic analysis,” *Journal of Cataract & Refractive Surgery*, vol. 36, no. 12, pp. 2105–2114, 2010.
- [12] ZEISS, *Cirrus HD-OCT. User Manual Addendum - Anterior Segment Imaging*. Model 4000.
- [13] M. Mita, G. O. Waring, and M. Tomita, “High-irradiance accelerated collagen crosslinking for the treatment of keratoconus: six-month results,” *Journal of Cataract & Refractive Surgery*, vol. 40, no. 6, pp. 1032–1040, 2014.
- [14] M. Doors, N. G. Tahzib, F. A. Eggink, T. T. Berendschot, C. A. Webers, and R. M. Nuijts, “Use of anterior segment optical coherence tomography to study corneal changes after collagen cross-linking,” *American journal of ophthalmology*, vol. 148, no. 6, pp. 844–851, 2009.
- [15] J. C. Yam, C. W. Chan, A. C. Cheng, *et al.*, “Corneal collagen cross-linking demarcation line depth assessed by visante oct after cxl for keratoconus and corneal ectasia,” *Journal of Refractive Surgery*, vol. 28, no. 7, pp. 475–481, 2012.
- [16] G. D. Kymionis, M. A. Grentzelos, A. D. Plaka, N. Stojanovic, K. I. Tsoulnaras, D. G. Mikropoulos, K. I. Rallis, and V. P. Kankariya, “Evaluation of the corneal collagen cross-linking demarcation line profile using anterior segment optical coherence tomography,” *Cornea*, vol. 32, no. 7, pp. 907–910, 2013.
- [17] G. D. Kymionis, M. A. Grentzelos, A. D. Plaka, K. I. Tsoulnaras, V. F. Diakonis, D. A. Liakopoulos, V. P. Kankariya, and A. I. Pallikaris, “Correlation of the corneal collagen cross-linking demarcation line using confocal microscopy and anterior segment optical coherence tomography in keratoconic patients,” *American journal of ophthalmology*, vol. 157, no. 1, pp. 110–115, 2014.
- [18] T. Seiler and F. Hafezi, “Corneal cross-linking-induced stromal demarcation line,” *Cornea*, vol. 25, no. 9, pp. 1057–1059, 2006.
- [19] E. B. Ozgurhan, B. I. Sezgin Akcay, Y. Yildirim, G. Karatas, T. Kurt, and A. Demirok, “Evaluation of corneal stromal demarcation line after two different protocols of accelerated corneal collagen cross-linking procedures using anterior segment optical coherence tomography and confocal microscopy,” *Journal of ophthalmology*, vol. 2014, 2014.
- [20] *OpenCV version 3.1: Image processing software*. December 2015.

- [21] E. Frank, M. Hall, G. Holmes, R. Kirkby, B. Pfahringer, I. H. Witten, and L. Trigg, “Weka,” *Data Mining and Knowledge Discovery Handbook*, pp. 1305–1314, 2005.
- [22] K. Kamiya, K. Shimizu, and F. Ohmoto, “Time course of corneal biomechanical parameters after laser in situ keratomileusis,” *Ophthalmic research*, vol. 42, no. 3, pp. 167–171, 2009.
- [23] A. Elsheikh and K. Anderson, “Comparative study of corneal strip extensometry and inflation tests,” *Journal of the Royal Society Interface*, vol. 2, no. 3, pp. 177–185, 2005.
- [24] G. Scarcelli, R. Pineda, and S. H. Yun, “Brillouin optical microscopy for corneal biomechanics,” *Investigative ophthalmology & visual science*, vol. 53, no. 1, pp. 185–190, 2012.
- [25] G. Scarcelli, S. Kling, E. Quijano, R. Pineda, S. Marcos, and S. H. Yun, “Brillouin microscopy of collagen crosslinking: Noncontact depth-dependent analysis of corneal elastic modulus,” *Investigative ophthalmology & visual science*, vol. 54, no. 2, pp. 1418–1425, 2013.
- [26] G. Scarcelli and S. H. Yun, “In vivo Brillouin optical microscopy of the human eye,” *Optics express*, vol. 20, no. 8, pp. 9197–9202, 2012.
- [27] M. Forbes, G. Pico, and B. Grolman, “A noncontact applanation tonometer: description and clinical evaluation,” *Archives of ophthalmology*, vol. 91, no. 2, pp. 134–140, 1974.
- [28] J. Liu, X. He, X. Pan, and C. J. Roberts, “Ultrasonic model and system for measurement of corneal biomechanical properties and validation on phantoms,” *Journal of biomechanics*, vol. 40, no. 5, pp. 1177–1182, 2007.
- [29] S. Wang and K. V. Larin, “Optical coherence elastography for tissue characterization: a review,” *Journal of biophotonics*, vol. 8, no. 4, pp. 279–302, 2015.
- [30] E. Spoerl, G. Wollensak, and T. Seiler, “Increased resistance of crosslinked cornea against enzymatic digestion,” *Current eye research*, vol. 29, no. 1, pp. 35–40, 2004.
- [31] J. Wang, T. L. Simpson, and D. Fonn, “Objective measurements of corneal light-backscatter during corneal swelling, by optical coherence tomography,” *Investigative ophthalmology & visual science*, vol. 45, no. 10, pp. 3493–3498, 2004.

- [32] G. Sallet, “Comparison of optical and ultrasound central corneal pachymetry,” *Bulletin-Société Belge d’Ophthalmologie*, pp. 35–40, 2001.
- [33] P. S. Zhao, T. Y. Wong, W.-L. Wong, S.-M. Saw, and T. Aung, “Comparison of central corneal thickness measurements by visante anterior segment optical coherence tomography with ultrasound pachymetry,” *American journal of ophthalmology*, vol. 143, no. 6, pp. 1047–1049, 2007.
- [34] M. Bechmann, M. J. Thiel, A. S. Neubauer, S. Ullrich, K. Ludwig, K. R. Kenyon, and M. W. Ulbig, “Central corneal thickness measurement with a retinal optical coherence tomography device versus standard ultrasonic pachymetry,” *Cornea*, vol. 20, no. 1, pp. 50–54, 2001.
- [35] S. M. Oud, M. Abdul Fattah, A. R. Dhaini, and S. Awwad, “Automated Detection and Measurement of Corneal Haze using Optical Coherence Tomography in Keratoconus Patients after Crosslinking,” *Proceedings of the 6th Annual AUB Biomedical Research Day*, 2016.
- [36] G. H. Dunteman, *Principal components analysis*. No. 69, Sage, 1989.
- [37] J. T. Kent, “Information gain and a general measure of correlation,” *Biometrika*, vol. 70, no. 1, pp. 163–173, 1983.
- [38] S. Fortmann-Roe, “Understanding the bias-variance tradeoff,” 2012.
- [39] R. R. Picard and R. D. Cook, “Cross-validation of regression models,” *Journal of the American Statistical Association*, vol. 79, no. 387, pp. 575–583, 1984.
- [40] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*, vol. 398. John Wiley & Sons, 2013.
- [41] B. Scholkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001.
- [42] J. Platt, “Sequential minimal optimization: A fast algorithm for training support vector machines,” 1998.
- [43] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014.
- [44] S. K. Pal and S. Mitra, “Multilayer perceptron, fuzzy sets, and classification,” *IEEE Transactions on neural networks*, vol. 3, no. 5, pp. 683–697, 1992.
- [45] I. T. Young, J. J. Gerbrands, and L. J. van Vliet, *Fundamentals of Image Processing*. Delft University of Technology, 2 ed., 2007.

- [46] OpenCV, *The OpenCV Reference Manual*, 3.0.0 ed., June 2014.
- [47] S. Suzuki *et al.*, “Topological structural analysis of digitized binary images by border following,” *Computer vision, graphics, and image processing*, vol. 30, no. 1, pp. 32–46, 1985.
- [48] D. L. Baggio, *OpenCV 3.0 Computer Vision with Java*. Packt Publishing Ltd, 2015.
- [49] I. Sobel and G. Feldman, “A 3x3 isotropic gradient operator for image processing,” *a talk at the Stanford Artificial Project in*, pp. 271–272, 1968.
- [50] O. Vincent and O. Folorunso, “A descriptive algorithm for sobel image edge detection,” in *Proceedings of Informing Science & IT Education Conference (InSITE)*, vol. 40, pp. 97–107, 2009.
- [51] S. Walfish, “A review of statistical outlier methods,” *Pharmaceutical technology*, vol. 30, no. 11, p. 82, 2006.
- [52] W. J. Dixon, F. J. Massey, *et al.*, *Introduction to statistical analysis*, vol. 344. McGraw-Hill New York, 1969.
- [53] T. E. Rhodes, O. Chumburidze, and S. Qiao, “Percentile methodology for probability distributions as applied to the representative scenario method,” *MIB Solutions Inc.*, 2015.
- [54] H. C. Thom, “A note on the gamma distribution,” *Monthly Weather Review*, vol. 86, no. 4, pp. 117–122, 1958.
- [55] I. Kotlarski, “On characterizing the gamma and the normal distribution,” *Pacific Journal of Mathematics*, vol. 20, no. 1, pp. 69–76, 1967.
- [56] J. Benesty, J. Chen, Y. Huang, and I. Cohen, “Pearson correlation coefficient,” in *Noise reduction in speech processing*, pp. 1–4, Springer, 2009.
- [57] P. Sedgwick *et al.*, “Pearsons correlation coefficient,” *Bmj*, vol. 345, no. 7, 2012.
- [58] R. Müller and P. Büttner, “A critical discussion of intraclass correlation coefficients. authors’ reply,” *Statistics in medicine*, vol. 16, no. 7, pp. 821–823, 1997.