UNIVERSIDADE D
COIMBRA

Margarida Sofia Silva de Andrade

Framework to Correct for Eye Motion Artefacts in Optical Coherence Tomography Imaging

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

September, 2022

University of Coimbra

Master in Biomedical Engineering

Framework to Correct for Eye Motion Artefacts in Optical Coherence Tomography Imaging

Margarida Sofia Silva de Andrade

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

Supervisors:
Ph.D. Pedro Guilherme da Cunha Leitão Dias Vaz
Ph.D. Danilo Andrade de Jesus

Coimbra, September 2022

Acknowledgements

First of all, I would like to thank Pedro Vaz, Danilo Jesus and Luisa Brea, for all the patience, help and guidance throughout this year. Thank you for always being available to answer all my doubts and questions. This work could not be done without your support.

I would like to thank all the LIBPhys-UC team, for making me feel welcomed and always being available to help whenever needed. I would also like to thank the BIGR team for receiving me in the Netherlands, offering important advise and making me feel integrated in such a short period of time.

To Ana Luisa, from the Unidade de I&D - Química-Física Molecular, for the guidance provided to produce the retinal layers. To Ana Batista, from CIBIT, for all the help with the OCT acquisition. And to Konstantinos, for the willingness to help and answering all my doubts.

To the amazing group of friends that I met in Coimbra, who helped me live these 5 years to the fullest: Ana, Rute, Ulyana, Beatriz, and specially Telmo, for all the support and long talks in these last 2 master years. You all know how important you are to me, thank you for always being there for me.

To Raquel, Kika and Calisto, thank you for listening to all my dramas, even when you did not understand everything I was talking about.

To Margarida and Gonçalo, thank you for all the patience, and all the memories we made together in these last years.

To all my family, aunts and grandparents, thank you for all the unconditional support.

Last, but certainly not least, I want to thank my parents and sister, for always believing in me, and for being my biggest supporters, without them this journey would not be possible.

To everyone who has helped me somehow throughout this journey, thank you all.

Abstract

Optical Coherence Tomography (OCT) is a noninvasive imaging technology used to obtain cross-sectional images of the retina. Despite being an imaging technique routinely used in clinical practice, correction of motion artefacts remains one of its major challenges. Involuntary eye movements during the OCT acquisition create substantial artefacts in the volumetric data which difficults its interpretation and analysis. Hence, the main goal of this project was to develop a framework for correcting intra-volume motion artefacts in OCT imaging. For that, OCT data with simulated motion artefacts from an eye phantom was obtained and used to study two image registration approaches.

The eye phantom was designed and created with the following structures: the lens (anterior segment), the vitreous body, and three posterior segment layers (the retina, the choroid, and the sclera). This eye phantom was used to acquire a synthetic OCT data set consisting of 156 volumes with intra-volume motion artefacts simulated by two external motors. In this set, two realistic involuntary eye movements were replicated: drifts and microsaccades. Their amplitude and frequency values were within the physiological range (microsaccades with amplitude between 0.25°and 1°, and frequency of 2 revolutions per minute (RPM); drifts with amplitude between 0.05°and 0.75°, and frequency between 0.02RPM and 0.08RPM). These volumes were used to study and compare the performance of two state-of-the-art registration approaches. Both algorithms were chosen to be intensity-based, in order to keep independence from engineered features. One of the algorithms used a conventional image processing, pair-wise approach, and compared pairs of B-scans on a multi-resolution level. The other algorithm used a deep leaning group-wise approach, which allowed it to preserve the retinal curvature better. Dice score and the Euclidean distance were used to objectively evaluate the algorithms. The algorithm with the best performance, the deep learning based model, was tested in a patient database consisting of 146 OCT volumes from 19 patients. The results were further analysed by comparing a pre-trained model on phantom data with a non pre-trained algorithm. The results showed that

the pre-training with the phantom OCT data improves the registration results on clinical data.

Overall, the developed eye phantom has shown to be representative of the main human eye structures allowing the acquisition of OCT data that mimics different involuntary eye movements. The created framework can be used to further improve and validate registration algorithms namely for microsaccades with large amplitudes, which has revealed to be more challenging in this project. In addition, the pre-trained algorithm showed improvements when compared to the non pre-trained algorithm, meaning that smaller data sets can be compensated by the usage of pre-training algorithms on similar registration tasks.

Key-words: Optical Coherence Tomography (OCT); Eye phantom; Involuntary eye movements; Artefacts correction; Deep-learning registration algorithm.

Resumo

Tomografia de Coerência Óptica (OCT) é a técnica de diagnóstico não invasiva utilizada para adquirir imagens de cortes transversais da retina. Apesar de ser uma técnica utilizada recorrentemente em prática clínica, a correção de artefactos do movimento nas imagens adquiridas é ainda um dos maiores desafios nesta área. Durante a aquisição de imagens de OCT, os movimentos involuntários do olho provocam artefactos substanciais, o que dificulta a interpretação e análise da informação adquirida. Desta forma, o principal objetivo deste projeto era desenvolver um algoritmo para corrigir artefactos do movimento para volumes de OCT. Para isso, foram adquiridos volumes de OCT com simulação de movimentos involuntários do olho, e foram ainda estudadas duas abordagens de registro de imagens.

O fantoma do olho foi desenvolvido com as principais estruturas do olho humano: a lente (segmento anterior), o humor vítreo e três camadas posteriores do olho: a retina, a coroide e a esclera. Este fantoma do olho foi utilizado para adquirir um data set sintético, que consiste em 156 volumes de OCT com artefactos do movimento simulados por dois motores externos. Neste data set foram replicados dois movimentos involuntários do olho: drifts e microsaccades. A sua amplitude e frequência está dentro dos valores fisiológicos (microsaccades com amplitudes entre 0.25°e 1°, e frequência de 2 revoluções por minuto (RPM); drifts com amplitudes entre 0.05°e 0.75°, e frequência entre 0.02RPM e 0.08RPM). Estes volumes foram utilizados para estudar e comparar a performance de duas abordagens de registo de imagem. Os dois algoritmos selecionados são algoritmos de análise de intensidade de imagem, para o registo de imagem ser independente de estruturas específicas. Um dos algoritmos utilizados usa um processamento de imagem convencional, com abordagem em pares de scans. O outro algoritmo estudado utiliza uma abordagem deep-learning, com uma abordagem em grupo de scans, o que permitiu preservar a estrutura da retina. A distância Euclideana e o coeficiente de Dice foram utilizados como métricas para avaliar os algoritmos. O algoritmo com a melhor performance, o modelo baseado na abordagem de deep-learning, foi testado num data set de pacientes, que consiste

em 146 volumes de OCT de 19 pacientes. Os resultados foram analisados comparando um algoritmo pré-treinado, em volumes adquiridos com o fantoma do olho, com um algoritmo sem pré-treino. Os resultados mostraram que o algoritmo pré-treinado melhora os resultados em imagens clínicas.

De um modo geral, o fantoma desenvolvido representa as estruturas essenciais do olho humano, permitindo a aquisição de volumes de OCT que simulam os movimentos involuntários do olho. O algoritmo criado pode ser utilizado para melhorar e validar diversos algoritmos de registo de imagem, nomeadamente para microsaccades com maiores amplitudes, que são os movimentos mais difíceis de corrigir. Para além disso, o algoritmo pré-treinado mostrou melhorias quando comparado com o algoritmo sem pré-treino, o que significa que data sets com poucos volumes podem ser melhorados com a utilização de algoritmos pré-treinados em tarefas semelhantes de registo de imagens.

Palavras-chave: Tomografia de Coerência Óptica (OCT); Fantoma do olho; Movimentos involuntários do olho; Correção de artefactos; Algoritmo de registo de imagem deep-learning.

Contents

List of Acronyms

$BaSO_4$ Barium sulfate

FD-OCT Fourier-domain optical coherence tomography

GCL Ganglion cell layer

INL Inner nuclear layer

IPL Inner plexifrom layer

IQR Interquartile range

MRI Magnetic resonance imaging

OCT Optical coherence tomography

ONL Outer nuclear layer

OPL Outer plexiform layer

PCL Polycaprolactone

PDMS Polydimethylsiloxane

PLA Polylactic acid

PVA Polyvynil alcohol

RNFL Retinal nerve fibre layer

RPE Retinal pigment epithelium

RPM Revolutions per minute

SD-OCT Spectral-domain optical coherence tomography

SLO Scanning laser ophtalmoscopy

SS-OCT Swept-source optical coherence tomography

TD-OCT Time-domain optical coherence tomography

TiO_2 Titanium dioxide

US Ultrasound

List of Figures

List of Tables

Chapter 1

Introduction

In this first chapter, the context and main objectives for this dissertation are stated, as well as the configuration of the document. Furthermore, the research team involved in this work is presented.

1.1 Motivation

Optical coherence tomography (OCT) is a noninvasive imaging technique, developed in the beginning of the 1990s [1]. Despite the fact that OCT technology is mostly used in the ophthalmological field, it has also been used in other medical fields, such as cardiovascular and dermatology [2]. With the introduction of the OCT, it became possible to obtain data from tissues that were previously inaccessible without performing the conventional histopathology, which involves tissue extraction and consequent microscopic analysis [3].

OCT is nowadays the gold standard modality for imaging the eye, and can be used to visualize the retinal layers and other structures of interest, with sufficient detail to quantify their morphological characteristics [4]. These characteristics can be linked to a number of diseases and bodily processes. For example, the thickness of specific retinal layers is an important marker to determine whether the eye is healthy or not. Also, different eye diseases can be identified and staged through changes in the thickness and shape of the retinal layers. As a consequence, many retinal diseases, such as glaucoma, the second leading cause of blindness worldwide [5, 6], are diagnosed based on retinal layers irregularities that can be seen and identified in OCT images [7].

Recent studies [8] show that almost a third of the OCT scans have artefacts. Among the different factors that can cause artefacts in an OCT of the retina, such as opacities in the crystalline lens or issues on the acquisition device or protocol, one of the most common is the

involuntary eye movements during the acquisition time. These artefacts present a problem in the visualization of the eye structures and, therefore, in the early detection and progression assessment of retinal diseases. Moreover, healthy subjects can also be wrongly diagnosed with diseases, which can lead to years of unnecessary treatments that can cause serious adverse effects [8].

In order to assist the disease diagnosis/staging using OCT scans, it is important to correct for eye motion artefacts. By correcting these artefacts, scans that were previously discarded due to the presence of irregularities can be used for ophthalmic care [9]. However, motion correction is still one of the biggest challenges in OCT imaging.

1.2 Objectives

The aim of this dissertation was to create a framework for motion correction in OCT scans. To achieve this goal, the work was divided into sequential parts:

- To develop an eye phantom that represents key structures of the human eye, such as the retinal layers.
- To use the eye phantom to acquire an OCT synthetic motion artifacts data set, replicating realistic eye motions and including still volumes to use as reference.
- To use the synthetic data to validate and further refine the results of state-of-the-art algorithms to correct motion artefacts.
  - Objectively quantify and compare the algorithms results compared with the ground truth in synthetic data.
  - Assess the results of the algorithm in a real retina OCT data set, comparing the results when pre-training in synthetic data with the results without pre-training.

The different steps of this project were performed in order to evaluate if using synthetic data, in pre-training of an algorithm, improves the performance of existent motion correction algorithms in the literature.

1.3 Dissertation's content

This dissertation is organized into six chapters.

1. **Introduction:** comprises the main goals and objectives of this work, as well as the research team involved in its development.

2. **Theoretical Background:** introduces the key topics used in this work. First, there is a description of the eye anatomy, followed by the involuntary eye movements and an introduction to the imaging modality used, the OCT. This chapter also contains the state of the art of eye phantoms and motion correction algorithms.

3. **Methods:** details the methods used in the different steps of this project. The chapter begins with the development of the eye model, followed by the OCT data acquisition, simulating the eye involuntary movements, and validating two state-of-the-art motion correction algorithms using the synthetic data set and a clinical data set.

4. **Results and discussion:** presents the results for the different steps of this project. First, the results from the eye model development, followed by the OCT scans acquisition. Second, it quantifies the results obtained using the two state-of-the-art motion correction algorithms. At last, it evaluates the results of the algorithm in a real retina OCT data set, and compares the results when the algorithm was pre-trained in synthetic data with the algorithm that was not pre-trained. Through the whole chapter there is a critical interpretation and discussion of the obtained results. The results are compared with the main goals of this work and the state of the art, when possible.

5. **Conclusion:** comprises the general conclusions of this work and the possible future improvements and research directions.

1.4 Project team

The research team that contributed to this project is presented in Table 1.1. This project was a collaboration between the Laboratory for Instrumentation, Biomedical Engineering and Radiation Physics - University of Coimbra (LIBPhys-UC) and the Biomedical Imaging Group Rotterdam (BIGR), Department of Radiology & Nuclear Medicine from Erasmus Medical Center (Erasmus MC) - The Netherlands.

This project also granted the opportunity to integrate the Erasmus + program, for a three-month internship at Erasmus University Medical Center, in Rotterdam, The Netherlands.

Table 1.1: Research team involved in this project.

Name	**Role**
Margarida Andrade [1,2]	Master student
Pedro Guilherme Vaz [1]	Technical supervisor
Danilo Andrade De Jesus [2]	Technical supervisor
Luisa Sánchez Brea [2]	Technical supervisor

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

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

Chapter 2

Theoretical background

This section describes the theoretical background needed for the development and comprehension of this thesis. First, an introduction to the eye anatomy is presented. Second, the physiological involuntary eye movements are explored. Then, a short reference to the imaging technique used, the Optical Coherence Tomography (OCT), is provided. Finally, the state of the art in eye phantoms and algorithms for correction of motion artifacts in OCT imaging is outlined.

2.1 Anatomy of the eye

The human eye is one of the most complex organs of the body. It is a sensory organ, where specialized neurons (photoreceptors) receive information in the form of light and convert it into signals recognizable by the nervous system. The eye has the approximated shape of a spherical ball. Based on a more accurate measurement, its dimensions are 24.2 mm (transverse) $\times$ 23.7 mm (sagittal) $\times$ 24.4 mm (axial), not showing major differences between sexes and age groups [10]. A sagittal section view of the eye (Figure 2.1) shows the different structures that conform it. Among these structures, three different layers can be distinguished.

The external layer is formed by the cornea and sclera. The cornea is the entry point of the light, and it has a protective function against infections and structural damage. The sclera also has a protective function, maintaining the eye shape and preventing damage, usually caused as a result of internal and external forces [12]. The intermediate layer is divided into two parts: the iris and ciliary body (anterior part), and the choroid (posterior part). Finally, the internal layer corresponds to the retina [11], where the light neurons are situated and the information is converted into nerve impulses that reach the brain via the optic nerve [12]. In the center of the

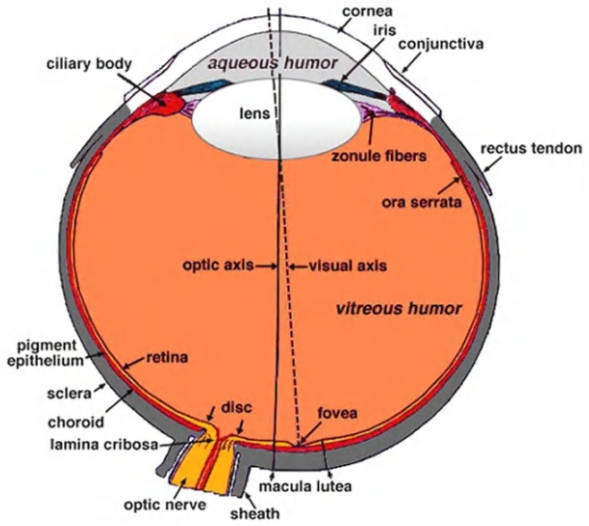


Figure 2.1: Sagittal section of the human eye [11].

retina there is the optic nerve, from where the the major blood vessels of the retina radiate. In the direction of the visual axis, there is an oval-shaped area, blood vessel-free, the fovea [13].

The human eye has a crystalline lens, located at roughly 3 mm inside the eye, behind the iris and suspended by ligaments connected to the ciliary body. The principal function of the crystalline lens is to transmit and focus light onto the retina [14]. The light pathway is depicted in Figure 2.2. It starts with the light entering the eye via its anterior components, the cornea and the crystalline lens, until it reaches the retina. Since the light has to reach the retina, the ocular structures in the visual pathway are transparent. This makes the retina more accessible through non-invasive image techniques.

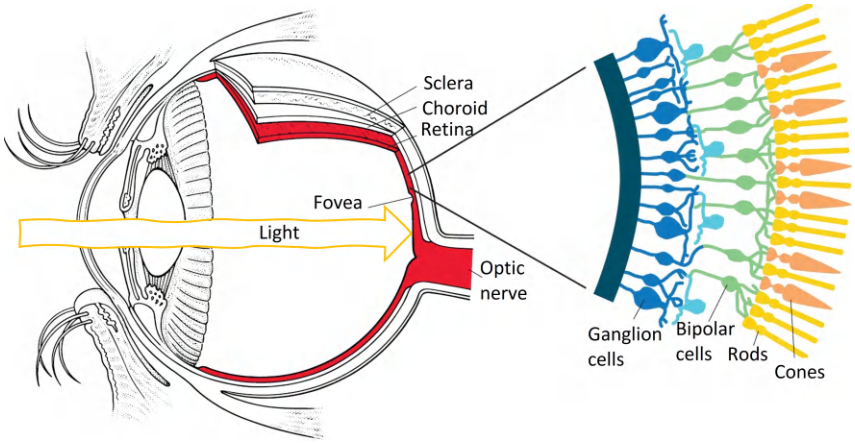


Figure 2.2: Visual pathway of the human eye. Adapted from Saey [15].

The neurons in the retina (Figure 2.2) can be divided into six classes: photoreceptors, bipolar cells, horizontal cells, amacrine cells, ganglion cells, and the Müllerian glia. These different cells

appear at different depths of the retina, dividing the tissue, with a thickness of roughly 200 μm, into 10 different layers, as it can be seen in Figure 2.3.

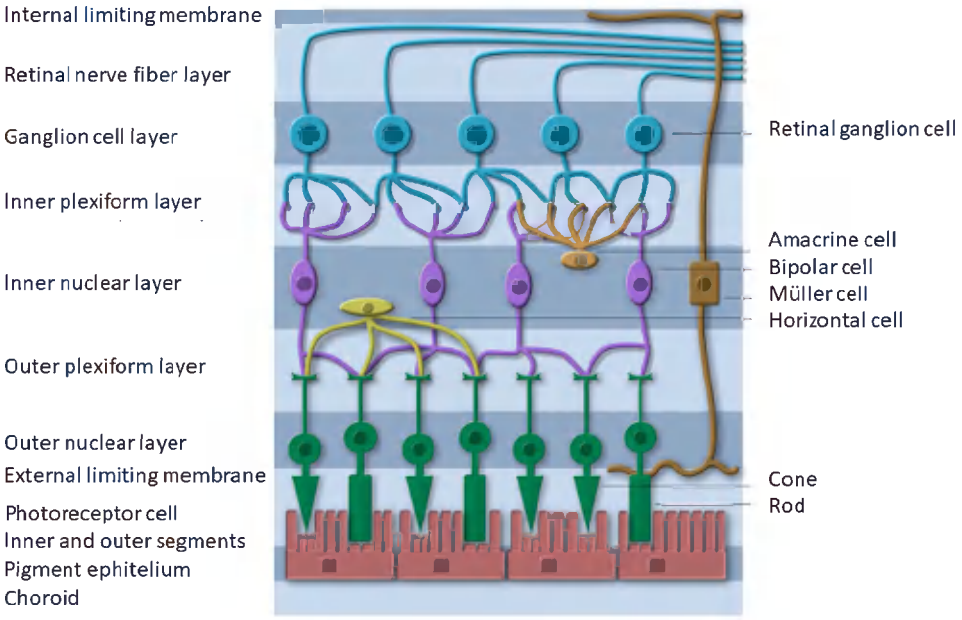


Figure 2.3: Retinal layers [16].

The human retina is supplied by two different blood circulations: the inner retina is supplied by the retinal circulation, and the outer retina is supplied by the choroidal circulation. Studies have shown that around 80% of all the sensory information is originated in the retina, highlighting the importance of the retinal function [17]. In virtue of that, and the fact that the retina has neurons (photoreceptors), the retina can also be referred to as an extension of the brain [18].

2.2 Involuntary eye movements

The retina is affected by three involuntary movements that occur during eye fixation: tremors, drifts, and microsaccades (see Figure 2.4). Besides these eye movements, blood pulsation and respiratory movements play a role in the eye's dynamics [19].

Tremors occur simultaneously with drifts, and they are movements with very small amplitudes, in the range of a single photoreceptor width (6 μm). Tremors are difficult to record accurately due to the fact that their frequency and amplitudes are in the noise range of the recording systems (e.g. OCT systems) [20].

Drifts occur simultaneously with tremors and between microsaccades. Drifts have bigger amplitudes than tremors, and their focal point can move across a dozen of photoreceptors.

Although it seems a random motion of the eye, it was discovered that drifts have a compensatory role in maintaining accurate visual fixation in the absence of microsaccades [20].

Finally, the microsaccades are normally described as fast changes of direction that occur during voluntary fixation. These have the biggest amplitudes, carrying the retinal image throughout several dozens to hundreds of photoreceptors at a frequency of 1-2Hz [21]. Microsaccades are essential to correct possible displacements in the eye position due to drifts. They tend to move the target back to the fovea when drifts carry the image away.

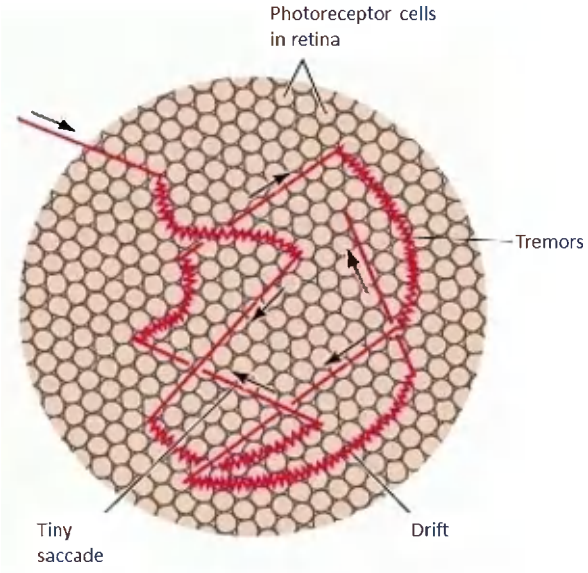


Figure 2.4: Representation of involuntary eye movements [22].

Besides the three fixational eye movements, pulsation and respiration can also induce movements. In these last two cases, the eye does a small translation only in the vertical axis, which did not happened in the previous described movements that were rotational. Although the pulsation and respiration motions are slow and with small amplitude, there is still distortion in OCT images.

2.3 Optical Coherence Tomography

Optical Coherence Tomography (OCT) is a non-invasive and non-contact imaging technique based on low coherence interferometry, that generates cross-sectional images with high resolution.

The first reported OCT system dates back to 1990 [2], and it had a focus on the ophthalmological field. Before that, there were no techniques that allowed the in-depth visualization of

the retina in a non-invasive way. As a consequence, the clinical application of OCT happened in a short period of years after its first demonstrations. Since then, the technique evolved and expanded to several different fields, but it is still mostly used in ophthalmology. OCT enables the visualization of several areas of the eye, such as the anterior segment, the tissue of the retina, and the optic nerve, all without performing tissue extraction, as it was done previously [23].

OCT systems have higher axial and lateral resolutions when compared to other medical imaging methods, such as ultrasound (US) or magnetic resonance imaging (MRI). The lateral resolution of an OCT is around 15-20 μm [24], and the axial resolution, much higher, is up to 3 μm, depending on the light source. However, the lateral and axial resolution of US and MRI is between 100 μm to 1 mm. On the contrary, OCT has a penetration depth in tissue of around 1 mm, whereas US has a higher penetration depth of 10 cm and MRI has a penetration depth of the entire body [25].

The principle behind OCT is light interference. The light from a low-coherence source is split into two pathways, one to the reference arm and the other to the sample arm. When the light exits the arm, the beam shape is controlled by optical components. In the reference arm, the light is reflected by a reference mirror and returns to the interference system, continuing the same path but from the opposite direction. In the sample arm, the process is similar. However the light beam is back scattered by the target sample instead of a reference mirror. Both returning lights recombine at the coupler and form an interference pattern [25].

There are two main OCT technologies, time-domain OCT (TD-OCT), which is not used anymore, and Fourier-domain OCT (FD-OCT). The FD-OCT can be divided into two categories, spectral-domain OCT (SD-OCT) and swept-source OCT (SS-OCT) [25]. In both FD-OCT methods all the light reflections are measured at the same time, which improves the sensitivity of the detection. Besides that, the reference arm is a static mirror and not a moving one, as it previously happened in the TD-OCT technology, which speeds up the data acquisition process.

The difference between SD-OCT and SS-OCT is that SD-OCT uses broad bandwidth light source and a spectrometer for analyzing interferences between the sample and reference beam using Fourier transform [26]. Contrarily, SS-OCT uses a narrow bandwidth light source, that can vary between different wavelengths over the time of acquisition. As a consequence, there is no need to use a spectrometer, since the light is already divided into a single spectral component. For this reason, SS-OCT enables faster acquisitions leading to less artefacts in scans. Figure 2.5 shows a representation of both systems.

Two orthogonal scanning directions can be identified in an OCT acquisition: fast and slow.

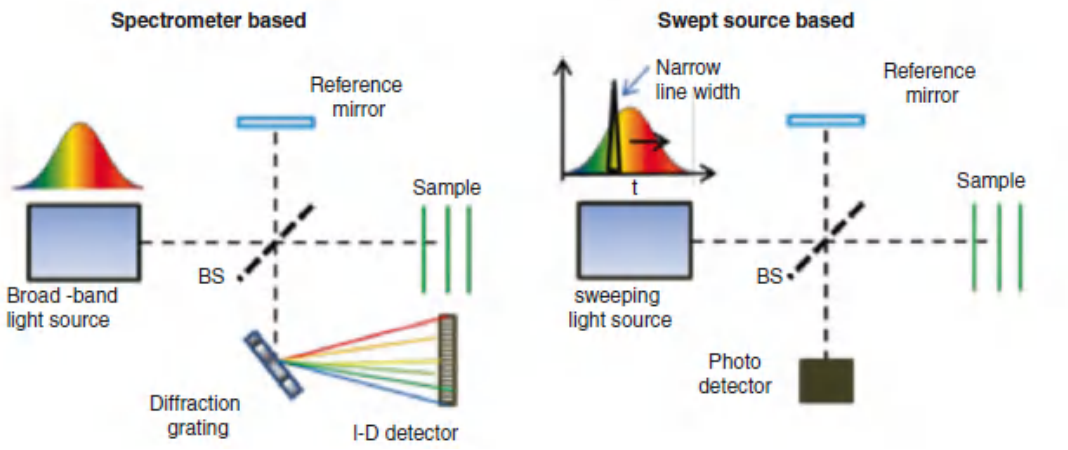


Figure 2.5: Representation of a SD-OCT (left) and a SS-OCT (right) system [23].

An A-scan is an one-dimensional amplitude scan, where the OCT signal is recorded across the sample surface. In order to create a B-scan, a two-dimensional scan, it is necessary to perform several A-scans along a scanning direction (fast axis) to obtain a B-scan. To reconstruct a three-dimensional OCT volume it is necessary to obtain consecutively several B-scans along the other scanning direction (slow axis). Figure 2.6 shows how an OCT volume is acquired. The motions artefacts in OCT imaging are usually between B-scans [27].

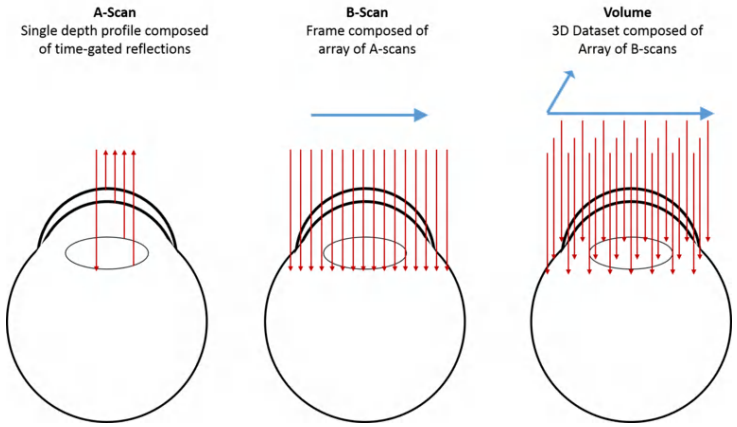


Figure 2.6: Schematic representation of OCT acquisition (A-scan, B-scan, and volume) [28].

There is also a different acquisition technique, en face OCT imaging. It produces frontal images of retinal layers at specific depth, also referred to in the literature as C-scans. This technique emerged through the development of software and data processing, and derives from the SD-OCT technique [29].

2.4 State of the art

2.4.1 Human eye phantoms

Throughout the years, the need for developing eye phantoms to mimic the real tissues and structures of the human eye has increased. As a consequence, the number of works describing human eye phantoms has increased in the past few years. Eye phantoms are designed to accomplish several objectives, such as instrument optimization, initial system testing, and performance comparison between different systems [30].

From the literature, the human eye phantoms can be roughly divided into three groups: tissue phantoms, blood flow phantoms, and mechanical phantoms.

Tissue phantoms

Tissue phantoms, as the one developed by Baxi *et al.* [31], are commonly used for performance evaluation of clinical OCT devices. In these phantoms, the main goal is to obtain an image of the phantom retina as similar as possible to the human retina. It is, therefore, especially important to replicate as many retinal layers of the eye as possible, with the correct thickness and optical properties. Some of the most complex eye phantoms have all the retinal layers, others, such as the one presented by Lee *et al.* [32], select only the most important ones.

The method of fabrication of these phantoms does not change significantly between studies. The most common process involves spin coating of polymers, such as Polydimethylsiloxane (PDMS) [31–35] or Epoxy Resin [36, 37], although PDMS is the most used. By spin coating the polymer mixture at different speeds it is possible to achieve different layer thickness for the desired layers [33].

Baxi *et al.* [31] phantom design replicates the retinal layers as well as the surface topography of the foveal pit (see Figure 2.7). This phantom was used as an assessment tool, to evaluate and standardize OCT performance. To replicate all the retinal layers with accuracy, the authors used thin scattering films of PDMS, fabricated layer by layer. The correct thickness was obtained using a spin coating protocol. The film thickness is dependent on the rotational speed, which was selected accordingly to each layer. Nano and microparticles, for example titanium dioxide (TiO_2) and barium sulfate ($BaSO_4$), were embedded in the PDMS mixture to replicate the retinal layers scaterring (optical) properties. After the fabrication of the last layer, a custom laser microetching technique, using a fiber laser, was used to cut in a fovea-like structure on the surface of the layers.

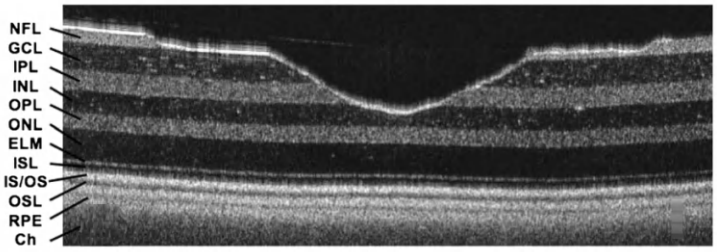


Figure 2.7: OCT image of the phantom, with all retinal layers replicated, using a spin coating technique, and the foveal pit, made using a laser microetching technique [31].

In order to use this phantom in an OCT system, the replicated parts were inserted into a water-filled model. The final result is shown in Figure 2.8.

Figure 2.8: Result phantom, with the replicated retinal eye layers embedded into the commercial human eye model [31].

There are other phantoms that also replicate the foveal pit, but using a different method, for example, the phantom made by Lee *et al.* [33], that can be seen in Figure 2.9. This phantom was made in three separated stages. In the first stage, the five top layers are replicated, from retinal nerve fibre layer (RNFL) to outer plexiform layer (OPL), including the foveal pit. The foveal pit was replicated using a glass mold inside the PDMS layers. The second stage replicates the layers from the outer nuclear layer (ONL) up to retinal pigment epithelium (RPE). The third and final stage replicates the choroid layer.

This method of construction was selected because it allowed the foveal pit to be restricted to the first five layers, making the phantom more realistic to the real human eye. Furthermore, by separating the bottom layers it is also possible to replicate more diseases states, which are more frequent in the layers near the RPE [33].

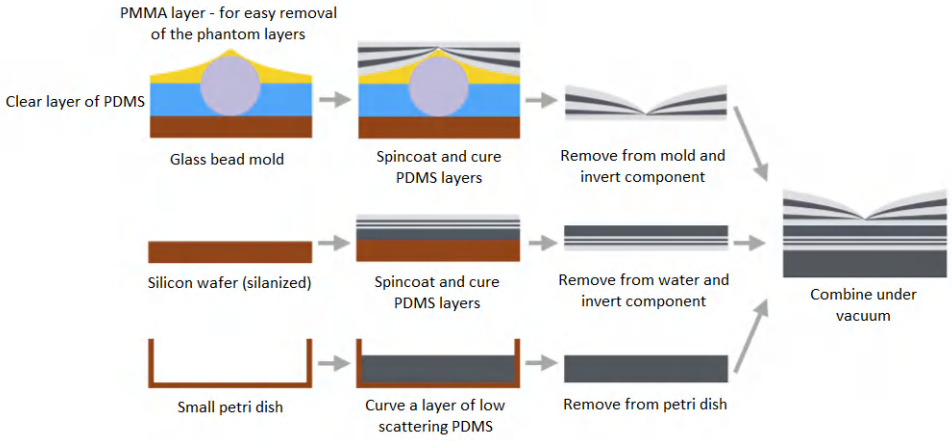


Figure 2.9: Schematics of the phantom created by Lee *et al.* [33].

The eye phantom designed by Fogli *et al.* [35] was made with the primary goal of validating new surgical procedures. Normally, when the goal is testing surgical techniques, the phantom used has more mechanical components. However, in this particular case, the phantom replicates the properties of the eye retinal layers using polymers and pays special attention to the humor vitreous, with the aim of reproducing the diffusion properties of the natural eye.

The phantom was obtained combining several retinal layers, the sclera, the choroid and the retina, as represented in the Figure 2.10 (a). These three layers were made using PDMS and polycaprolactone (PCL), to replicate the mechanical properties of the respective layers. In this work, the vitreous humor was made of gelatin in water mixed with a polyvinyl alcohol (PVA) solution at different percentages, which allowed for a more accurate replication of the diffusion properties of the vitreous humor. The result phantom is shown in Figure 2.10 (b), with the three retinal layers put together, from sclera to retina, and the vitreous humor on top.

(a) (b)

Figure 2.10: Eye phantom for ophthalmologic surgeries where (a) is the sketch of the different layers and (b) is the end-result of the proposed phantom [35].

Blood flow phantoms

Retinal blood vessels are commonly used to study some retinal diseases. To simulate this type of structure, retinal blood flow phantoms are often used. These phantoms normally do not focus on the retinal layers as much as the previous, since their main goal is the visualization of the retinal blood vessels and their irrigation.

The phantom designed by Lee *et al.* [32] uses microfluidic circulation channels by making patterns in silicon wafers using a photolithographic process. The structure of the microfluidic channels was designed by the authors, with the goal of representing the superficial and the deep retinal vessels. After this, the channels were imprinted into silicon wafers. This is followed by pouring the mixture of PDMS and TiO_2 into each mold. Finally, the phantom is detached from the mold after the curing process. The overall process is summarized in Figure 2.11.

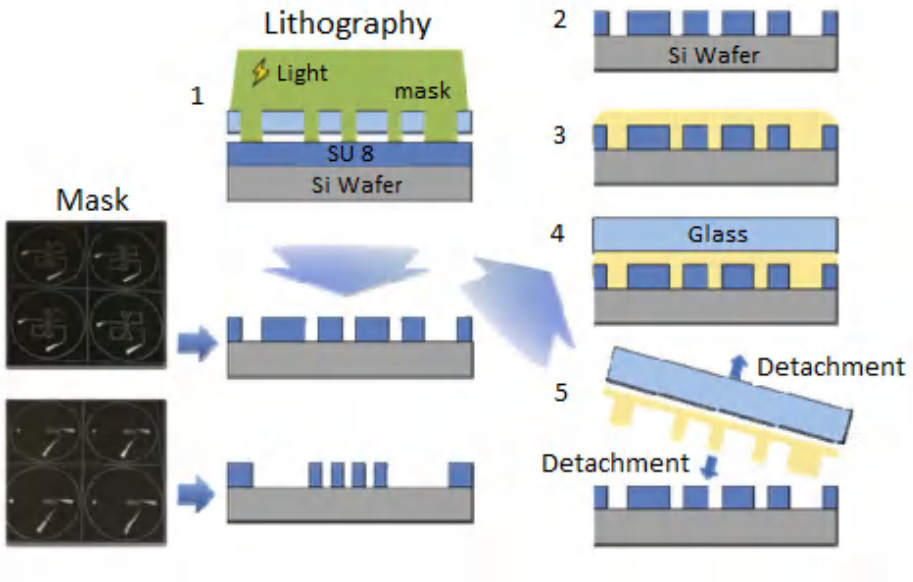


Figure 2.11: Steps to create microfluidic channels [32].

This phantom also replicates some retinal layers, such as the ganglion cell layer (GCL), the inner plexifrom layer (IPL), and the inner nuclear layer (INL). These layers were obtained using a mixture of TiO_2 and PDMS, as in the previous work [31], and were positioned on top of each other, creating a multilayer film.

To assemble the whole phantom together, the multilayer film was positioned between the microfluidic channel layers. There was also the need to incorporate tubes to circulate the blood mimicking fluid through the system. The final phantom design is presented in Figure 2.12.

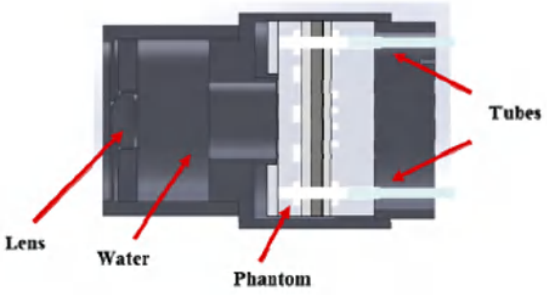


Figure 2.12: Blood phantom with retinal layers [32].

The design published by Braaf *et al.* [38] is a simpler model (Figure 2.13), with the purpose of calibrating and validating flow measurements. This model has a lens in the front part of the container as the focusing optics. Regarding the flow of the phantom, it uses a syringe to pump swine whole blood into the phantom, simulating the blood flow in the retina. Instead of

replicating the retina layers, a Teflon slab was used to act as homogeneous scatter. Having a Teflon slab, with a flat surface, allows for a controlled change in the incidence angle of the OCT beam.

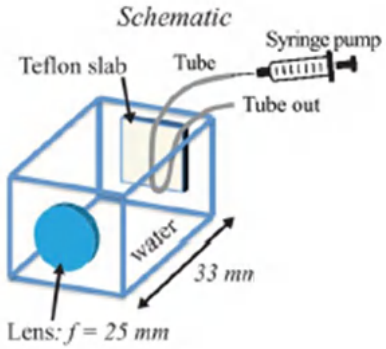


Figure 2.13: Blood flow phantom using a syringe pump [38].

Mechanical phantoms

When the main goal is the instrument calibration, for example, the calibration of an OCT device, a mechanical phantom is used. The mechanical phantom made by Wang *et al.* [39] was produced in order to simulate the structure of the eye and its refractive power, offering a tool for OCT quality control and calibration. In this phantom, the optical structure of the eye is replicated using two convex lens, fiber filaments to test the alignment of the OCT scan, and a microsphere-embedded phantom. This structures can be seen with more detail in Figure 2.14.

The light goes through the first lens and is focused on the back surface of the second one. The width of the beam of light focused on the second lens is used to test the angular field of view. To test the lateral and axial resolution of the OCT device, this phantom has a microsphere-embedded phantom with microspheres of 1μm diameter.

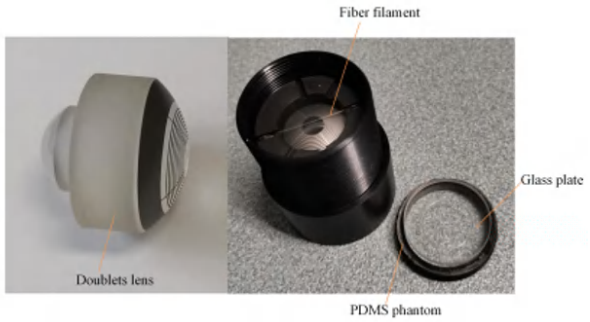


Figure 2.14: Phantom for OCT quality control and calibration [39].

Mechanic phantoms can also be used for eye surgery simulation. The eye model developed by Tanaka *et al.* [34] has the main goal of quantifying the performance of manual and robotic microcannulation, a surgical procedure which involves drug injection in a retinal vein. This specific phantom contains a retinal vein model, a force sensor, a sclera model, and a mechanism that simulates the rotational motion of the eye. The schematic representation of this model, with all the mentioned components, is shown in Figure 2.15.

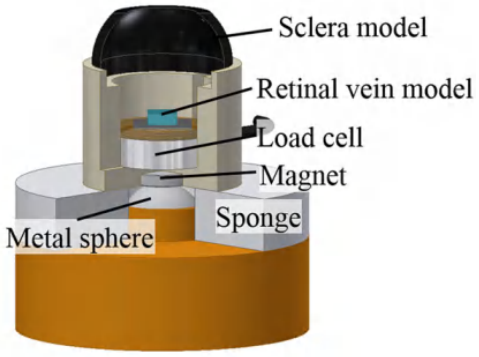


Figure 2.15: Eye model for eye surgery simulation [34].

The sclera model is made of natural rubber, simulating the mechanical properties and size of the human sclera. This part of the phantom has a hole on the top to insert the tools needed to perfrom microcannulation. The retinal vein model is made of microchannels to simulate the diameter of the target retinal vein. These channels were produced using PDMS and a photolithography procedure, as previous phantoms.

2.4.2 OCT motion correction algorithms

One of the biggest challenges in OCT imaging are the artefacts caused by involuntary eye movements, which can make visualization and quantification of structures unreliable. The correction of these artefacts is still a hot topic in biomedical engineering [19].

In general, motion correction algorithms can be divided into methods that process all the image using intensity metrics (intensity-based registration), and methods that use specific structures in the image to perform the registration (feature-based registration) [19]. There are also algorithms that combine both approaches. Although intensity-based approaches are more general, the point-based approach has lower computational complexity and higher robustness and accuracy rates [40].

Thanks to the improvements in OCT technology in recent years, in this work, as in most recent works [41–43], the authors assume that the 2D slices (B-scans) from an OCT volume

along the fast direction do not have motion artefacts.

Intensity-based registration

One common application of intensity-based registration is to sequentially align the B-scans of a volume to the previous, motion-free, ones. The work by Cheng *et. al* [41] applies this method so that, once an OCT 3D volume is motion-free, several volumes can be aligned using their *en face* projections. The first step of this work is to detect the B-scans which have motion artefacts. Assuming pixels from static regions of the 3D volume present small decorrelation values whereas pixels from the faster moving regions of the 3D volume yield high decorrelation values, it is possible to detect the B-scans with movement calculating the decorrelation value for overlapped B-scans. Each B-scan which was affected by eye movement is aligned (translated) with the previous B-scan, assumed to be artefact-free. The alignment is calculated minimizing the difference between the two B-scans analysed. Since a full search in the whole B-scan is time consuming, this work uses a diamond search strategy [44], which only computes the decorrelation values for a subset of points.

Although a significant amount of research has focused on developing the registration algorithms for medical imaging, there are fewer examples of open-source and ready-to-apply methods. With the goal of making an open source and user-friendly application for registration algorithms, the SimpleElastix software was developed by Martsal *et al.* [45] as an extension of SimpleITK that allows the user to configure and run Elastix [46] (a toolbox for image registration) easily. Elastix and SimpleElastix are highly configurable, enabling different registration strategies. Among many other medical imaging applications, both Elastix and SimpleElastix have been successfully adapted and applied to OCT images [4,47]. These algorithms use an iterative pair-wise approach, and they treat each 2D registration as a minimization problem. The minimization of the cost function is made by an iterative optimization method, using a cross-correlation metric.

Most of the OCT registration algorithms are based on the registration of consecutive images, B-scans, which can lead to errors when inspecting the whole OCT volume. This drawback can be tackled using group-wise approaches, such as in the work proposed by Ntatsis *et al.* [48] The proposed unsupervised learning registration framework corrects motion in the retinal volume by decomposing the task into multiple group-wise registrations of consecutive B-scans. For each analysed B-scan, translation in the x and y directions are predicted. The model architecture is a 3D adaptation of the DenseNet [49] and, similar to [4, 47], the optimized metric is pair-wise cross-correlation. The network is fully-convolutional, so it can perform inference in the entire

OCT volume for any number of B-scans. The evaluation metrics showed improvements for all OCT volumes, although in some cases a number of mis-registrations were observed. Figure 2.16 shows the central slow-axis B-scan of the same volume at different steps of the registration algorithm. It can be observed how the artifacts are minimized, until the last B-scan which is almost artefact-free.

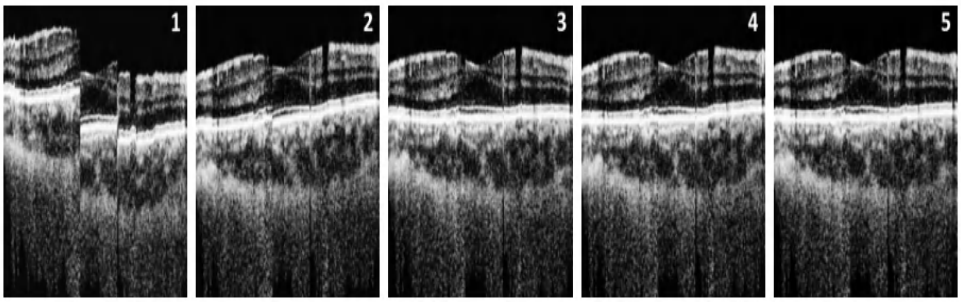


Figure 2.16: Central slow axis B-scan correction throughout the recurrent iterations, used in the algorithm. The numbers on the corner of each scan corresponds to the iteration number [48].

Feature-based registration

The work done by Ricco *et al.* [9] presents a two-step registration algorithm starting with the correction of small deformations, as tremors and drifts, followed by abrupt deformations, such as microsaccades. The proposed method corrects for transverse motion artefacts by registering the en face OCT to an scanning laser ophtalmoscopy (SLO) image of the retina. Since the main observable features in 2D retinal images are the vessels, vessel detection is the first step of this method. This is performed in both SLO and OCT *en face* images, so that the vasculature can serve as reference for the registration, using the technique of Periaswamy and Farid [50]. This is a technique developed for the registration of medical images (not specific for OCT) that show different pixel intensity throughout the image. Once the OCT *en face* image is aligned with the SLO reference image, the microsaccades within the OCT will be corrected by finding the horizontal shift at each pixel of the OCT en face scan that shows the best correspondence when aligning with the result from the drifts and tremors corrections and the SLO image. This work only corrects the horizontal microsaccades, and does not address axial movements. The results from this method are shown in Figure 2.17.

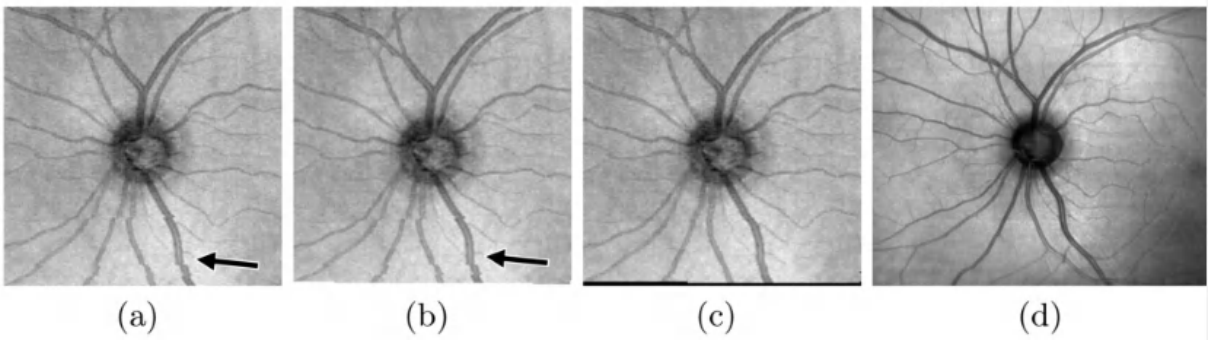


Figure 2.17: Results of motion artefact correction. (a) *En face* image of an uncorrected scan (b) Result image for correction for tremors and drifts, which still shows discontinuities from microsaccades (c) Result image after microsaccade correction, which matches the SLO reference image, shown in (d) [9].

Registration algorithms combining both approaches

Combining both types of registration approach is uncommon in the literature. One of few examples is the algorithm proposed by Pan *et al.* [40], which uses both intensity-based an feature-based techniques to perform a 3D volume registration. The proposed method can be divided into three steps: preprocessing, feature design and, correspondence detection.

In the preprocessing, seven retinal layers were segmented using a graph search-based method. After this, the B-scans are flattened to ignore the curvature of the retina in the next steps of the algorithm. In the feature design step, three different types of structures are detected: surface-based, intensity-based and vessel-like. The surface-based structure is defined by the difference between the intensity of voxels, those voxels that are on the surface have a higher intensity than voxels not on the surface. The intensity-based region is based on the mean intensity value of each respective layer, so that different mean intensities will describe different retinal layers. The correspondence detection is made through detection of vessel structures. The vessel-like features are detected by the different voxels in a horizontal direction of a B-scans. Blood vessels have a hyper reflective area in the inner retina, and the respective shadows produce hypo reflective areas in the outer retina, therefore the vessel information can be extracted using a projection image of an OCT volume.

This method was tested using data from healthy and severely diseased subjects. In Figure 2.18, the results from the normal eyes data set are shown, there is a reference image, without artefacts, the original image, without correction, and the image after the motion correction. In Figure 2.19, the results from subjects with lesions are presented, there is also a template image,

a reference image without artefacts, an image without correction, and a final image after the algorithm is applied. The results suggest that this algorithm overcomes some problems of other registration methods and consistently produce high registration accuracy.

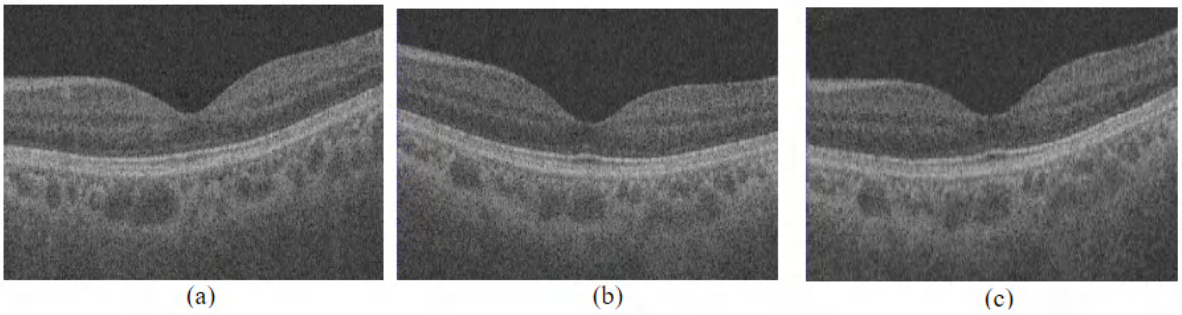


Figure 2.18: Results of feature-based 3D registration for an healthy eye where (a) is the reference image, (b) is the image without correction, and (c) is the corrected image [40].

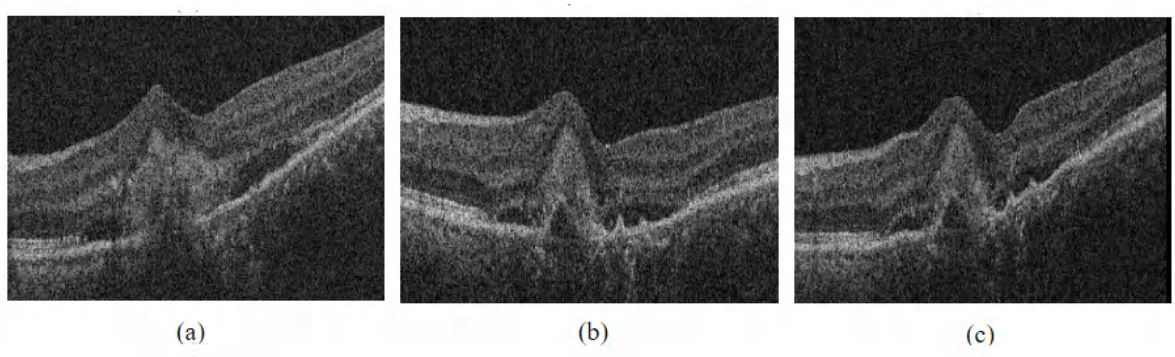


Figure 2.19: Results of feature-based 3D registration for the data set with subjects that present several lesions, where (a) is the reference image, (b) is the image without correction, and (c) is the corrected image [40].

Chapter 3

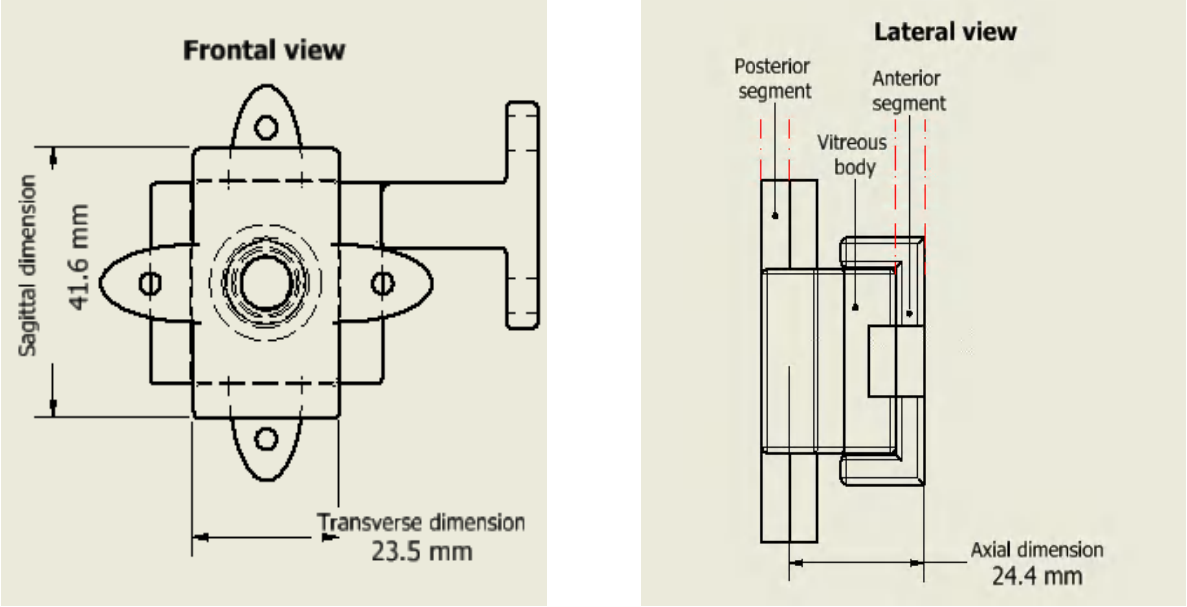
Methods

This chapter describes the methods used in the project. First, the design of a human eye model is described. Second, the data set acquisition is depicted, and lastly, the motion correction algorithms are presented.

3.1 Human eye phantom

A phantom eye model was developed to acquire OCT scans with simulated eye motions. Building it allowed the characteristics reproduced in it to be chosen accordingly to the main goals of this project. This phantom was designed to be a modular piece, which allows small alterations such as the diameter of the eye, the number of retinal layers or thickness. The model is also easily reproducible, since all the applied techniques are relatively simple and the material cost is low. Furthermore, an advantage when comparing to other models is that this model is non-degradable and does neither accumulate dust nor deformate over time.

In order to get a realistic model, the axial dimension was 24.4 mm, the averaged axial dimension of a human eye [10]. All other dimensions (23.5 mm (transverse) $\times$ 41.6 mm (sagittal)) of the designed phantom do not correspond to the real eye dimensions due to design adaptations. In Figure 3.1 there is frontal and lateral projection view of the eye phantom with the respective dimensions.



(a) (b)

Figure 3.1: Frontal (a) and lateral (b) views of the eye phantom with the respective dimensions.

The mechanical parts were designed with the Autodesk Inventor software [51], and 3D printed using the Ultimaker 2+. The software Ultimaker Cura [52] was used to slice the phantom parts made in Autodesk Inventor, converting the STL files (from Autodesk Inventor) into gcode (to use in the printer).

The material selected to print the phantom parts was the Polylactic Acid (PLA). With the development of 3D printing, PLA has become one of the most used materials, as it has a low melting point, presents good layer adhesion, and has high resistance [53]. For this specific application, a waterproof and watertight material was needed, which was able to hold the water without absorbing it into the walls, characteristics that the PLA material presented. The PLA filament used was from Ultimaker B.V., Geldermalsen, The Netherlands, with a density of 1.24 g/cm^3, white pearl color, and a printing temperature between 195 °C and 240 °C [54].

The next sub-sections present a description of the 3 parts in which the phantom can be divided: the anterior segment, the vitreous body, and the posterior segment.

3.1.1 Anterior segment

The anterior part of the phantom was developed using a plano-convex lens, with a focal distance of 20 mm, that simulates both the cornea and crystalline of the human eye [55, 56]. The selected lens (LA1074-B - N-BK7) was produced by Thorlabs [57], with a back focal distance

of 17 mm, a diameter of 12.7 mm, and a functional wavelength between 650 and 1050 nm.

Furthermore, this anterior segment of the model needed a support to make the assemble with the other parts of the phantom. The support was designed to have bolt insertions, needed to connect the different parts and to accommodate the lens. To prevent the lens from getting damaged, an O-ring was added inside the support piece. This way the lens does not touch the PLA support, preventing its coating from getting scratched. Figure 3.2 shows a 3D sketch of this part of the model.

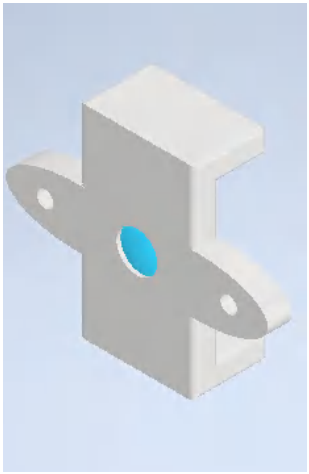


Figure 3.2: Sketch of the lens support. The lens is represented in blue in the middle of the piece.

3.1.2 The vitreous body

The central part of the model corresponds to the vitreous body, which constitutes the largest part of the phantom. Similar to the anterior part, the shape of the central part was made to accommodate bolts to assemble the anterior and posterior segments of the model.

Additionally, water was used to simulate the vitreous humor of the human eye since both materials have approximately the same refraction index [32]. As a consequence of using water, it was necessary to seal this part of the model using two O-rings. These were placed in the front and back of the vitreous body part, to stank the water inside the model without compromising other components. Figure 3.3 shows a 3D sketch of the vitreous body part. In the front part of this design it is also possible to see the spot where the O-ring is placed.

3.1.3 Posterior segment

The third, and posterior part of the model corresponds to the retina, choroid, and the sclera. In the human eye, its layers cover the whole sphere. On the contrary, in the phantom the retinal

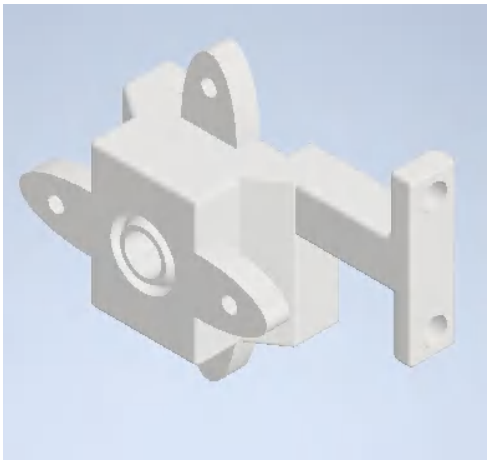


Figure 3.3: Sketch of the central part of the model with the designed space to insert the O-rings.

layers are only in the fundus of the model. This does not represent a problem in comparison with scans from real patients, since the OCT scan only acquires information in a restricted field of view.

The human eye has a great number of different retinal layers, however, in this phantom, the ones reproduced were the retina, the choroid, and the sclera. The retina was reproduced as a single layer, not taking into account the different scatters in the sub layers of the retina. This may produce differences in the results obtained when compared to human eye scans, and, consequently, represent a limitation of the eye phantom.

Following previous works in the literature, the layers were fabricated using a mixture of PDMS and TiO_2 (scattering agent) [58]. These are relatively cheap materials and easy to maneuver. The different layers were obtained using different percentages of TiO_2 in the mixture, as depicted in Table 3.1. To create the thin layers needed for this application, a spin coating protocol was used, which allowed to have the same thickness through the whole layer. The target thickness for each layer and the rotational speed used in the spin coater are also shown in Table 3.1.

Table 3.1: Characteristics of the three replicated layers - the percentage of TiO_2 in PDMS, the layers thickness and the rotational speed used in the spin coater.

Replicated layer	**TiO_2 (% in PDMS) [33]**	**Thickness (μm) [31]**	**Rotational speed (RPM) [33]**
Retina	0.6	200	500
Choroid	0.025	400	240
Sclera	2.5	700	140

For each of the three retinal layers, the following protocol was applied:

1. Weight the mass of TiO_2 according to the concentrations of the layer.
2. Weight the respective mass of elastomer base (PDMS).
3. Mix the elastomer with the TiO_2 homogeneously.
4. Place the mixture into an ultrasonic bath for 1 hour.
5. Add the curing agent with a ratio of 10:1, mix until there are no bubbles present.
6. Pour the mixture into the spin coater at the respective speed for 3 min.
7. Leave the layer to dry off for 24 hours.

Each layer was added on top of the previous one, avoiding the problem of detaching thin layers and possibly breaking them. Figure 3.4 shows the 3D sketch of the layer support, also used in the spin coater to produce the layers directly in the model part.

Figure 3.4: Sketch of the support where the layers were made - this support was also placed in the spin coater.

Figure 3.5 presents the final 3D sketch of the whole eye model, with the 3 modular parts: the anterior part (white), the intermediate part (blue), and the posterior part with the layers support (dark gray).

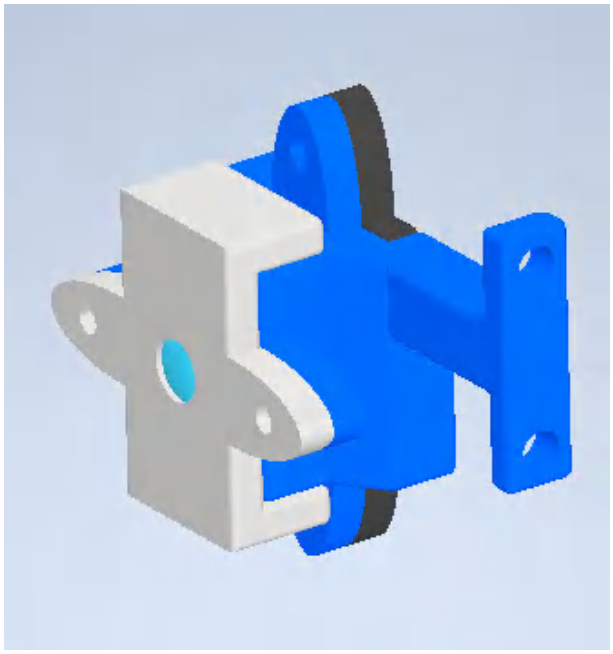


Figure 3.5: 3D sketch of the whole eye model - the anterior part is represented in white, the middle part in blue, and the posterior part in dark gray.

3.2 Data acquisition

In this step of the project, a data set with OCT scans obtained using the developed phantom was built (synthetic OCT data set).

The data set includes static scans of the model, without any movement, as well as moving scans. The moving scans were obtained with different patterns of movement, replicating different involuntary eye movements.

3.2.1 OCT system

The scans acquired in this work were obtained using an Optical Coherence Elastography (OCE) system, based on a swept-source OCT, from Coimbra Institute for Biomedical Imaging and Translational Research (CIBIT) [59].

In Figure 3.6 there is a schematic of the OCE system used. It consists of a SS-OCT combined with a piezoelectric actuator, for the dynamic mechanical excitation of the sample, or a 10 MHz ultrasound A-scan probe (Imasonic, Voray sur L'Ognon, France). This SS-OCT is based on a swept-source laser (Axsun, Excelitas Technologies Corp., Mississauga, Canada), that emits at a central wavelength of 1060 nm, with a bandwidth of 110 nm and a repetition rate of 100 kHz. The laser source also includes a fiber-based Mach-Zehnder interferometer to provide evenly spaced wavenumbers in the space output. This allows to sample the OCT interferogram linearly and

use a direct Fourier processing technique.

The light used by the OCT is split into the sample and reference arms. In the sample arm, the light is delivered and collected from the sample using a optical fiber coupler and a long working distance microscope objective. In the reference arm, the light is reflected from a stationary mirror. The optical setup of the sample arm can be changed when the imaging samples have optical elements.

The OCT scans are acquired in a standard acquisition model: axial A-scans, in the sagittal plane, cross-sectional B-scans, in the transversal plane, and volumetric acquisitions. In the acquisition protocol, axial scans are repeated 512 times at a given location, this corresponds to a scanning time of 5.12 ms for each B-scan. The eye phantom was positioned in position S of Figure 3.6, in the sample arm box.

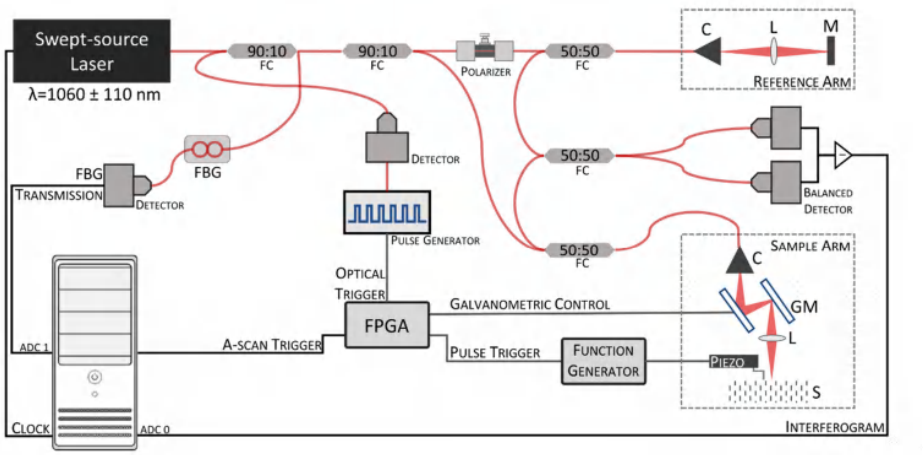


Figure 3.6: Schematic representation of the SS-OCE system [59].

Figure 3.7 shows the experimental set-up during the acquisition of an OCT volume. Due to limitations in the optical setup, specifically the use of a microscope objective, the scans were obtained using only the layers of the phantom and not the whole model. The use of the anterior and middle parts of the phantom presented a problem in the visualization of the layers, due to the refraction of the lens and the large distance from the layers to the OCT objective. Nevertheless, as can be seen in Section 4.2, the obtained scans have a clear correspondence with the expected data.

3.2.2 Artefacts replication

The artefacts/movement simulation was made using two different motors. One motor did rotational movements, simulating drifts and microsaccades, whilst another motor did vertical

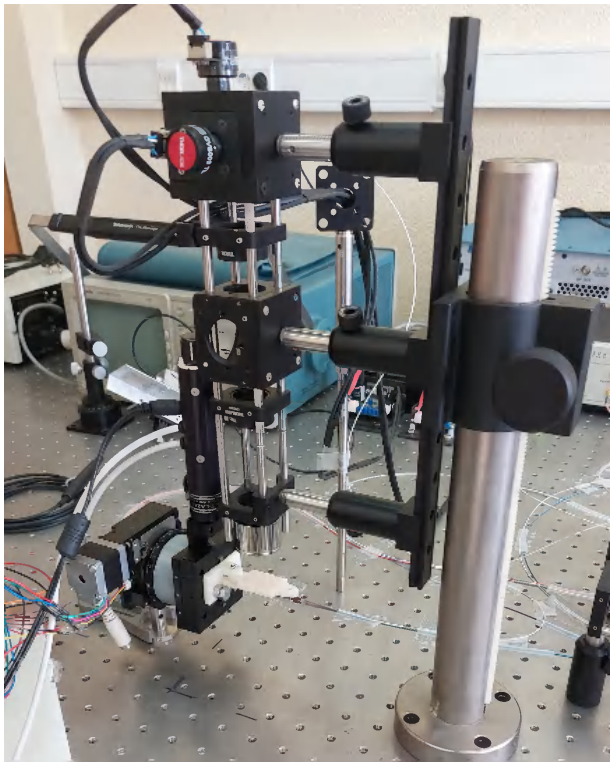


Figure 3.7: Acquisition of one volumetric scan with the phantom layers (posterior segment).

movements, simulating the respiratory movements.

The drifts and microsaccades were simulated using a stepper motor (Zaber RSW-60 rotary platform) with Arduino Motor Shield Rev 3 [60]. This motor is controlled using a specific library from Arduino, the Stepper library [61]. In Figure 3.8 there is a schematization of the processes done in the Arduino to control the stepper motor. The circle represents the stepper motor, starting at the initial position, the respective arrows represent the two different movements replicated: first a microsaccade, with a bigger amplitude, then a drift, in the opposite direction with a smaller amplitude, followed by a microsaccade, in the same direction as the previous drift, and a final drift in the opposite direction. When replicating these eye involuntary movements, there was the need to ensure the eye model remained in the field of view of the acquisition. For this reason, the movement pattern, described before, was used to make sure the model did not leave the acquisition field.

The second motor was used to simulate the respiratory movements. In this case, the motor used was a Zaber miniature linear actuator [62]. This was programmed using the Zaber motion library [63]. This library sent a command to make vertical movements with an amplitude of 50 μm. As happened in the Zaber rotatory platform, the motor was programmed in order to keep the phantom centered during the whole acquisition. In order to this, the motor did half of the

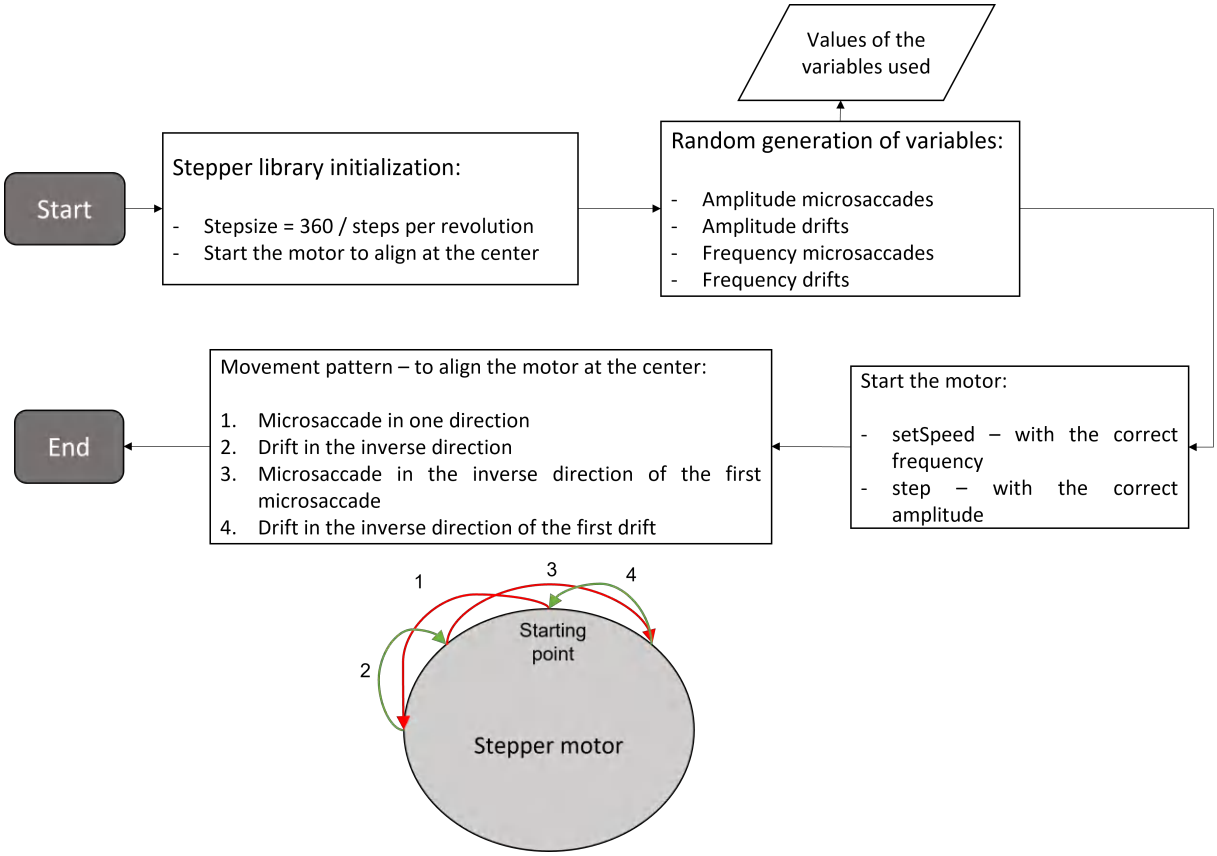


Figure 3.8: Flowchart of the Arduino code to control the stepper motor. Red arrows indicate microsaccades and green arrows indicate drifts.

movement amplitude downwards, followed by the full movement amplitude upwards, and half of the movement amplitude downwards to make two respiratory movements. This movement pattern was replicated throughout the whole acquisition time.

Having in mind the real pattern of involuntary eye movements, as shown in Figure 3.9, patterns of a microsaccade, followed by a drift were simulated. These patterns are represented in Table 3.2. In order to obtain a random data set, a large range of possible artefacts were simulated. The simulated movements were randomly chosen within an interval, shown in Table 3.2, representing the physiological movements. The amplitudes and frequencies used to program the motors were calculated taking into account the displacement of the eye model to the center of the motor. Moreover, the extra side of the model, with the appearance of an arm, was added to the middle part of the eye model to connect to the motor. Due to maximum limitations in the rotational frequency of the motor, the maximum frequency was of 2.4 revolutions per minute (RPM). For this reason, the speed for the microsaccades is constant, at maximum speed of 12^{o}/s. The linear actuator was used to simulate the respiratory movements with an amplitude of 50μm.

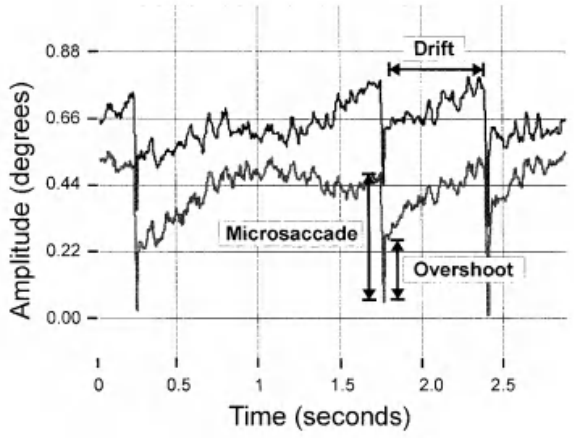


Figure 3.9: Fixation pattern with fast microsaccades interrupted by drifts [64].

Table 3.2: Movement pattern used in the Arduino software to replicate the involuntary eye movements.

Movement pattern	**1º Microsaccade**
	Duration of 25 ms – constant
	Speed of 12º/s – constant
	Movement amplitude between 0,25º and 1º - randomly selected
	2 º Drift
	Duration between 0,5 s and 1,5 s – randomly selected
	Speed between 0,1º/s and 0,5º/s – randomly selected
	Movement amplitude dependent on speed and movement duration (0,05º and 0,75º)

3.2.3 Data set description

Synthetic data set

Using the movement pattern described in Table 3.2, 5 types of acquisitions were designed to compose the data set. First, there were static scans, acquired with the phantom in a static position. These scans are artefact-free and were used as reference volumes to evaluate the algorithms' performance. Second, there were scans acquired with only one type of movement, either microsaccades or drifts, throughout the whole acquisition time, with repetitive movements. Third, there were scans acquired with the movement pattern from Table 3.2, that means these scans will have, at least, one drift and one microsaccade during one acquisition. Fourth, scans were acquired in which the movement started after initializing the acquisition. This means that the first part of the volume was static and the second part had either a microsaccade and a drift or just one of them.

At last, there were scans acquired with non-physiological movements. That means that

the movement pattern was the same as the pattern from Table 3.2, however the movement amplitudes and duration used were larger (amplitudes up to 12°and duration up to 3 s), to simulate extreme movements.

Table 3.3 has the number of volumes acquired for each type of movement simulation. In total, the data set has 156 volumes with artefacts from the simulation of eye involuntary movements. These volumes were the volumes this work aimed to correct.

Table 3.3: Data set description with number of volumes acquired for each type of movement simulated.

Data set type	**Num. of volumes**
Static volumes	9
Volumes with repetitive movements	20
Volumes with the movement pattern	96
Volumes where the movement pattern starts after the acquisition	10
Volumes with extreme movements	30

Clinical data set

The clinical OCT data set, used in this work for post-training evaluation of the deep-learning algorithm, was the data set used in the work from Ntatsis *et al.* [48]. This data set was acquired at Centro Hospitalar São João using Spectralis S2000 Spectral Domain OCT from Heidelberg Engineering. It consists of 37 eyes of 19 subjects, with 4 volumes per eye, making a total of 146 volumes (two volumes were excluded due to different acquisition resolution).

3.3 OCT motion correction

The next step in this work was to correct the previously acquired scans. In order to do this, two algorithms were used: a general purpose conventional registration algorithm (SimpleElastix [45]), and a group-based deep-learning algorithm, previously published by Ntatsis *et al.* [48]. Both chosen approaches are intensity-based methods, removing the need for engineered features.

3.3.1 Preprocessing

The raw information from the OCT volumes was preprocessed, at an initial phase. The first step was a data centring of the data. Following this, the DC components and low frequencies were removed using an interferogram. As part of this preprocessing stage, there was also the need to correct pixel displacements, due to the OCT scanning process of the system used. When acquiring the OCT volumes, there was an acquisition of a background volume. This volume was subtracted to the phantom OCT volumes, in order to remove the background noise. The final step of the preprocess was the data normalization. The values were normalized between 0 and 255. The processes preformed in this stage are schematized in Figure 3.10.

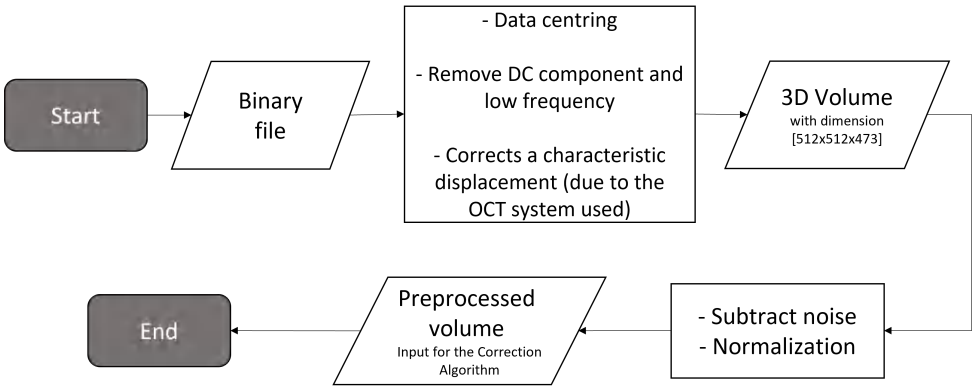


Figure 3.10: Flowchart of the preprocessing steps.

3.3.2 Conventional registration

As described in the state of the art (section 2.4), the SimpleElastix software, developed by Martsal *et al.* [45], was used to correct the movement artefacts in the data set. This framework was made for intensity-based registration of medical images, and it is an extension of SimpleITK [65], an open source image analysis toolkit.

This registration algorithm involves two input images: a fixed image, artefact-free, and a moving image, with motion artefacts. Mathematically, the moving image is deformed to match the fixed imaged, by finding the transformation coordinates necessary to do that. The goal is to correct the motion artefacts present in the moving image.

In an OCT volume, each B-scan was registered in a sequential way. That is, a total of $N-1$ pair-wise 2D registrations are performed, with N the number of B-scans in the volume. We can denote B_n each original B-scan of the volume (n in $[1..N]$), and R_m each registered scan of the registered volume (m in $[2..N]$, since the first B-scan is not registered). The fixed and moving

images are updated at each iteration, starting with the first pair of B-scans of the volume (B_1 and B_2, respectively). Once this first pair is registered, the fixed image for the next iteration will be the registered moving image (R_2). As a consequence, the scans registration is based on the previously registered scan. Finally, the next moving image will be the next B-scan of the original (unregistered) volume (B_3). Figure 3.11 represents the processes involved in the registration using SimpleElastix.

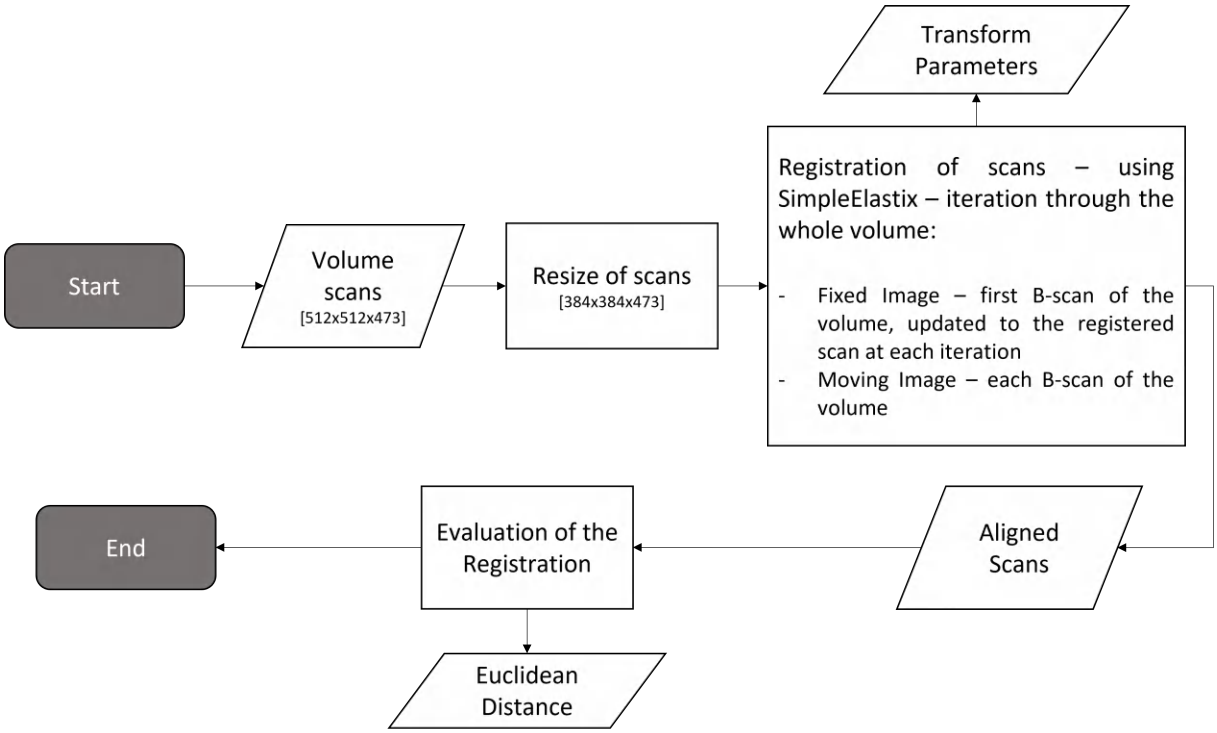


Figure 3.11: Flowchart of the registration process using SimpleElastix [45].

To configure SimpleElastix, there is the need to define the parameter map. This is a collection of key-value pairs that configure the components of the registration, and any settings they might require, automatically. Only input images and output options need to be specified separated. In order to configure the registration process to the specific OCT scans obtained, it was necessary to test different parameters used to perform the registration, in order to obtain the best possible result.

In the SimpleElastix documentation [45], the parameters, which have the most influence in the performance of the registration, are outlined. Using a multi-resolution registration, that uses a 2D translation, improves the chances of finding the optimal solution. To find the optimal translation, the sampler, which is responsible for finding the location to evaluate in the input images, needs to be defined. The image sampler, was random coordinate, which means that

each pixel of the image has the same probability of being chose, and a pixel may be chosen more than once. The number of resolutions is one of the parameters that can be altered. In general 4 is sufficient however if the fixed imaged is quite different from the moving image, as happens in some scans, this number can be increased to 5. The number of spatial samples defines the number of pixels analysed in each image. The maximum number of iterations is the number of iterations done analysing the image pixels before delivering the registered result. For moving images with major differences to the fixed image, this number needs to be increased. Finally, one of the most important parameters to define is the type of correction that will be applied to the images. Translation transform enables only corrections in the x and y axis, while Euler transform corrects also rotations. Since the data set included both type of movements it is expected that the registration using Euler transform will produce better results. However, since the rotation movements are small, it is interesting to see if Translation transform is able to correct them. Table 3.4 shows the parameters used, and its respective values.

Table 3.4: Different parameters used for the parameter map in SimpleElastix registration.

Parameters	**Transform**	**Nº of resolutions**	**Maximum nº of iterations**	**Nº of spatial samples**
v1	Euler	4	128	1024
v2	Euler	5	256	2048
v3	Translation	4	128	1024
v4	Translation	5	256	2048

3.3.3 Deep-learning registration

The second registration approach explored in this work was an unsupervised deep-learning algorithm, previously described by Ntatsis *et al.* [48]. When compared to conventional approaches, such as SimpleElastix, a deep-learning algorithm performs the registration in a shorter period of time. The deep-learning model learns its parameters from the data, which did not happen using SimpleElastix. Additionally, the deep-learning model is group-wise, it uses groups of different sizes, although not the complete volume at the same time, whereas the conventional algorithm is pair-wise, uses only pairs of images.

The OCT volumes, with dimensions of 512 $\times$ 512 $\times$ 473, were used as input to train the algorithm. Based on the work from Ntatsis *et al.* [48], the volumes were splitted into individual

fast axis B-scan images. Each B-scan was then resized to a dimension of 384 × 384 pixels. Then, each volume, with 473 fast axis B-scans, was divided into sub-volumes with 144 fast axis B-scans each. These changes were made to match the algorithm required input size.

The motion artefacts, present in the OCT volumes, are corrected using a group-wise registration procedure, within a group of adjacent fast axis B-scans. The algorithm output is the equivalent motion-corrected group of adjacent fast axis B-scans.

The algorithm framework, from Ntatsis *et al.* [48], consists of four main components:

- a trainable model, which is a 3D adaptation of the DenseNet architecture [49] - this model produces a pair of translation parameters for each input fast-axis B-scan;
- a cumulative summation operation, along the two translation directions, applied to the output translation parameters from the trainable model - this makes the network to learn relative translations;
- a zero-mean operation, along each translation direction - this addresses the illposedness of the group-wise registration problem;
- a spatial transformation module, which receives the initial unregistered group, with the translation parameters, and outputs the registered group via bicubic interpolation.

In Figure 3.12, there is a flowchart representing all the steps to train this network: the first steps correspond to the data set adaptation, that was necessary in order to use the volumes as input for the network, and the last steps correspond to the deep-learning algorithm framework [48].

The algorithm was trained in an unsupervised way, through the maximization of the mean pair-wise cross-correlation of the registered group. Although this model, due to its fully-convolutional design, can be used for any number of B-scans, it was trained with the same 144 B-scans, as done by Ntatsis *et al.* [48], in order to perform comparisons between the obtained results.

The optimization problem of deep-learning algorithms rose with the increase number of new algorithms in the literature [66]. The most common optimization technique is parameter optimization, different parameters are used to obtain an optimized algorithm. However, in the last years, a data-centric approaches started to be applied [67]. In this, different types of data sets are used as input of an algorithm. Then, by analysing the results, it is possible to see which data set version is better suited for each algorithm. Data-centric practices outline

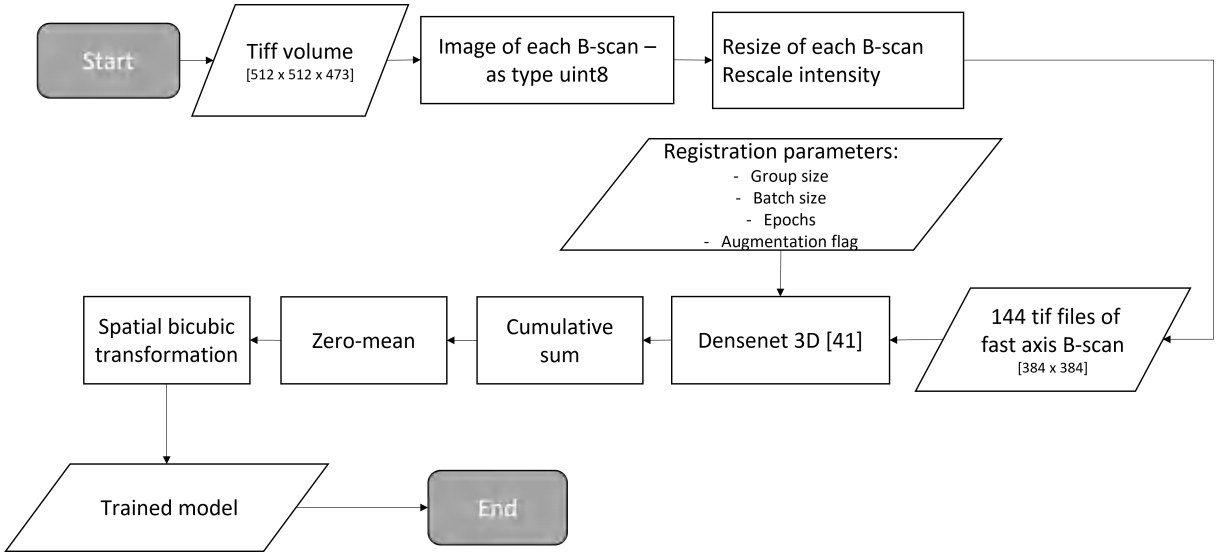


Figure 3.12: Flowchart of the registration process using the deep-learning registration algorithm.

the importance of having a good data set because without it even the best algorithms cannot perform well. With this practice the goal is to obtain a better data set, by systematically change or enhance the data set used to improve the model performance.

This work presents, in a first stage, a data-centric approach where five different data sets were used to see which data set produced the best training results. The characterization of each data set is shown in Table 3.5. For the training of the algorithm, each data set was divided into three groups: a train set, with 80% of the data; a validation set, with 10% of the data; and, finally, a test set, with 10% of the data.

In a second stage, having selected the best data set version, there was a parameter optimization, a model-centric approach. The parameters used in each version are shown in Table 3.6. These parameters were selected based on the values used in the work from Ntatsis *et al.* [48].

After using the best data set from Table 3.5 with the best parameters from Table 3.6, the best trained model was obtained. The performance of this model was tested using a clinical data set of OCT volumes with artifacts. These results were compared with the ones obtained by Ntatsis *et al.* [48] where the same framework was trained and evaluated only using the physiological data set.

3.3.4 Application on retinal OCT data

At last, the best model developed in this work, was retrained with the clinical OCT volumes, starting with a pre-train model, obtained with the phantom OCT volumes. This technique is referred to as transfer learning, where a model which was trained on one task, in this case the

Table 3.5: Characterization of the different data sets used for the deep-learning algorithm training.

Data set version	**Num. volumes**	**Data order**	**Load strategy**
1	100*	Random order of volumes with 144 B-scans from the final part of the volume	Load all the data set to the gpu in the start
2	100*	Random order of volumes divided into sub-volumes with 144 B-scans each – using the whole volume	Load all the data set to the gpu in the start
3	136*	Random order volumes with 144 B-scans from the whole volume with 10 volumes with extreme movements	Load all the data set to the gpu in the start
4	381	Random order volumes with 144 B-scans from the whole volume	Load to gpu in each epoch
5	466	Random order volumes with 144 B-scans from the whole volume with extreme movements	Load to gpu in each epoch

* These versions had a limited amount of volumes due to GPU memory limitations, from the computer used. Versions 4 and 5 corrected this issue by loading the data to gpu in each epoch.

Table 3.6: Characterization of the different parameters used in different training versions of the deep-learning algorithm, with the data set selected previously.

Parametrization version	Group size	Batch size	Augmentation flag	Epochs
1	4	8	False	15
2	4	8	True	15 + 15**
3	8	8	False	15
4	8	8	True	15 + 15**

** In these versions, the network was pre-trained. This means that the first 15 epochs correspond to the epochs from the train from version 1 and version 3 respectively.

synthetic OCT data set, is then used to re-train on a second task, the clinical OCT data set. Transfer learning allows a rapid model optimization and improved performance when training the second task. This technique is normally used when there is a limited amount of training data [68].

Using this technique, the goal was to assess if by using a pre-trained model, the results can be improved and, consequently, the addition of synthetic data to real OCT data set is an advantage when working with deep-learning algorithms, specially to correct motion artefacts.

3.3.5 Algorithms evaluation

The acquired data set contained static volumes, which were artefact-free. Having these type of volumes allowed to use comparative metrics, such as the Euclidean distance and the Dice score. In addition to that, using these metrics, it was possible to evaluate the performed registration in comparison to an artefact-free B-scan. In order to allow for a comparison between the two different registration algorithms, the evaluation metrics used were the same in the two cases.

Two different quantitative metrics were implemented to evaluate the output results: Euclidean distance and Sørensen–Dice coefficient. Besides this, the registered volumes were qualitatively analysed. For the qualitative metrics, a fast axis B-scan (artefact-free) was compared to the registered B-scan, for the registered volumes from the conventional registration algorithm. This results were discussed in Section 4.3.1. For the deep-learning algorithm registered volumes, a B-scan from a static volume was overlapped on a B-scan after the registration process and the visual results were discussed in Section 4.3.2.

The Euclidean distance is the most commonly used metric to find the difference between both images [69]. It calculates the square root of the sum of the absolute differences between

two feature points as follows:

$$\text{Euclidean distance} = \sqrt{(\sum (u_i - v_i)^2)} \tag{3.1}$$

where u_i and v_i represent the feature points, from two different images, and i between 0 and 144, the number of columns of the analysed B-scans.

To perform this operation it was used the public application programming interface SciPy [70]. All the B-scans from each volume were analysed. In each B-scan, the selected feature points were the pixels from each column with the higher intensity. For this specific registration problem, the lower the Euclidean distance, the better the performed registration. In a perfect registration the value should be 0.

The Sørensen–Dice coefficient can be used to evaluate the similarity between two images. This coefficient is calculated using the ratio of the overlapped regions between two binary images as follows:

$$\text{Dice score} = \frac{2\,|A| \cap |B|}{|A| + |B|} \tag{3.2}$$

where A and B are the two analysed images. Before calculating the Dice score, the intensities in the analysed volumes were normalized between 0 to 1, and the intensities above 0.2 were labelled as 0 and the rest as 1.

As opposed to the synthetic data set, the clinical data set did not have a ground truth, or reference scan. For this reason, the application of the deep-learning model in clinical OCT data could not be quantified using the previously metrics. As a consequence, the mean pair-wise cross-correlation, previously used as well in the work from Ntatsis *et al.* [48], was used to determine the quality of the registration.

The mean pair-wise cross-correlation is calculated for two different images, I_1 and I_2, using Equation 3.3:

$$CC(I_1, I_2) = -\frac{1}{n} \sum_{x,y} \frac{1}{\sigma_1 \sigma_2} (I_1(x,y) - \mu_1)(I_2(x,y) - \mu_2) \tag{3.3}$$

where n is the total number of pixels, μ is the mean intensity and σ is the standard deviation of each of the two images.

Chapter 4

Results and Discussion

The results presented in this chapter follow the same order of the methods presented in Chapter 3. First, it presents the results from the development of the eye model. Second, it shows the data set acquired with the eye model, represented by an example B-scan for each type of movement replicated. Next, it presents the quantitative and qualitative evaluation of the obtained results using the registration algorithms. Finally, it shows a comparison between the results from previous studies, using clinical OCT data, and the results obtained with this data set. This is done in order to achieve the goal of understanding if by using a pre-trained model to correct OCT scans, the results are improved when comparing with a non pre-trained model. Throughout the chapter there is a critical analysis of the results, and, when possible, they are compared to existing works in the literature.

4.1 Development of the eye model

Using the 3D sketches previously shown in Section 3.1, and produced using Autodesk Inventor Software, the mechanical parts of the model were 3D printed. Figure 4.1 shows the 3D printing process.

Taking into account the specifications of this project, refered in Section 3.1, there was the need to test the plastic material before printing the final designs. From the literature, it is known that the chosen material, PLA, has waterproof and watertight characteristics, however the material was tested before printing the final designs.

Since 3D printing is characterized by layer deposition of the material, it was necessary to test the pieces watertight and waterproof capacities. The eye model was filled with water and left for 5 days in order to check its properties. No leaks were found during that period. After

that period, the eye model was opened and it was still full of water. The 3D printing parameters used for the final phantom model are shown in Table 4.1.

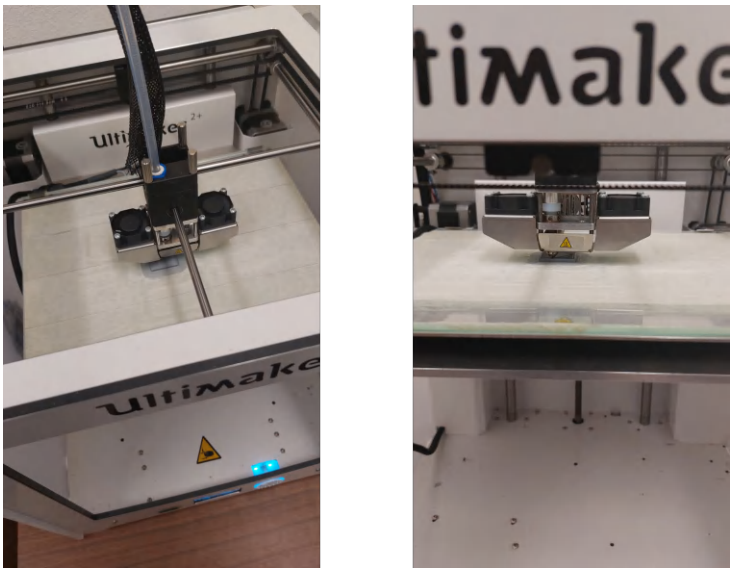


Figure 4.1: Illustration of the 3D printing process using Ultimaker 2+.

Table 4.1: 3D printing parameters.

Printing temperature	200 °C
Build plate temperature	60 °C
Wall thickness	2 mm
Infill	40 %

The eye model was developed in different stages. The final result, shown in Figure 4.2, is a modular piece, made of three parts. Being a modular piece allows the model to be altered, *e.g* to simulate eye diseases, which is an advantage compared to other phantoms presented in the literature [31, 38]. Moreover, the developed phantom is non-degradable. This means that it can be used for long periods of time (weeks/months) without the need to substitute any part, this is also an advantage. An assessment of the phantom properties along time, to ensure good performance and determine its life time, is out of scope of this project. Nevertheless, these tasks could be done in future work.

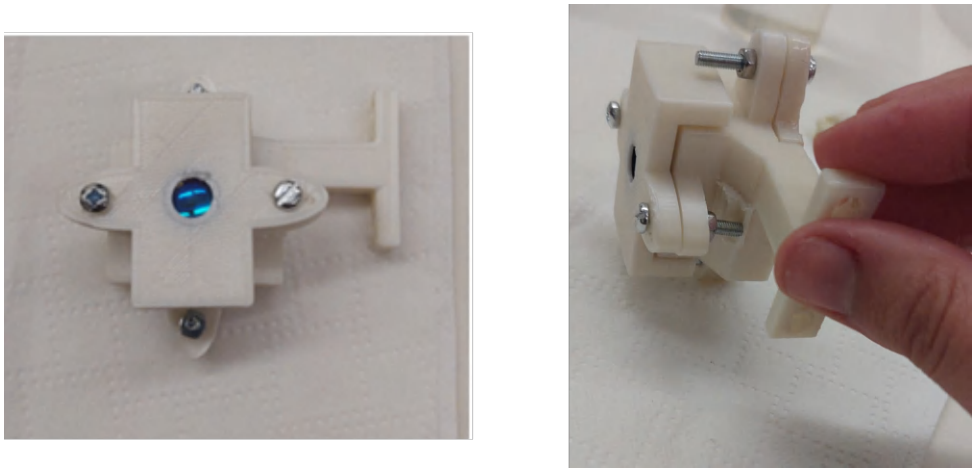
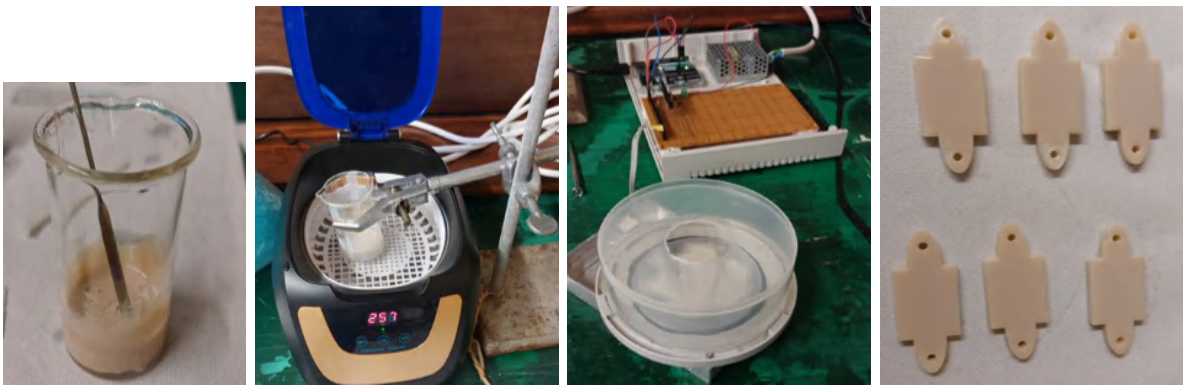


Figure 4.2: Frontal (left) and lateral (right) view of the final eye model.

4.1.1 Posterior segment

As described in Section 3.1.3, the layers, in the posterior segment, were replicated using a spin coating protocol. The key steps of the protocol used to produce the posterior segment layers are shown in the next Figure 4.3.



(a) (b) (c) (d)

Figure 4.3: Key steps of the spin coating protocol to produce the retinal layers (a) Mixture of the elastomer base with TiO_2; (b) Ultrasonic bath for 1 hour; (c) Pour the mixture into the spin coater for 3 minutes; (d) End result layers.

The first step (Figure 4.3 (a)) was to mix the elastomer base with TiO_2. The mixture presents a beige color due to the use of TiO_2, as scattering agent. The central layer, the choroid, is the layer with the less percentage of TiO_2, and therefore this is the mixture with the lighter tone. In the next step, the ultrasonic bath is used to make sure the mixture is homogeneous (Figure 4.3 (b)). This step prevents the concentration of the scattering agent in large particles. Finally, the mixture is poured into the spin coater (Figure 4.3 (c)). To use the correct amount

of mixture, to prevent waste of the materials used, the mixture was dropped using a syringe. The end final retinal phantoms, with the 3 different layers in each piece, are presented in Figure 4.3 (d).

However, the reproduced layers may present a limitation for this phantom. In this model, only three layers of the human eye were reproduced, in particular, the retina was reproduced as a single layer. The human retina, as seen in Section 2.1, is a complex tissue that can be subdivided into 10 different layers, where each one of these layers has a thickness between 13 and 102 μm [71]. Since the anatomical thickness of individual retinal layers is small (up to 18 μm [72]), it was challenging to effectively measure the correct mass of TiO_2, for the production of each layer, and the risk of producing an uneven layers was high. Moreover, it was not possible to produce, nor replicate, the thinner retinal layers because of the limited maximum speed of the spin-coater device. For these reasons, the replicated retina, in this eye phantom, was considered as a single layer in the phantom. As a consequence, the type of data expected is simpler than the data from real retinas, where all the sub layers are present, which is also the function of a phantom, to have simple and controlled data.

4.2 Data set description

As described in Section 3.2.1, the OCT system that was used in this work has a long working distance microscope objective. Therefore, it was necessary to change the optical distance of the OCT system to acquire images of the back of the phantom. Since that was not possible, the data set, described in Section 3.2.3, and shown in the next sections, was acquired using only the last part of the developed phantom, the posterior segment layers.

The simulated movements were in the range of real involuntary eye movements and the values used in each acquisition were generated randomly. In total, 156 volumes with movement artefacts were acquired. For each acquired volume, the amplitudes and frequencies of the microsaccades and drifts were stored. These values are presented in the intervals in Table 4.2.

Figure 4.4 shows a B-scan from a static volume and the respective replicated layers. From top to bottom, this Figure shows the first replicated layer, the retina, followed by the choroid, with the least concentration of scattering agent, and, lastly, the sclera. Visually there is not a significant difference between the concentration of scattering agent in the first layer to the third layer.

Table 4.2: Interval range of the amplitudes and frequencies, of microsaccades and drifts, for each data set type (for volumes with repetitive movements and extreme movements there is only one type of movement simulated).

Data set type	**Microsaccade amplitude (°)**	**Microsaccade frequency (RPM)**	**Drift amplitude (°)**	**Drift frequency (RPM)**
Volumes with repetitive movements	0.08 - 1.20	0.25 - 2	-	-
Volumes with the movement pattern	0.40 - 0.98	2	0.07 - 0.69	0.02 - 0.08
Volumes where the movement pattern starts after the acquisition	0.33 - 0.96	2	0.12 - 0.38	0.04 - 0.08
Volumes with extreme movements	4.00 - 12.00	1 - 2	-	-

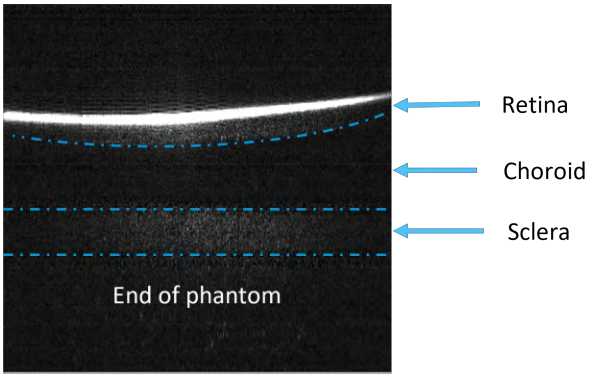


Figure 4.4: B-scan with the replicated layers.

In Figure 4.5, there is a representation of each type of scan available in this data set. All images show are B-scans from the slow axis, where the movement artefacts are visible.

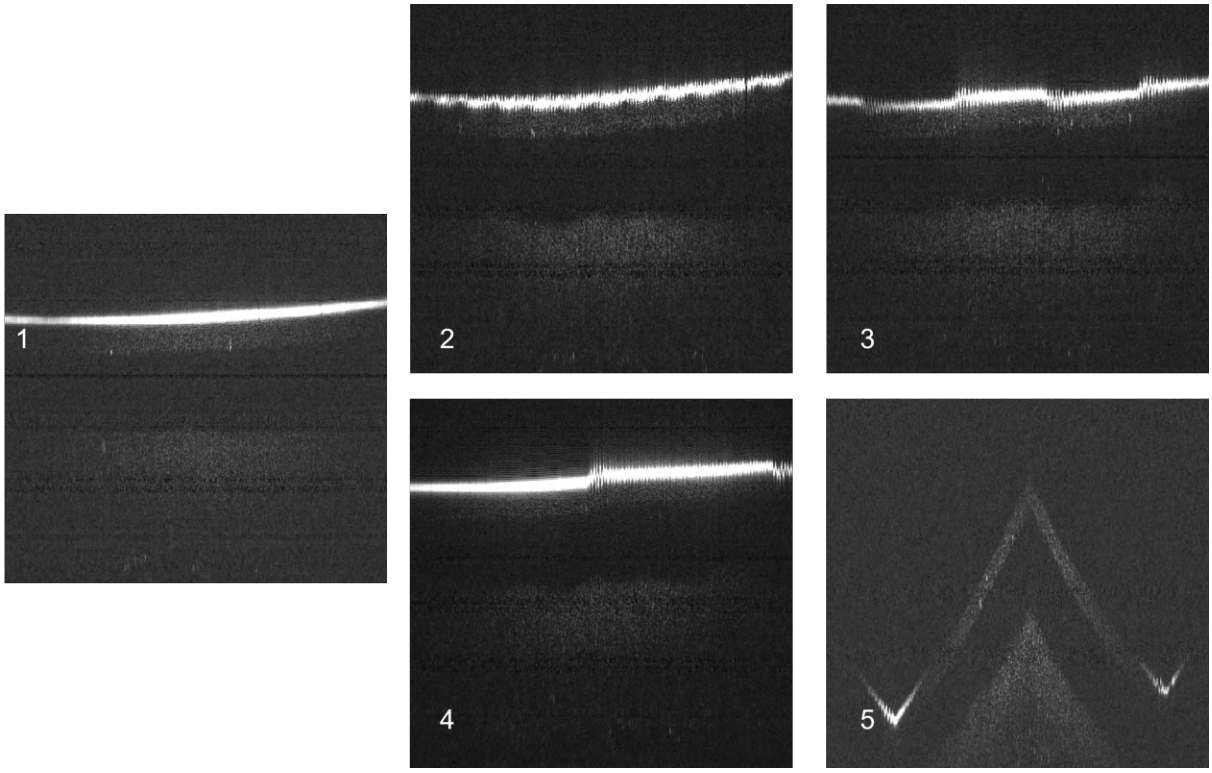


Figure 4.5: Data set overview, slow axis B-scans - (1) Static B-scan; (2) B-scan with repetitive movements; (3) B-scan with the movement pattern; (4) B-scan where the movement starts after the start of the acquisition; (5) B-scan with extreme movements.

As explained before, the volumes were acquired using only the posterior segment part of the model. This allowed the visualization of the three layers reproduced, however the contrast between layers is highly reduced. When comparing a B-scan from a human eye with a B-scan from the model developed there are some important differences, specially in the layers contrast, as can be seen in Figure 4.6.

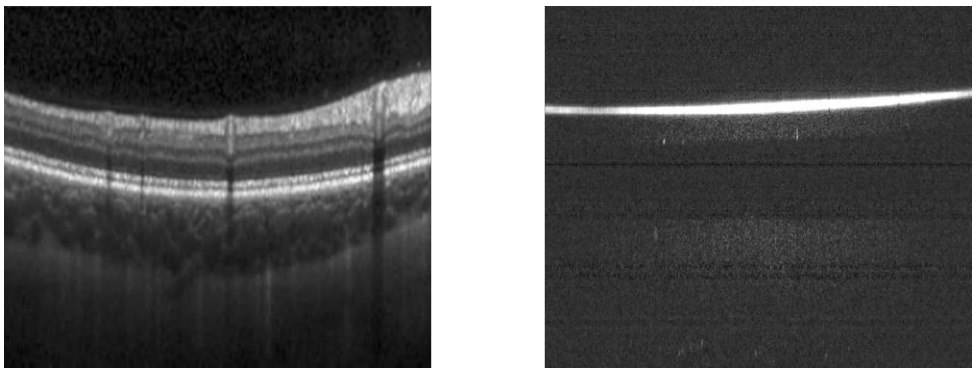


Figure 4.6: B-scan acquired from a human eye (left) [73] and from the developed model (right).

The difference between contrasts can be explained due to several reasons. Firstly, the OCT systems used are different, the volumes obtained in this work were not acquired using a com-

mercial OCT. Secondly, the image postprocessing is also different, there are multiple different techniques, including contrast enhancement strategies that could have been applied to the real OCT scan.

Lastly, the interface is different. In the acquisition with a human eye, the first interface is between vitreous humor (refractive index between 1.334 and 1.337 [74]) and the first retina layer (refractive index between 1.330 and 1.370 [75]). However, in the developed model, the first interface is between air (refractive index of 1.00) and the first phantom layer, since it was not possible to acquire scans with the whole developed phantom. When calculating the ratio of reflected intensity, using Equation $R = \left(\frac{n_1-n_2}{n_1+n_2}\right)^2$, where n_1 and n_2 are the respective refractive indexes of the two structures, for the B-scan with the human eye the value is R = 0.000029 and from the phantom B-scan is R = 0.022182 (for the refractive indexes in intervals the value used was the mean value). Analysing these values, the phantom B-scan has a higher reflectance, which can explain the high intensity in the first interface. The high intensity of the first layers makes the visualization of the other layers difficult, even with the image in logarithmic scale.

When comparing the two different B-scans, it is also noticeable that the human eye layers have a bigger curvature than the scan from the eye model. This happens due to the technique used to produce the different layers. The spin coater produced thin homogeneous layers, however, the curvature of the layers produced was not a parameter monitored and therefore the layers do not replicate the real curvature of the human eye.

4.3 OCT Motion correction

This section presents the results and discussion of the motion correction algorithms that were tested to correct for the simulated eye movements.

4.3.1 Conventional registration

This section presents the registration results using SimpleElastix. Tables 4.3 and 4.4 show the results for both the Euclidean distance and Dice score, respectively, for four different versions of parameters using the SimpleElastix registration, shown in Table 3.4. The test set used for the SimpleElastix is the same test set used to test the deep-learning trained models (10% of the OCT volumes acquired with movement artefacts), in the next Section. Using the same data set allowed to make comparisons between the results obtained. Since the results do not follow a normal distribution, the statistical measures used to evaluate the metrics were the median and

Table 4.3: Descriptive statistics for Euclidean distance, for the conventional registration algorithm, SimpleElastix, in pixels.

SimpleElastix	**Euclidean Distance**	
	Median	**IQR**
v1	0.36	0.16
v2	0.32	0.11
v3	0.59	2.85
v4	0.42	0.51

interquartile range (IQR). To obtain a better visualization of the obtained results, the Euclidean distance and Dice score box-plots of each the analysed versions are shown in Figures 4.7 and 4.8. In these Figures, the green lines inside the boxes represents the median value, and the bullet points represent outliers.

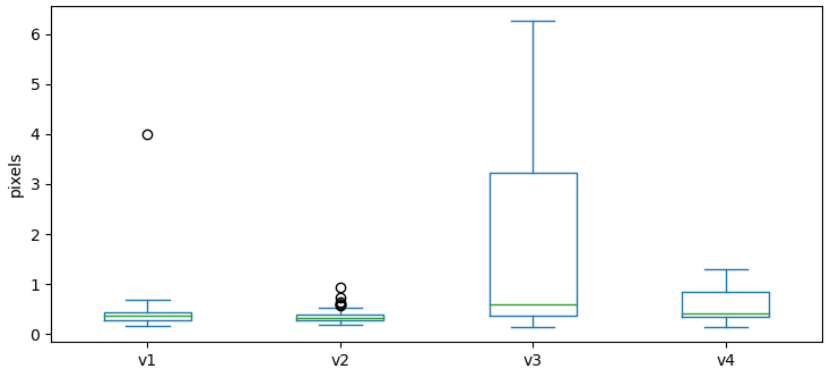


Figure 4.7: Box-plot representation for Euclidean distance, for the conventional registration algorithm, SimpleElastix.

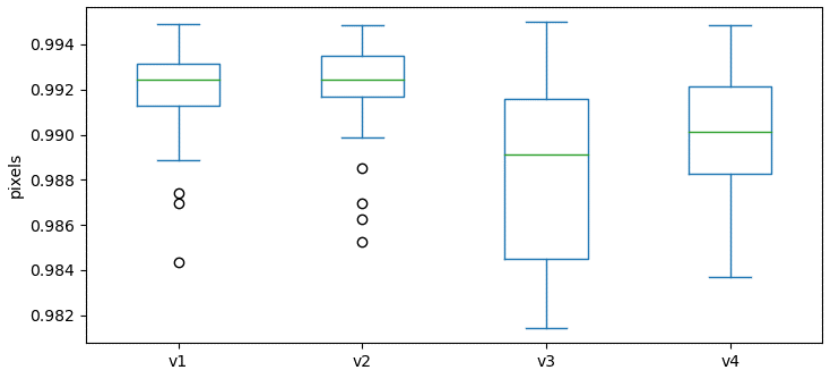


Figure 4.8: Box-plot representation for Dice score, for the conventional registration algorithm, SimpleElastix.

Table 4.4: Descriptive statistics for Dice score for the conventional registration algorithm, SimpleElastix, in pixels.

SimpleElastix	Dice Score	
	Median	**IQR**
v1	0.9924	0.0019
v2	0.9924	0.0018
v3	0.9891	0.0071
v4	0.9901	0.0039

Regarding the Euclidean distance, the best results occur for version 2. Although the median values can be considered similar for all the different versions, the interquartile interval present the smallest value for version 2. This means that the distribution of results is more stable than in the other versions. When analysing the box-plot, in Figure 4.7, version 2 also presents the best representation of results. This representation shows that version 3 is clearly the worst version, with a wide range of results.

The Dice score median results are also very similar between the different versions, moreover, for version 1 and 2 the median is the same. Therefore, the best version was determined analysing the interquartile interval, which has the smallest value for version 2. The box-plot representation, in Figure 4.8, validates the conclusions made before. The results from version 1 and 2 are very similar, however version 2 has less outliers, and therefore is considered the best version.

In Figures 4.9, 4.10 and 4.11, some examples of the registered scans using SimpleElastix are shown. Each Figure has, on the left, the fast axis B-scan, on the center, the uncorrected slow axis B-scan, and on the right the registered slow axis B-scan. There is one example for each type of movement present in the data set. Only the results from version 2 are shown, since this is the best version.

Figure 4.9 shows the registration output for a volume that was acquired with the movement pattern, described in Section 3.2.2, during the whole acquisition time. For this type of artefact, the artefacts are mostly corrected, visually the B-scan only present a slight artefact, represented with the red arrow. However, it is clear, when comparing to the fast axis B-scan, that the registered B-scan loses the curvature of the retinal layers.

Figure 4.10 shows the registration output for a volume that was acquired with the movement pattern, described in Section 3.2.2, but as opposed to what happened in the previous Figure, the movements only start after the beginning of the acquisition. In the registered B-scan there

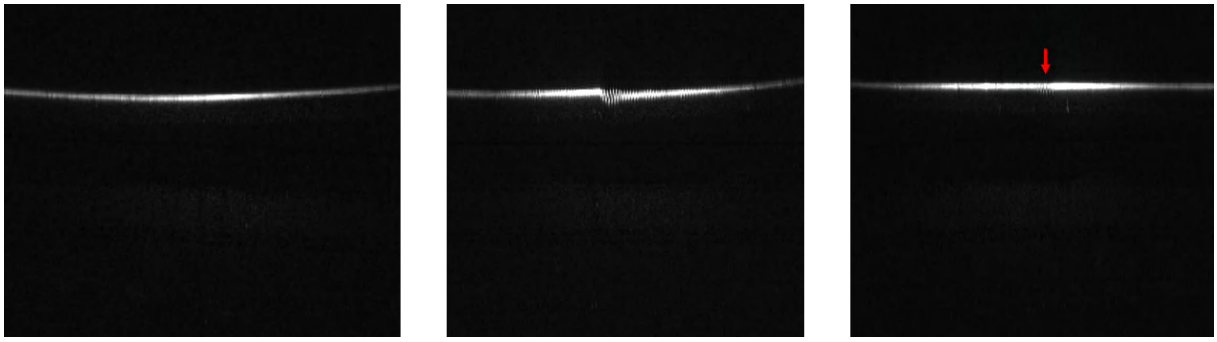


Figure 4.9: SimpleElastix output, from a volume acquired with the movement pattern - fast axis B-scan (left), unregistered slow axis B-scan (center), registered slow axis B-scan (right).

are still some visible artefacts that were not totally corrected. These artefacts appear to be microsaccades, since it occurs twice in the B-scan, represented with red arrows, and it is a short duration artefact. Therefore, for this type of movement the registration was not as successful as it was for the previous movements, and these volumes can be considered as the most challenging volumes to correct.

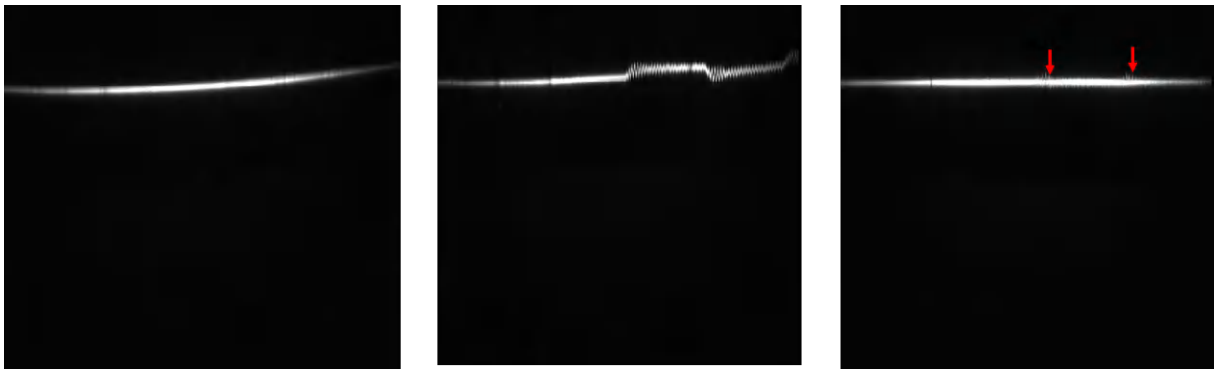


Figure 4.10: SimpleElastix output, from a volume acquired with the movement pattern after the acquisition starts - fast axis B-scan (left), unregistered slow axis B-scan (center), registered slow axis B-scan (right).

At last, Figure 4.11 shows the registration results for a volume acquired with repetitive movements, as described in Section 3.2.2. In this registered B-scan, the artefacts seem to have been attenuated in amplitude, however, there are still some artefacts visible. This means that the registration is not as successful as it was for the first type of movements. Despite this, the artefacts are reduced, in terms of amplitude, in a significant way.

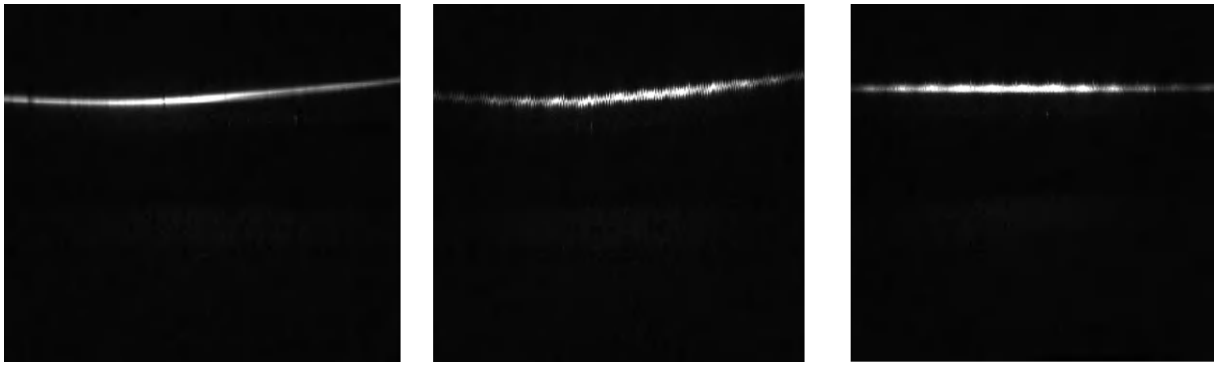


Figure 4.11: SimpleElastix output, from a volume acquired with repetitive movements - fast axis B-scan (left), unregistered slow axis B-scan (center), registered slow axis B-scan (right).

4.3.2 Deep-learning registration

This section outlines the different results based on the deep-learning algorithm developed by Ntatis *et al.* [48]. First, different data sets were used for training the the algorithm which led to different trained models and, therefore, different registration results.

Table 4.5 shows the Euclidean distance for the five different versions obtained using the data-centric approach (see Section 3.3.3), followed by Table 4.6, which shows the Dice score results for the same trained models. As happened in the previous Section (Section 4.3.1), these results do not follow a normal distribution, and therefore the statistics analysed were the median and interquartile range (IQR). In Figures 4.12 and 4.13 there is a box-plot representation. The green line represented in each box is the median, and the bullet points are outliers.

Table 4.5: Descriptive statistics for Euclidean distance, using the data-centric approach, in pixels.

Models	**Euclidean Distance**	
	Median	**IQR**
Data set v1	4.59	0.25
Data set v2	4.65	0.28
Data set v3	5.08	1.98
Data set v4	4.54	0.04
Data set v5	1.77	1.67

Table 4.6: Descriptive statistics for Dice score, using the data-centric approach, in pixels.

Models	Dice Score	
	Median	**IQR**
Data set v1	0.9815	0.0100
Data set v2	0.9845	0.0159
Data set v3	0.9774	0.0042
Data set v4	0.9904	0.0083
Data set v5	0.9735	0.0094

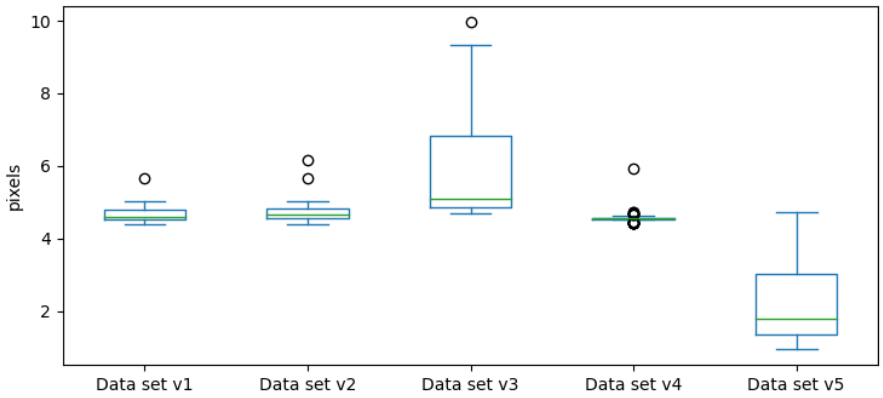


Figure 4.12: Box-plot representation for Euclidean distance, using the data-centric approach.

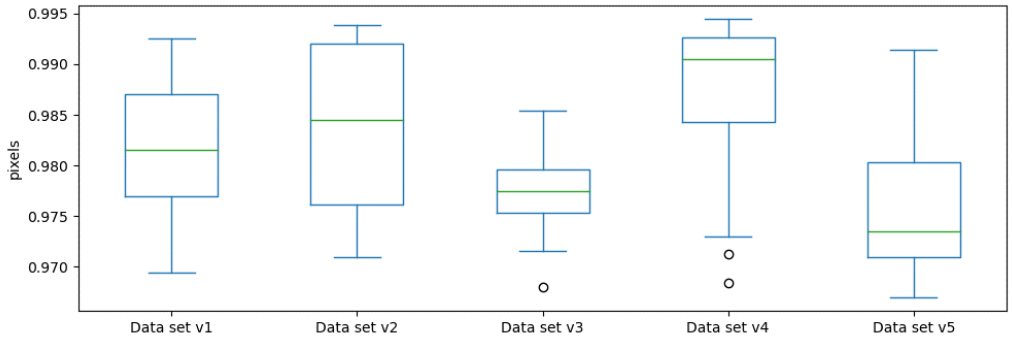


Figure 4.13: Box-plot representation for Dice score, using the data-centric approach.

On a first analysis of Table 4.5, the best version is Data set 5, where the minimum value for the Euclidean distance is significantly lower than the other results. However, when analysing Table 4.6, with the Dice score results, Data set 5 presents the worst results. Due to the discrepancy between the two metrics, and the substantial difference in the Euclidean distance for Data set 5, the registration volumes were observed. In Figure 4.14, it is possible to see that the

registered B-scan, from Data set 5, has all the artefacts, which means that the correction was not done properly in this version. The registered B-scan was overlapped with a reference B-scan (represented in red).

The training and validation curves from Data set 5, represented in Figure 4.15, show that the training did not happened. The metric used to evaluate learning in this algorithm is a loss function, and therefore the goal is to minimize the loss function over the training. In Data set 5, that does not happen, which implies that there was no model training and, therefore, the results obtained with Data set 5 are not valid. The reason behind this is not due to the loading strategy, which is the same as Data set 4, where it did not produce errors. The hypothesis behind this result is that the extreme movements, introduced in the training data set of Data set 5, degrades the registration. When analysing the training plot from Data set 3, where the data set also contained extreme movements, (see Figure 4.25), the training curve does not present the expected tendency. Therefore, the inclusion of extreme movements in the training data set of the motion correction algorithm may deteriorate the registration results. After excluding the results from Data set 5, there was the need to evaluate the remaining four versions, regarding the data-centric approach.

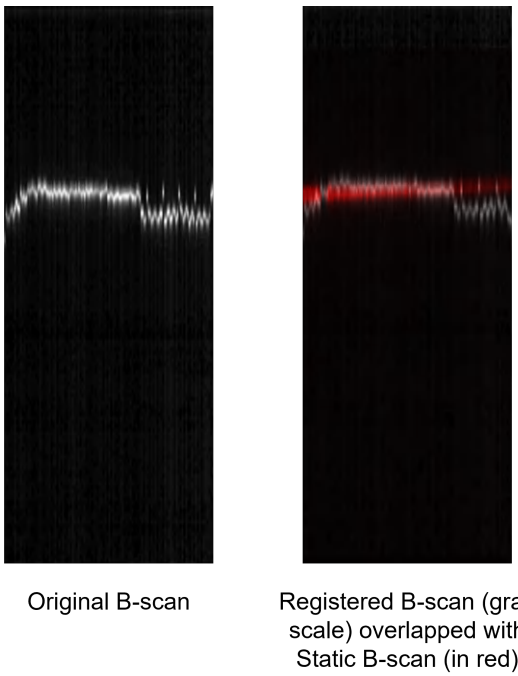


Figure 4.14: Original B-scan, with artefacts compared with a registered B-scan, from Data set 5.

When analysing the median results for the remaining 4 versions (Data set 1-4), the best result, for the Euclidean distance and Dice score, is Data set 4. For the Euclidean distance,

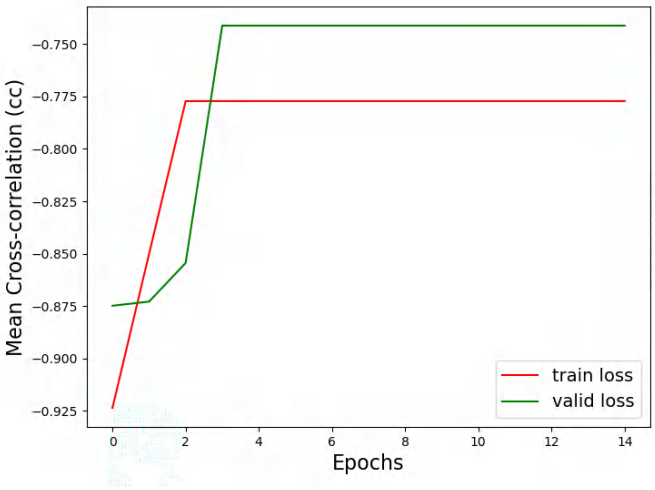


Figure 4.15: Training and validation curve of Data set 5.

this version also has the the smallest interquartile interval, which shows the small variability of results. For the Dice score, although the interquartile interval is not the smallest for Data set 4, the median result is the best of the remaining models.

In the box-plot representations, in Figures 4.12 and 4.13, as expected from the results, Data set 4 has the best results for the Euclidean distance. In the representation, this version has the smallest distribution of results, as expected. For the Dice score, Data set 4 has the best median result, although this version has a bigger interquartile interval. Analysing both metrics, from the different registration versions, Data set 4 is the best registration model, consequently, this is the data set used in the model training for the parameter optimization, in the next step of this project.

Figures 4.16, 4.17 and 4.18 present the visualization of the registration results for the four versions obtained with the data-centric approach. Each Figure is from one different type of movement present in the data set. The registered B-scans have a static B-scan, a reference image (represented in red), overlapped, in order to perform better comparisons of the obtained results. There are no major differences between the three type of movements analysed. For all the different movements, the registered B-scan from Data set 4 is the scan which showed the best registration results. In this version, the retinal layers from the registered scan are almost perfectly aligned with the reference retinal layers, and the motion artefacts are removed. All B-scans shown are slow axis B-scans.

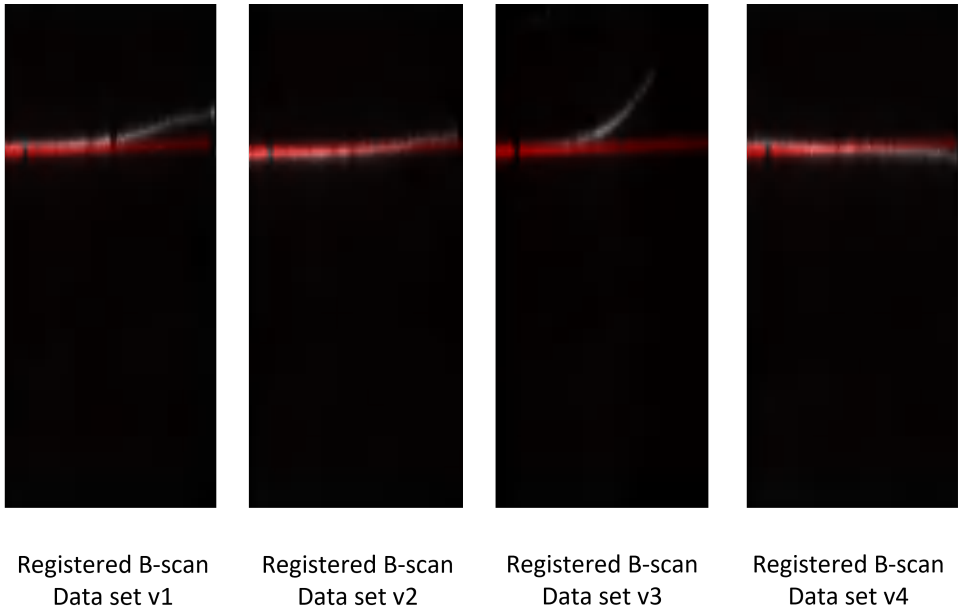


Figure 4.16: Registration results for a volume acquired with the movement pattern, for the four different versions of data set trained models, using different data sets.

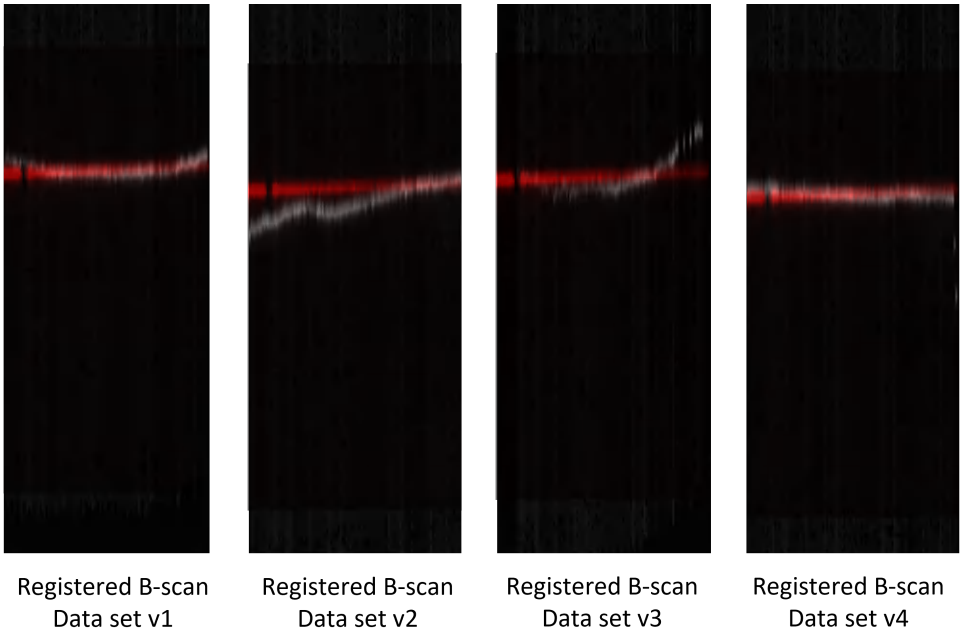


Figure 4.17: Registration results for a volume acquired with the movement pattern, that starts after the start of the acquisition, for the four different versions of data set trained models, using different data sets.

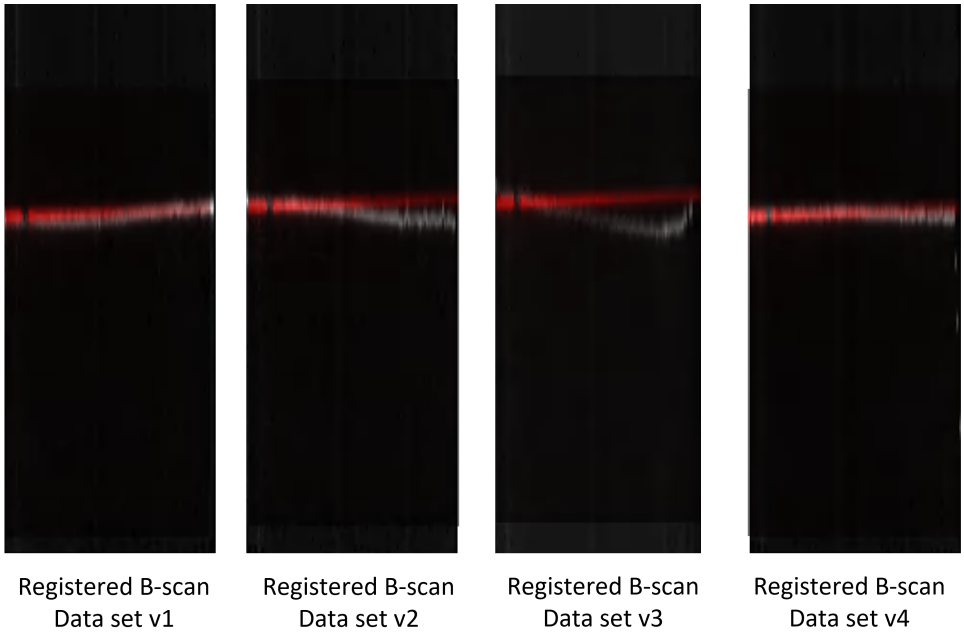


Figure 4.18: Registration results for a volume acquired with repetitive movements, for the four different versions of data set trained models, using different data sets.

The second approach (model-centric) used was based on parameter optimization. This approach was used to understand if different parameters would lead to better registration results. The different parameters used are shown in Table 3.6. Table 4.7 shows the Euclidean distance for the four different versions analysed, followed by Table 4.8, which shows the Dice score results for the same trained models. The respective box-plot representations are shown in Figures 4.19 and 4.20.

Table 4.7: Descriptive statistics for Euclidean distance, using the parameter optimization approach, in pixels.

Models	**Euclidean Distance**	
	Median	**IQR**
Parametrization v1	6.75	3.36
Parametrization v2	4.82	1.00
Parametrization v3	4.39	0.04
Parametrization v4	4.22	0.27

Table 4.8: Descriptive statistics for Dice score, using the parameter optimization approach, in pixels.

Models	Dice Score	
	Median	**IQR**
Parametrization v1	0.9735	0.0037
Parametrization v2	0.9778	0.0078
Parametrization v3	0.9881	0.0081
Parametrization v4	0.9890	0.0048

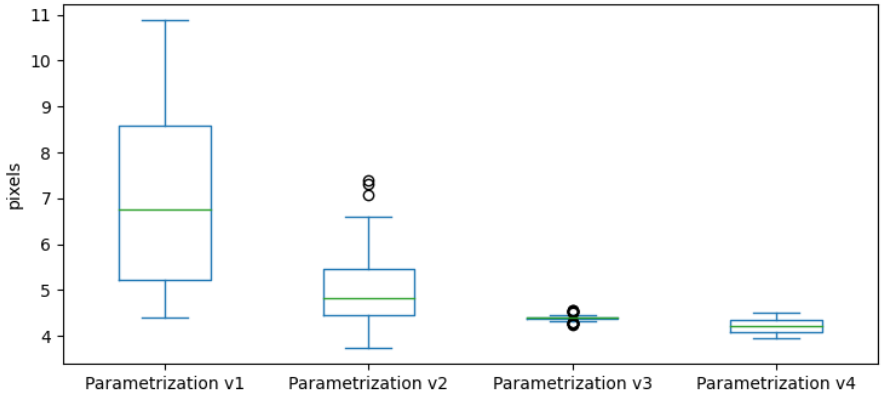


Figure 4.19: Box-plot representation for Euclidean distance, using the model-centric approach.

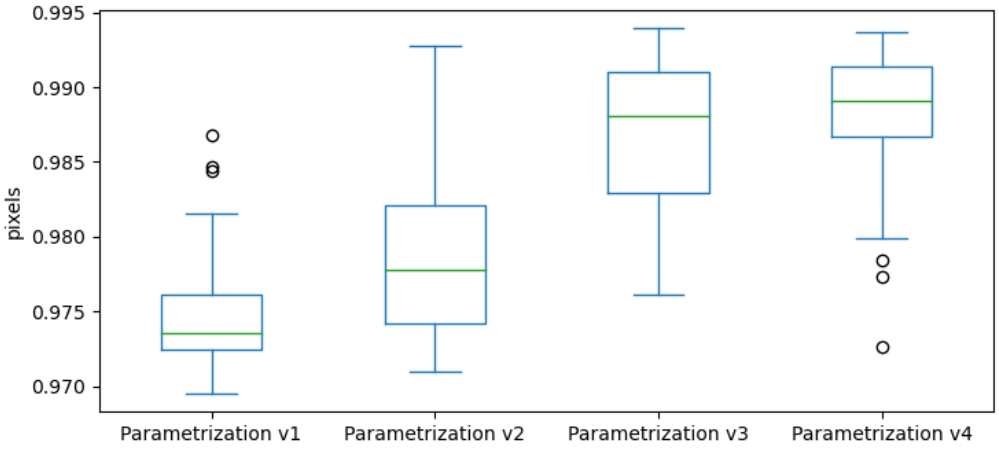


Figure 4.20: Box-plot representation for Dice score, using the model-centric approach.

The results from Parametrization 1 are, significantly, worse, when compared to the results from the data-centric approach, and therefore did not present any improvement on the results obtained. When analysing the remaining versions, Parametrization 2 has worse results than

Parametrization 3 and 4. These last two versions do not present major differences. For the Euclidean distance, the median is better for Parametrization 4, however, the interquartile interval is bigger for this version. When analysing the results from the Dice score, specially the visualization, in Figure 4.20, there are hardly any differences between this versions. The trained model from Parametrization 4 was selected as the best for the smaller interquartile interval for the Euclidean distance.

The Figures 4.21, 4.22, and 4.23 show the registration results for the four different versions analysed. There is one example B-scan for each of the three different movements present in the data set used. Volumes where the movement starts after the start of the acquisition (Figure 4.22) proved to be more difficult to correct. The registered B-scans (represented in a gray scale) are overlapped with a static B-scan (represented in red). As expected from the metric results, for all the different type of movements, Parametrization 1 and 2 had the worst registration results. These versions correct some movement artefacts, however the retinal layers suffer deformation in these versions. Analysing the results from Parametrization 3 and 4, the results seem better for Parametrization 3. In this version, not only the artefacts are corrected, but the correspondence between the registered B-scan and static scan is the best from all the analysed versions.

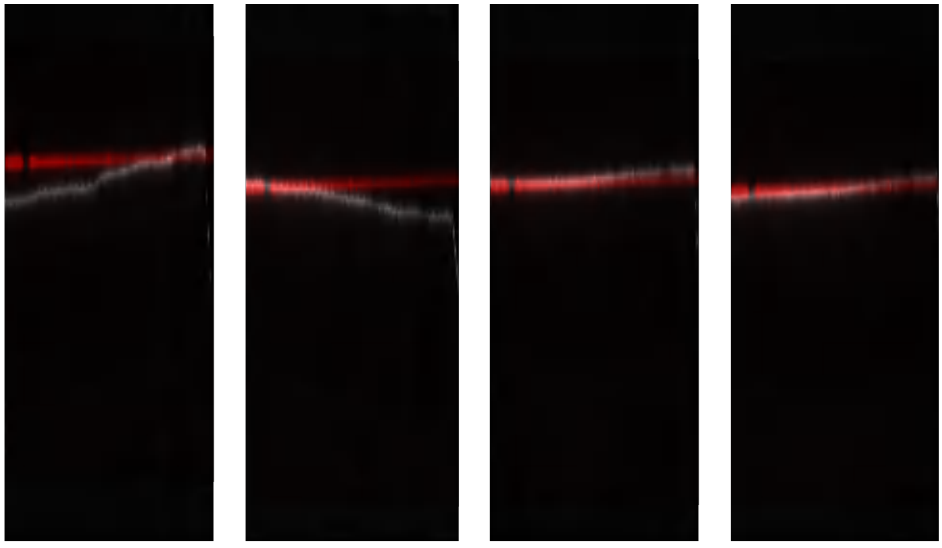


Figure 4.21: Registration results for a volume acquired with the movement pattern, for the four different versions of parametrization models, using different parameters when training the algorithm.

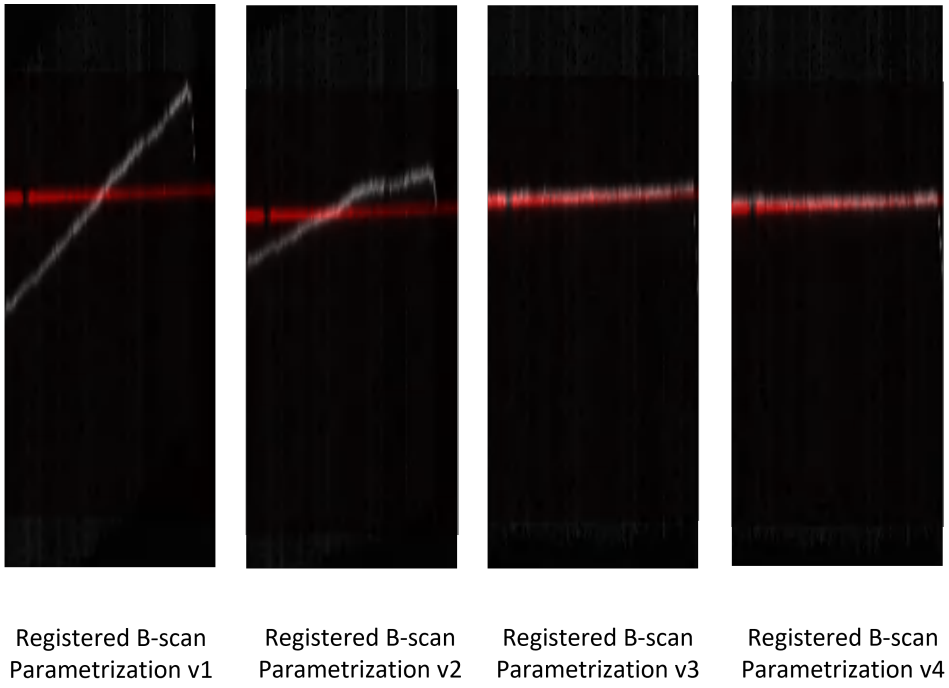


Figure 4.22: Registration results for a volume acquired with the movement pattern, that starts after the start of the acquisition, for the four different versions of parametrization models, using different parameters when training the algorithm.

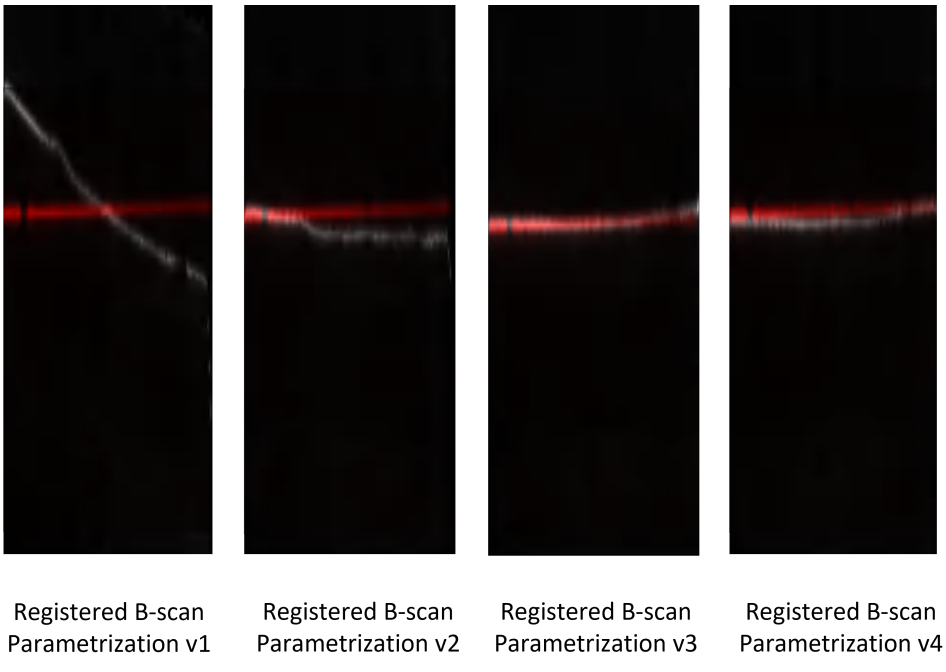
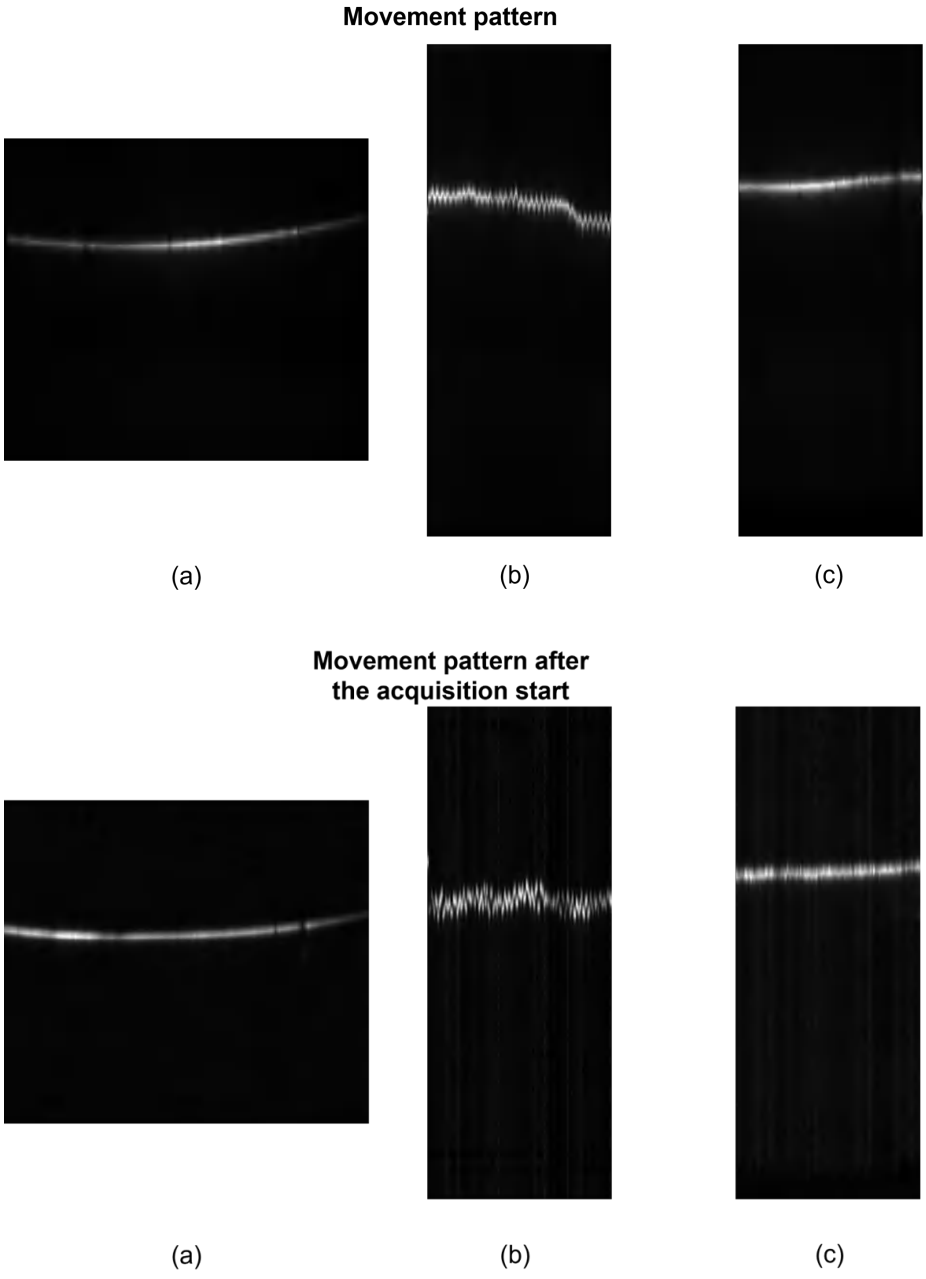


Figure 4.23: Registration results for a volume acquired with repetitive movements, for the four different versions of parametrization models, using different parameters when training the algorithm.

When comparing the best version from the data-centric approach (Data set 4) with the best

version from the model-centric approach (Parametrization 3), the last presents better results for the Euclidean distance, however for the Dice score, Data set 4 has better results, although the difference between results is small. As a conclusion, the best trained model was obtained with the Data set 4 and Parametrization 3. This trained model was the model used to obtain the next registered B-scans, Figure 4.24, where there is a fast axis B-scan (artefact-free) and a comparison, for each type of movement, between the initial B-scan, with artefacts and the registered B-scan, output of the trained model.



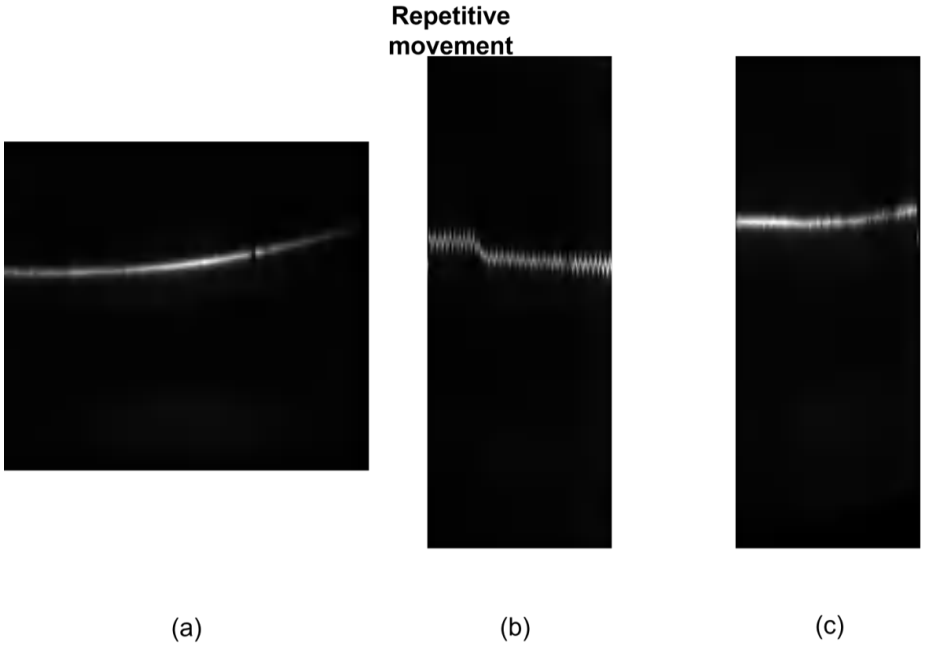


Figure 4.24: Registration results, for the best trained model, for the three different types of movements - (a) Fast axis B-scan, (b) Original (uncorrected) slow axis B-scan, (c) Registered slow axis B-scan.

In general, all the different types of movements show reasonable results for the best registration model. The left images represent the fast axis B-scan for all the volumes analysed. The two right images show the slow axis B-scans (before and after the registration) from the volumes, from top to bottom, acquired with the movement pattern, acquired with the movement pattern starting after the start of the acquisition and acquired with repetitive movements. For these registered B-scans, the volumes also maintain the layer structure.

The learning and validation curves for the nine different models trained are presented in Figure 4.25, including Data set 5, which was previously discussed. The training curves were plotted with the same intervals for the mean cross-correlation, expect for Data set 3 (Figure 4.25 (c)) and Data set 5 (Figure 4.25 (e)), which had training curves with values out of the normal interval for the other trained models.

The best trained model, with Data set 4 and Parametrization 3, has its training curves represented in Figure 4.25 (h). In this representation, the learning curves show the learning performance of the algorithm over the training epochs. In this case, the goal is to minimize the loss function over training. Both training and valid loss curves decrease to a point of stability, and with a small gap between curves, which indicates a good fit for this trained model.

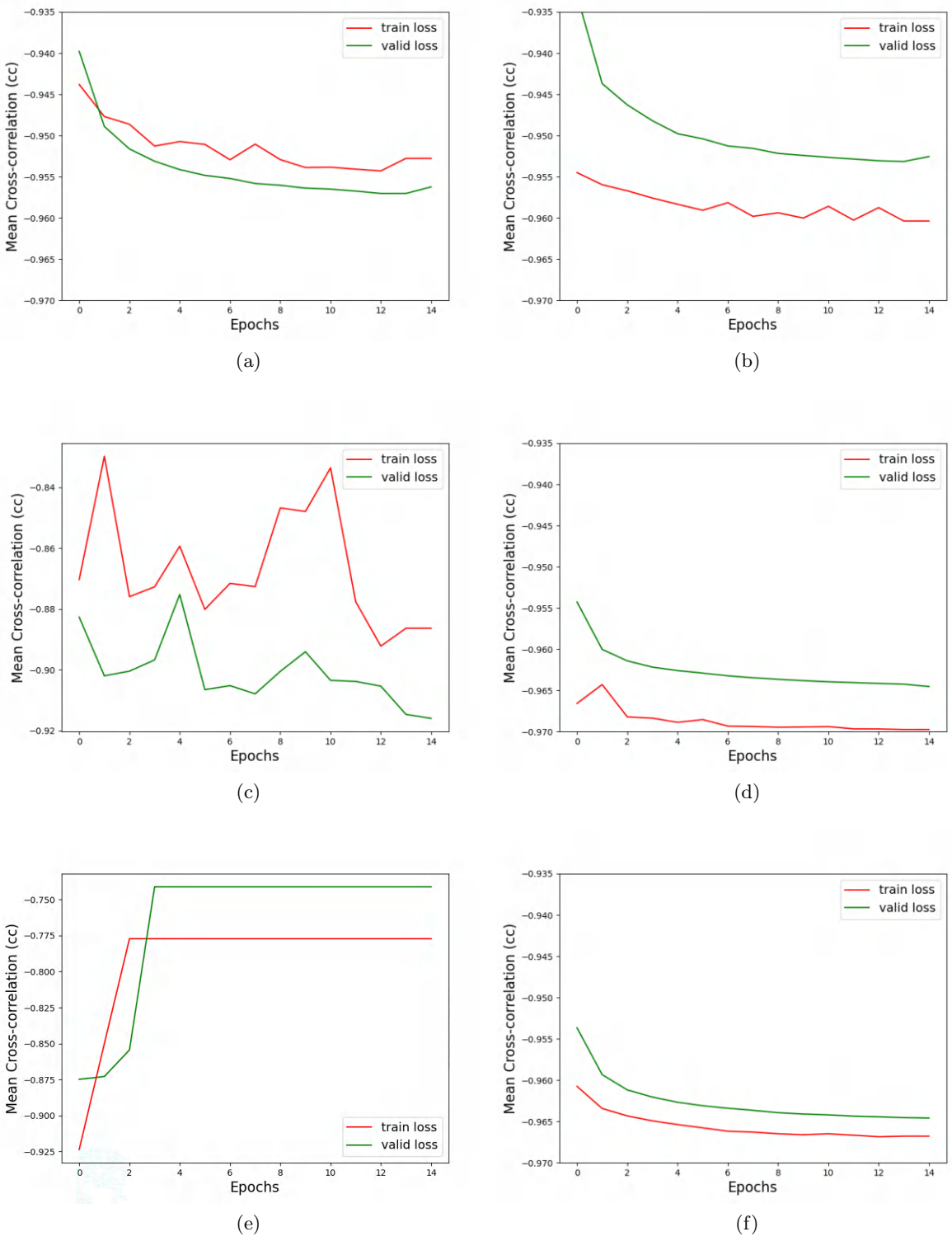
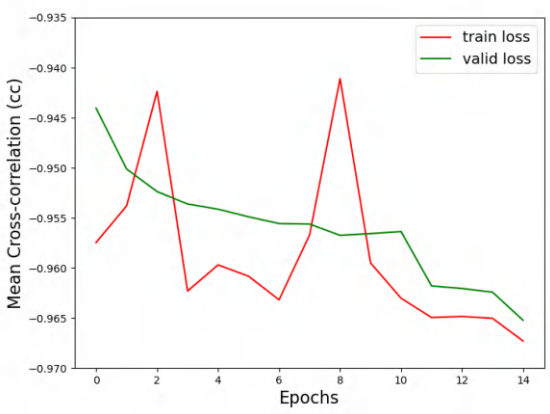
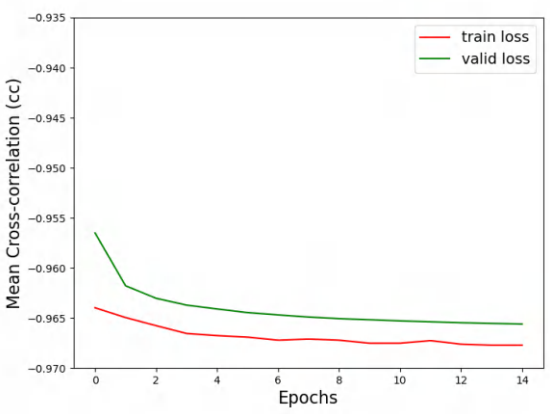


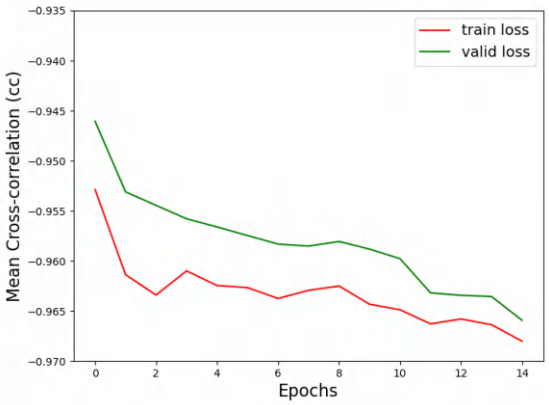
Figure 4.25: Training and validation curves for the different trained models - (a) to (e) Data set 1 - 5 and (f) to (i) Parametrization 1 - 4.



(g)



(h)



(i)

Figure 4.25: Training and validation curves for the different trained models - (a) to (e) Data set 1 - 5 and (f) to (i) Parametrization 1 - 4.

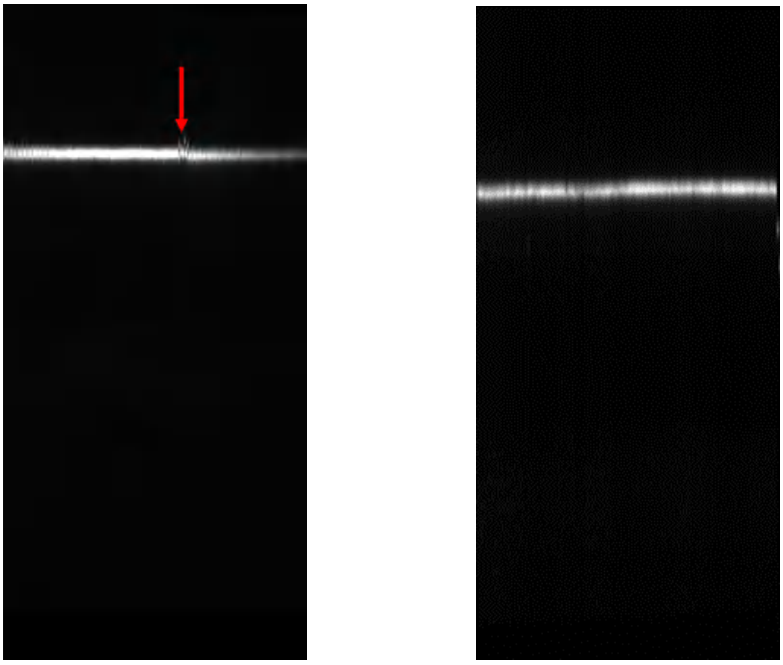
4.3.3 Comparison between registration approaches

The first major difference between the two strategies used was the registration time. Using SimpleElastix the registration took about 30 minutes per volume, 63.5 hours for all the volumes registered, whereas the deep-learning algorithm took around 24 hours to perform the model training, and just a few minutes to run the model for the test set. Furthermore, using the deep-learning model, the model learns the data in which is trained on, specially important in optical imaging.

When analysing the results obtained with the two approaches, there are some differences. The registration results from SimpleElastix depend strongly on the type of movements to be corrected. Some of the largest amplitude movements, that create larger artefacts, were more difficult to correct. In addition to that, the curvature of the interfaces was altered by the registration algorithm.

On the contrary, the registration results from the deep-learning model show no major differences between the correction of the different types of artefacts, the majority of artefacts present in the acquired data set are successfully corrected. These registered B-scans also maintain the curvature of the retinal layers, which did not happen in the results from SimpleElastix, this may happen as a consequence of the group-wise approach used in the deep-learning algorithm.

Figure 4.26, shows side-by-side a registered B-scan obtained using SimpleElastix (Figure 4.26 (a)) and a registered B-scan using the deep-learning algorithm (Figure 4.26 (b)). The registered B-scan from SimpleElastix was cropped, in the area analysed in the deep-learning B-scan and to its dimensions as well. In this B-scan from SimpleElastix it is possible to see some artefacts which are not corrected, indicated with the red arrow.



(a) (b)

Figure 4.26: Registered slow axis B-scan - (a) Registered B-scan using SimpleElastix; (b) Registered B-scan using the deep-learning algorithm.

4.3.4 Application on retinal OCT data

In the last part of the project, the Parametrization 3 model (the best trained model obtained previously) was re-trained in clinical OCT data. This model was used as pre-train for a model trained in clinical OCT data.

Table 4.9 shows the cross-correlation values on the clinical OCT test set. The results are presented for the model trained on clinical OCT test set (1), for the model trained on the phantom OCT test set (2) and for the model pre-trained on the phantom OCT test set and trained on the clinical OCT test set (3). This means that the network weights do not start at 0, as happened in the models without pre-training, the training weights for this model start with the values calculated for the best model trained in OCT synthetic data.

Analysing Table 4.9, it is possible to see that the algorithm trained in OCT data acquired with the eye model had the worst performance among the three models. Even though it is the worst result from the three, the algorithm trained in eye model OCT volumes still corrects part of the motion artefacts that clinical OCT volumes present. When comparing the metrics results, the pre-trained model presented better results than the non pre-trained model.

Table 4.9: Mean cross-correlation results for the different algorithms.

	Mean cross-correlation
(1) Algorithm trained in clinical OCT data	-0.8079
(2) Algorithm trained in the synthetic OCT data set	-0.7167
(3) Algorithm trained in clinical OCT data and pre-trained in the synthetic OCT data set	-0.8473

Visually there are almost no differences between the three models, as shown in Figure 4.27. The first B-scan, is a slow axis uncorrected B-scan, followed by the registered B-scan from the algorithm trained in clinical OCT data, the registered B-scan from the algorithm trained in the synthetic data set, and at last, the registered B-scan from the algorithm pre-trained in the synthetic data set and trained in the clinical data set. Although the differences are small, it is possible to see that the registered B-scan obtained with the model trained in the synthetic data set (third B-scan) still presents some artefacts, specially visible in the last retinal layer. Between the registered B-scan from the non-pre trained algorithm (second B-scan) and the registered B-scan from the pre-trained algorithm (fourth B-scan) the differences are even smaller. There are small irregularities at the end of the second B-scan, which are corrected in the B-scan from the pre-trained algorithm. As a conclusion, even though there are just small differences, the pre-trained algorithm shows improvements when compared to the non pre-trained algorithm.

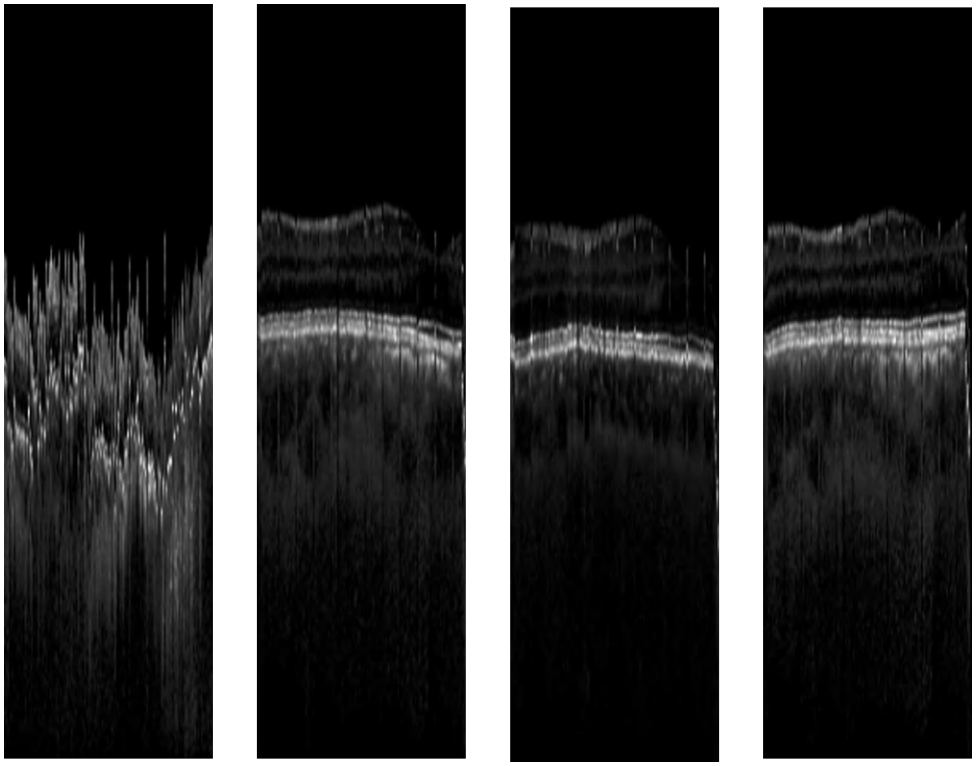


Figure 4.27: Registration results for the three different algorithms - from left to right, slow axis unregistered B-scan, slow axis registered B-scan, for the algorithm trained in clinical OCT data, slow axis registered B-scan, for the algorithm trained in synthetic data, and slow axis registered B-scan, for algorithm pre-trained in synthetic data.

Chapter 5

Conclusions

The correction of motion artefacts is still one of the major challenges in ophthalmic imaging, namely in OCT. Some of the aspects that make OCT motion correction so difficult are the lack of a ground truth, the heterogeneity of data sets (due to differences among individual retinas, but also to hardware/software differences in the commercial devices), and the scarcity of representative data.

This project aimed to present an automatic approach for OCT intra-volume motion correction. To that end, an eye phantom was designed and created. The eye model was developed successfully, with key representative eye structures. This model was then used to acquire OCT scans. These scans were representative of realistic involuntary eye movements, imitating the artefacts that can be seen in clinical data. Then, these artificial scans, with known motion parameters and ground truth, were used to objectively evaluate two state-of-the-art algorithms for motion correction. Both algorithms were chosen to be intensity-based, in order to keep independence from engineered features. One of the algorithms used a conventional image processing, pair-wise approach, and compared pairs of B-scans on a multi-resolution level. The other algorithm used a group-wise approach, which allowed it to preserve the retinal curvature better. A limitation of conventional group-wise approaches is the computation time. However, the chosen algorithm used deep learning to overcome this limitation. Lastly, the deep learning algorithm was evaluated in clinical OCT data. The results were compared with and without pre-training in synthetic data, and the results showed that the pre-training improves the registration results.

The work developed during this project tackled some difficulties. During the development of the eye model, the main challenge was to connect the different parts without obstructing the acquisition field. This was solved by using an hallow design, with the use of O-rings. In addition, the eye phantom replicates only three layers in the posterior eye segment, which is

a much more complex structure. However, the replication of the smallest structures, such as the retinal layers, was difficult due to the used technique. Some hindrances appeared also after obtaining the scans from the phantom. The main difficulty was to preprocess the phantom B-scans in a correct way so that these could resemble retinal OCT images and be used in the registration algorithms. This was linked to the high contrast the phantom OCT scans presented, that made the visualization of the different replicated layers harder than anticipated.

When analysing the registration results, both approaches presented a satisfactory result. Most of the artefacts were corrected, although, it was noticeable through the various registered B-scans that the artefacts with bigger amplitudes (microsaccades) were the most challenging to correct. In addition, the volumes where the movement started after the acquisition, also proved to be the most difficult to correct.

The use of a pre-trained model was able to improve the results previously obtained by those published in the literature [48]. It is known that pre-training improves results in several applications, however the importance of pre-training in real data or synthetic is not known. Therefore, this is a promising result, where the use of synthetic data as pre-training improved the registration results of clinical data. However, further studies in this approach are needed to understand the importance of a pre-train in the different data sets.

5.1 Future work

While this was not a limitation for our application, the eye model developed in this work does not represent all the structures in a real human eye, and could be further improved. Apart from the structures replicated in this work, the eye vessels could be reproduced in a future work. This would add an interesting characteristic to the model, since it would allow the simulation of the eye blood circulation. Additionally, the retinal layers and fovea pit could be reproduced, improving the similarity between the model and the human eye.

The developed eye model was only used to acquire OCT scans in the OCT system described in Section 3.2.1. Using the model in several commercial OCT devices would allow comparisons between devices. Furthermore, it would provide insight of the behaviour of the registration algorithms when data with different post-processing methods is used.

The results of this work show that the artefacts with bigger amplitudes are harder to correct. In future works this problem can be addressed using an improved data set, for example using a larger data set through data augmentation. Besides, different techniques of model optimization

can be tested to see whether there are improvements in the results or not.

The comparison between a conventional registration algorithm and a deep-learning algorithm, approached in this work, can be further analysed with different data sets. Although this work shows that both algorithms present a good registration result, it could be interesting to see the algorithms performance in data sets with pathological images. This could lead to important conclusions on the behaviour of the algorithms for a bigger range of clinical patients, not only focused on B-scans from the retinal layers.

Finally, the importance of pre-trained algorithms for image registration is a relatively new research question. The result from this work shows that there is some improvement when using a pre-trained model. However, there are some different options that could be further explored. Future works could focus on the amount of eye phantom data needed to have a positive impact in OCT registration, the use of larger data sets could improve the final results.

Bibliography

[1] D. Huang *et al.*, "Optical coherence tomography," *Science*, vol. 254, pp. 1178–1181, Nov. 1991.

[2] C. E. Psomadakis *et al.*, "Optical coherence tomography," *Clinics in Dermatology*, vol. 39, pp. 624–634, July 2021.

[3] H. E. Bedell and S. B. Stevenson, "Eye movement testing in clinical examination," *Vision Research*, vol. 90, pp. 32–37, 2013. Testing Vision: From Laboratory Psychophysical Tests to Clinical Evaluation.

[4] M. F. Shirazi *et al.*, "Multi-modal and multi-scale clinical retinal imaging system with pupil and retinal tracking," *Scientific Reports*, vol. 12, no. 1, pp. 1–12, 2022.

[5] A. Bekkers *et al.*, "Microvascular damage assessed by optical coherence tomography angiography for glaucoma diagnosis: a systematic review of the most discriminative regions," *Acta ophthalmologica*, vol. 98, no. 6, pp. 537–558, 2020.

[6] K. Allison, D. Patel, and O. Alabi, "Epidemiology of glaucoma: The past, present, and predictions for the future," *Cureus*, vol. 12, p. e11686, Nov. 2020.

[7] A. El-Baz, "Classification of retinal diseases based on oct images," *FBL*, vol. 23, no. 2, pp. 247–264, 2018.

[8] A. Bayer and A. Akman, "Artifacts and anatomic variations in optical coherence tomography," *Turk. J. Ophthalmol.*, vol. 50, pp. 99–106, Apr. 2020.

[9] S. Ricco *et al.*, "Correcting motion artifacts in retinal spectral domain optical coherence tomography via image registration," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*, pp. 100–107, Springer Berlin Heidelberg, 2009.

[10] I. Bekerman, P. Gottlieb, and M. Vaiman, "Variations in eyeball diameters of the healthy adults," *Journal of Ophthalmology*, vol. 2014, pp. 1–5, 2014.

[11] H. Kolb, E. Fernandez, and R. Nelson, eds., *Gross Anatomy of the Eye*. Salt Lake City (UT): University of Utah Health Sciences Center.

[12] C. E. Willoughby *et al.*, "Anatomy and physiology of the human eye: effects of mucopolysaccharidoses disease on structure and function - a review," *Clinical & Experimental Ophthalmology*, vol. 38, pp. 2–11, Aug. 2010.

[13] H. Kolb, *Simple Anatomy of the Retina*. University of Utah Health Sciences Center, Salt Lake City (UT), 1995.

[14] S. Ji *et al.*, "A bio-inspired polymeric gradient refractive index (grin) human eye lens," *Opt. Express*, vol. 20, pp. 26746–26754, Nov 2012.

[15] T. H. Saey, "Seeing the light: Transforming nerve cells into photoreceptors to restore vision," *Science News*, vol. 187, pp. 22–26, May 2015.

[16] A. Cerveró, A. Casado, and J. Riancho, "Retinal changes in amyotrophic lateral sclerosis: looking at the disease through a new window," *Journal of Neurology*, vol. 268, pp. 2083–2089, Dec. 2019.

[17] G. D. Hildebrand and A. R. Fielder, "Anatomy and physiology of the retina," in *Pediatric Retina*, pp. 39–65, Springer Berlin Heidelberg, Aug. 2010.

[18] M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208, 2010.

[19] L. S. Brea *et al.*, "Review on retrospective procedures to correct retinal motion artefacts in OCT imaging," *Applied Sciences*, vol. 9, p. 2700, July 2019.

[20] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel, "The role of fixational eye movements in visual perception," *Nature Reviews Neuroscience*, vol. 5, pp. 229–240, Mar. 2004.

[21] R. G. Alexander, S. L. Macknik, and S. Martinez-Conde, "Microsaccade characteristics in neurological and ophthalmic disease," *Frontiers in Neurology*, vol. 9, Mar. 2018.

[22] D. Purves *et al.*, eds., *Neuroscience*. Sinauer Associates, 6 ed., Jan. 2018.

[23] S. Aumann *et al.*, *Optical coherence tomography (OCT): Principle and technical realization*. Springer International Publishing, 2019.

[24] F. Prati *et al.*, "Expert review document on methodology, terminology, and clinical applications of optical coherence tomography: physical principles, methodology of image acquisition, and clinical application for assessment of coronary arteries and atherosclerosis," *European Heart Journal*, vol. 31, pp. 401–415, 11 2009.

[25] D. P. Popescu *et al.*, "Optical coherence tomography: fundamental principles, instrumental designs and biomedical applications," *Biophysical Reviews*, vol. 3, pp. 155–169, Aug. 2011.

[26] A. Akman, A. Bayer, and K. Nouri-Mahdavi, eds., *Optical Coherence Tomography in Glaucoma*. Springer International Publishing, 2018.

[27] R. J. Zawadzki *et al.*, "Adaptive optics-optical coherence tomography: optimizing visualization of microscopic retinal structures in three dimensions," *Journal of the Optical Society of America A*, vol. 24, p. 1373, Apr. 2007.

[28] G. Wildeman and H. Dhalla, "What is oct and how can it help ophthalmologists acquire high resolution information on ocular tissue?," Dec 2016. Available at `https://www.leica-microsystems.com/science-lab/what-is-oct-and-how-can-it-help-ophthalmologists-acquire-high-resolution-information-on-ocular-tissue/`.

[29] I. Wong *et al.*, "En-face optical coherence tomography in the diagnosis and management of age-related macular degeneration and polypoidal choroidal vasculopathy," *Indian Journal of Ophthalmology*, vol. 63, no. 5, p. 378, 2015.

[30] B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," *Journal of Biomedical Optics*, vol. 11, no. 4, p. 041102, 2006.

[31] J. Baxi *et al.*, "Retina-simulating phantom for optical coherence tomography," *Journal of Biomedical Optics*, vol. 19, p. 021106, Sept. 2013.

[32] H.-J. Lee *et al.*, "Retina phantom for the evaluation of optical coherence tomography angiography based on microfluidic channels," *Biomedical Optics Express*, vol. 10, p. 5535, Oct. 2019.

[33] G. C. F. Lee *et al.*, "Fabrication of healthy and disease-mimicking retinal phantoms with tapered foveal pits for optical coherence tomography," *Journal of Biomedical Optics*, vol. 20, p. 085004, Aug. 2015.

[34] S. Tanaka *et al.*, "Quantitative assessment of manual and robotic microcannulation for eye surgery using new eye model," *Int. J. Med. Robot.*, vol. 11, pp. 210–217, June 2015.

[35] G. Fogli *et al.*, "New eye phantom for ophthalmic surgery," *J. Biomed. Opt.*, vol. 19, p. 068001, June 2014.

[36] U. Sukowski *et al.*, "Preparation of solid phantoms with defined scattering and absorption properties for optical tomography," *Phys. Med. Biol.*, vol. 41, pp. 1823–1844, Sept. 1996.

[37] P. Lemaillet and J. C. Ramella-Roman, "Dynamic eye phantom for retinal oximetry measurements," *J. Biomed. Opt.*, vol. 14, p. 064008, Nov. 2009.

[38] B. Braaf *et al.*, "A neural network approach to quantify blood flow from retinal OCT intensity time-series measurements," *Scientific Reports*, vol. 10, June 2020.

[39] H. Wang *et al.*, "Model eye tool for retinal optical coherence tomography instrument calibration," *Journal of Innovative Optical Health Sciences*, vol. 14, p. 2150010, Apr. 2021.

[40] L. Pan *et al.*, "OCTRexpert: A feature-based 3d registration method for retinal OCT images," *IEEE Transactions on Image Processing*, vol. 29, pp. 3885–3897, 2020.

[41] J. Cheng *et al.*, "Motion correction in optical coherence tomography for multi-modality retinal image registration," in *Proceedings of the Ophthalmic Medical Image Analysis Third International Workshop*, University of Iowa, Oct. 2016.

[42] Y. Chen *et al.*, "Three-dimensional eye motion correction by lissajous scan optical coherence tomography," *Biomedical Optics Express*, vol. 8, p. 1783, Feb. 2017.

[43] M. F. Kraus *et al.*, "Quantitative 3d-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization," *Biomedical Optics Express*, vol. 5, p. 2591, July 2014.

[44] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," *IEEE Transactions on Image Processing*, vol. 9, no. 2, pp. 287–290, 2000.

[45] K. Marstal *et al.*, "SimpleElastix: A user-friendly, multi-lingual library for medical image registration," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, June 2016.

[46] S. Klein *et al.*, "elastix: A toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196–205, 2010.

[47] L. S. Brea *et al.*, "Motion correction in 3d-oct data by intensity-based image registration: an evaluation study," *Investigative Ophthalmology & Visual Science*, vol. 60, no. 11, pp. PB0177–PB0177, 2019.

[48] K. Ntatsis *et al.*, "Motion correction in retinal optical coherence tomography imaging using deep learning registration," in *Medical Imaging 2022: Image Processing* (I. Išgum and O. Colliot, eds.), SPIE, Apr. 2022.

[49] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," *CoRR*, vol. abs/1608.06993, 2016.

[50] S. Periaswamy and H. Farid, "Elastic registration in the presence of intensity variations," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 865–874, July 2003.

[51] "Inventor software: Get prices & buy official inventor 2023," 2022. Available at https://www.autodesk.com/products/inventor/overview.

[52] "Ultimaker cura: Powerful, easy-to-use 3d printing software." Available at https://ultimaker.com/software/ultimaker-cura.

[53] S. Farah, D. G. Anderson, and R. Langer, "Physical and mechanical properties of PLA, and their functions in widespread applications — a comprehensive review," *Advanced Drug Delivery Reviews*, vol. 107, pp. 367–392, Dec. 2016.

[54] "Ultimaker pla material: Highly versatile, easy to print." Available at `https://ultimaker.com/materials/pla`.

[55] S. Ji *et al.*, "A bio-inspired polymeric gradient refractive index (GRIN) human eye lens," *Optics Express*, vol. 20, p. 26746, Nov. 2012.

[56] H. Gross, F. Blechinger, and B. Achtner, *Survey of optical instruments*. Wiley-VCH Verlag GmbH & Co. KGaA, 4th edition ed., 2008.

[57] "La1074-b n-bk7 plano-convex lens, ø1/2", f = 20 mm, ar coating: 650 - 1050 nm." Available at `https://www.thorlabs.com/thorproduct.cfm?partnumber=LA1074-B`.

[58] G. J. Greening *et al.*, "Characterization of thin poly(dimethylsiloxane)-based tissue-simulating phantoms with tunable reduced scattering and absorption coefficients at visible and near-infrared wavelengths," *Journal of Biomedical Optics*, vol. 19, p. 115002, nov 2014.

[59] A. Batista *et al.*, "Swept-source phase-stabilized optical coherence tomography setup for elastography,"

[60] T. T. Arduino, "Motor shield rev3 : Arduino documentation." Available at `https://docs.arduino.cc/hardware/motor-shield-rev3`.

[61] "Stepper - Arduino Reference." Available at `https://www.arduino.cc/reference/en/libraries/stepper/`.

[62] "Zaber Technologies." Available at `https://www.zaber.com/products/linear-actuators/T-LA/specs?part=T-LA28A`.

[63] "Python - zaber motion library (ascii)." Available at `https://www.zaber.com/software/docs/motion-library/ascii/references/python/`.

[64] F. Møller *et al.*, "Binocular quantification and characterization of microsaccades," *Arbeitsphysiologie*, vol. 240, pp. 765–770, Sept. 2002.

[65] B. C. Lowekamp *et al.*, "The design of SimpleITK," *Frontiers in Neuroinformatics*, vol. 7, 2013.

[66] D. Soydaner, "A comparison of optimization algorithms for deep learning," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, p. 2052013, Apr. 2020.

[67] S. E. Whang *et al.*, "Data collection and quality challenges in deep learning: A data-centric ai perspective," 2021.

[68] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, May 2016.

[69] K. Kavitha, B. Sandhya, and B. Thirumala, "Evaluation of distance measures for feature based image registration using AlexNet," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, 2018.

[70] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[71] A. M. Bagci *et al.*, "Thickness profiles of retinal layers by optical coherence tomography image segmentation," *American Journal of Ophthalmology*, vol. 146, pp. 679–687.e1, Nov. 2008.

[72] Q. Wang *et al.*, "Thickness of individual layers at the macula and associated factors: the beijing eye study 2011," *BMC Ophthalmology*, vol. 20, Feb. 2020.

[73] P. P. Srinivasan *et al.*, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomedical Optics Express*, vol. 5, p. 3568, Sept. 2014.

[74] N. R. Raia *et al.*, "Characterization of silk-hyaluronic acid composite hydrogels towards vitreous humor substitutes," *Biomaterials*, vol. 233, p. 119729, Mar. 2020.

[75] J. Schnekenburger *et al.*, "Refractive index properties of the retina accessed by multi-wavelength digital holographic microscopy," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXVI* (T. G. Brown and T. Wilson, eds.), SPIE, Feb. 2019.