Optical coherence tomography and deep learning for ophthalmology

by

Morgan Lindsay Heisler

MASc, Simon Fraser University, 2017

BASc (Hons.), Simon Fraser University, 2015

Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in the

School of Engineering Science

Faculty of Applied Sciences

SIMON FRASER UNIVERSITY

Spring 2020

Approval

Name:	**Morgan Lindsay Heisler**
Degree:	**Doctor of Philosophy**
Title:	**Optical coherence tomography and deep learning for ophthalmology**
Examining Committee:	**Chair:** Bonnie Gray Professor
	Marinko V. Sarunic Senior Supervisor Professor
	Mirza Faisal Beg Supervisor Professor
	Eduardo V. Navajas Supervisor Clinical Assistant Professor
	Pierre Lane Internal Examiner Associate Professor of Professional Practice
	Yali Jia External Examiner Associate Professor of Ophthalmology and Biomedical Engineering School of Medicine Oregon Health & Science University
Date Defended/Approved:	April 6, 2020

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

Robust quantitative tools require large data sets for testing efficacy and accuracy, which is especially true when using machine learning and neural networks. However, large datasets with corresponding manual annotations are uncommon with state-of-the-art imaging systems, particularly in the medical field. Ophthalmology is one such field, for which recent imaging advances allow clinicians to use multiple imaging modalities to diagnose and monitor patients. Optical coherence tomography (OCT) has become an integral imaging modality in ophthalmic clinics due to its non-invasive nature and ability to acquire micrometer scale sub-surface images of ophthalmic tissue. In this thesis, several different techniques to mitigate the need for large annotated datasets when translating machine learning tools to an ophthalmic clinic are evaluated. First, the concept of transfer learning is assessed through fine-tuning networks trained on a different domain (adaptive optics scanning laser ophthalmoscopy) to the domain of interest (adaptive optics OCT) to detect cone photoreceptors. Second, both adversarial and semi-supervised learning are investigated which allow for unlabelled data to be used in the training process. Finally, the more challenging task of diagnostics with limited data was investigated using diabetic retinopathy OCT Angiography data and an ensemble of networks. Through these investigations, the utility of transfer learning, adversarial and semi-supervised learning, and ensembling is shown for small ophthalmic datasets.

Keywords: optical coherence tomography; ophthalmology; image processing; angiography; machine learning; deep neural networks

To my parents for their endless love, support, and encouragement.

Acknowledgements

I would like to express my deepest gratitude to my senior supervisor, Dr. Marinko V. Sarunic, whose expertise, understanding, and patience added considerably to my graduate experience. I appreciate his confidence and trust in my abilities while pushing me outside my academic comfort zone. I would also like to sincerely thank Dr. Mirza Faisal Beg for all of his expertise in medical image processing which was crucial to the completion of this research. His passion and enthusiasm for the research made every interaction an enjoyable experience. I would also like to acknowledge Dr. Eduardo V. Navajas, who shared invaluable medical expertise and direction for these projects.

I am very grateful to be part of the Biomedical Optics Research Group (BORG) at SFU, and to have worked with so many great students both past and present. I am especially grateful to Dr. MyeongJin Ju, Dr. Sieun Lee and Dr. Yifan Jian for their mentorship over the years. They have demonstrated some of the traits that I can only strive for in future endeavours: Dr. Ju's dedicated work ethic, Dr. Lee's eye for detail, and Dr. Jian's willingness to make the work environment as enjoyable as possible. Additionally, I have had the opportunity to mentor so many great students, and would like to thank Mr. Arman Athwal, Mr. Francis Tran and Mr. Julian Lo especially for teaching me as much as I taught them.

Lastly, I would like to thank my family for their endless support and encouragement in all my endeavours. I couldn't have done it without you.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AO	Adaptive Optics
AO-OCT	Adaptive Optics Optical Coherence Tomography
AO-SLO	Adaptive Optics Scanning Laser Ophthalmoscopy
BMO	Bruch's Membrane Opening
CAM	Class Activation Map
CNN	Convolutional Neural Network
CRA	Central Retinal Artery
DCP	Deep Capillary Plexus
DNN	Deep Neural Network
DVC	Deep Vascular Complex
DR	Diabetic Retinopathy
ETDRS	Early Treatment Diabetic Retinopathy Study
FA	Fluorescein Angiography
FAZ	Foveal Avascular Zone
FOV	Field of View
GAN	Generative Adversarial Network
GAP	Global Average Pooling
ICP	Inner Capillary Plexus
ILM	Inner Limiting Membrane
INL	Inner Nuclear Layer
IOP	Intra-Ocular Pressure
IPL	Inner Plexiform Layer
OCT	Optical Coherence Tomography
OCTA	Optical Coherence Tomography Angiography
ONH	Optic Nerve Head
ONL	Outer Nuclear Layer
OPL	Outer Plexiform Layer
RGC	Retinal Ganglion Cell
RNFL	Retinal Nerve Fiber Layer
RPC	Radial Peripapillary Capillaries
SAO	Sensorless Adaptive Optics
SCP	Superficial Capillary Plexus

SLO	Scanning Laser Ophthalmoscopy
SVP	Superficial Vascular Plexus
SSADA	Split-spectrum amplitude-decorrelation angiography
SVM	Support Vector Machine
U-Net	U Network
VF	Visual Field

Chapter 1.

Introduction

1.1. Overview

The current renaissance of artificial intelligence, and in particular machine learning, is driving great progress and innovation across many fields. We now live in an age where computer memory is inexpensive and computer processing speeds are continually progressing. Combined with new algorithms and easier access to vast amounts of digitized data, these factors have allowed machine learning to be integrated into daily life – from finding the best route home, to sorting pictures using facial recognition, and even song recommendations tailored to individual taste. Machine learning is even being applied to medical applications.

Many areas within the healthcare industry are now of great interest to those researching the power of machine learning: from diagnosis and prognosis to drug development and epidemiology. There is significant interest in using this technology to reduce costs, improve screening in remote areas without medical specialists and answer questions to better understand diseases. With significant potential to transform the medical landscape, one area where machine learning is gaining particular traction is ophthalmology.

Vision is one of the five senses, and is one that many people take for granted. Two of the most common eye diseases that cause blindness are diabetic retinopathy and glaucoma. Simulated examples of the impacts on vision from these diseases are shown in Figure 1.1. Diabetic Retinopathy (DR) causes a partial blurring or patchy loss of vision as shown in Figure 1.1B. Patients with glaucoma on the other hand, may experience loss of peripheral vision, called tunnel vision, as depicted in Figure 1.1C. As our population ages, the number of people affected by these diseases is expected to rise significantly. The National Eye Institute estimates that the number of people who will have DR will nearly double from 7.7 million to 14.6 million and the number of people with glaucoma will more than double from 2.7 million to 6.3 million from 2010 to 2050 in the United States [1]. This increases the need for better diagnostic tools in which to enable

ophthalmologists to detect these diseases earlier and with better confidence. A brief overview of retinal anatomy and a more detailed description of DR and glaucoma is provided in Chapter 2.1.

Figure 1.11 Representative images of the Indian Ocean seen by someone with A) normal vision, B) diabetic retinopathy, and C) glaucoma.

Over the past two decades, the development of visible and near-infrared retinal imaging technology has grown rapidly. One of the dominant imaging modalities is Optical Coherence Tomography (OCT), which has revolutionized clinical diagnostic ophthalmic imaging. OCT provides a detailed volumetric view of the retina for clinicians to identify the structural hallmarks of diseases such as DR and glaucoma. OCT images are used to assess the need for treatments (surgical, intravitreal injection, laser, etc.) and afterwards to evaluate the results and monitor changes. Additionally, imaging blood flow is also important as abnormal circulation is the leading cause of irreversible blindness in diseases such as DR. A recently commercialized variant of OCT, OCT Angiography (OCTA), allows for detailed visualization of the retinal microvasculature and is becoming more prevalent in clinics worldwide. Additionally, adaptive optics OCT, which provides better lateral resolution to better visualize cellular structures such as the photoreceptors, is also gaining interest from ophthalmologists for its ability to visualize structures *in vivo* which were previously only available in ex vivo tissue through histology. An overview of OCT, OCTA and AO-OCT is provided in Chapter 2.2. However, as OCT images are inherently only qualitative, there is an unmet need amongst ophthalmic clinicians for automatic quantitative analysis.

Robust quantitative tools require large data sets for testing their efficacy – this is especially true for machine learning tools. However, large datasets with manual segmentations are uncommon with new and state-of-the-art imaging systems, such as the ones developed in BORG. The small data set size of clinical prototypes presents a

unique challenge when translating to a clinical setting. As such, the topic of this thesis is focused on translational image analysis for newly developed OCT technologies, using machine learning.

1.2. Machine Learning in Ophthalmology

The unique properties of the eye make it suitable for non-invasive optical imaging; however, quantification of pathological features is generally requested by the clinicians for diagnosis and monitoring. Quantification and analysis of different morphometric properties generally requires segmentation of the features of interest, such as the individual light sensitive cells (cone photoreceptors) or the retinal microvasculature (capillaries). As manual segmentations are both subjective and laborious, automated methods are desired. Although traditional mathematical methods work on the specific data for which they were designed, reliance on *ad hoc* rules and specific algorithmic parameters does not allow for alternative imaging conditions, such as different resolutions, areas within the retina, and imaging modalities.

A new alternative approach to medical image segmentation and analysis is using machine learning where a model is learned directly from data points. While traditional statistical approaches such as regression are forms of machine learning, they require the manual extraction of features of interest. Deep neural networks (DNNs) have emerged as an excellent machine learning alternative where the network is able to serve as both a feature extractor and classifier. A more detailed background of DNNs is presented in Chapter 2.3. This allows for a higher degree of adaptability as the same machine learning algorithm can be re-purposed by using different training images given a sufficiently large training set. Several of these networks have shown high performance for many different image analysis tasks, including ophthalmic applications [2]. For example, DNNs have been used in retinal fundus images for the segmentation of retinal blood vessels [3], and in OCT images for the segmentation of retinal layers [4], [5] and microvasculature [6] or lack thereof [7]. Multiple review articles on deep learning in ophthalmology are available in the Literature, for example [2], [8], [9].

Machine Learning is also a useful tool for classifications, and its primary use in the medical field is for diagnostic purposes. Where machine learning diagnostics could be of potential benefit in ophthalmology is in screening where traditional approaches

require large amounts of manual labour and place a huge financial burden on the healthcare systems. Deep learning systems for diagnosing diabetic retinopathy [10]–[13], glaucoma [10], [14], age-related macular degeneration [10], [15], [16] and retinopathy of prematurity [17] have been created that show excellent diagnostic performance for the imaging modality on which they were trained. [2]

Massive strides in making AI tools available to consumers and clinicians have been made in the past couple years. For example, the IDx-DR, which received marketing authorization by the U.S. Food and Drug Administration a little over a year ago, is an autonomous AI diagnostic system for diabetic retinopathy [18]. It is the first device of its kind in ophthalmology, or any field of medicine, and paved the way for future AI-based medical technologies to get FDA approval. As such, this is an exciting time to be working in the cross-section of AI and ophthalmology as this new industry is still in its infancy and there are many unmet needs on which to work.

1.3. Outline

The remaining chapters of this thesis are organized as follows. Chapter 2 presents background information on topics discussed in this thesis including: the anatomy of the human eye and on the diseases of interest, the imaging modalities used, and machine learning techniques and methods used for overcoming small datasets. Chapter 3 details the use of transfer learning for retraining a DNN originally trained on confocal AO-SLO images in order to detect cones in AO-OCT images with a different field-of-view (FOV). This chapter demonstrates the ability of transfer learning to effectively apply knowledge from one domain, with a larger dataset to a similar domain where manually labeled data is scarce. Chapter 4 demonstrates the use of a semi-supervised Generative Adversarial Network (GAN) to segment the peripapillary tissue and optic nerve head (ONH). In this chapter, we also provide a method for automatically segmenting the termination of the Bruch's membrane, an important landmark in the parameterization of the ONH. In Chapter 5, the focus shifts from segmentation to classification and ensemble-based methods are compared for classifying DR with a relatively small dataset. Finally, Chapter 6 discusses future research directions to advance the analysis of human ophthalmic imaging.

1.4. Contributions

At the early stage of my research career, my work was mainly focused on clinical data acquisition of retinal images using prototype OCT instruments and developing image processing algorithms using traditional techniques to visualize and quantify the data. As our prototype OCT systems were situated in a glaucoma clinic, much of the work I did focused around this topic [A2, A7-8, A10, A18]. This is where I was able to integrate myself into the clinic to truly understand how our prototype machines either met needs that the commercial ones at the Eye Care Center (ECC) did not have (ex. the 1060nm system was better able to see the choroid [A8, A10]) or fill a need for a specialized commercial machine which the ECC did not have (ex. the anterior segment system [A7] or OCTA capabilities [A2]). I published a case series on how our novel swept-course anterior segment OCT system allowed for visualization of focal fluid in overfiltering blebs, leading to better transconjunctival suture placement and patient outcomes [A7]. Additionally, and more importantly, this experience allowed me to see what image processing techniques were still needed to further meet the needs of the clinicians, which fueled my research in OCTA processing.

Through my work in the glaucoma clinic and mentorship from glaucoma specialist Dr. Paul Mackenzie, I was also exposed to the new and exciting research domain of OCTA. I worked with a prototype OCT instrument to acquire OCTA images in control and patient populations, and led the technical image analysis of ONH OCTA images looking closely at the radial peripapillary capillaries (RPCs), the results of which were published in a journal paper [A2]. Also using the early OCTA instrument, I performed data acquisition myself, as well as trained and managed the clinical data acquisition and processing of OCTA data for control subjects [A1], and other diseases [A4, A16, A20]. This research was extended to an ex-vivo retinal imaging [A13, A15] application under the supervision of Dr. Dao-Yi Yu at the University of Western Australia.

By working closely with the clinicians mentioned above, as well as retina specialist Dr. Eduardo Navajas, two needs became apparent: 1) better quality *en face* OCTA angiograms, and 2) a more automated process for quantifying the data was needed for accurate clinical diagnosis. My main first-authored paper from my Master's degree [A11] described and demonstrated the development of an automated method for registration and averaging of serially acquired OCTA images. The improved visualization

of the capillaries with this tool is continuing to be investigated to enable robust quantification and study of minute changes in retinal microvasculature in the future. Also at that time, through working with a PhD exchange student, a process for segmenting the retinal vasculature in our OCTA data using Deep Neural Networks (DNNs) was developed which led to a co-first authored paper [A3]. This was my first introduction to machine learning and DNNs.

In the transition to my PhD research, my focus also transitioned to using DNNs to aid in our high resolution imaging with adaptive optics (AO) integrated OCT. I made a significant contribution to the translation of the AO-OCT prototype instrument to the clinic. In addition to running and managing the training of research assistants for data acquisition, I was actively involved in the processing of images and reports of the patient images [A5, A6, A9]. Through this experience, I was motivated to start the process of developing tools for quantifying the photoreceptors in the AO-OCT images. Due to the novel nature of the prototype AO-OCT instrument, there did not exist a large database from which to train a network. As such, I investigated a technique called transfer learning to mitigate this issue. A dataset from the Literature consisting of images from a different imaging modality, AO Scanning Laser Ophthalmoscopy (AOSLO), was used to initially train a neural network to segment cones. Then, specific sections of the network (also known as layers) were re-trained using our limited amount of AO-OCT training data [A12]. The details of this project and its results are provided in Chapter 3

Through this work, a need to further investigate different techniques to mitigate the need for large datasets became evident. A problem facing the machine learning community is the need for a vast amount of data for supervised training. As such, I investigated two methods as potential solutions to this issue: transfer learning (as discussed in the previous paragraph) and semi-supervised generative adversarial networks. For the above-mentioned transfer learning project with the AO-OCT instrument, although few labeled images were available, a larger data set from a similar imaging technique was publically available making transfer learning a good choice. For the semi-supervised GAN project however, we possessed a large amount of data from my aforementioned work in the glaucoma clinic, but there were few manual segmentations for training a neural network. This is an especially common situation with OCT datasets where commercial devices have been acquiring data from patients for years, but there aren’t enough human resources to go through the data to provide

professional level segmentations or classification. A more detailed account of the proposed research, targeting semi-supervised learning for the analysis of OCT images used in glaucomatous patients, is provided in Chapter 4.

The diagnostic capability of OCTA as it relates to DR in the Literature has thus far focused mainly on manually created parameters based on a priori knowledge of the disease pathophysiology. One advantage of deep learning is that it allows the machine to extract features for classification, and therefore identify potential information from the images that may be discarded otherwise. I evaluated the role of ensemble learning techniques with deep learning in classifying diabetic retinopathy in OCTA *en face* images and their corresponding co-registered structural images, described in more detail in Chapter 5.

Although the main contributions as detailed in this thesis pertain to the three journal publications that are described in Chapters 3-5, the scope and body of work done during my graduate career exceeds this. I am first author or co-first author on 5 papers, and a co-author on 17 papers. During my graduate studies I have given 5 podium presentations at international conferences. All peer-reviewed journal papers on which I am an author are listed in Appendix 1 [A1-A22]. These publications represent my contributions to the development and clinical translation of novel imaging instruments, the acquisition and analysis of images from patients groups, and ultimately to the deep learning research which is the focus of this dissertation.

Chapter 2. Background[1]

2.1. Ophthalmic Diseases

2.1.1. Eye Anatomy

The eye is a complex organ that allows us to perceive and convert light to electrical signals that the brain can interpret. Figure 2.1 presents a simple schematic of a human eye.

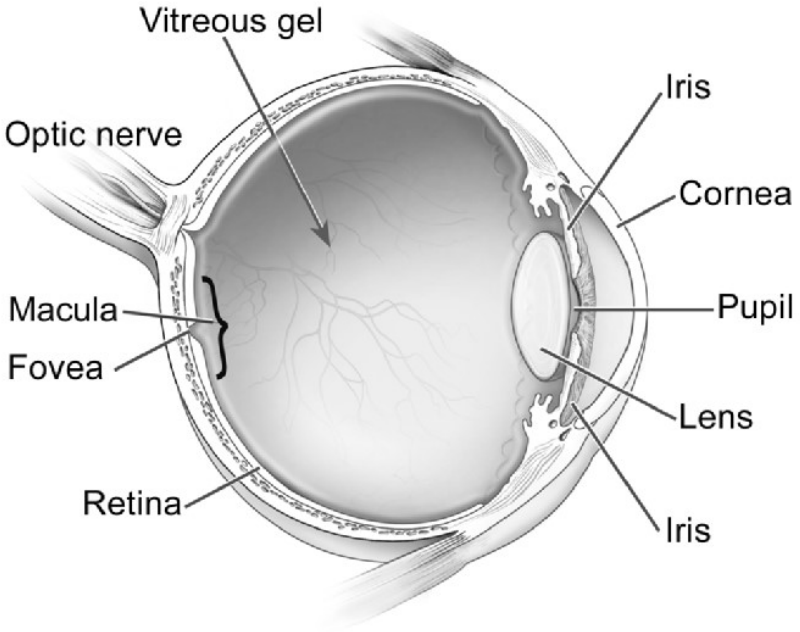

Figure 2.1 Eye diagram [credit: National Eye Institute, National Institutes of Health]

Briefly, light incident on the eye is focused by the cornea and lens onto the retina. The eye is roughly 25mm in diameter, and the retina, located at the back of the eye contains cell layers that detect light, perform some processing on the information, and

[1] This section has been modified from my previously published Master's thesis [Morgan Heisler, "Clinical optical coherence tomography angiography registration and analysis," April, 2017.]

transmit electrical signals to the brain via the neurons of the optic nerve. The region where the nerve fibre bundles exit the eye is the optic nerve head (ONH) and has significance for glaucoma as mentioned in Section 2.1.3, along with the surrounding peripapillary region. The rest of the report will focus mainly on the retina, a cross sectional diagram of which is shown in Figure 2.2.

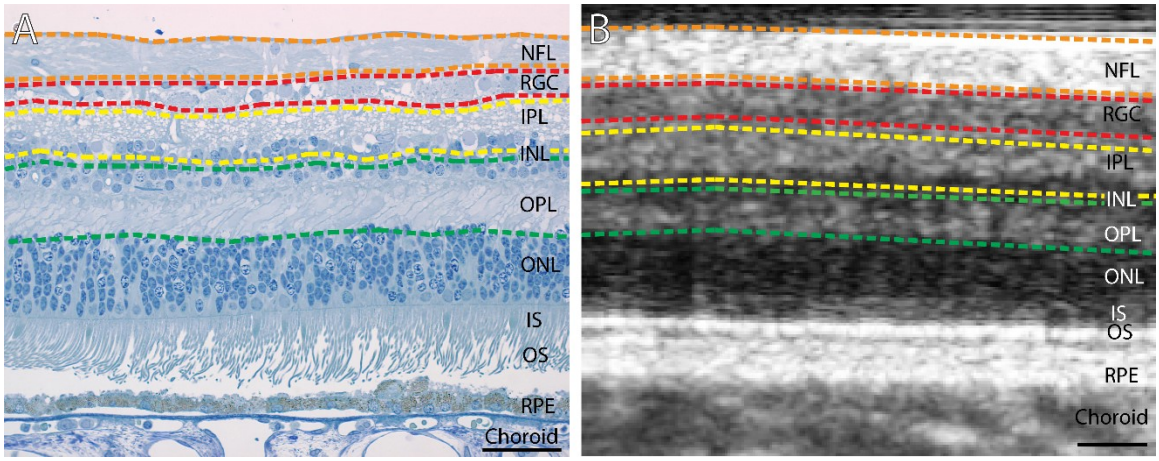

Figure 2.2 **A portion of the human retina. Transverse histological retinal section was stained with toluidine blue (A), and a B-scan image was acquired using OCT (B) to illustrate the various retinal layers at the eccentricity located 3 mm superior to the optic disk. Colored dashed lines demarcate the retinal vascular layers. *Orange dashed lines* indicate the RPC layer; *red dashed lines*, SVP layer; *yellow dashed lines*, ICP layer; *green dashed lines*, DCP. *Scale bar*: 50 µm. Image from** [19]**.**

There are two sources of blood supply to the human retina: the central retinal artery (CRA) and the choroidal blood vessels. The inner retina is bounded by the inner limiting membrane (ILM) anteriorly and the inner nuclear layer (INL) posteriorly, and it gets nourished by blood that travels through the CRA from the optic nerve head. The choroidal blood vessels feed the outer retina, particularly the photoreceptors, which are arranged in an approximately hexagonal grid in the outer retina. Chapter 3 of this thesis investigates the analysis of cone photoreceptors.

The retinal vasculature emerging from the CRA follows the patterns as shown in the retinal fluorescein angiography (FA) image in Figure 2.3, where the vessels radiate outward from the ONH and curve towards and around the fovea. These vessels supply four layers of capillary networks which are the radial peripapillary capillaries (RPCs), the superficial vascular plexus (SVP), intermediate capillary plexus (ICP), and deep capillary

plexus (DCP) [20]. The RPCs are the most superficial layer of capillaries lying in the inner part of nerve fiber layer (NFL), and feed the superficial nerve fibres surrounding the ONH. The SVP, which is supplied by the central retinal artery and composed of various vessel types resides in the retinal ganglion cell (RGC) layer. Below this is the ICP, which is supplied by vertical anastomoses from the SVP and contains the capillaries encapsulated by the anterior boundary of the IPL and anterior half of the INL. Lastly, the DCP is also supplied via vertical anastomoses of the SVP and is considered to be the vessels contained within the posterior half of the INL and the OPL. These layers are shown in cross-sectional images in Figure 2.2. For visualization purposes, these networks are sometimes summed together to form the superficial vascular complex (SVC), comprised of the SVP and RPC, and the deep vascular complex (DVC), comprised of the ICP and DCP [20].

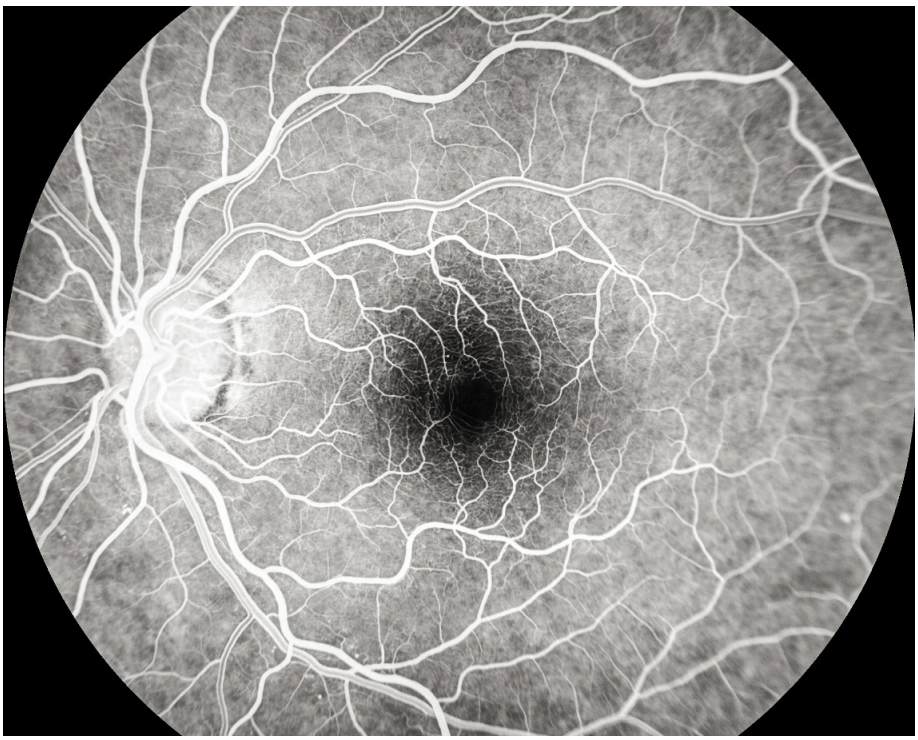

Figure 2.3 Fundus photograph showing fluorescein imaging of the major arteries and veins in a human left eye.

2.1.2. Diabetic Retinopathy

Diabetic retinopathy (DR) is the most prevalent retinal vascular diseases worldwide, affecting a third of people with diabetes [21]. It is a leading cause of adult blindness, responsible for 15-17% of cases of blindness in the western world [22]. The

pathophysiology of DR is closely related to its deleterious effect on the inner retinal microcirculation which includes altered vascular permeability and capillary bed closure [23], [24]. Diabetic retinopathy has many different categorization schemes, however one of the most clinically useful and commonly used is the 5 stage International Clinical Diabetic Retinopathy categorization of no retinopathy, mild, moderate, severe non-proliferative retinopathy, and proliferative retinopathy [25]. Retinal ischemia secondary to capillary non-perfusion has been observed in the early stages of diabetic retinopathy and has been correlated to disease severity and progression [26]. Findings such as decreasing macular capillary density and enlargement of the perifoveal zone [27] have been correlated with the severity of vision loss [28], [29]. Thus, early detection and reliable quantification of these microvascular changes may play a role in predicting visual morbidity and improve the management of DR.

2.1.3. Glaucoma

Glaucoma is a group of optic neuropathies that are the leading cause of irreversible blindness worldwide [30] and the second most common cause of blindness in the developed world [31]. Symptoms include the gradual loss of peripheral vision followed by the loss of central vision if the disease is not detected and properly managed. As the peripheral vision loss is gradual and often imperceptible by patients until later stages of the disease, permanent damage can occur if proper screening for early detection is not done. The pathophysiology of glaucoma is complex and characterized by the time-dependent loss of retinal ganglion cells (RGCs) and their accompanying axons [32]. Indices that are currently used to quantify and evaluate progression of glaucomatous optic neuropathy include visual field testing, nerve fibre layer thickness, ganglion cell layer with inner plexiform layer (GCIPL) and ganglion cell complex parameter analysis, optic nerve head morphometrics such as Bruch's membrane opening minimum rim width, and measurement of lamina cribrosa thickness [33].

2.1.4. Summary

Vision threatening diseases, such as diabetic retinopathy and glaucoma, have characteristic changes that can be distinguished using ophthalmic imaging techniques. In the next section, we discuss some of these techniques.

2.2. Retinal Imaging Systems

Retinal imaging has become a standard of care in ophthalmic clinics. While different techniques may be used depending on the pathology, optical coherence tomography has become a gold-standard for many ophthalmologists. Variants of this imaging modality, such as adaptive optics OCT and OCT angiography, have been able to provide clinicians with even more capabilities. A brief overview of the ophthalmic imaging techniques used in this thesis is provided here.

2.2.1. Optical Coherence Tomography

Optical Coherence Tomography (OCT) is a non invasive, *in vivo* imaging method that uses a low coherence interferometric technique to produce cross-sectional and volumetric images. The cross-sectional images, termed B-scans, are comprised of several adjacent depth profiles (A-scans). Several adjacent B-scans can be acquired to form a 3D volume of the tissue being imaged. Although the penetration depth is on the order of millimeters in tissue, the axial resolution is typically between 1-10 microns.

Swept source OCT has been gaining in popularity recently due to the increase in scanning, and therefore acquisition, speeds. These systems use a swept source laser typically centered ~1 μm to sweep across different wavelengths and a photodiode balanced detector to detect the interference pattern caused by the path mismatch of the backreflected light from the reference and sample arms. Through calculating the Fourier transform of the interference fringes, the axial location of the light scattered from the tissue can be determined. All OCT systems used in this thesis are swept source.

2.2.2. Adaptive Optics OCT

When imaging the retina, light must pass through the cornea and lens of the subject, which can be afflicted with imperfections resulting in aberrations. These aberrations are different for each patient and must be corrected for cellular imaging. One such way to do this is Adaptive Optics (AO), whereby an adaptive element in the optical imaging system can adjust the light to compensate for imperfections and enable higher

lateral resolution. This technique has enabled the imaging of the living retina at a cellular resolution, something which was previously only possible using *ex vivo* methods. Recently, AO has been combined with OCT to enable high resolution cross-sectional imaging of the human eye. For example, AO-OCT systems can provide high-resolution *en face* images of the photoreceptor mosaic, but with the added benefit of high axial resolution which allows for cross-sectional tomography images in which the different outer retinal layers can be clearly delineated [34]. The technique used in our lab, termed sensorless AO-OCTA, allows for image-based aberration correction and does not require the use of a Shack–Hartmann wavefront sensor, which is used in more traditional AO techniques.

2.2.3. Optical Coherence Tomography Angiography

For the past 30 years, fluorescein angiography has been the gold standard modality for assessing retinal vascular diseases [17] but vessel leakage and excessive choroidal fluorescence affects the ability of FA to visualize the retinal microcirculation. FA is an invasive procedure that requires venipuncture and the administration of exogenous contrast agents. The injected dye can cause nausea, vomiting, skin discolouration, pruiritis and in rare cases death and anaphylaxis [46]. Moreover, these techniques only provide 2D information (*en face* views of the vasculature). Therefore, there is a clinical demand for a non-invasive approach to provide visualization of the microvasculature within the retina layers.

Optical Coherence Tomography Angiography is an emerging imaging modality with which the retinal circulation can be visualized by computing the OCT speckle difference between adjacent B-scans on a pixel-by-pixel basis. Variants of OCTA methods have been described in recent review articles [35]–[38]. The speckle variance approach to OCTA has been evaluated against standard invasive techniques such as Fluorescein Angiography [39], [40], in which only the superficial capillaries can be distinguished due to excessive choroidal fluorescence[41], and *ex vivo* histological analyses [19], [40], [42], [43].

2.2.4. Summary

This section presented a background to OCT, and the adaptive optics and angiography variants that are used in this thesis. While the imaging techniques mentioned in this section enable high resolution visualization of the retinal layers and vasculature, they are qualitative without further analysis. The next section describes machine learning methods in the Literature for quantitative analysis, while the remaining chapters utilize these topics and provide additional depth.

2.3. Machine Learning Techniques

The imaging modalities mentioned in the previous section allow for detailed visualization of the retina, however quantification is generally requested by the clinicians for diagnosis and monitoring. Machine learning has become an extremely powerful tool for the quantification of biomedical images [2], [8], [9] and will be used in this work. Therefore, a brief background is provided here.

2.3.1. Machine Learning Versus Deep Learning

Machine learning is a subset of artificial intelligence where systems use statistical models to improve their performance on a specific task. The algorithm uses input data so that it automatically learns and improves from experience without being explicitly programmed. Many machine learning algorithms rely on manually extracted features for training, such as support vector machines (SVMs) and random forest classifiers.

Deep learning is a subset of machine learning in which a model learns to perform classification tasks directly from the input data (images, video, text, or sound). A DNN is a neural network with multiple intermediate layers, often termed ‘hidden’ layers, between the input and output layers. There are typically many hidden layers in a deep neural network, hence the word “deep”, which allow for gradually more abstract and higher-level representation within the network [9]. These models can achieve state-of-the-art accuracy and have been shown to outperform classic image processing methods. DNNs are trained by using large sets of labeled data, as performance improves as the dataset size increases, and neural network architectures that learn features directly from the

data without the need for manual feature extraction. The differences are shown below in Figure 2.4.

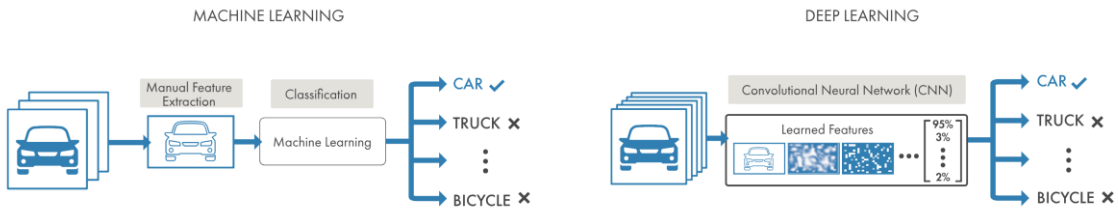

Figure 2.4 Comparing a machine learning approach to categorizing vehicles (left) with deep learning (right). Image from [44]**.**

2.3.2. Network Architectures

There are many different network architectures in the literature, and each one was created with a specific purpose in mind. The two architectures that we will focus on in this body of work are: U-Network (UNET) and Generative Adversarial Networks (GANs). A brief description of each is given here.

U-Net is a fully convolutional network originally developed for biomedical image segmentation. It consists of a contraction path and an expansion path. The contraction path consists of convolutional and max pooling blocks to extract features then the expansion path contains up-convolution and convolution layers to recover the original size of the input image. To recover some localization of features lost in the contraction path, features are concatenated at the same level connecting the contraction and expansion path. [45]

Generative Adversarial Networks are a type of generative network where two networks, a generator and a discriminator, are used to produce data [46]. In this framework, the generator tries to maximize the probability of fooling the discriminator into marking its outputs as real; whereas, the discriminator provides feedback to guide the generator into producing outputs that are more realistic.

2.3.3. Class Activation Maps

In traditional machine learning, parameters that are easily understood by clinicians are used where their relationship to the disease pathophysiology is (often)

known. For example, the area of the foveal avascular zone (FAZ) may increase with severity due to increased capillary drop out. However, as deep learning chooses its own parameters which are not generally exported as an output, tools such as a class activation map (CAM) [47] can be used to show the location most pertinent for the DNN classification. A CAM for a particular category indicates the discriminative image regions used by the CNN to identify that category. The procedure for generating these maps is shown in Figure 2.5 using a global average pooling (GAP) layer. In brief, the last convolution layer of the network is followed by a GAP layer then one fully connected layer with a softmax activation function that would yield the predicted image category. Each output activation map from the convolutional layer corresponds to an assigned weight from the GAP layer which is connected to the predicted image category. In order to obtain the class activation map, the sum of each weight multiplied by the activation map is computed.

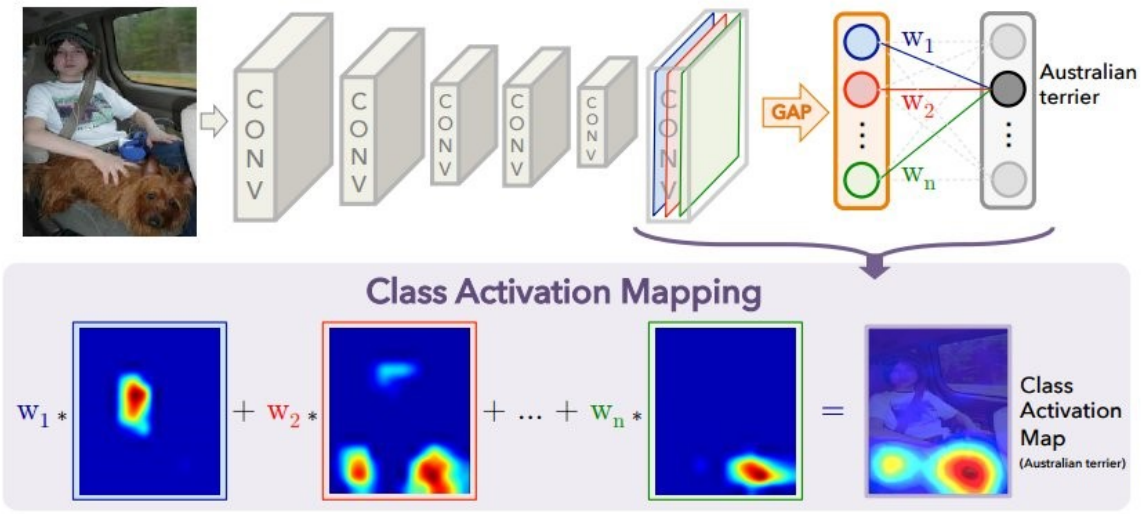

Figure 2.5 The procedure for generating CAMs from networks with GAP layers. Image from [48]**.**

2.3.4. Methods to overcome a small amount of labeled data

In cases where a large dataset of manually marked images does not exist, the construction of a DNN on an inadequate amount of data can have a negative impact on performance by causing overfitting. One method of addressing a small dataset is to use data from a similar domain, a technique known as transfer learning. Transfer learning has proven to be highly effective when faced with domains with limited data. Instead of

training a new network, by fixing the weights in certain layers of a network already optimized to recognize general structures from a larger dataset, and retraining the weights of the non-fixed layers, the model can recognize features with appreciably fewer examples. This was the method used in our paper on segmenting cone photoreceptors in AO-OCT data [34]. However, for the glaucoma pipeline, a large dataset of similar manually marked data is not easy to find, rather we have several data points with manual segmentations and a large amount of data without manual demarcations. Therefore, another method that may be more appropriate to the dataset we plan to use for the glaucoma classification is semi-supervised learning.

Semi-supervised learning is a technique in which both labeled and unlabeled data are used to train a classifier. This type of classifier takes a small portion of labeled data and a much larger amount of unlabeled data from the same domain. The goal is to combine these sources of data to train a DNN, which would be more accurate than training on the small amount of labeled data alone.

Ensemble methods, whereby multiple weak learners or different neural networks are combined, have shown great results when applied to imbalanced or small data sets. By aggregating multiple, diverse, and accurate component networks, the variance in the data is reduced to achieve greater predictive accuracy. For training component neural networks, the most prevalent ensemble approaches are Bootstrap Aggregating (Bagging) and Boosting which are algorithms that determine the training sets of component networks. Bagging is a method based on bootstrap sampling (sampling with replacement) that generates a number of training sets from an original training set and trains a component neural network on each sampled dataset. Boosting generates a series of component neural networks whose training sets are determined by the performance of previous ones. Incorrect predictions are more heavily emphasized in the training of later networks. The networks are then combined typically by majority voting, which can be used for segmentation networks as well as classification networks. Another method of combining multiple networks is stacking, whereby the networks are combined by a meta-classifier. This meta-classifier is typically a fully-connected neural network and allows for more complex, non-linear combinations of the network features.

2.3.5. Summary

Machine learning is becoming a large part of our daily life, and can aid our quantitative analysis of different images. In the next Chapter, I will discuss how we can use machine learning to help aid in the analysis of AO images of cone photoreceptors.

Chapter 3. Automated Identification of Cone Photoreceptors in Adaptive Optics Optical Coherence Tomography Images using Transfer Learning[2]

3.1. Abstract

Automated measurements of the human cone mosaic requires the identification of individual cone photoreceptors. The current gold standard, manual labeling, is a tedious process and can't be done in a clinically useful timeframe. As such, we present an automated algorithm for identifying cone photoreceptors in adaptive optics optical coherence tomography (AO-OCT) images. Our approach fine-tunes a pre-trained convolutional neural network originally trained on AO scanning laser ophthalmoscope (AO-SLO) images, to work on previously unseen data from a different imaging modality. On average, the automated method correctly identified 94% of manually labeled cones when compared to manual raters, from twenty different AO-OCT images acquired from five normal subjects. Voronoi analysis confirmed the general hexagonal-packing structure of the cone mosaic as well as the general cone density variability across portions of the retina. The consistency of our measurements demonstrates the high reliability and practical utility of having an automated solution to this problem.

3.2. Background

Adaptive optics (AO) techniques have been used to facilitate the visualization of the retinal photoreceptor mosaic in ocular imaging systems by improving the lateral resolution [49]–[51]. AO techniques have been combined with scanning light ophthalmoscope (AO-SLO) [52], flood illumination ophthalmoscopy (AO-FIO) [53], [54] and optical coherence tomography (AO-OCT) [55]. Similar to AO-SLO and AO-FIO, AO-OCT systems can provide high-resolution *en face* images of the photoreceptor mosaic in

---

[2] This work has been published in *Biomedical Optics Express*. [Morgan Heisler, Myeong Jin Ju, Mahadev Bhalla, Nathan Schuck, Arman Athwal, Eduardo V. Navajas, Mirza Faisal Beg, and Marinko V. Sarunic, "Automated identification of cone photoreceptors in adaptive optics optical coherence tomography images using transfer learning," Biomed. Opt. Express 9, 5353-5367 (2018)]

which the cones appear as bright circles surrounded by dark regions [56]–[58], but with the added benefit of high axial resolution which allows for cross-sectional tomography images in which the different outer retinal layers can be clearly delineated.

High-resolution images acquired using AO-assisted imaging systems have been used to investigate various changes in the appearance of the photoreceptor mosaic in both normal eyes [59]–[62] and eyes with degenerative retinal diseases, such as cone-rod dystrophy [63]–[65], retinitis pigmentosa [66]–[68] and occult macular dystrophy [69], [70]. Although these images can be qualitatively useful, quantitative analysis of different morphometric properties of the mosaic is generally preferred and requires identification of individual cones.

As manual segmentations are both subjective and laborious, several automated methods for detecting cones in AO images using a variety of traditional image processing techniques such as local intensity maxima detection [71]–[74], graph-theory and dynamic programming (GTDP) [75], and estimation of cone spatial frequency [76]–[78] have been developed. Although these mathematical methods work for the specific data for which they were designed, reliance on *ad hoc* rules and specific algorithmic parameters does not allow for alternative imaging conditions, such as different resolutions, areas within the retina, and imaging modalities.

A new alternative approach is using deep convolution neural networks (CNNs) where features of interest are learned directly from data. This allows for a higher degree of adaptability as the same machine learning algorithm can be re-purposed by using different training images given a sufficiently large training set [79]. Several of these networks have shown high performance for many different image analysis tasks, including ophthalmic applications [80]. For example, CNNs have been used for the segmentation of retinal blood vessels [81], [82], and detection of diabetic retinopathy [10] in retinal fundus images, and classification of pathology [83] or segmentation of retinal layers [84] and microvasculature [6] in optical coherence tomography (OCT) images. More recently, a CNN using a large dataset of manually marked images for training, has been developed to identify cones in AO-SLO images [85].

Using supervised deep learning approaches for quantification of the photoreceptors requires manually marked images. Unfortunately, a large dataset of manually marked

images from an AO-OCT system does not currently exist and the construction of a model on an inadequate amount of data can have a negative impact on performance by causing overfitting. One method of addressing a small dataset is to use data from a similar domain, a technique known as transfer learning. Transfer learning has proven to be highly effective when faced with domains with limited data [86], [87]. Instead of training a new network, by fixing the weights in certain layers of a network already optimized to recognize general structures from a larger dataset, and retraining the weights of the non-fixed layers, the model can recognize features with appreciably fewer examples [88]. In this study, we present an effective transfer learning algorithm for retraining a CNN originally trained on manually segmented confocal AO-SLO images in order to detect cones in AO-OCT images with a different field-of-view (FOV). Three different transfer learning techniques were applied and compared against manual raters.

3.3. Materials and Methods

3.3.1. AO-OCT Dataset

All AO-OCT subject recruitment and imaging was performed at the Eye Care Centre of Vancouver General Hospital. The project protocol was approved by the Research Ethics Boards at the University of British Columbia and Vancouver General Hospital, and the experiment was performed in accordance with the tenets of the Declaration of Helsinki. Written informed consent was obtained by all subjects. The 1060nm, 200 kHz swept source AO-OCT system used in this study is similar to a previously described prototype [57], but a fixed collimator was used in place of a variable collimator. The axial resolution defined by the -6 dB width was measured to be 8.5μm in air (corresponding to a resolution of 6.2μm in tissue (n = 1.38)), and the transverse resolution with a 5.18mm beam diameter incident on the cornea was estimated to be 3.6μm assuming a 22.2mm focal length of the eye and refractive index of 1.33 for water at 1.06μm. Images were acquired with the system focus placed at the photoreceptor layer using GPU-based real-time OCT B-scan images. Wavefront distortion correction was realized by optimizing the shape of a deformable mirror in the system using a Sensorless AO (SAO) technique.

Four different locations (centered at ~3:5°, ~5°, ~6:5°, and ~8°) temporal to the fovea were imaged in five subjects. For each retinal location, five AO-OCT volumes

(1.25° x 1.25° FOV) were acquired with 200 x 200 sampling density in a second. From the acquired volumes, a single volume with the least motion artifact was chosen and used for the rest of the analysis. Each AO-OCT image was resized to 400 x 400 pixels using bicubic interpolation to allow for the 'non-cone' space to be at least 1 pixel. An example of the data acquired from one subject is shown in Figure 3.1.

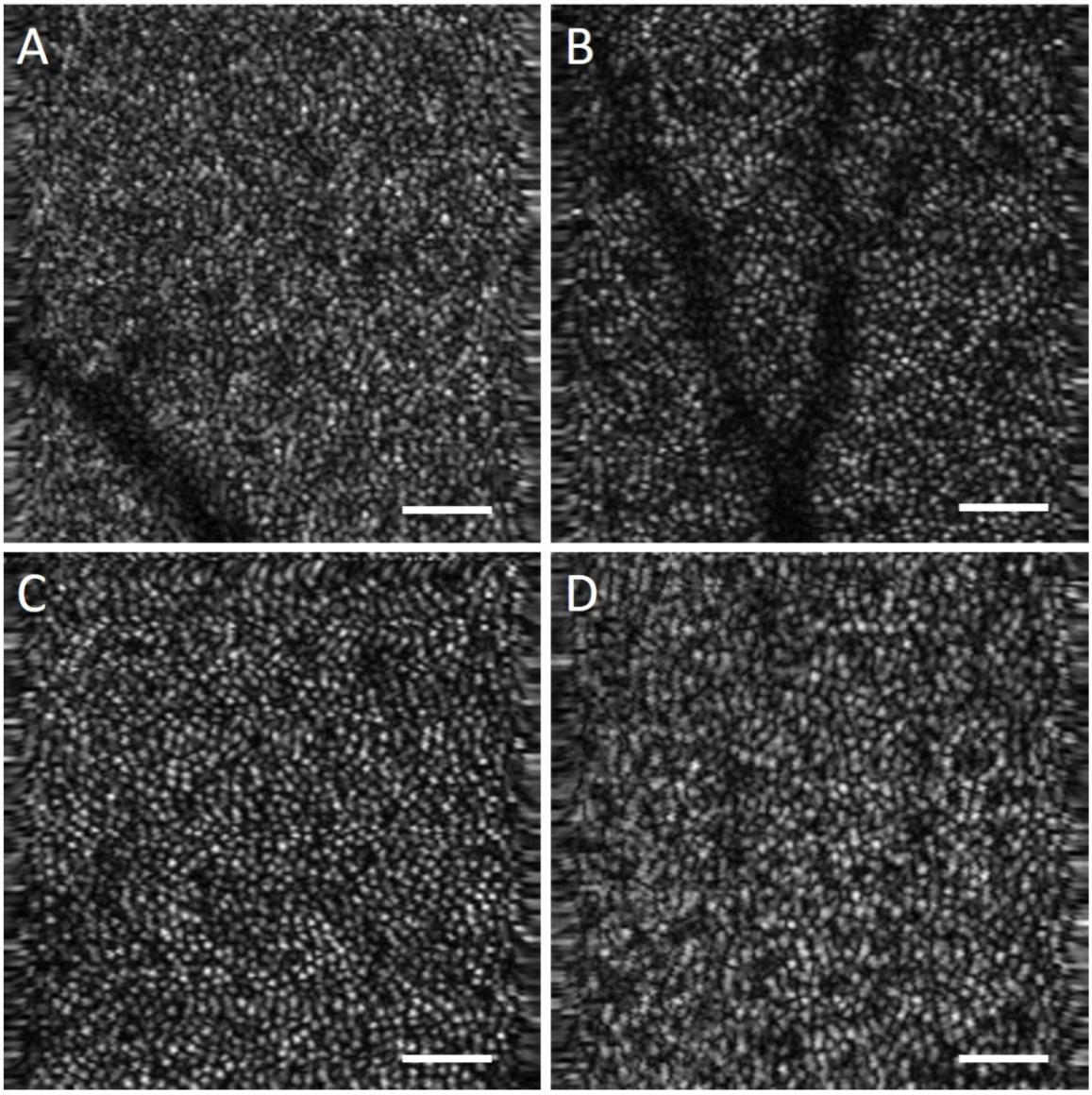

Figure 3.1 Data acquired from a 26 year old female control subject at retinal eccentricities of (a-d) ~3.5°, ~5°, ~6.5°, and ~8° respectively. Scalebar 50 µm.

The center of each cone photoreceptor was manually segmented using a Wacom Intuos 4 tablet and free image processing software [GNU image manipulation program (GIMP)] in all 20 AO-OCT images. One manual rater (Rater A) segmented data from

three subjects (12 images) and another manual rater (Rater B) segmented the data from the remaining two subjects (8 images). For analysis of inter-rater agreement, two AO-OCT images from a normal subject not included in retraining the network were segmented by both raters and compared to the CNN output.

3.3.2. AO-SLO Dataset

The AO-SLO dataset used to implement the initial conditions for the convolutional neural network was obtained from Ref. [85], and consisted of 840 confocal AO-SLO images acquired at 0.65° from the center of fixation, as well as the corresponding manual segmentations for the center of each cone. Each of the images within this dataset were extracted from a 0.96° x 0.96° FOV image, resulting in a FOV ranging from ~0.20° x 0.20° to ~0:25° x 0.25° from 21 subjects (including 20 normal subjects and 1 subject with deuteranopia) [38].

3.3.3. Data Pre-processing

The image acquisition protocols and processing strategies of the AO-SLO and AO-OCT systems are quite different. In particular, the data from the AO-OCT system has a 5-6.25 times larger FOV. Therefore, data augmentation was performed on the AO-SLO data used for training the base network in order to improve the similarity between the two datasets. The data was sub-sampled at a rate of 1.5, 2, 2.5 and 3 to be a closer representation of the resolutions used for AO-OCT imaging.

Image patches were then extracted to use as inputs for training the network. As the manual segmentation protocol did not include non-cone locations, these were extracted using the protocol in [85], [89]. In brief, a Voronoi diagram was constructed using the manually labeled cones as the center of each Voronoi cell. The boundaries were then assumed to be non-cone pixels, and a single point along each boundary was randomly chosen and placed in the non-cone set. A 33x33 window was then placed over each cone and non-cone location and used as input to the network. Locations closer than 16 pixels to the edge were discarded and not used. Because the image on the whole was imbalanced (there were far fewer pixels categorized as cone locations than non-cone locations) we welcomed the imbalance in the network and did not attempt to correct for this.

3.3.4. Network Training Methods

The convolutional neural network used in this experiment is a slightly modified Cifar network taken from the AO-SLO cone segmentation paper [85]. The details of the network architecture are given in Table 3-1. Inputs to this network are 33x33x1 feature maps centered on a cone or non-cone pixel with the corresponding binary label. The final fully connected layer provides a score for each class (cone and non-cone), which are input into a soft-max layer that outputs the probability of the original center pixel belonging to each class.

Table 3-1. CNN Architecture

Layer	Type	Input Size	Filter Size	Stride
1	Convolutional	33×33×1	5×5×32	1
2	Batch Normalization	33×33×32	-	-
3	Max Pooling	33×33×32	3×3	2
4	ReLu Activation	16×16×32	-	-
5	Convolutional	16×16×32	5×5×32	1
6	Batch Normalization	16×16×32	-	-
7	ReLu Activation	16×16×32	-	-
8	Average Pooling	16×16×32	3×3	2
9	Convolutional	8×8×32	5×5×32	1
10	Batch Normalization	8×8×64	-	-
11	ReLu Activation	8×8×64	-	-
12	Average Pooling	8×8×64	3×3	2
13	Fully Connected	4×4×64	4×4×64	-
14	Batch Normalization	1×1×64	-	-
15	ReLu Activation	1×1×64	-	-
16	Fully Connected	1×1×64	1×1×64	-
17	Soft Max	1×1×2	-	-

A similar network was also used in [90] to incorporate confocal AO-SLO and split detector AO-SLO image pairs. In brief, separate paths were used for layers 1-15 for the confocal and the split detector images, after which a concatenation layer combined the two 1 x 1 x 64 vectors output from the confocal and split detector paths into a single 1 x 1 x 128 vector that was fed into the rest of the network.

The experiments reported in this paper use two different methods to modify the AO-SLO network to segment cone photorecepters in AO-OCT images: transfer learning and fine-tuning. In both methods, certain layers of the base network trained on AO-SLO data were set to be non-trainable and the rest of the layers were then retrained using the AO-OCT data. In the first experiment, transfer learning was used so that only the classifier would be retrained and the first 15 layers of the base network were set to be non-trainable. The other two experiments used fine-tuning, where the base network was frozen before the second and third convolutional layers (layer 5 and 9, respectively) and the remaining trainable weights were subsequently retrained using the AO-OCT data. There was an average of 35,494 cone AO-OCT patches and 82,867 non-cone AO-OCT patches used on average for fine-tuning, and 336,280 AO-SLO patches used for the initial training. The batch size for training the base network was set to 100, and the maximum number of epochs was set to be 50 with an early stopping parameter set to when the validation loss hadn't decreased in 4 epochs. For the transfer learning techniques, the batch size was decreased to 32. The learning rate was 0.001 and the weight decay was set at 0.0001 for both initial training and fine tuning. In general, the learning rate for fine-tuning should have been lower to preserve learned features, but this learning rate produced the best results. Binary cross-entropy was used as the loss function for all networks.

Five-fold cross-validation on all manually segmented AO-OCT images was performed. The 20 original images were divided into 5 sets, so that all images from the same subject were placed into the same set. Images from four of the subjects were used to train the network, and images from the remaining subject were used to test the network. This procedure was repeated five times with a different test subject each time.

The CNN based detection method was implemented in TensorFlow and the Keras API [91] using Python 3.5.4. We ran the algorithm on a desktop PC with an i7-6700K CPU at 4.0 GHz, 16 GB of RAM, and a GeForce GTX 780 Ti X GPU. The average run time for segmenting a new image after training was 20 seconds.

3.3.5. Performance Evaluation Methods

A number of quantitative metrics were used to determine the effectiveness of the convolutional neural network. Probability maps were generated by extracting 33x33 pixel patches from each pixel location in the image as inputs to the trained network. The outputs of the network, the probabilities that each pixel was centered on a cone, were then arranged to generate a probability map the same size as the original image. These probability maps were binarized using Otsu's method [45] and the centroid of any 4-connected components were taken to be the centers of cones. Any pixels within half the input size (16 pixels) to the edge of the input image were discarded from analysis. Within this implementation, True Positive (TP) results will indicate that the data was located within 0.5 of the median spacing between manually marked cones to a manually marked cone, False Positive (FP) results will indicate that the automatically detected cones were not matched to a manually detected cone, and False Negative (FN) results will indicate that manually marked cones did not automatically match detected cones. Given these definitions, Dice's coefficient, Sensitivity, and False Discovery Rate are defined in Equations (1-3) respectively.

$$Dice's\ Coefficient = \frac{2TP}{TP + FN + TP + FP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$False\ Discovery\ Rate = \frac{FP}{TP + FP} \tag{3}$$

3.3.6. Cone Mosaic Analysis Methods

Voronoi diagrams were automatically constructed from each automated cone mosaic to calculate density, area, and proportion of hexagonal Voronoi domains, which are indicators of the regularity of the cone packing arrangement [92]–[94]. To analyze the regularity of the cone mosaics, the images were grouped together by their retinal

eccentricities where Area 1 was closest to the fovea (~3.5°), and Area 4 was furthest (~8°). Cone density was defined as the ratio of the number of bound Voronoi cells in an image to the summed area of the bound Voronoi cells. To calculate the proportion of 6-sided Voronoi domains, the number of Voronoi cells with six sides was divided by the total number of bound Voronoi cells within an image. The number of neighbours was calculated as the mean number of sides of all bound Voronoi cells in an image. Similarly, the Voronoi cell area was calculated as the mean area of the bound Voronoi cells in an image. An example summary of this analysis is shown in Figure 3.2 along with the original AO-OCT image. The Voronoi boundary map (green) and the automated centres of each cone (magenta) are shown in Figure 3.2B, the number of neighbours map where each Voronoi cell is shaded depending on the number of neighbours is shown in Figure 3.2C and the Voronoi cell area map where each Voronoi cell is shaded based on the cell area is shown in Figure 3.2D.

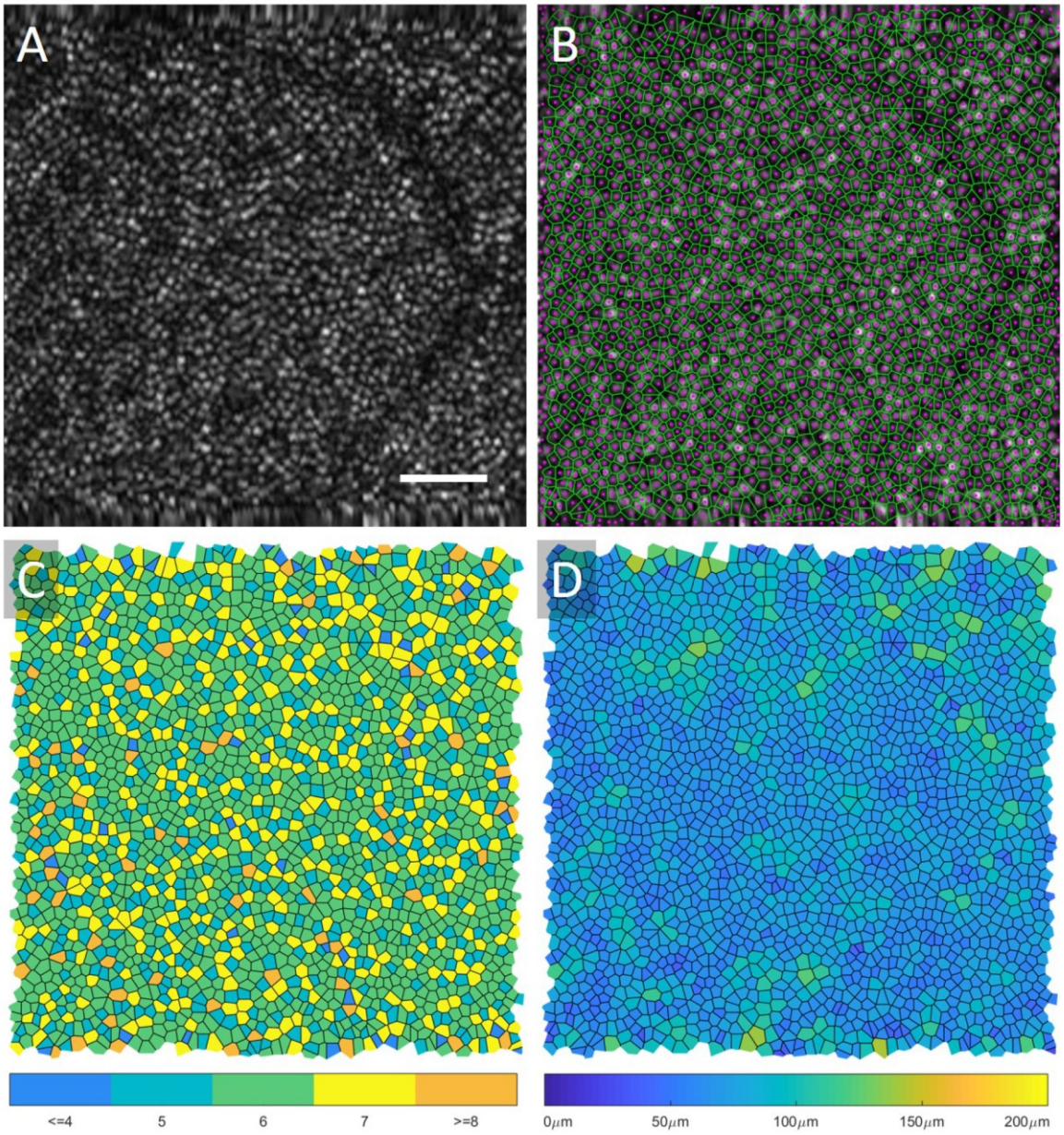

Figure 3.2 **An original AO-OCT image taken at ~6:5_ retinal eccentricity is displayed in (a), and the center of the cones (magenta) and Voronoi map (green) is overlaid onto the image in (b). In (c) the Voronoi cells are shaded based on the number of neighbours, and in (d) the cells are shaded based on their area. Scalebar 50 μm.**

3.4. Results

3.4.1. Performance Evaluation

The performances of the automated algorithms in comparison to manual grading are summarized in Table 3-2, including training on a network with randomly initialized weights. For the case of the randomly initialized weights, only AO-OCT images were used to train the network. The results from all four methods are within the standard deviations across all three quantitative measurements. A trend of over-segmenting in the methods which retrained more weights (Fine- Tuning Layer 5 and Layer 9) can be seen in the lower resultant sensitivities and slightly better false discovery rates. These numbers are slightly worse than the original network trained only on confocal AO-SLO data, where the sensitivity was 0.989 ± 0.012, the false discovery rate was 0.008 ± 0.014, and the Dice's Coefficient was 0.990 ± 0.010 [38].

Table 3-2. Average performance of the automated methods with respect to manual marking

	Sensitivity	False Discovery Rate	Dice's Coefficient
Transfer Learning	0.940 ± 0.041	0.079 ± 0.037	0.929 ± 0.018
Fine-Tuning (Layer 5)	0.936 ± 0.046	0.062 ± 0.038	0.935 ± 0.017
Fine-Tuning (Layer 9)	0.936 ± 0.049	0.073 ± 0.043	0.930 ± 0.018
Random Initialization	0.942 ± 0.034	0.093 ± 0.038	0.923 ± 0.017

Figure 3 displays the results of the automated algorithms in comparison to manual grading for one of the AO-OCT datasets. In the marked images, a green point indicates an automatically detected cone that was matched to a manually marked cone (true positive), a yellow point indicates a cone missed by the automatic algorithm (false negative), and a red point indicates an automatic marking with no corresponding manually marked cone (false positive). From the images we can see that the methods performed quite well and that the methods that retrained more weights in the network produced more automated cone locations, resulting in fewer missed cones (hence less yellow locations) and more red locations, than the transfer learning method which only retrained the classifier.

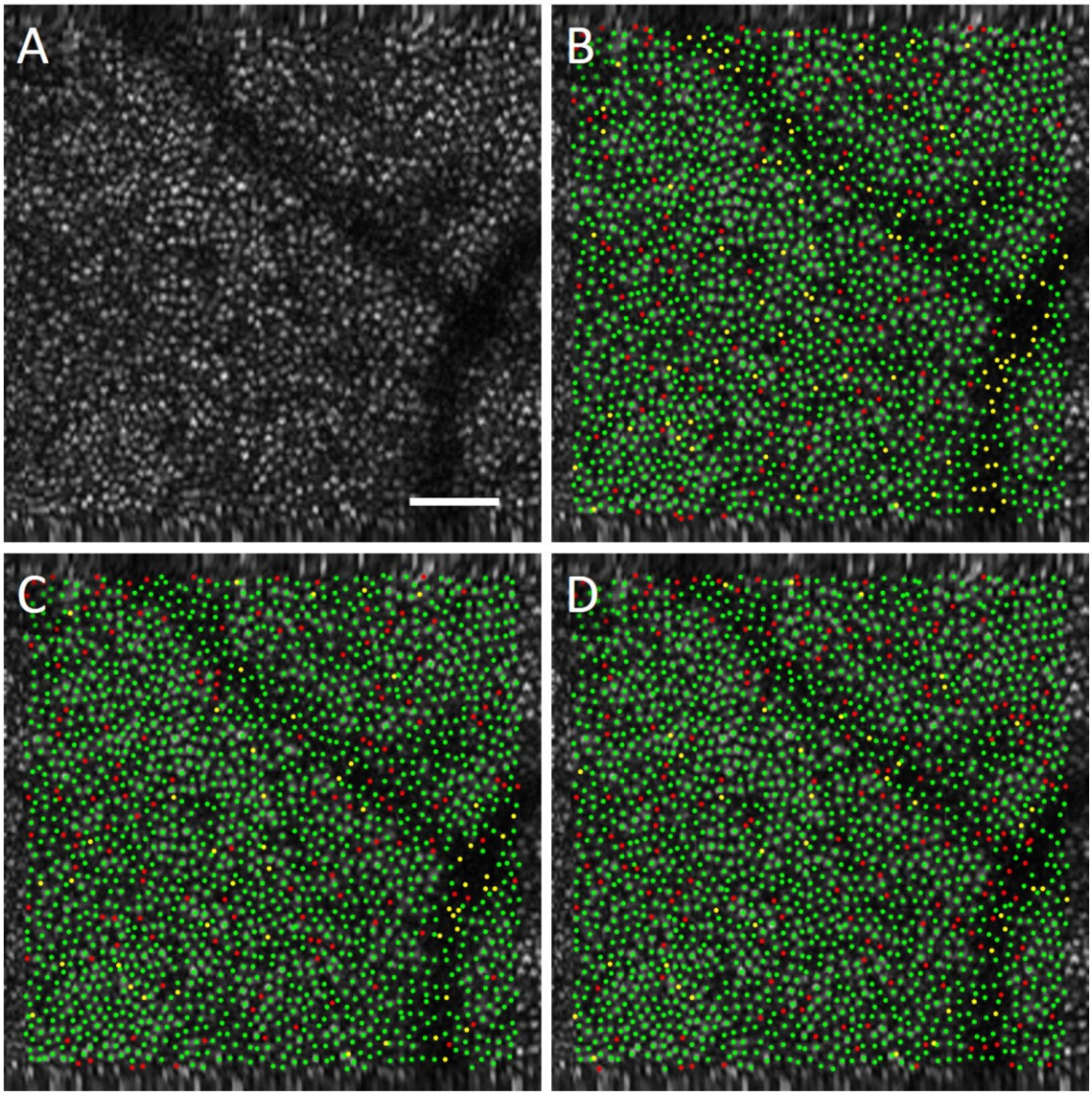

Figure 3.3 **Comparison of automated results in (a) an AO-OCT image from the (b) transfer learning, (c) fine-tuning (Layer 5) and (d) fine-tuning (Layer 9) methods. In the marked images, a green point indicates an automatically detected cone that was matched to a manually marked cone (true positive), a yellowpoint indicates a cone missed by the automatic algorithm (false negative), and a red point indicates an automatic marking with no corresponding manually marked cone (false positive). Scalebar 50 μm.**

3.4.2. Inter-rater Agreement

As previously mentioned in Section 3.3.1, two AO-OCT images from a normal subject not included in training the network were segmented by both raters for inter-rater analysis. For the AO-SLO images in [85], the manual rater segmentation quality was

high because the FOV was small, the images highly sampled, and the cones hence had a high contrast. The AO-OCT images used in this report had a larger FOV, and hence a lower sampling density. Consequently, the contrast of the cones was poorer, leading to minor disagreement in segmentation even between the manual graders. The inter-rater performance can be seen in Figure 3.4 along with the original AO-OCT image and a comparison of Rater A to the CNN output. As can be seen in Figure 3.4b,e, Rater B found more cones than Rater A did in the majority of the image area, with the exception of areas in the blood vessel shadow where they were markedly more conservative. Similarly, the CNN found more cones than Rater A and in fact was more similar to Rater B as shown in Table 3-3. All inter-rater measurements are within the standard deviation of the automated results, suggesting that the automated segmentation is comparable to that of a human rater.

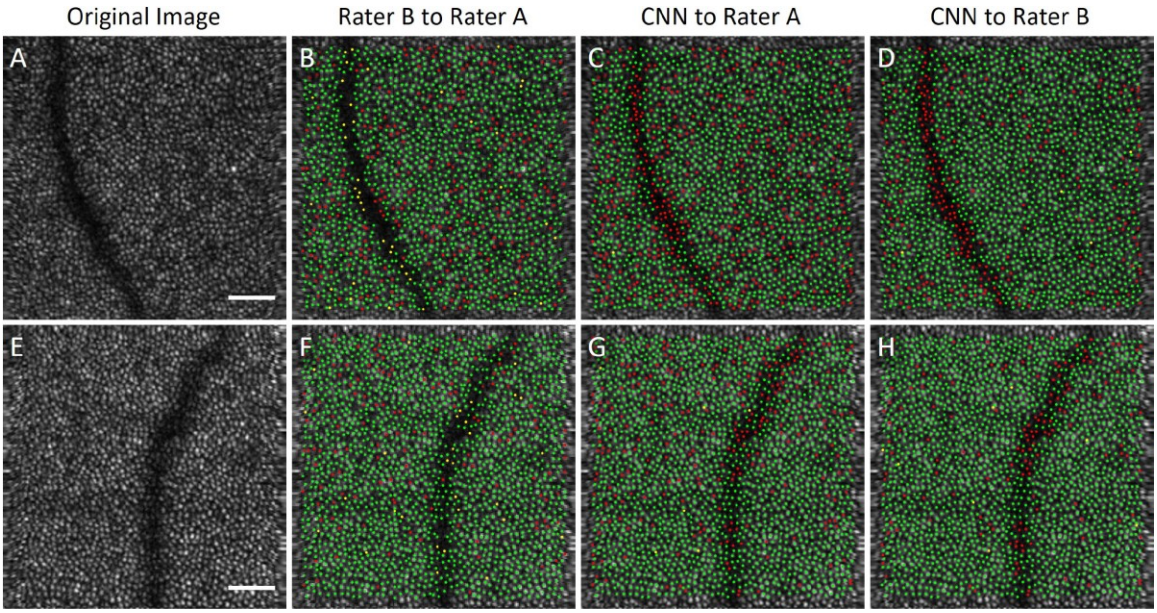

Figure 3.4 **Comparison of manual results in two AO-OCT images (a) and (e) to a second rater (b,f) and the CNN (c-d,g-h). In the manually marked comparison images (b,f), a green point indicates a cone marked by Rater A that was matched to a cone marked by Rater B (true positive), a yellow point indicates a cone missed by Rater B (false negative), and a red point indicates a marking by Rater B with no corresponding cone marked by Rater A (false positive). In the comparison images to the CNN (c-d,g-h), a green point indicates a cone marked by a Rater that was matched to a cone marked by the CNN (true positive), a yellow point indicates a cone missed by the CNN (false negative), and a red point indicates a marking by the CNN with no corresponding cone marked by the Rater (false positive). Scalebar 50 µm.**

Table 3-3. Average performance of both raters and the automated methods with respect to manual marking

	Sensitivity	False Discovery Rate	Dice's Coefficient
Inter-rater	0.978 ± 0.007	0.109 ± 0.043	0.932 ± 0.027
CNN to Rater A	0.998 ± 0.002	0.177 ± 0.066	0.902 ± 0.039
CNN to Rater B	0.997 ± 0.001	0.099 ± 0.033	0.946 ± 0.018

3.4.3. Cone Mosaic Analysis

From the performance evaluation of the three different transfer learning methods, we chose the results from the Fine-Tuning (Layer 5) method to further analyze as it had the highest Dice's Coefficient. The results are summarized in Table 3-4 for all cone analysis parameters. The proportion of hexagonal cells and mean number of neighbours remained relatively constant over the areas imaged, while a general trend of the mosaic becoming less dense further from the fovea can be observed.

Table 3-4. Cone Mosaic Measurements by Area

	Area 1 (~3.5°)	Area 2 (~5°)	Area 3 (~6.5°)	Area 4 (~8°)
Automated Results				
Cone Density (cones/mm^2(×1000))	15.49 ± 1.02	13.15 ± 0.88	11.95 ± 0.53	10.96 ± 0.39
Percent 6-Sided (%)	47.71 ± 2.56	47.42 ± 3.30	48.40 ± 1.36	46.89 ± 1.33
Number of Neighbors	5.87 ± 0.34	5.89 ± 0.33	5.87 ± 0.34	5.87 ± 0.36
Voronoi Cell Area (μm^2)	52.50 ± 11.22	61.91 ± 12.82	67.94 ± 13.84	73.98 ± 14.72
Manual Results				
Cone Density (cones/mm2(×1000))	16.12 ± 1.31	12.86 ± 0.88	11.77 ± 0.89	10.71 ± 1.09
Percent 6-Sided (%)	47.39 ± 1.58	44.01 ± 1.35	45.87 ± 1.71	43.77 ± 0.92
Number of Neighbors	5.87 ± 0.35	5.86 ± 0.35	5.86 ± 0.32	5.85 ± 0.36
Voronoi Cell Area (μm2)	50.66 ± 18.05	63.34 ± 18.02	69.20 ± 21.61	76.54 ± 23.80

3.5. Discussion

In this work, we investigated the use of transfer learning techniques using an automatic CNN based method for detecting cone photoreceptors in AO-OCT images. Using manually marked images from a confocal AO-SLO system to initialize the weights of our network, we have demonstrated retraining a CNN to extract features of interest and classify cones in previously unseen images from an AO-OCT imaging system. We tested our method on images of various retinal eccentricities and showed that our

method had good agreement with the current gold standard of manual marking. In addition, we used various morphometric cone mosaic measurements to show quantitative agreement with measurements from AO-SLO systems.

As shown in Table 3-2, performance of the CNN based algorithms were comparable to the current gold standard of manual grading which suggests that the automated segmentation is comparable to that of a human rater. This is highly encouraging, because traditional methods of photoreceptor cone segmentation heavily utilize modality and FOV specific *ad hoc* rules, which limit their application to other imaging protocols and require further algorithm modification and development for new imaging protocols. All that was needed to adapt the algorithm from confocal AO-SLO to AO-OCT was the corresponding training dataset. As we can see from the Table, the results from only using AO-OCT images and random weight initialization are comparable to using fine-tuning and transfer learning methods. We postulate that this is due to the large number of AO-OCT training data (118,361 patches) and that for a CNN, recognizing simple shapes like high contrast cones is a relatively straightforward task. As such, Figure 3.5 shows how the number of training patches affects the performance of the different CNN algorithms. Random initialization performed poorly for 1000 patches, was close to the transfer learning methods at 2000 patches, but ultimately fit within the standard deviation at 3000 patches and above, whereas the transfer learning methods were stable from 1000 patches. This has important implications as we look to use CNNs for pathological images where the signal-to-noise ratio is lower and there are much fewer datasets from which to train.

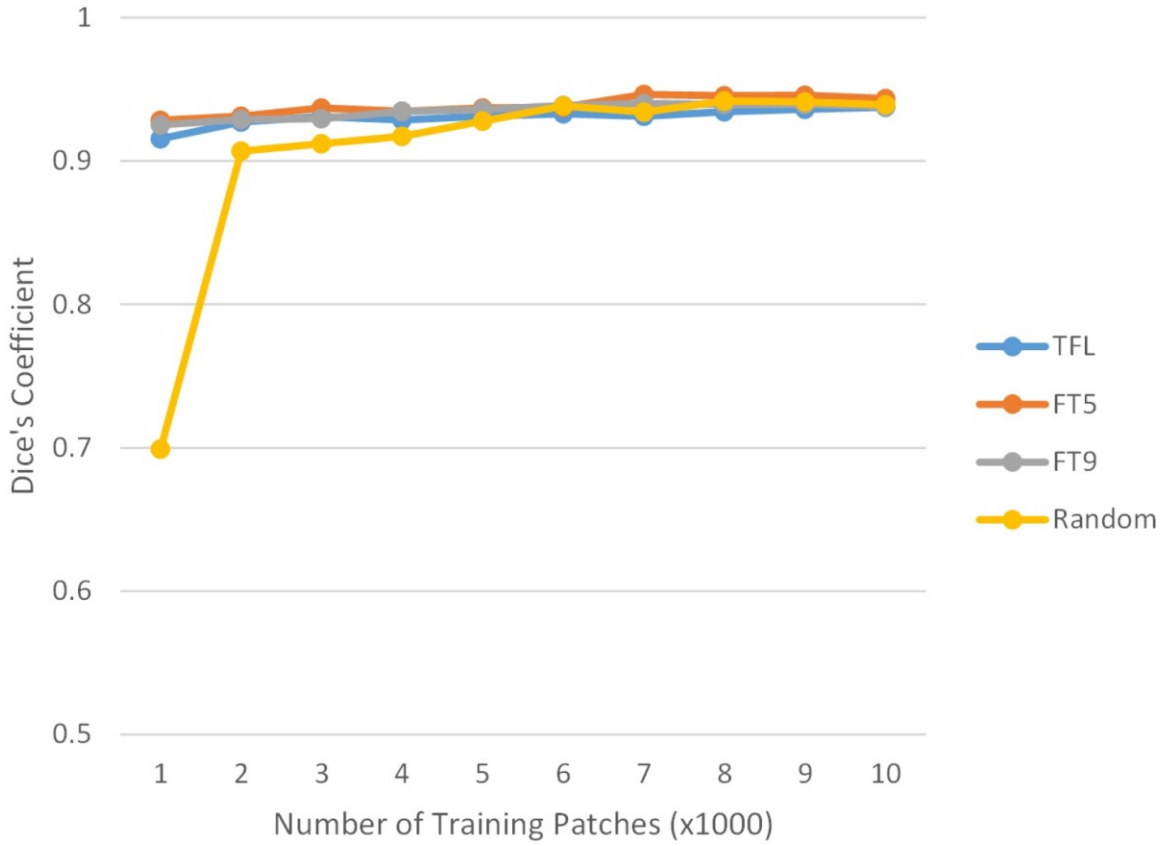

Figure 3.5 How the number of training patches affects the performance of the different CNN algorithms. Random weight initialization (Random) is comparable to the Transfer Learning (TFL), Fine-Tuning Layer 5 (FT5), and Fine-Tuning Layer 9 (FT9) methods at 3,000 training patches.

Additionally, the algorithm's output measurements are congruent with AO-SLO data from the Literature [95]. The proportion of hexagonal cells remained relatively constant over the areas imaged, which is consistent with the Literature although the overall values are slightly lower than reported for AO-SLO data (52.6 ± 6.56 at 3.5° and 50.9 ± 7.32 at 8°) [95]. Similarly, the mean number of neighbours also remained relatively constant over the areas imaged, which is consistent with the Literature although the overall values are slightly lower than reported for AO-SLO data (52.6 ± 6.56 at 3.5° and 50.9 ± 7.32 at 8°) [95]. The general trend of the mosaic becoming less dense further from the fovea was also observed and is shown in Figure 3.6 which compares our cone density measurements to measurements found in the Literature [96]–[98] for histology and AO-SLO data. Datapoints from Ref. [96] were extracted from the figures in the paper. As shown, our data follows the general trend for the eccentricities imaged.

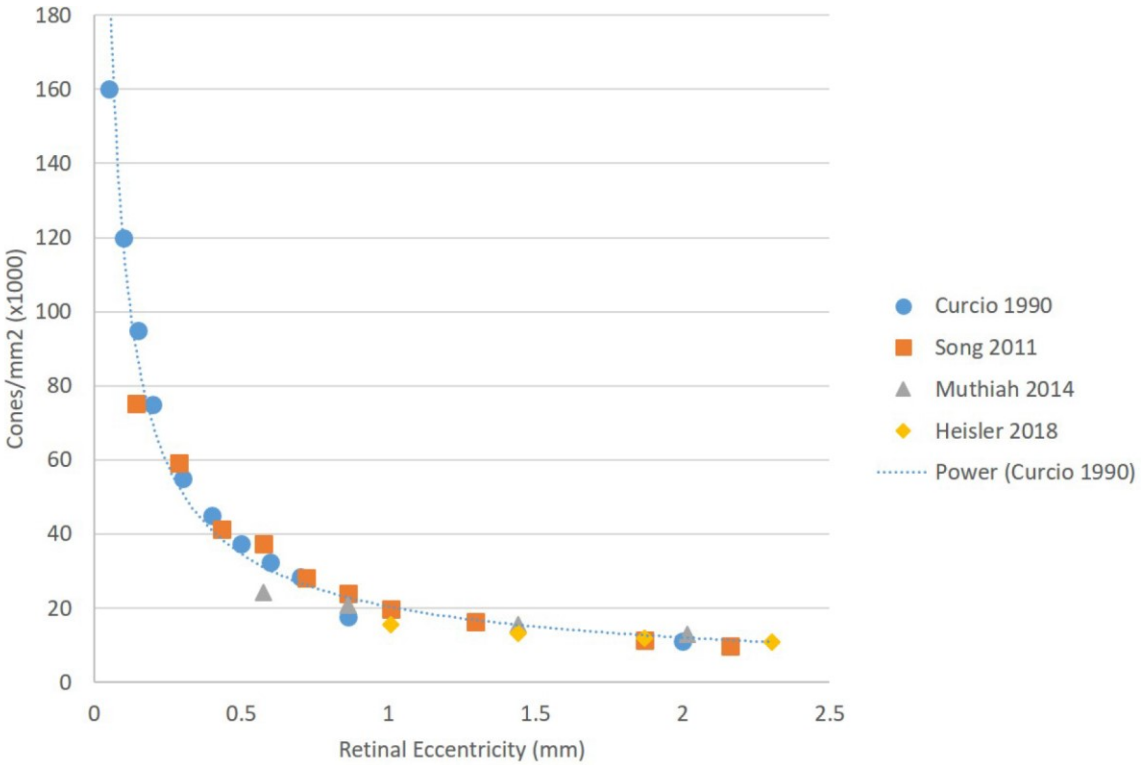

Figure 3.6 Comparison of cone density measurements from the Literature to the cone density measurements from the AO-OCT system. The gold standard of histology [49] is shown with a trendline, as well as measurements from two AO-SLO systems [50, 51].

There is a trend towards overestimating cones in the algorithms where more layers were frozen with the AO-SLO initialized weights (Fine-Tuning (Layer 9) and Transfer Learning) as shown in the higher false discovery rate in Table 3-2 although the true positive rate was better. This is also reflected in Figure 3.3, where the Fine-Tuning (Layer 5) results show less false positive cones. The majority of false positives was in regions below blood vessels. This could be improved by including more training data specifically around regions of vessels. This could either be done by acquiring more data around blood vessels, or using data augmentation techniques to modify the current dataset. Alternatively, the CNN could be combined with other pre-processing steps, such as blood vessel segmentation (for example, using OCTA as shown in Figure 3.7) to identify and remove the regions below vessels from the quantitative analysis. Moreover, poor inter-observer agreement can negatively affect the performance of learning based methods such as CNN. Utilization of datasets graded by multiple observers, for example both Rater A and Rater B, could further improve performance.

There is also space for improvement in all of the proposed algorithms mentioned in this work. The network architecture was chosen as it was the one published with the open-source confocal AO-SLO dataset, and was not optimized for the AO-OCT data. Additionally, the hyper-parameters were empirically chosen to provide good performance for both the fine-tuning and transfer learning methods. It is possible these parameters could be further optimized to provide better performance. Additionally, applying further custom pre-processing and post-processing steps such as multi-acquisition registration and averaging [99] may improve the results presented here.

Previously demonstrated in the Literature was confocal AO-SLO to split aperture AO-SLO; although different, there were similarities in field of view. In this report, transfer to a completely different imaging modality was demonstrated. Furthermore, the differences in scale of the PRs across retinal eccentricities while maintaining good sensitivity and false discovery rate is a significant achievement and demonstration of CNN usefulness. In general, for the cone photoreceptor mosaic, the features have a relatively well defined underlying structure, and a regularity for which CNNs are well suited. This implies that this transfer learning technique could be used to analyze images from commercial flood fundus photoreceptor images as well. There exists potential for CNNs to help unite images from different modalities with or without AO enabled imaging techniques for understanding photoreceptor changes in disease.

An example of this kind of multi modality imaging is shown in Figure 3.7 where structural OCT, OCTA and AO-OCT were used. Though the individual AO-OCT images used in this report are currently considered to be large field of view when imaging photoreceptors, being able to view wide field structural images is necessary for locating areas of interest. As the data diversity in this study is low, further studies looking at various pathologies using wide field structural OCT cross sections and OCTA to observe microvasculature changes and locate smaller regions of interest to then image with the higher resolution AO-OCT, using the CNN for quantification of the cone mosaic and vasculature [6], would be pertinent for proving clinical utility.

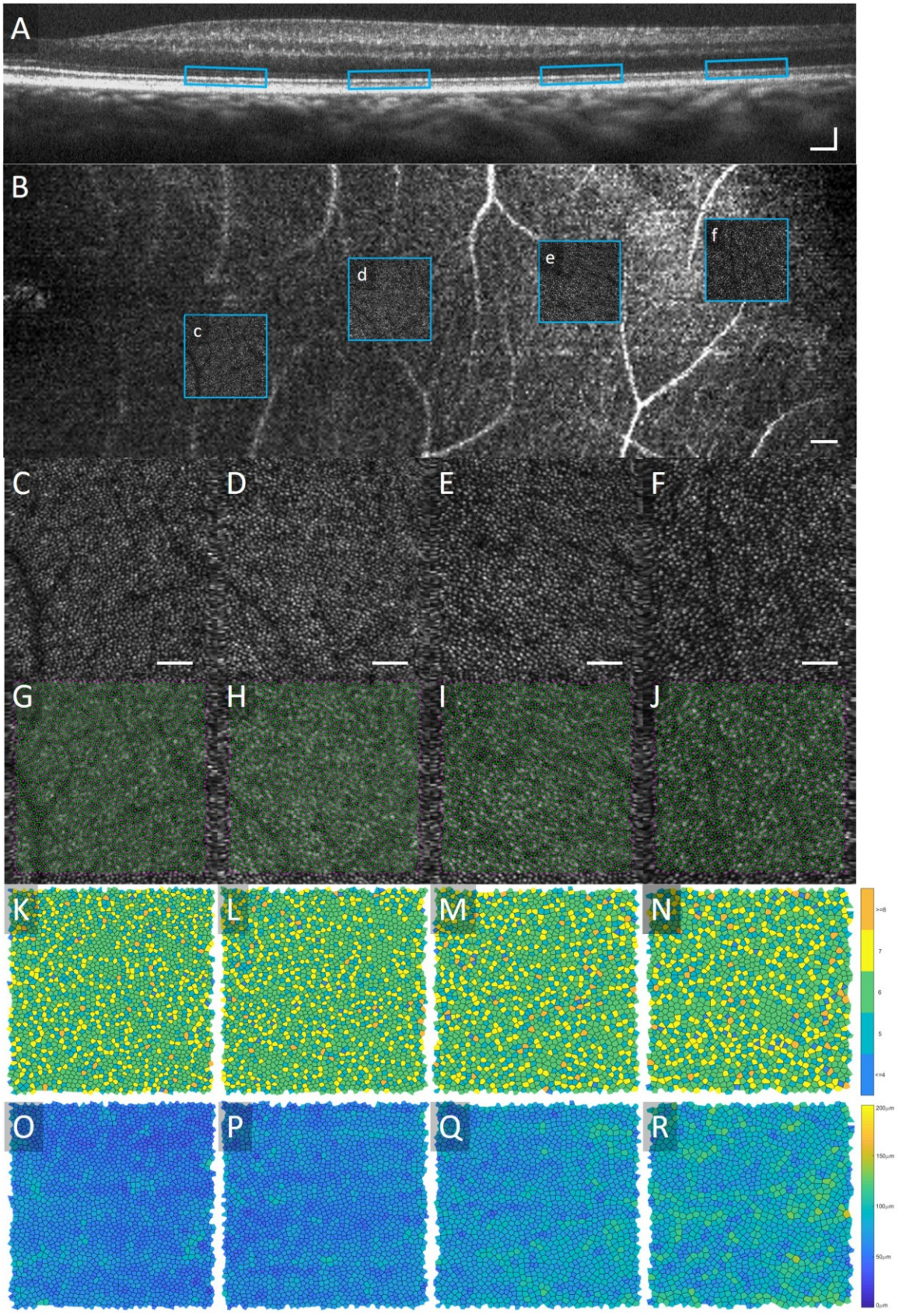
A
B
c
d
e
f
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R

Figure 3.7 **Results from a 22 year old male subject. Colocalization of the Areas 1 and 2 to a (a) widefield OCT B-scan [scalebar 100 μm] and (b) widefield AO-OCTA *en face* view [scalebar 100 μm] are shown. The original AO-OCT images from Areas 1-4 are shown in (c-f) [scalebar 50 μm], automated segmentation results and the Voronoi diagrams are in (g-j), the Voronoi cells in (k-n) are shaded based on the number of neighbours, and in (o-r) the cells are shaded based on their area.**

3.6. Conclusion

CNNs provide the opportunity to adapt to changing conditions without having to adjust *ad hoc* rules, but instead retraining the network on a sufficiently large database. In this paper, we experimentally demonstrated three different transfer learning methods to identify the cones in a small set of AO-OCT images using a base network trained on AO-SLO images which all obtained results similar to that of a manual rater. Using the results from the Fine-Tuning (Layer 5) method, we calculated four different cone mosaic parameters which were similar to results found in AO-SLO images showing the utility of our method.

3.6.1. Contributions

The details of the methods were conceived and designed by myself. I acquired the data for all subjects except one (as it was myself), as well as processed all the images. All machine learning scripts were written by me, but I did use MATLAB scripts from the open source repository and from Dr. Myeong Jin Ju for the performance evaluation and Voronoi analysis, respectively. Additionally, I wrote the manuscript which was subsequently published.

Chapter 4. Semi-Supervised Deep Learning Based 3D Analysis of the Peripapillary Region

4.1. Abstract

Optical coherence tomography (OCT) has become an essential tool in the evaluation of glaucoma, typically through analyzing retinal nerve fiber layer changes in circumpapillary scans. Three-dimensional OCT volumes enable a much more thorough analysis of the optic nerve head (ONH) region, which may be the site of initial glaucomatous optic nerve damage. Automated analysis of this region is of great interest, though large anatomical variations and the termination of layers make the requisite peripapillary layer and Bruch's membrane opening (BMO) segmentation a challenging task. Several machine learning-based segmentation methods have been proposed for retinal layer segmentation, and a few for the ONH region, but they typically depend on either heavily averaged or pre-processed B-scans or a large amount of annotated data, which is a tedious task and resource-intensive. We evaluated a semi-supervised adversarial deep learning method for segmenting peripapillary retinal layers in OCT B-scans to take advantage of unlabeled data. We show that the use of a Generative Adversarial Network and unlabeled data can improve the performance of segmentation. Additionally, we use a Faster R-CNN architecture to automatically segment the BMO. The proposed methods are then used for the 3D morphometric analysis of both control and glaucomatous ONH volumes to demonstrate the potential for clinical utility.

4.2. Introduction

Glaucoma is a leading cause of irreversible blindness worldwide [30] and the second most common cause of blindness in the developed world [31]. It is characterized by the degeneration of retinal ganglion cells and the loss of their axons [32], which manifests as narrowing of the neuroretinal rim and structural remodeling of the optic nerve head (ONH)[100]. Current clinical indices for evaluating the progression of glaucomatous optic neuropathy include assessment of the cup-to disc ratio and neuroretinal rim area in fundus photographs, visual field testing, and retinal nerve fiber layer (RNFL) analysis using optical coherence (OCT) images. Although optical coherence tomography (OCT) is mainly used for sectoral analysis of the RNFL tissue

thickness through circumpapillary scans, recent advances in OCT imaging have enabled the acquisition of high-resolution 3D images from which morphometric measurements and new biomarkers based on anatomical landmarks can be derived. One such landmark, the Bruch's Membrane Opening (BMO), which has been shown to be more reliable than disc margin-based rim evaluation of fundus photography[100]–[102]. As such, its segmentation has become a crucial step to enable the detection of ONH shape parameters.

Several traditional processing tools have been employed to automatically segment the ONH layers and extract ONH morphological parameters [4], [103]–[106]. The high degree of variability in the appearance of this region, such as the large range of peripapillary tissue thickness and surface regularity due to the region's dense vasculature, age and disease severity, makes accurate modeling particularly challenging. Additionally, the ONH cup can be shallow or deep greatly changing the shape of the prelaminar neural tissue. These factors, in addition to complications arising from typical OCT imaging conditions, such as different fields of view or shadowing artifacts, can yield traditional algorithms clinically unviable without the dependence on several ad hoc parameters to account for different imaging cases. In our previous works using volumetric images of the ONH and graph-cut segmentation [107]–[112], extensive manual corrections of the automated layer segmentations were sometimes necessary before quantification, especially in high myope, pathological, or poor quality images.

Recently, deep learning has emerged as a potential solution to previously inflexible retinal segmentation methods. Several studies have shown the successful segmentation of retinal layers from macular OCT images[113]–[118] as well as ONH OCT images[119]–[121]. Of note, Fang et al.[84] used a patch-based convolutional neural network to provide an initial segmentation which is refined using a graph based approach. More recently, Roy et al.[114] designed a fully convolutional network, ReLayNet, to segment retinal layers and macular edema in retinal OCT images. Both methods were applied and tested only on macula scans. Devalla et al.[119], recently improved upon their previous patch-based ONH segmentation method[120], by using a modified U-Net achieving great results on the limited dataset. However, these methods were only validated on 200 B-scans (including the paper's Appendix) through the center of the ONH and not extended to the extraction of 3D parameters in the acquired OCT volume. More recently Zang et al. [121] presented a method by which a U-Net based

neural network and graph search were combined to delineate the boundary of the optic disc as well as the retinal boundaries. This method was tested on both control and glaucomatous OCT volumes and achieved good results when trained on 10,000+ scans.

In this study, we present a Generative Adversarial Network (GAN) based method for segmenting the peripapillary retinal layers. Although amassing a large amount of OCT images is typically not difficult in a clinical setting, attaining the corresponding manual annotations can be tedious and resource intensive, therefore having a large number of scans without annotations is not helpful in fully supervised methods. Through the use of semi-supervised (SS) learning, we demonstrate methods to train a network with little annotated training data while taking advantage of unlabelled scans. Additionally, as the BMO must be segmented in order to extract quantitative parameters, we present results using a Faster R-CNN for automatically segmenting the BMO.

4.3. Materials and Methods

The use of the data for this study was approved from the ethics review boards of Simon Fraser University (SFU) and from the University of British Columbia (UBC). The study was conducted in accordance with the guidelines of the Declaration of Helsinki, and informed consent form was obtained from each participant.

4.3.1. Database Demographics

All volumes in this study were previously acquired for our group's previous studies[107], [109], [122]–[124]. A total of 63 subjects (122 eyes) were used for this study: twenty-one healthy controls (42 eyes, mean age ± standard deviation = 33.95 ± 11.37 years), and forty-two patients with glaucoma (80 eyes, mean age = 65.59 ± 8.91 years). A total of 326 OCT volumes were used in this study. A diagnosis of open-angle glaucoma was made clinically by a fellowship-trained glaucoma specialist based on conventional examination including dilated stereoscopic examination of the optic nerve, analysis of stereo disc photography, and typical reproducible Humphrey SITA-Standard white on white visual field abnormality. In addition, severity of glaucomatous visual field loss was quantified by visual field mean deviation (MD) values.

As the purpose and parameters was different for each of the previous studies from which this data set was generated, the availability of manually corrected segmentations was different for each volume and ranged from no segmentations, only layer segmentations, only BMO segmentations, and both layer and BMO segmentations. A breakdown of the participants' demographics for each dataset is shown in Table 4-5.

Table 4-5. Dataset Demographics for the varying levels of segmentation

Peripapillary Layer Segmentation Dataset	**Controls**	**Glaucomatous**
Number of subjects (eyes)	13 (25)	16 (30)
Number of volumes	128	69
Number of B-scans	46,287	26,354
Age (± std)	36.80 ± 6.42	68.10 ± 8.44
Axial Length (± std)	24.02 ± 1.08	25.70 ± 2.03
VF MD (± std)	-0.92 ± 0.48	-9.95 ± 8.62
Male: Female	6:7	9:7
BMO Segmentation Dataset		
Number of subjects (eyes)	16 (30)	34 (65)
Number of volumes	145	136
Number of Radial frames	11,600	10,880
Age (± std)	32.30 ± 9.12	63.69 ± 13.71
Axial Length (± std)	24.87 ± 1.24	25.58 ± 1.65
VF MD (± std)	-1.00 ± 0.77	-9.68 ± 8.97
Male: Female	8:8	19:15
Unlabeled Dataset		
Number of subjects (eyes)	4 (8)	13 (22)
Number of volumes	8	22
Number of B-scans	2,662	6,796
Age (± std)	44.75 ± 16.31	58.41 ± 13.82
Axial Length (± std)	25.02 ± 1.34	24.91 ± 1.36
VF MD (± std)	-1.28 ± 0.65	-12.30 ± 10.16
Male: Female	1:3	7:8

4.3.2. OCT Volume Acquisition & Processing

Details on the OCT system used in this study has been previously published [122]. In brief, a custom-built swept-source OCT system with a center wavelength of 1.06 μm and 100 kHz sweeping frequency was used to image the ONH region. The acquired three-dimensional (3D) images consisted of 400 B-scans, each with 400 A-scans, and 1024 pixels per A-scan. The imaged region in physical space spanned an axial depth of 2.8 mm and a square area of 5 × 5 to 8 × 8 mm^2 dependent on the axial length of the

eye and scan angle. Resulting voxel dimensions were 2.7 μm in the axial direction and ranged from 12.5 to 20 μm in the lateral direction.

Axial displacement caused by involuntary axial eye motion during image acquisition was corrected using cross-correlation between adjacent frames, which were subsequently cropped to be 640 pixels. Each B-scan used for input to the network was averaged with the preceding and following 5 B-scans. Additionally, three-dimensional bounded variation smoothing was applied to the volumes used for BMO segmentation in order to reduce the effect of speckle while preserving and enhancing edges.

4.3.3. Ground Truth Labels

Automated layer segmentations of the Inner limiting membrane (ILM), the posterior boundary of nerve fiber layer (NFL), Bruch's membrane (BM), Bruch's membrane opening, and the choroid–sclera boundary (CS boundary) were generated in 3D using a graph-cut algorithm [108], [125]. The automated segmentation result was examined and corrected by trained research engineers in Amira (version 5.2; Visage Imaging, San Diego, CA, USA) or ITK-SNAP (version 3.2).

The BMO, defined as the termination point of the high-reflectance BM/retinal pigment epithelium (RPE) complex, was segmented on 80 radial slices extracted from the smoothed volume, intersecting at the approximate center of the BMO and spaced at a constant angle of 2.25°. The ONH is a relatively radially symmetric structure, and radial slices provide a more consistent cross-sectional view of the BMO compared to the raster scan pattern in which the volumes were acquired. For input into the BMO Segmentation Network, the BMO was considered to be the 50x50 pixel area centered on the manual BMO segmentation point.

4.3.4. Layer Segmentation Network

Architecture
A Pix2Pix GAN[126] based approach was used with the layer segmentation. For the generator network, we used a U-NET based segmentation network[45] which is comprised of a contracting and expanding path connected by skip connections as denoted in Figure 4.1. The contracting path contains 4 blocks consisting of a 2D convolutional layer (3x3), dropout layer, 2D convolutional layer (3x3) and max pooling

layer. Similarly, the expanding path also contained 4 blocks consisting of a 2D transpose convolutional layer (2x2, stride 2x2) concatenated with the corresponding layer from the contracting path, a 2D convolutional layer (3x3), dropout layer, and 2D convolutional layer (3x3). To compare the effect of semi-supervised learning on a standard U-Net and on a GANS, the output layer was slightly different for each network. For the U-Net implementation, the output layer was implemented as a 1×1 convolution layer with 5 filters corresponding to the number of regions segmented and a sigmoid activation. Sigmoid was used over softmax as it produced better segmentation performance. Pixel-wise classifications were made based on the highest probability. For the GANs implementation, to make the output layer differentiable it was implemented as a 1×1 convolution layer with 1 filter, where the output is integer encoded to the number of regions segmented and a linear activation.

The discriminator model is implemented as a PatchGAN [126]. The output of the network is a single feature map of manual/automated predictions corresponding to a patch size of 16x16 on the original image that is then averaged to give a single score.

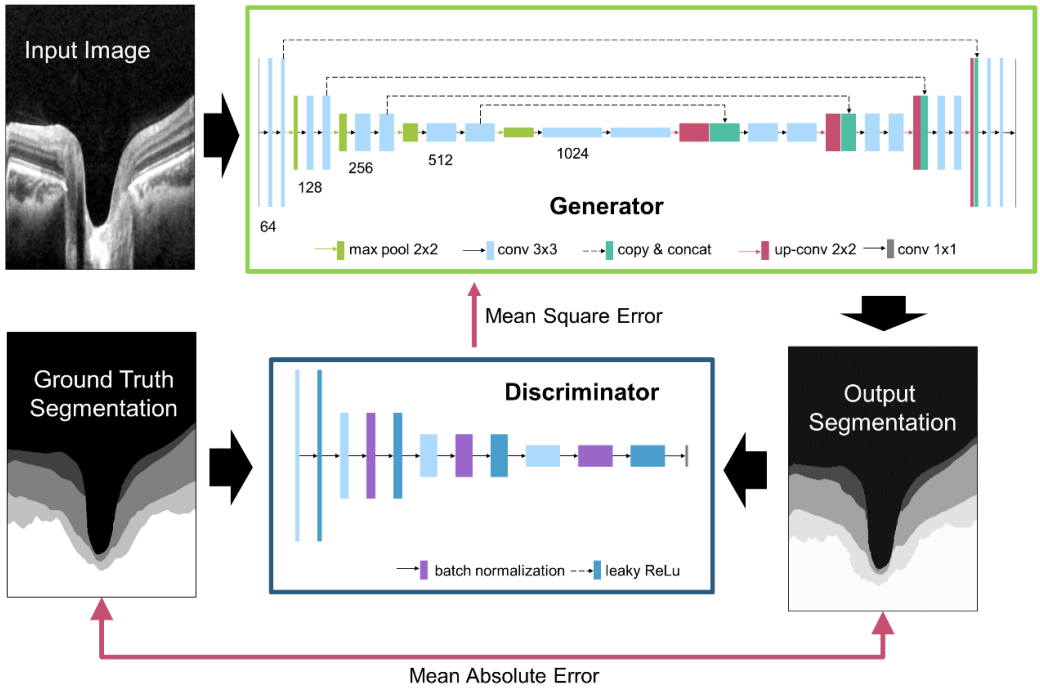

Figure 4.1 Network architecture for adversarial layer segmentation of ONH peripapillary layers.

Training

Data was augmented with random horizontal flips, cropping from 0 to 10%, linear contrast stretching from 0.75 to 1, and rotations from -20 to +20 degrees. The Adam optimizer with a learning rate of 1e-4 was used and the loss function was categorical crossentropy. Additionally, a batch size of 1 was used and the maximum epochs was set to 50 with a callback set to stop training if the validation loss hadn't improved in 5 epochs. To test the effects of semi-supervised learning on the layer segmentation, the GAN was first trained in a fully supervised fashion on the available labeled dataset. Subsequently, the network was fine-tuned using a 10:1 split of labeled and unlabeled data. When training the network using the unlabeled data, only the generator weights were updated and the predicted output of the network was rounded to the nearest integer before being used as the pseudo-label. The standard U-Net without adversarial learning was not trained with pseudo-labels as it did not converge.

4.3.5. Bruch's Membrane Opening Segmentation Network

Architecture

The architecture for this network was based on the Faster R-CNN [127] architecture. A ResNet 50 backbone was used, and the ImageNet pretrained weights were loaded. The output feature map was then used as an input for a region proposal network (RPN). Each point on the feature map is considered an anchor, and as the OCT radial frames were approximately the same size and all ground truth inputs were square, we only used a 1:1 ratio anchor of size 50x50. The RPN consisted of a 3x3 convolutional layer connected to two 1x1 convolutional layer output channels for classification and box-regression.

After regions have been proposed, region of interest (ROI) pooling is performed. The output is then fed through three 3x3 convolutional layers, averaged with a 7x7 filter and flattened. The final step is a softmax function for classification and linear regression to fix the boxes' location. A high-level overview of this network is shown in Figure 4.2.

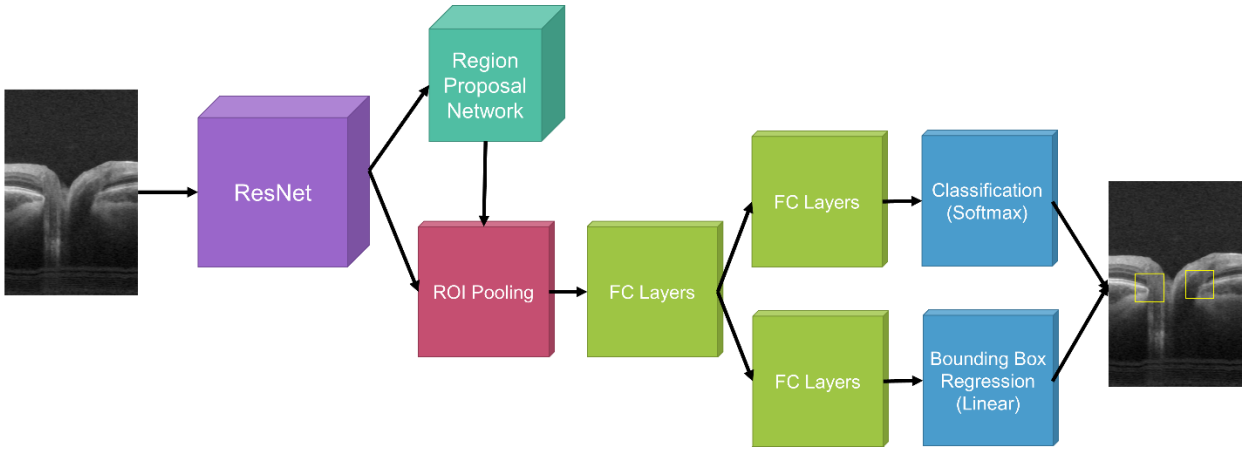

Figure 4.2 **High level overview of Faster-RCNN architecture used for BMO detection. The yellow boxes on the rightmost image represent the bounding boxes detected for the BMO, where the center of the box corresponds to the BMO.**

Training

For training, all anchors are separated into BMO and non-BMO patches based on the Intersection over Union (IoU). Anchors that overlap a manual segmentation box with an IoU larger than 0.5 are considered BMO and anchors with an IoU less than 0.1 are considered non-BMO. A mini batch was chosen to be 256 of these anchors, equally split between BMO and non-BMO classes. Non-maximum suppression is applied to ensure there is no overlapping for the proposed regions.

The Adam optimizer with a learning rate of 1e-5 was used to train the RPN and classifier layers (labelled FC Layers in Figure 4.2). The loss function was defined as the addition of the losses for the classification and bounding box regression. The classification loss was the log loss over the BMO and non-BMO classes. The box regression loss was the smooth L_1 loss for the coordinates. Three-fold cross-validation was performed on this dataset. Effort was made to keep the three different training sets equal in glaucoma/control ratio while maintaining separation of subjects in the training set to ensure that they were not in the test set.

4.3.6. Clinical Parameters

Four boundaries were extracted using the layer segmentation network: the inner limiting membrane (ILM), posterior boundary of the NFL, posterior boundary of the BM

and Choroid-Sclera (CS) boundary. To extract these boundaries, the segmentations for each B-scan in a volume were grouped into a volume. The largest 3D connected component corresponding to each tissue was assigned to that tissue and holes within the connected component were filled by the corresponding tissue. The boundaries were taken to be the first pixel in the axial direction for each tissue. Three shape characteristics were measured using these extracted layers and BMO points as previously described[112]: NFL thickness, choroidal thickness, and BMO area.

Nerve fiber layer thickness was measured at each pixel of the posterior RNFL surface as the closest distance to the ILM surface. Similarly, choroidal thickness was measured at each pixel of the posterior choroid boundary (CS boundary) as the closest distance to the BM surface. For statistical analysis, the thickness measurements were averaged over an elliptical annulus, inwardly bounded at 0.75 mm from BMO and outwardly bounded at 1.75 mm from BMO. This provided a level of anatomical consistency in averaging measurements over multiple eyes with different image and BMO sizes.

To quantify the BMO shape, segmented points on the radial frames were first transformed back to the volume scans. Erroneous points were then eliminated by removing any segmentations more than one standard deviation away from the mean axial position of the segmented BMO points. An ellipse was then fitted to the segmented BMO points by first finding the best-fit plane using principal component analysis (PCA) and fitting an ellipse to the projection of the BMO points on the plane by least-squares criterion. Bruch's membrane opening area was calculated from the fitted ellipse.

4.3.7. Analysis

The Dice similarity coefficient was used to measure the spatial overlap between the manual and automated layer segmentation. It is defined between 0 and 1, where 0 represents no overlap and 1 represents complete overlap. The Dice similarity coefficient was calculated for each tissue as follows

$$Dice(X,Y) = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (4)$$

where *X* denotes the set of pixels corresponding to the tissue in the manual segmentation, while Y denotes the set of pixels corresponding to the tissue in the automatically segmented image.

Mean Average Precision (mAP) was used to measure the accuracy of the BMO detection network. Average precision was calculated as

$$AP = \sum_{n} (R_n - R_{n-1}) P_n, \tag{5}$$

where P_n and R_n are the precision and recall at the n^{th} threshold. A prediction was considered positive if the IoU ≥ 0.5.

Clinical parameters for the volumes which had both BMO and layer manual segmentations were presented in a table including the mean and standard deviation for both manual and automated measurements. Paired, two-tailed Student's t-Tests were run to compare the means of the parameters. Bland-Altman plots were also used to evaluate the agreement between manual and automated methods. The differences between manual and automated measures were plotted against the average of both measures. The mean and standard deviation (SD) of the differences, mean of the absolute differences and 95% confidence intervals (± 1.96 SD) were calculated. Statistical significance was set at $P<0.05$ for all the tests performed.

4.4. Results

4.4.1. Layer Segmentation

To test the effects of pseudo labelling and adversarial training, the networks were trained with 800 and 8000 B-scans from 22 volumes (7 subjects, 9 eyes). The corresponding validation set was 200 and 2,000 B-scans respectively, from 4 volumes (2 subjects, 2 eyes), to total 1,000 and 10,000 B-scans used for training and validation purposes. In order to compare the effects of the different architecture and training schemes, no post-processing was done on the network outputs. The mean DICE coefficients are reported in Table 4-6. A total of 57,319 B-scans were used for testing purposes. As shown the semi-supervised (SS) Pix2Pix GAN has a higher Dice value for

the regions inside the retina and choroid. The Dice values were slightly worse for the vitreous and scleral regions in this scheme, but this was mainly due to fluctuations in noise in these regions and were simple to remove in the post-processing steps described in Section Clinical Parameters 4.3.6. Therefore, the SS Pix2Pix GAN was chosen for parameterization in Section 4.4.3.

Table 4-6. Mean DICE Coefficient for 57,319 B-scans before post-processing. The SS Pix2Pix GAN refers to the semi-supervised approach. Bolded values represent the best Dice Coefficient for that region out of all methods.

	Vitreous-ILM	ILM-RNFL	RNFL-BM	BM-CS Boundary	CS Boundary - End
U-Net					
1,000 labels	0.9291	0.7049	0.7903	0.6181	0.9359
10,000 labels	0.9694	0.8799	0.9319	0.8828	**0.9644**
Pix2Pix GAN					
1,000 labels	0.9696	0.8816	0.9296	0.8629	0.9614
10,000 labels	**0.9730**	0.8918	0.9338	0.8799	0.9641
SS Pix2Pix GAN					
1,000 labels	0.9621	0.8865	0.9308	0.8711	0.9594
10,000 labels	0.9701	**0.8958**	**0.9394**	**0.8840**	0.9626

4.4.2. BMO Segmentation

The mAP for the BMO segmentation network was 0.8547 for glaucomatous (n=21,596 manual BMO points) and 0.9567 for control subjects (n=23,052 manual BMO points). Qualitatively, the network performed quite well. Example B-scans showing BMO segmentations of both control and glaucomatous subjects are shown in Figure 4.3.

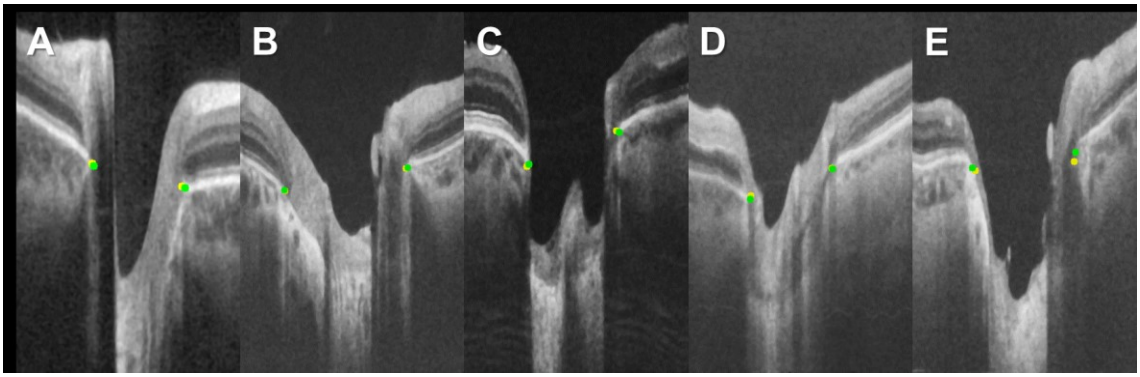

Figure 4.3 Examples of BMO segmentations on control (A,B) and glaucomatous subjects (C,D,E). Ground truth labels are shown in green, with automated segmentations in yellow.

4.4.3. Parameters

The clinical parameters extracted from the datasets are shown in Table 4-7. There was no statistical significance between the manual and automated measurements for the BMO Area, though both thickness measurements showed significantly thinner measurements. However, the coefficient of determination was above 0.97 showing excellent correlation.

Table 4-7. Mean values for the clinical parameters using both manual and automated methods.

	Mean Manual (±std)	Mean Automated (±std)	P-value	R^2
NFL Thickness (n=102)	82.65±16.06 µm	80.27±14.89 µm	< 0.05	0.971
Choroid Thickness (n=102)	147.26±57.41 µm	140.68±55.75 µm	< 0.05	0.980
BMO Area (n=281)	2.62±1.26 mm^2	2.64±1.24 mm^2	0.14	0.980

Example BMO segmentations with corresponding ground truth segmentations are overlaid on the sum-voxel, *en face* view of the OCT volumes in Figure 4.4. Manual segmentations (purple dots), the fit ellipse (green circle) and fit ellipse center (green star) as well as the automated segmentations (yellow dots), the fit ellipse (blue circle) and fit ellipse center (blue star) are shown for a control (A), high-myope control (B) and glaucomatous (C) subject.

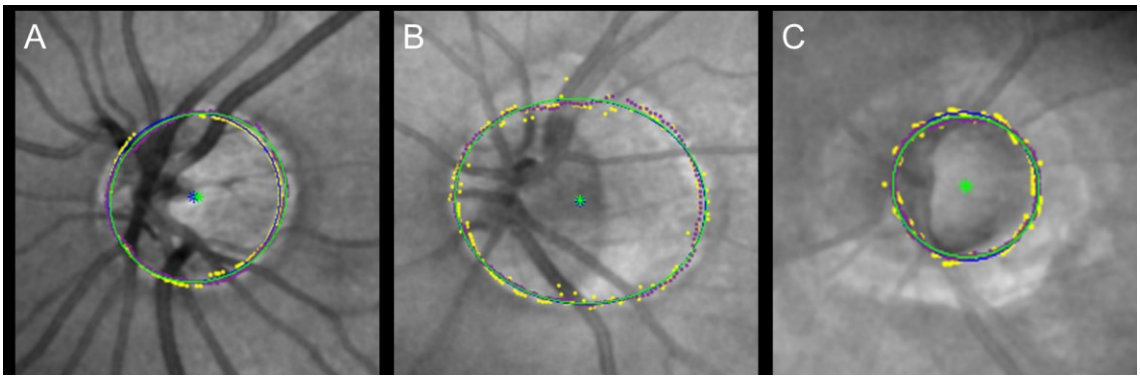

Figure 4.4 **Examples of manual and automated BMO segmentations. Manual segmentations (purple dots), the fit ellipse (green circle) and fit ellipse center (green star) as well as the automated segmentations (yellow dots), the fit ellipse (blue circle) and fit ellipse center (blue star) are shown for a control (A), high myope control (B) and glaucomatous (C) subject.**

Example automated RNFL and choroidal thickness measurements from the semi-supervised pix-to-pix GAN method are overlaid on the sum-voxel, en face view of the OCT volumes in Figure 4.5. Qualitatively, the RNFL thickness maps follow the characteristic pattern expected in the control eyes, whereas the glaucomatous eye (Figure 4.5.C) exhibits a much thinner RNFL, which is congruent with the pathophysiology of glaucoma. Additionally, the choroid is a highly vascular layer and large differences in thickness can be seen in all three eyes with no apparent pattern.

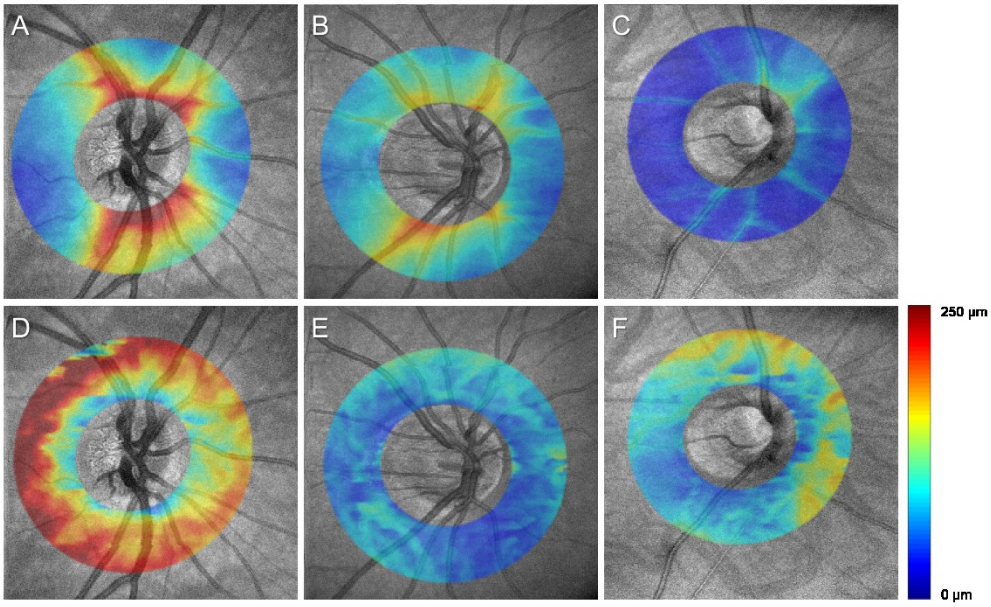

Figure 4.5 **Example images of the RNFL (A-C) and corresponding choroidal thickness (D-F) for a young control (A,D), myopic control (B,E) and glaucomatous (C,F) eye. Thickness measurements are inwardly bounded 0.25mm from the best fit BMO ellipse and outwardly bounded at 1.25mm from the best fit BMO ellipse.**

Scatter plots showing the Pearson's correlation coefficient for glaucomatous and control data points are shown in Figure 4.6. All clinical parameters showed good correlation between those extracted from automated and manual segmentations.

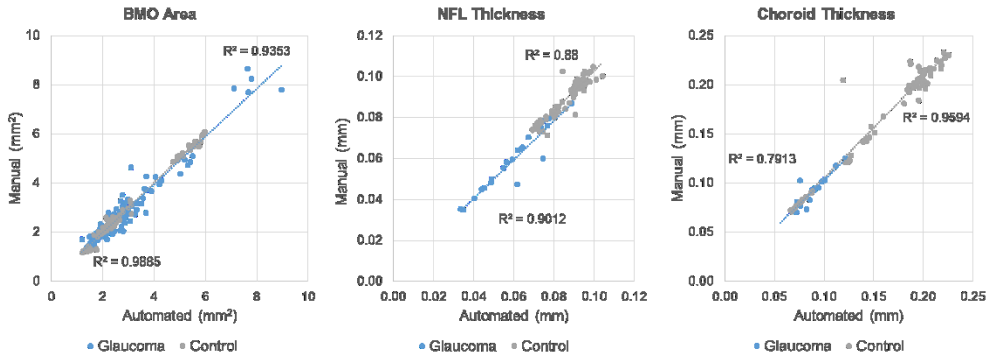

Figure 4.6 **Scatter plots for the BMO area (top), mean NFL thickness (middle,) and mean choroid thickness (bottom).**

Bland-Altman plots for the clinical parameters further confirm the high reliability of the automated measurements and are shown in Figure 4.7.

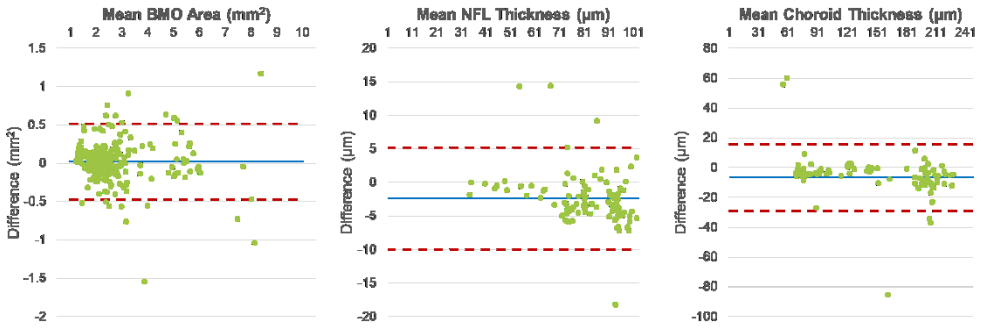

Figure 4.7 Bland-Altman Plots for the clinical parameters extracted from the datasets.

4.5. Discussion

Training neural networks to segment diseased data generally requires a large amount of manually annotated ground truth images for fully supervised learning. Although a general U-Net architecture did show good performance when trained on 1000 B-scans (the equivalent of 2.5 OCT volumes), significantly better performance was shown for both adversarial training and fine-tuning using pseudo-labels when using a small dataset. This is particularly useful for medical imaging modalities, as curating expertly segmented scans is a finite and highly limited resource. Importantly, the results demonstrate that the benefits to the semi-supervised approaches with 10x more data is only an improvement of 1-2% in performance.

Fine-tuning the network using pseudo-labelling generally improved the Dice scores for the regions of interest; however, it should be noted that the regions above the ILM and below the choroid-scleral boundary did get slightly worse after pseudo-labelling as indicated in Table 4-6. This may be due to the reinforcement of poor segmentations during the pseudo-labelling training, particularly in regions of noise such as above the ILM and below the choroid-sclera boundary. As these errors generally present themselves as smaller pockets fully encased by the correctly segmented region, they are easily filtered out during the post-processing steps described in Section 4.3.6.

It is important to address that although the Dice Similarity Coefficient reported above for the peripapillary layers was satisfactory, it did not reach as high as other papers have reported for similar layer segmentation. This may be due to several factors. First, the variation in the anatomical shape and retinal layers in the region near the ONH is much higher than in the macula. Additionally, the manual segmentations provided for

layer segmentation were completed for studies that did not consider the region inside the BMO. Therefore, manual raters were told not to correct inside the optic cup as it was not used for parameterization, leading to Dice scores inside the optic cup to be significantly lower. As such, the clinical measurements provide a better idea of the accuracy of the networks.

The clinical thickness parameters extracted from the automatically segmented volumes were shown to be slightly thinner than the manually segmented volumes. There was however, excellent correlation with the ground truth parameters, suggesting this was a small systematic difference. Additionally, through the bland-altman analysis, we see that the majority of datapoints fall within the limits of agreement. Although the thickness values were averaged to calculate a single score for ease of comparison, further 3D analysis could be done using these automated segmentations.

We were also able to show how a Faster R-CNN could be used to detect the BMO in radial OCT scans. From both the mAP and coefficient of determination values, BMO segmentation was shown to be better on control data than on glaucomatous data. The reason for this may be due to the larger number of control radial frames for training or the poorer quality of the glaucomatous dataset. However, most of the erroneous BMO segmentations were easily eliminated during post-processing (Section 4.3.6) leading to no significant difference in the BMO area parameter when comparing manual and automated methods as seen in Table 4-7, with good correlation for both glaucoma (R^2 = 0.9353) and control eyes (R^2 = 0.9885) as shown in Figure 4.6.

Future works using this pipeline may also include looking at the 3D BMO minimum rim width (BMO-MRW), a parameter that has been shown to be useful in discriminating preperimetric and perimetric glaucoma [128]. Additionally, the methods described in this paper are readily translatable to more of the retinal layers which may also be of interest to clinicians, such as the ganglion cell layer. Furthermore, though the parameters were extracted from 3D volumes, the layer segmentations were still done on 2D scans. Future work could be done to extend the work into 3D segmentations[129], allowing the network to take advantage of the volumetric nature of the images.

4.6. Conclusion

In this study, we presented a Generative Adversarial Network based method for segmenting the peripapillary ONH layers which outperformed the vanilla U-Net. Through the use of pseudo-labelling, B-scans that did not have a corresponding manual segmentation were still able to be used and provided a further increase in performance. A Faster R-CNN was also used to segment the BMO from the volumes, allowing for comparison of volumetric parameters. The BMO area was shown to have no statistically significant difference, while the thickness parameters were slightly under segmented but highly correlated.

4.6.1. Contributions

The details of the application of this architecture were conceived and designed by myself. Additionally, the idea to use and implement a Faster RCNN for the BMO segmentation was mine after previous attempts of other networks course projects with Mr. Julian Lo. The data used for this work was acquired for other studies, the majority of which I was a co-author as I either acquired and processed the data or managed the acquisition and processing of the data. Additionally, I wrote the manuscript which has been submitted for publication.

Chapter 5. Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography[3]

5.1. Abstract

To evaluate the role of ensemble learning techniques with deep learning in classifying diabetic retinopathy (DR) in optical coherence tomography angiography (OCTA) images and their corresponding co-registered structural images. Four hundred and sixty-three volumes from 380 eyes were acquired using the 3x3mm OCTA protocol on the Zeiss Plex Elite system. *En face* images of the superficial and deep capillary plexus were exported from both the OCT and OCTA data. Component neural networks were constructed using single data-types and fine-tuned using VGG19, ResNet50, and DenseNet architectures pre-trained on ImageNet weights. These networks were then ensembled using majority soft voting and stacking techniques. Results were compared to a classifier using manually engineered features. Class activation maps (CAMs) were created using the original CAM algorithm and Grad-CAM. The networks trained with the VGG19 architecture outperformed the networks trained on deeper architectures. Ensemble networks constructed using the four fine-tuned VGG19 architectures achieved accuracies of 0.92 and 0.90 for the majority soft voting and stacking methods respectively. Both ensemble methods outperformed the highest single data-type network and the network trained on hand-crafted features. Grad-CAM was shown to more accurately highlight areas of disease. Ensemble learning increases the predictive accuracy of CNNs for classifying referable DR on OCTA datasets. As the diagnostic accuracy of OCTA images is shown to be greater than the manually extracted features currently used in Literature, the proposed methods may be beneficial towards developing clinically valuable solutions for DR diagnoses.

---

[3] This work has been submitted to *Translational Vision Science & Technology*. [Morgan Heisler, Sonja Karst, Julian Lo, Zaid Mammo, Timothy Yu, Simon Warner, David Maberley, Mirza Faisal Beg, Eduardo V. Navajas, and Marinko V. Sarunic, " Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography," *TVST* (2020)]

5.2. Background

Diabetic retinopathy (DR) is a leading cause of blindness in the working age population[130] and of an estimated 285 million people worldwide with diabetes mellitus, over one third have signs of DR[21]. As patients with DR may be asymptomatic, even in late stages of the disease, it is recommended that any patient diagnosed with diabetes be screened regularly for signs of DR to palliate visual loss[24]. Retinal microvasculopathy such as microaneurysms and capillary occlusion and nonperfusion, is generally observed first[131], [132] followed by secondary inner retinal degeneration[133], [134]. Optical Coherence Tomography Angiography (OCTA) is an emerging technology that is able to provide both vascular information for detecting signs of microvasculopathy and structural information through its inherently co-registered OCT volumes to detect neurodegeneration. The majority of publications investigating the diagnostic capability of OCTA as it relates to DR focus on manually created parameters based on a priori knowledge of the disease pathophysiology. Such morphometric and functional parameters can be quite useful in classifying diseased and non-diseased retinas and techniques using manually engineered features are considered traditional machine learning. However, in recent years a trend in Deep Learning has been to identify where potential information from the images that may be discarded (or not readily observed by human perception) otherwise can be detected and used for classification with Convolutional Neural Networks (CNNs).

Convolutional Neural Networks learn through stochastic optimization, hence they are inherently limited in performance due to the high variance in predictions that results from sensitivity to small fluctuations in the training set leading to overfitting[135]. As such, large datasets are generally desired. Although large diabetic retinopathy databases (DRIVE[136], STARE[137], etc.) are publically available, they are comprised of fundus photographs, an imaging modality that does not have the ability of OCT/OCTA images to provide depth resolved images of the various retinal layers. An alternative approach to reducing the variance is to combine multiple, diverse, and accurate models to achieve greater predictive accuracy[138]. This is termed ensembling, and in general, a neural network ensemble is constructed in two steps: training a number of component neural networks and then combining the component predictions[139]. For training component neural networks, the most prevalent ensemble approaches are Bootstrap

Aggregating (Bagging) and Boosting which are algorithms that determine the training sets of component networks. Bagging[140] is a method based on bootstrap sampling[141] (sampling with replacement) that generates a number of training sets from an original training set and trains a component neural network on each sampled dataset. Boosting[142]–[144] generates a series of component neural networks whose training sets are determined by the performance of previous ones. Incorrect predictions are more heavily emphasized in the training of later networks. The networks are then combined typically by majority voting, which can be used for segmentation networks[145] as well as classification networks. Another method of combining multiple networks is stacking, whereby the networks are combined by a meta-classifier. This meta-classifier is typically a fully-connected neural network and allows for more complex, non-linear combinations of the network features.

In healthcare applications, identifying the underlying features through which the algorithm classifies disease, in addition to the quantitative algorithmic performance, is important to promote physician acceptance[2]. As such, methods to visualize the areas of images most responsible for the CNNs classification are gaining popularity. Class Activation Maps (CAMs)[146] are a common method where a heat map is generated by projecting the class specific weights of the output classification layer back to the feature maps of the last convolutional layer, thereby highlighting important regions for predicting a particular class. This method has been used in ophthalmic application previously to confirm CNN decision was based off the anterior chamber angle in categorizing angle closure[147], areas of OCT B-scans associated with various diagnoses[148], [149] and areas of segmentation error[150], and area of OCT *en face* images associated with the diagnosis of glaucoma[151]. There exists several variants of this method that build off of the original CAM paper[146], including: Grad-Cam[47], Guided Grad-Cam[47], Guided Grad-Cam++[152], and GAIN[153].

In this paper we use ensemble learning techniques together with CNNs to classify referable DR, using OCT and OCTA images. The results of the deep learning algorithms will be compared to manually extracted features. Additionally, we will show how CAMs can be used to aid in the interpretation of the CNN classification.

5.3. Materials and Methods

5.3.1. Patients

This study adhered to the tenets of the Declaration of Helsinki and was approved by the Research Ethics Boards of the University of British Columbia and Simon Fraser University. Patients with diabetes mellitus type 1 or 2, and any diabetic retinopathy severity level, as well as controls were included in the study. Patients were excluded if they had substantial media opacity which would preclude successful imaging, active inflammation, structural damage to the center of the macula, substantial non-diabetic intraocular pathologies, or any intraocular surgery with the exception of cataract surgery. A total of 380 eyes were examined from 242 subjects. 224 eyes were classified as non-referable DR, and the other 156 were classified as having referable DR by a trained ophthalmologist using fundus photography. Referable DR was classified as having more than mild non-proliferative DR or any stage DR with diabetic macular edema[18]. The mean age of patients with referable and non-referable DR was 59.3 ± 11.7 years and 58.8 ± 17.4 years, respectively.

5.3.2. Optical Coherence Tomography Data

Patients were imaged using the Zeiss PlexEite 9000 (Zeiss Meditec. Inc, Germany) with an A-scan rate of 100Khz. OCT angiography was computed using the OCT- microangiography complex algorithm (OMAG). Data was acquired as a 3x3mm volume centered on the foveal avascular zone (FAZ). A sampling rate of 300x300 was used which corresponds to a distance of 10 microns between scanning locations. Each B-scan location was scanned a total of 4 times. The A scan depth is reported as 3mm in tissue with an optical axial resolution of 6.3 μm and a transverse resolution of 20 μm. Scans were only included in the study if the system specified signal strength was 7 or higher. Figure 5.1 shows representative OCT and OCTA *en face* images for a severe NPDR subject.

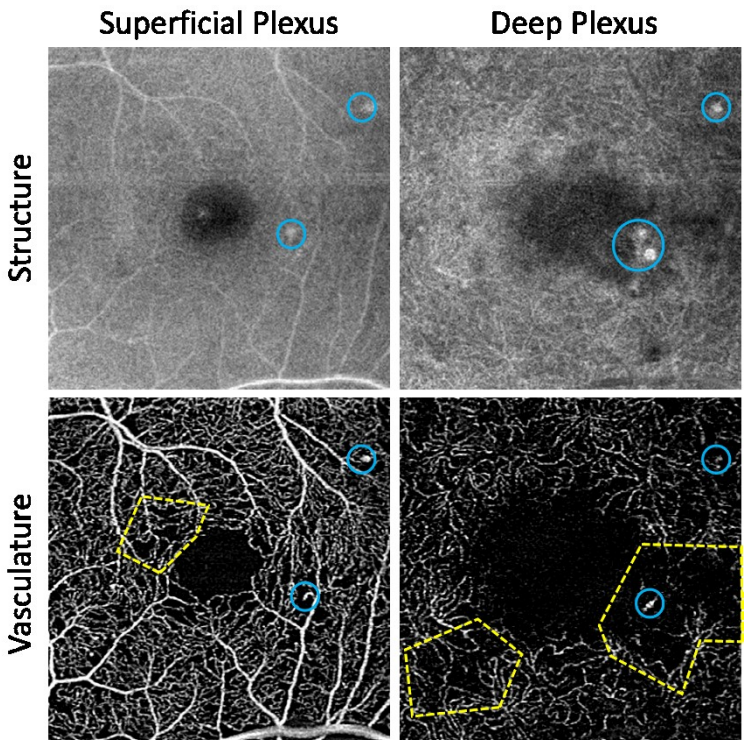

Figure 5.1 **Comparison of clinical features seen on both OCT and OCTA *en face* images of a proliferative DR patient. Dilated capillaries/microaneurysms (blue circles) are clearly visible in the superficial and deep capillary plexus of both OCT and OCTA images. Areas of capillary dropout (outlined in yellow) are more clearly seen in OCTA images, though the deep structural OCT image also shows areas of lower intensities in the larger area of nonperfusion.**

5.3.3. Manual Feature Extraction

En face OCTA images were extracted from the superficial and deep vascular complexes and projection artifacts were removed using the in-built system software before being exported. Images were then segmented using a separate vessel segmentation DNN[154] in order to calculate the handcrafted features. Methods for feature extraction have been previously reported[155]–[157], but are explained here in brief for completeness. Seven FAZ morphometric parameters were calculated from the vessel segmentation network results: area, perimeter, acircularity index, maximum and minimum diameter, axis ratio and eccentricity. The FAZ was found as the largest connected non-vessel area. The centroid for this area was then used to determine the perimeter, and maximum and minimum diameter. Acircularity index was defined as the ratio of the perimeter of the FAZ to the perimeter of a circle with equal area. Axis ratio

was the ratio of the maximum FAZ diameter to the minimum FAZ diameter, and eccentricity was calculated as the eccentricity of the ellipse made by the minimum and maximum diameters.

Five vascular parameters were also extracted from the superficial vessel segmentation network results: whole image density, inner density, central density, skeleton density and fractal dimension. Before quantification, the vessel segmentation network result was binarized using a threshold of 0.5. Whole image density was then calculated as the proportion of measured area occupied by pixels which were classified by the algorithm as a vessel. Central density was calculated as the density within the center 1mm circle, and inner density as the density in the ring between 1mm and 3mm from the center.

5.3.4. Diagnostic Network Architectures

Three different CNN architectures were used in this paper: VGG19, ResNet50, and DenseNet. Each network was loaded with the pre-trained weights on the ImageNet dataset and truncated at the deepest convolutional layer. A global average pooling (GAP) layer was then appended followed by a dense layer with two outputs.

For inputs, each base was trained with four different single data-type *en face* images extracted from the OCTA and OCT volumes. From the OCTA, both *en face* superficial and deep plexus images were extracted. Similarly, from the structural OCT volume, both *en face* superficial and deep plexus images were extracted.

The various networks were then combined to be of the configurations in Figures 5.2 and 5.3 in order to evaluate voting and stacking. For voting as in Figure 5.2, we implemented a majority soft voting scheme by averaging out the probabilities calculated by individual networks. While voting is the most common aggregation method in classification tasks, it only considers linear relationships among classifiers. Stacking as in Figure 5.3, is another ensemble technique where the meta-classifier is able to learn complex associations. As ensemble networks perform best when the networks are diverse and accurate, the best performing trained network for each input type was chosen for this architecture. The input to the meta-classifier was the concatenation of last convolutional layer of each chosen component network. This was followed by a GAP

layer, a dense layer with 1024 nodes and relu activation, a dense layer with 512 nodes and relu activation, and a final dense layer with 2 nodes for classification. The weights of all the trained convolutional bases were frozen while the meta-classifier was trained.

To compare the diagnostic capability of a feature agnostic CNN to the manually extracted features, the twelve manual features were also fed into a classifier. The classifier comprised of a multi-layer perceptron with two hidden layers of 12 and 6 nodes, respectively and one binary output for referable DR or non-referable DR. A threshold of 50% was applied to the output probabilities to determine the classification.

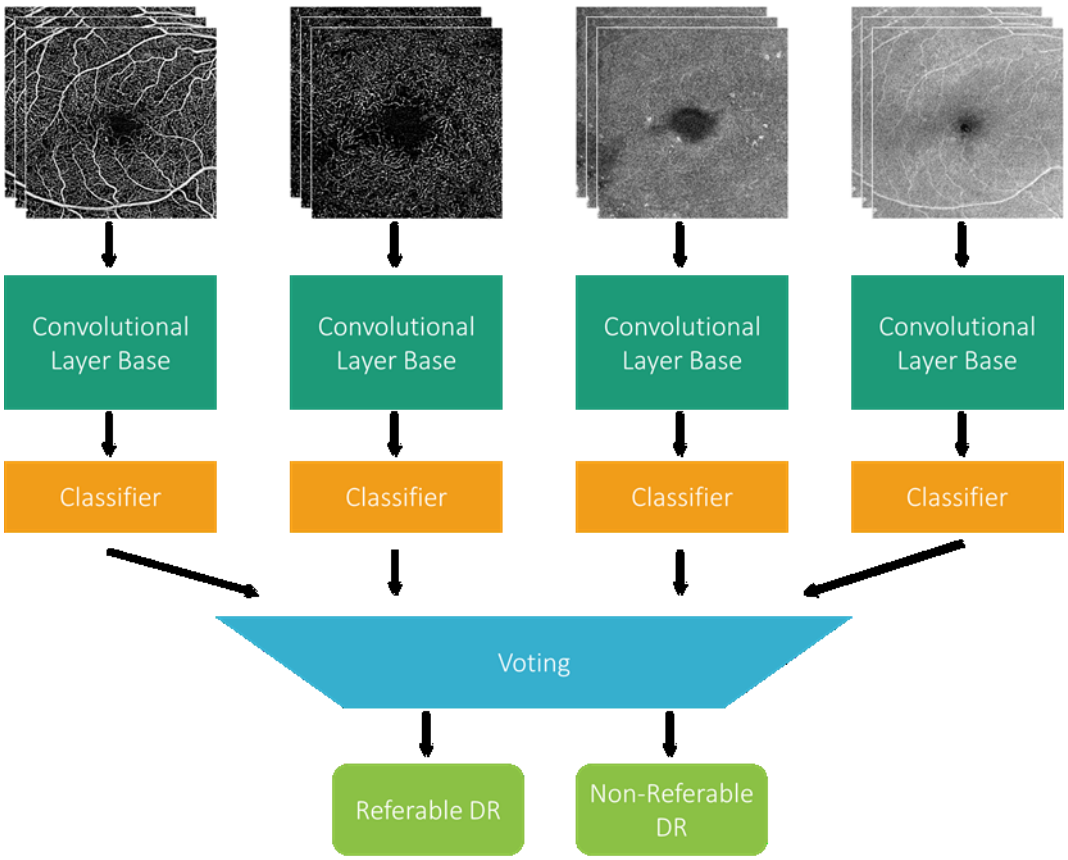

Figure 5.2 **Example of the majority voting ensemble method for combining classification results from multiple component networks. The component networks were previously trained on superficial and deep plexus *en face* images of OCT and OCTA volumes separately.**

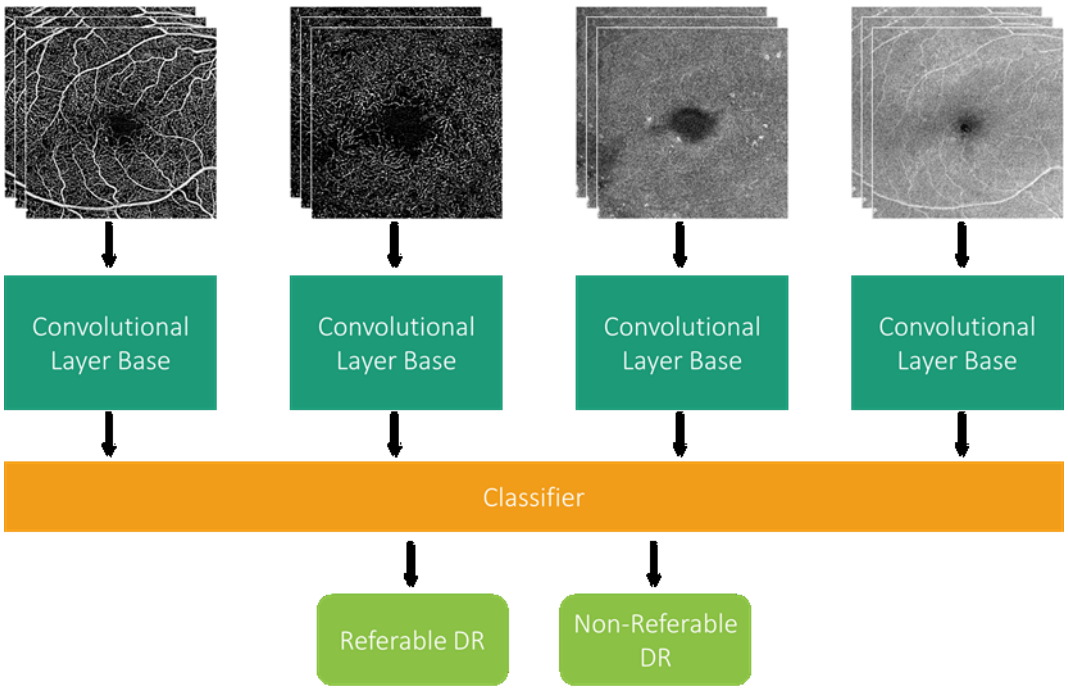

Figure 5.3 **Example of the Stacking Ensemble Method for combining classification results from multiple component networks. The component networks were previously trained on superficial and deep plexus *en face* images of OCT and OCTA volumes separately and the weights were frozen while the meta-classifier was trained.**

5.3.5. Experimental Settings

The CNN based detection method was implemented in Keras using the Tensorflow backend and Python 3.5.4. We ran the algorithm on a desktop PC with an i7-6700K CPU at 4.0 GHz, 16 GB of RAM, and a GeForce GTX 1060 GPU. 5-fold cross validation was performed on each configuration, where the data was split 60% for training, 20% for validation and 20% for test. Care was taken to ensure eyes from the same subject were only included in one of either the training, validation or testing datasets. Initially all weights in the convolutional layer base were frozen and just the two new layers comprising of the classier were trained. This was done for 10 epochs, with a learning rate of 0.00001, batch size of 8, and two callback functions set to only save the best network and to stop training if the validation loss hadn't improved after 5 epochs. Then, all weights were unfrozen and the network was retrained for 20 epochs with the same callbacks and learning rate. As suggested in the Literature[148], training from scratch on OCT images may be preferable as many of the low-level filters in networks

pre-trained on natural images are tuned to colors and OCT images are monochromatic. However, retraining the entire network pre-initialized on ImageNet provided us with better performance than training from uninitialized weights, likely due to our significantly smaller dataset. Data augmentation techniques were also used with random rotations ([-5°, +5°]), zoom (≤ 20%), height and width shift (≤ 10%) and both horizontal and vertical flipping set. In response to the unbalanced classes used for training, class weights were also assigned to the loss function to mitigate any undue bias towards the class with more training data.

5.3.6. Model Visualization

In this paper, we will compare two class activation maps: the original class activation map[146] and a variant termed Grad-Cam[47]. Class activation maps were used to help visualize the areas of the image which were most helpful in determining the classification. The original CAM method did this by modifying the network architecture to add a global average pooling layer, followed by a dense layer to the convolutional network base. Then, then CAM was calculated as a weighted sum of the feature maps per class. Grad-CAM performs a similar function by using the gradients of any class, flowing into the final convolutional layer to produce the localization map. For our purpose we have chosen to only propagate positive gradients for positive activations[158].

5.4. Results

5.4.1. Manual Features

The mean values of the manually extracted features are shown in Table 5-1. Fifteen volumes were removed from the manual feature analysis due to poor segmentation resulting in inaccurate parameterization. Two-tailed t-tests indicate that all the manually extracted features in the dataset are statistically different between means, except the axis ratio. As such, all features except axis ratio were included in the manual feature classifier.

Table 5-8. Mean Values of Manually Extracted Parameters

	Non-Referable DR (± std)	Referable DR (± std)	p-value
FAZ Parameters			
Area (mm^2)	1.340 (0.825)	1.915 (1.204)	$p < 0.01$
Perimeter (mm)	7.591 (5.070)	10.199 (5.741)	$p < 0.01$
Acircularity Index	1.835 (0.661)	2.072 (0.665)	$p < 0.01$
Minimum Diameter (mm)	0.997 (0.242)	1.088 (0.293)	$p < 0.01$
Maximum Diameter (mm)	1.545 (0.341)	1.867 (0.534)	$p < 0.01$
Axis Ratio	1.654 (0.822)	1.762 (0.429)	$p = 0.11$
Eccentricity	0.562 (0.152)	0.625 (0.147)	$p < 0.01$
Vascular Parameters			
Vessel Density	0.449 (0.041)	0.384 (0.044)	$p < 0.01$
Inner Density	0.463 (0.042)	0.395 (0.046)	$p < 0.01$
Central Density	0.261 (0.062)	0.207 (0.065)	$p < 0.01$
Skeleton Density	0.062 (0.007)	0.051 (0.007)	$p < 0.01$
Fractal Dimension	1.883 (0.014)	1.861 (0.017)	$p < 0.01$

5.4.2. Diagnostic Network Results

Table 5-2 shows the accuracy of the single input networks. Interestingly, the VGG19 networks achieve the best accuracy for all four inputs. Additionally, the superficial structural images achieve the worst performance when compared to the other input image types. Conversely, the deep structural images achieve the highest accuracy out of all the single networks when using the VGG19 architecture.

Table 5-9. Accuracy of Single Input Networks

	Deep Structural (± std)	Superficial Structural (± std)	Deep Vascular (± std)	Superficial Vascular (± std)
VGG19	87.45 (2.98)	77.57 (2.57)	85.56 (2.33)	85.76 (2.86)
ResNet50	77.76 (5.72)	67.81 (6.46)	79.25 (3.76)	76.92 (5.18)
DenseNet	71.70 (1.83)	64.51 (4.35)	76.07 (5.54)	81.70 (5.68)

Table 5-3 reports the accuracy, sensitivity and specificity of the ensembled networks as well as the network classifier the manual features. Both ensemble methods achieved higher accuracy than the manual feature classifier.

Table 5-10. Comparison of Ensembled Networks to Manual Feature Classifier

	Majority Voting	Stacking	Manual Features
Accuracy	92.00 (1.92)	89.86 (2.55)	83.10 (4.89)
Sensitivity	90.41 (6.23)	87.38 (5.85)	69.26 (9.02)
Specificity	93.33 (5.18)	92.09 (5.16)	78.42 (6.32)

Table 5-4 reports the accuracy, sensitivity and specificity of a 3-channel input ensemble network and 3-channel input VGG19 network. As the superficial structural images showed the lowest diagnostic accuracy in the single networks, and we used networks including pre-trained ImageNet weights which require a 3-channel input, the networks in these results did not include the superficial structural images. Additionally, as majority voting outperformed the stacking method this was chosen for comparison to a standard single VGG19 network.

Table 5-11. Comparison of 3 Channel Input Networks

	Majority Voting	VGG-19
Accuracy	90.71 (1.65)	87.70 (3.41)
Sensitivity	93.32 (5.34)	94.20 (3.13)
Specificity	87.74 (4.88)	80.53 (6.60)

5.4.3. Model Visualization

A comparison of representative CAM and Grad-CAM visualizations are shown in Figure 5.4 for the case of a subject with DR. As shown, the Grad-CAM image is better able to focus on features associated with the diseased regions. The vessel thickening above the FAZ, and region of capillary dropout to the left of the FAZ are shown to be more predictive of disease than the relatively normal looking vasculature more peripherally. An additional diabetic patient (severe DR) is shown in Figure 5.5 along with the Grad-Cam images for the superficial and deep plexus of both structural OCT and OCTA images. These images highlight that for each input image, the networks are searching for distinctly different features for classification. In the structural images (Figure A-B), the most attention is given to regions of fluid and the hard exudates surrounding that region. For the OCTA images, the region of greatest influence is

centered on the larger microaneuryms in the images. Conversely, for the control images shown in Figure 5.6, the whole parafoveal zone is shown to be influential.

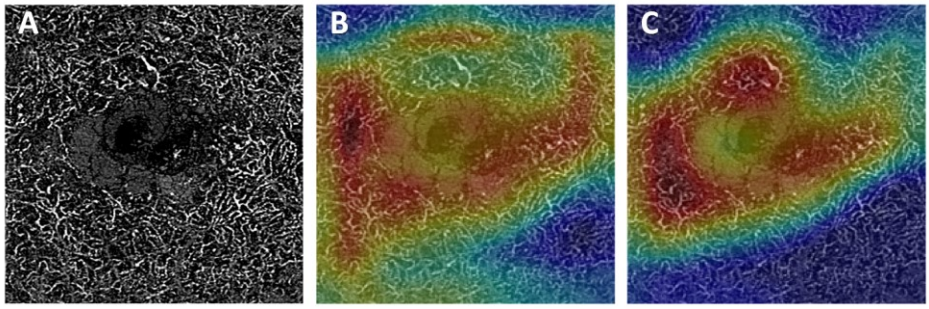

Figure 5.4 **A deep plexus *en face* image of a DR subject (A) and the corresponding heat maps using the original CAM method (B) and Grad-CAM (C). As shown by the smaller, more focal regions of warmer colors, the Grad-CAM image is able to localize on areas of disease better than the original CAM method. The vessel thickening above the FAZ, and region of capillary dropout to the left of the FAZ are shown to be more predictive of disease than the relatively normal looking vasculature more peripherally.**

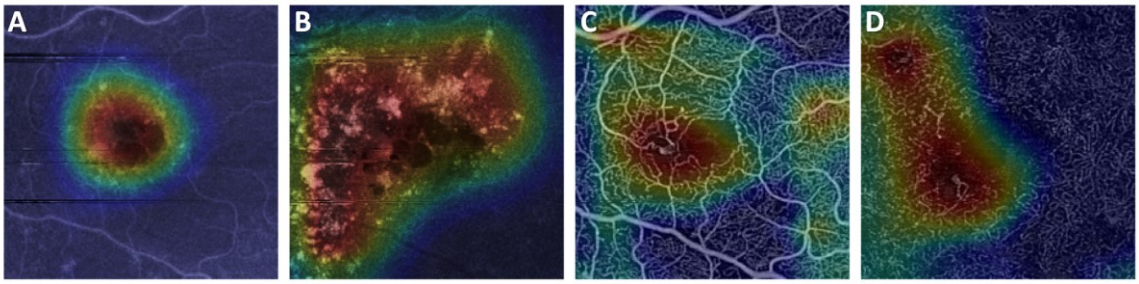

Figure 5.5 **Grad-CAMs for the superficial structural image (A), deep structural image (B), superficial angiography image (C) and deep vasculature image (D) of a Severe DR patient. Hard exudates and regions of fluid are highlighted in the structural images. Microaneurysms and regions of capillary dropout are highlighted in the vascular images.**

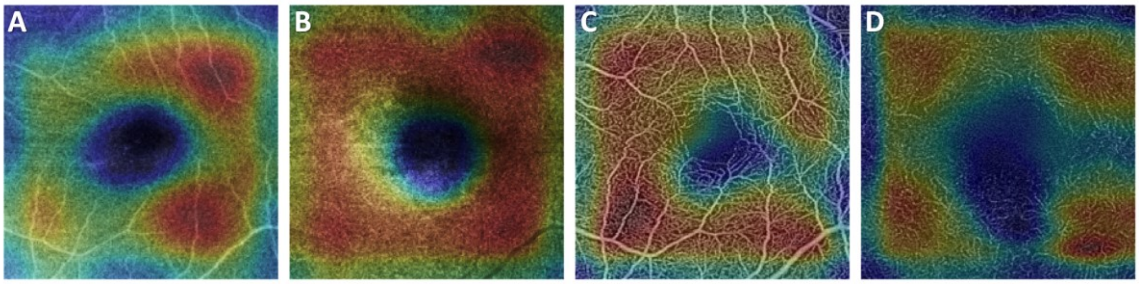

Figure 5.6 **Grad-CAMs for the superficial structural image (A), deep structural image (B), superficial angiography image (C) and deep vasculature image (D) of a Control patient. Regions of higher uniform intensity in the structural images, and regions of normal vasculature tend to have a greater effect on the classification.**

In cases of misclassification, the Grad-CAMs resembled the typical pattern of the incorrect classification. For a referable NPDR patient, as in Figure 5.7, the Grad-CAMs show the non-referable characteristic pattern of a brighter parafoveal zone with the FAZ being less influential. For a case of mild non-referable DR, more focal regions of potential vessel dropout on the temporal side of the fovea are shown to be of high importance for the images which were misclassified as referable DR (Figure 5.8 F-H).

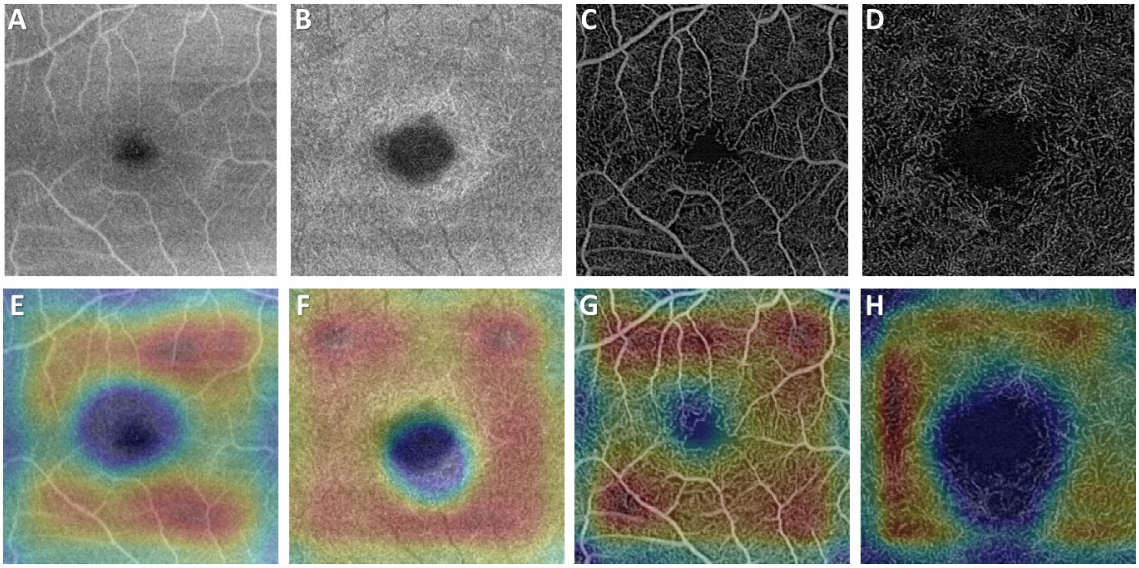

Figure 5.7 **The superficial structural image (A), deep structural image (B), superficial angiography image (C) and deep vasculature image (D) of a referable NPDR patient which was misclassified as non-referable DR and the corresponding Grad-CAMs (E-H). The typical non-referable DR pattern of a brighter parafoveal region is observed in all Grad-CAMs.**

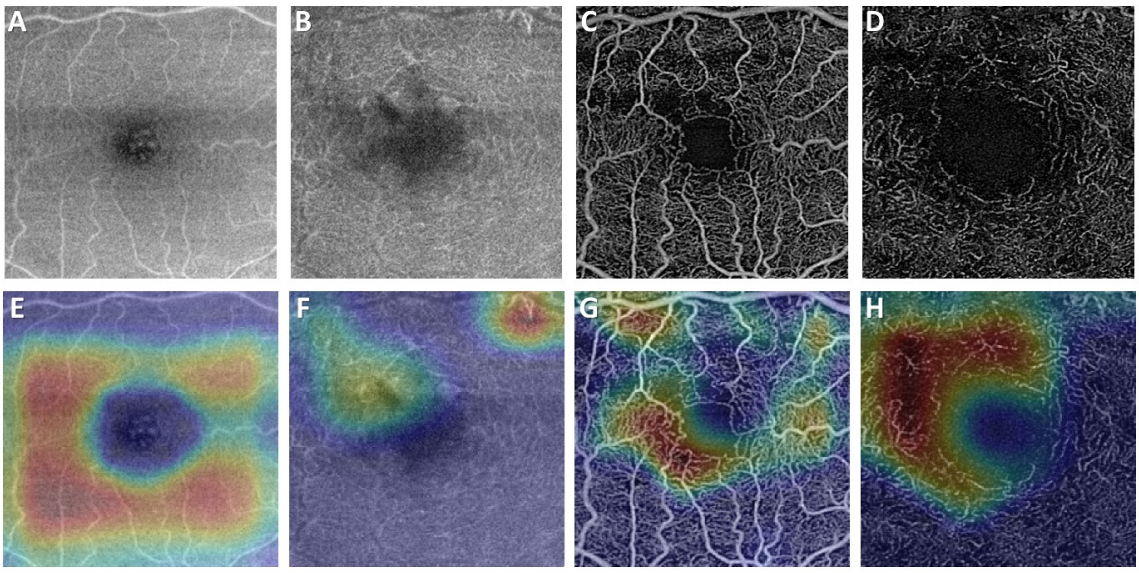

Figure 5.8 **The superficial structural image (A), deep structural image (B), superficial angiography image (C) and deep vasculature image (D) of a non-referable mild DR patient which was misclassified as referable DR and the corresponding Grad-CAMs (E-H). Focal regions of potential vessel dropout on the temporal side of the fovea are shown to be of high importance in the images which were misclassified as referable DR (F-H), whereas the superficial structural image shows the characteristic parafoveal pattern as it was the only image correctly classified.**

5.5. Discussion

As retinal imaging systems continue to improve, so does our ability to see hallmark features of DR. Optical coherence tomography angiography allows clinicians to view depth-resolved sections of the retina for both structural clues, as well as vascular. As a result, OCTA may enable accurate detection of DR if the right features are used for classification. In this paper, we compare the classification accuracy of features automatically learned from single plexus *en face* images, combinations of these learned features, and hand-crafted features. Insight to the features learned by the CNN are highlighted through CAM heat maps.

It was shown in Tables 5.2 and 5.3 that a combination of diverse component networks provided a higher accuracy than single component networks alone. Both majority voting and stacking techniques yielded higher accuracies. For the component networks, which were trained on *en face* images from a single plexus, the shallower

network of VGG19 performed better on this task than the deeper state-of-the-art networks. While there is some precedent for this[159], it is expected that the ResNet or DenseNet architectures would perform better with more data. Although pre-trained ResNet18 ImageNet weights are not currently available for use with keras, this architecture may perform more similarly or better than the VGG19 architecture. The superficial structural images had the worst performance, which may be due to their relatively homogenous appearance, even in diseased states. This lack of texture can make it difficult for CNNs to learn features. Additionally, the deep structural images appear to achieve accuracies on par or better than the vascular images, suggesting that OCT structural information should also be taken into account while analyzing OCTA volumes.

There exists strong evidence in the Literature that hand-crafted features extracted from OCT and OCTA images are able to differentiate between grades of DR[160]–[162]. One study[163] that looked at combining hand-crafted features from both the superficial and deep plexus resulted in an overall accuracy, sensitivity and specificity of 94.3, 97.9, and 87.0, when classifying between controls and mild DR patients. This paper extracted parameters for the vessel density, blood vessel calibre and width of the FAZ and used an SVM for the classifier. Another recent paper[164] uses both deep learning and manually extracted features to detect DR in OCT. For their network, they combined both handcrafted features and learned features to differentiate between grades 0 and 1 DR with an accuracy, sensitivity and specificity of 0.92, 0.90 and 0.95, respectively. It is important to note that with manually extracted features there is the ability for errors to propagate where errors which arise during the segmentation and parameterization phase to affect the classification. Future work could use an ensemble of both hand-crafted features and learned features to enhance performance.

To compare whether ensemble techniques achieved better performance than a standard CNN, the three grayscale images with the highest diagnostic accuracy in Table 5.2 were chosen to create an RGB image as input for a VGG-19 network. When compared to an ensemble network with the same inputs, the ensemble network achieved a higher accuracy and specificity as shown in Table 5.4. Additionally, although the superficial structural image had a significantly worse diagnostic accuracy on its own, the fact that the 4 channel input ensemble network outperformed the 3 channel input

shows that it still holds important information for the classification of DR and should be considered in future works.

The Class Activation Maps, and Grad-CAM in particular, showed good localization of the biomarkers associated with DR, thereby increasing the interpretability of the CNN results. In correctly classified referable DR images, areas of hard exudates, DME, microaneurysms and capillary dropout show higher activation. The interpretability, and subsequent physician trust in the networks, could also be increased by utilizing Grad-CAM in the choice of layers in which to ensemble as done in a recent paper by Liu et al[165]. Grad-CAMs were created for each feature layer in networks trained to detect pseudo-progression of glioblastoma multiform, and a team of three specialists chose the most discriminating layer from each network with which to ensemble. This technique could be used in future work to increase both algorithm performance and physician uptake of the innovation.

While this study demonstrates the ability of CNNs to classify DR in OCTA with high accuracy there are notable limitations. First, the use of ensemble learning methods greatly increases the computational cost as it requires the training of multiple networks. This increases both training time as well as the data size of the final model, though this could be partially alleviated through training the component networks in parallel. Another limitation includes the restricted dataset size, however this limitation was mitigated through the use of fine-tuning, data augmentation, and class-weighting of the loss function. The authors note that improved performance could be achieved through a larger dataset. Furthermore, the dataset only included images with a signal strength of 7 or above, which is sometimes infeasible in patients with pathology. The dataset also consists of images from only one machine, thereby potentially limiting the network's performance on other OCTA machines. Additionally, as only one clinician was used for grading and disagreement between clinicians can sometimes occur, inaccurate diagnoses may have impacted the results. Future work could endeavor to further stratify the DR classification.

5.6. Conclusion

To the best of the authors' knowledge, this is the first paper to classify DR using OCTA *en face* images, as opposed to manually extracted features. There is still ongoing

research as to how to fully utilize the data rich volumes while producing results for clinicians which are clinic-friendly and not overwhelming due to data overload. By utilizing the whole en-face image as opposed to several manually extracted features which attempt to capture relevant information, the DNN can use more information to decide what is truly relevant. As such, the diagnostic accuracy of *en face* images for this application is higher than manually extracted features, which we believe has significant clinical impact as OCTA continues to rise in popularity among clinicians.

5.6.1. Contributions

This study was done in close collaboration with our clinical counterparts, especially Dr. Sonja Karst and Dr. Eduardo Navajas who provided invaluable expertise when designing the experiment. I was involved in this study from the start, where we initially used Zeiss AngioPlex data from another clinic, which I extracted from the clinical imaging machine and curated. The details of the methods, including the use of ensemble learning, were conceived and designed by myself. I wrote the code to extract the majority of the manual parameters once segmented by a blood vessel segmentation network. All machine learning scripts were written by me, but I did supervise Mr. Timothy Yu in his exploration of optimizing hyperparameters and network architectures for the manual parameter network. Additionally, I wrote the manuscript which was submitted to *Translational Vision Science & Technology* and has been accepted for publication.

Chapter 6. Conclusions

6.1. Summary

The main purpose of this thesis is to investigate translational image analysis, from an engineering laboratory to clinical application, for newly developed OCT technologies using machine learning. The small size of data sets acquired with clinical prototype imaging instruments presents a unique challenge when translating to a clinical setting, especially when using deep learning which typically requires a vast amount of labeled data for supervised training. The works described in Chapters 3-5, describe the evaluation of several different techniques to mitigate the need for large annotated datasets.

First, the basic concepts of the pertinent eye anatomy, ophthalmic imaging techniques, and machine learning were described. An extensive overview of all the potential subtopics for each category is beyond the scope of a thesis, but the relevant information as it pertains to this thesis was described. Multiple review articles on deep learning in ophthalmology are available in the Literature, for example [2], [8], [9].

The first study looked into the concept of transfer learning through fine-tuning networks trained from a different domain (adaptive optics scanning laser ophthalmoscopy) to the domain of interest (adaptive optics OCT) to detect cone photoreceptors. Through this study we were able to show that transfer learning with a smaller dataset performed comparably to a network trained solely on the data of interest with a much larger dataset. As the field moved away from patch-based segmentation networks to fully convolutional neural networks with the wide-spread adoption of U-Net, so did we with our next investigations, applying transfer learning where applicable.

Using volumes of the optic nerve head, we were able to show that adversarial learning and pseudo-labelling advanced the performance of a vanilla U-Net for the retinal layers. These techniques are important for clinical imaging analysis, as it is typically more difficult (resource intensive) to obtain high quality annotations than it is to obtain the images themselves. With the pseudo-labelling technique we were able to still use all the acquired images, even without manual segmentations. Along with a Faster R-CNN approach to segment the Bruch's membrane opening, we were able to apply these

techniques to the 3D OCT volumes and show good correlation between automated and manual clinical parameters.

Finally, the more challenging task of diagnostics with limited data was investigated using diabetic retinopathy OCTA data. OCTA volumes are inherently extremely data rich as they are not only 3D volumes themselves but are processed from 3D OCT intensity images, which also provide features of interest when studying diabetic retinopathy. As such, ways to use all the information contained in these volumes is of high interest. Through ensembling networks for four extracted *en face* images, we were able to use a relatively small amount of data to diagnose referable and non-referable diabetic retinopathy.

In summary, the major contributions of this work are (1) the translation of machine learning / deep learning based tools to the clinic and (2) the development and demonstration of approaches using transfer learning and semi-supervised learning to reduce the amount of clinical data required for training artificial neural networks.

6.2. Future Work

While several deep learning techniques were investigated for the translation of state-of-the-art OCT systems to the clinic, there exist several topics that can be considered for future work.

The results presented in Chapter 3 for the automatic segmentation of retinal cones was applied to control (non-diseased) image data. A first and natural extension of the cone counting algorithm is to apply it to pathological data. Figure 6.1 shows a representive *en face* image acquired from a patient with Acute Zonal Occult Outer Retinopathy (AZOOR). When compared to the cone mosaic of the control subjects in Chapter 3, there are obvious regions where the homogeneity of the mosaic is disturbed. Although an AZOOR patient is shown as an example, there exist several pathologies which affect the photoreceptor layer such as age-related macular degeneration, Stargardt's disease, retinitis pigmentosa, and achromotopsia. Our unique position of having an AO-OCT system in a retinal specialist’s clinic puts us at an advantageous spot to investigate DNN methods for the analysis of pathological datasets. Future technical work for this project could also be to update the patch-based method to a fully-

convolutional network, like the one described by Hamwood et al. [166], thereby allowing the analysis of an image with just one inference.

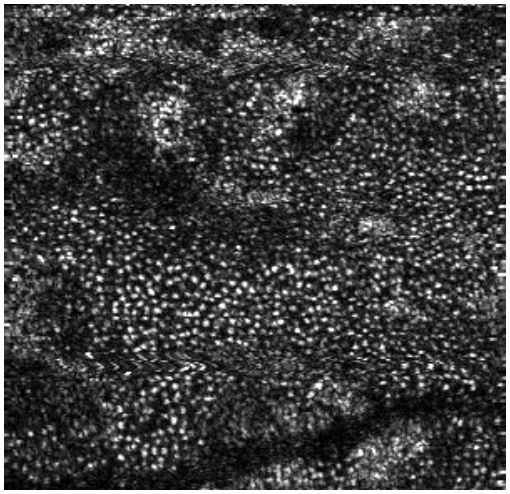

Figure 6.1 ***En face*** **image of a patient with Acute Zonal Occult Outer Retinopathy (AZOOR). Patches of typical cones are interspersed with patches of pathological areas.**

Additionally, extending on the theme of research presented in Chapter 5, there are several other ways one could approach the diagnosis of diabetic retinopathy using OCTA. For example, it is hypothesized that early changes in diabetic retinopathy occur in the deep capillary plexus of the retinal periphery. As such, using wider fields of view may be of benefit. Furthermore, class activation maps are a great start to aiding in the clinical translation of Deep Learning technology, though they may be insufficient. Therefore, another approach may be to use segmentation and identification networks such as Mask R-CNN to detect the same clinical landmarks clinicians use for fundus images (cotton-wool spots, venous beading, intraretinal microvascular abnormalities, micro aneurysms, etc.) but in OCT/OCTA volumes. This would enable to same criteria to be used for classification and may ease the adoption of such technology into the clinic. This would be a challenging technical task however as some features such as micro aneurysms have been shown to be undetectable by manual raters in OCTA images, and may be even more difficult to detect in larger fields of view if high resolution is not maintained.

Furthermore, as the tools and techniques discussed in this thesis worked towards solving the limited data case, it should be acknowledged that as these instruments continue to collect more data other questions can begin to be addressed. One in

particular is how the data and deep learning tools can be used for prognostics. As longitudinal image data for individual patients, along with pertinent demographic and progression details are curated, we should start to see a rise in studies looking to predict which patients will progress faster than others, or which patients will respond to certain treatments. These questions cannot currently be answered with limited datasets however, and may also require further improvements to the current hardware.

The examples of future work thus far have been relatively small steps from the current state of the work in the field. Looking further into the future, machine learning truly has the potential to aid in much more than the segmentation of image features or even the diagnosis of diabetic retinopathy. As the algorithms and hardware continue to improve, one could imagine a world where every imaging system comes equipped with its own suite of AI tools that not only aid in the diagnosis of diseases, but also aid in the individualized treatment plan and prognosis of patients.

References

[1] "Understanding the Epidemiology of Vision Loss and Impairment in the United States - Making Eye Health a Population Health Imperative - NCBI Bookshelf." [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK402366/. [Accessed: 22-Feb-2020].

[2] D. S. W. Ting *et al.*, "Artificial intelligence and deep learning in ophthalmology," *British Journal of Ophthalmology*. 2019.

[3] M. Melinscak, P. Prentasic, and S. Loncaric, "Retinal Vessel Segmentation Using Deep Neural Networks," in *International Conference on Computer Vision Theory and Applications (VISAPP 2015)*, 2015, pp. 577–582.

[4] S. Li, D. Cunefare, C. Wang, R. H. Guymer, L. Fang, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express*, 2017.

[5] A. Shah, L. Zhou, M. D. Abrámoff, and X. Wu, "Multiple surface segmentation using convolution neural nets: application to retinal layer segmentation in OCT images," *Biomed. Opt. Express*, 2018.

[6] P. Prentasic *et al.*, "Segmentation of the Foveal Microvasculature Using Deep Learning Networks," *J. Biomed. Opt.*, vol. 21, no. 7, p. 075008, Jul. 2016.

[7] Y. Guo, A. Camino, J. Wang, D. Huang, T. S. Hwang, and Y. Jia, "MEDnet, a neural network for automated detection of avascular area in OCT angiography," *Biomed. Opt. Express*, vol. 9, no. 11, p. 5147, Nov. 2018.

[8] A. Lee, P. Taylor, and J. Kalpathy-Cramer, "Machine Learning Has Arrived!," *Ophthalmology*, vol. 124, no. 12. Elsevier Inc., pp. 1726–1728, Dec-2017.

[9] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, "Artificial intelligence in retina," *Progress in Retinal and Eye Research*, vol. 67. Elsevier Ltd, pp. 1–29, 01-Nov-2018.

[10] D. S. W. Ting *et al.*, "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *JAMA - J. Am. Med. Assoc.*, 2017.

[11] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA - J. Am. Med. Assoc.*, 2016.

[12] M. D. Abràmoff *et al.*, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Investig.*

Ophthalmol. Vis. Sci., 2016.

[13] R. Gargeya and T. Leng, “Automated Identification of Diabetic Retinopathy Using Deep Learning,” *Ophthalmology*, 2017.

[14] Z. Li, Y. He, S. Keel, W. Meng, R. T. Chang, and M. He, “Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs,” *Ophthalmology*, 2018.

[15] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, “Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks,” *JAMA Ophthalmol.*, 2017.

[16] F. Grassmann *et al.*, “A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography,” *Ophthalmology*, 2018.

[17] J. M. Brown *et al.*, “Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks,” in *JAMA Ophthalmology*, 2018.

[18] M. D. Abràmoff, P. T. Lavin, M. Birch, N. Shah, and J. C. Folk, “Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices,” *npj Digit. Med.*, vol. 1, no. 1, p. 39, Dec. 2018.

[19] P. E. Z. Tan *et al.*, “Quantitative Comparison of Retinal Capillary Images Derived By Speckle Variance Optical Coherence Tomography With Histology.,” *Investig. Ophthalmol. Vis. Sci.*, vol. 56, no. 6, pp. 3989–3996, 2015.

[20] J. P. Campbell *et al.*, “Detailed Vascular Anatomy of the Human Retina by Projection-Resolved Optical Coherence Tomography Angiography,” *Sci. Rep.*, 2017.

[21] J. W. Y. Yau *et al.*, “Global prevalence and major risk factors of diabetic retinopathy,” *Diabetes Care*, vol. 35, no. 3, pp. 556–564, 2012.

[22] S. Resnikoff *et al.*, “Global data on visual impairment in the year 2002.,” *Bull. World Health Organ.*, vol. 82, no. 11, pp. 844–51, 2004.

[23] M. M. Nentwich and M. W. Ulbig, “Diabetic retinopathy - ocular complications of diabetes mellitus.,” *World J. Diabetes*, vol. 6, no. 3, pp. 489–99, 2015.

[24] D. A. Antonetti, R. Klein, and T. W. Gardner, “Diabetic Retinopathy,” *N Engl J Med*, vol. 366, pp. 1227–1239, 2012.

[25] C. P. Wilkinson *et al.*, “Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales,” *Ophthalmology*, vol. 110, no. 9,

pp. 1677–1682, Sep. 2003.

[26] G. Bresnick, G. De Venecia, F. Myers, J. Harris, and M. Davis, “Retinal ischemia in diabetic retinopathy,” *Arch. Ophthalmol.*, vol. 93, no. 12, pp. 1300–1310, 1975.

[27] G. A. Lutty, “Effects of diabetes on the eye,” *Investig. Ophthalmol. Vis. Sci.*, vol. 54, no. 14, 2013.

[28] O. Arend, S. Wolf, A. Harris, and M. Reim, “The Relationship of Macular Microcirculation to Visual-Acuity in Diabetic-Patients,” *Arch. Ophthalmol.*, vol. 113, no. 5, pp. 610–614, 1995.

[29] C. Balaratnasingam *et al.*, “Visual Acuity Is Correlated with the Area of the Foveal Avascular Zone in Diabetic Retinopathy and Retinal Vein Occlusion,” *Ophthalmology*, 2016.

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

[31] H. A. Quigley and A. T. Broman, “The number of people with glaucoma worldwide in 2010 and 2020.,” *Br. J. Ophthalmol.*, vol. 90, no. 3, pp. 262–267, Mar. 2006.

[32] Y. H. H. Kwon, J. H. H. Fingert, M. H. H. Kuehn, and W. L. M. L. M. Alward, “Primary open-angle glaucoma.,” *N. Engl. J. Med.*, vol. 360, no. 11, pp. 1113–1124, Mar. 2009.

[33] E. J. Lee, T.-W. Kim, M. Kim, and H. Kim, “Influence of lamina cribrosa thickness and depth on the rate of progressive retinal nerve fiber layer thinning.,” *Ophthalmology*, vol. 122, no. 4, pp. 721–729, Apr. 2015.

[34] M. Heisler *et al.*, “Automated identification of cone photoreceptors in adaptive optics optical coherence tomography images using transfer learning.,” *Biomed. Opt. Express*, vol. 9, no. 11, pp. 5353–5367, Nov. 2018.

[35] M. S. Mahmud *et al.*, “Review of speckle and phase variance optical coherence tomography to visualize microvascular networks.,” *J. Biomed. Opt.*, vol. 18, no. 5, p. 50901, 2013.

[36] A. Zhang, Q. Zhang, C.-L. Chen, and R. K. Wang, “Methods and algorithms for optical coherence tomography-based angiography: a review and comparison,” *J. Biomed. Opt.*, vol. 20, no. 10, p. 100901, 2015.

[37] I. Gorczynska, J. V. Migacz, R. J. Zawadzki, A. G. Capps, and J. S. Werner, “Comparison of amplitude-decorrelation, speckle-variance and phase-variance OCT angiography methods for imaging the human retina and choroid,” *Biomed. Opt. Express*, vol. 7, no. 3, pp. 911–942, 2016.

[38] Y. Jia *et al.*, “Split-spectrum amplitude-decorrelation angiography with optical coherence tomography.,” *Opt. Express*, vol. 20, no. 4, pp. 4710–25, 2012.

[39] J. Xu *et al.*, “Retinal angiography with real-time speckle variance optical coherence tomography,” *Br. J. Ophthalmol.*, pp. 1–5, 2015.

[40] Z. Mammo *et al.*, “Quantitative Noninvasive Angiography of the Fovea Centralis Using Speckle Variance Optical Coherence TomographySpeckle Variance Optical Coherence Tomography of Macula,” *Investig. Ophthalmol. Vis. Sci.*, vol. 56, no. 9, pp. 5074–5086, 2015.

[41] K. R. Mendis *et al.*, “Correlation of histologic and clinical images to determine the diagnostic value of fluorescein angiography for studying retinal capillary detail.,” *Invest. Ophthalmol. Vis. Sci.*, vol. 51, no. 11, pp. 5864–5869, Nov. 2010.

[42] G. Chan *et al.*, “In vivo optical imaging of human retinal capillary networks using speckle variance optical coherence tomography with quantitative clinico-histological correlation,” *Microvasc. Res.*, vol. 100, pp. 32–39, 2015.

[43] P. K. Yu *et al.*, “Label-Free Density Measurements of Radial Peripapillary Capillaries in the Human Retina,” *PLoS One*, vol. 10, no. 8, p. e0135151, 2015.

[44] MathWorks, “Deep Learning.” [Online]. Available: https://www.mathworks.com/discovery/deep-learning.html. [Accessed: 26-Feb-2019].

[45] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *CoRR*, vol. abs/1505.0, 2015.

[46] I. J. Goodfellow *et al.*, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, 2014, vol. 3, no. January, pp. 2672–2680.

[47] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017.

[48] B. Zhou, A. Khosla, L. A., A. Oliva, and A. Torralba, “Learning Deep Features for Discriminative Localization.,” *CVPR*, 2016.

[49] D. R. Williams, “Imaging single cells in the living retina,” *Vision Research*, vol. 51, no. 13. pp. 1379–1396, 2011.

[50] J. Liang, D. R. Williams, and D. T. Miller, “Supernormal vision and high-resolution retinal imaging through adaptive optics,” *J. Opt. Soc. Am. A*, vol. 14, no. 11, p. 2884, Nov. 1997.

[51] M. Pircher and R. J. J. Zawadzki, “Review of adaptive optics OCT (AO-OCT):

principles and applications for retinal imaging.," *Biomed. Opt. Express*, vol. 8, no. 5, pp. 2536–2562, May 2017.

[52] A. Roorda, F. Romero-Borja, W. J. D. III, H. Queener, T. J. Hebert, and M. C. W. Campbell, "Adaptive optics scanning laser ophthalmoscopy," *Opt. Express*, vol. 10, no. 9, pp. 405–412, May 2002.

[53] D. Merino, J. L. Duncan, P. Tiruveedhula, and A. Roorda, "Observation of cone and rod photoreceptors in normal subjects and patients using a new generation adaptive optics scanning laser ophthalmoscope," *Biomed. Opt. Express*, vol. 2, no. 8, p. 2189, Aug. 2011.

[54] J. Liang, D. R. Williams, and D. T. Miller, "Supernormal vision and high-resolution retinal imaging through adaptive optics," *J. Opt. Soc. Am. A*, vol. 14, no. 11, p. 2884, Nov. 1997.

[55] B. Hermann *et al.*, "Adaptive-optics ultrahigh-resolution optical coherence tomography.," *Opt. Lett.*, vol. 29, no. 18, pp. 2142–2144, Sep. 2004.

[56] O. P. Kocaoglu, Z. Liu, F. Zhang, K. Kurokawa, R. S. Jonnal, and D. T. Miller, "Photoreceptor disc shedding in the living human eye," *Biomed. Opt. Express*, vol. 7, no. 11, p. 4554, Nov. 2016.

[57] M. J. Ju, M. Heisler, D. Wahl, Y. Jian, and M. V. Sarunic, "Multiscale sensorless adaptive optics OCT angiography system for in vivo human retinal imaging," *J. Biomed. Opt.*, vol. 22, no. 12, p. 1, 2017.

[58] M. Salas *et al.*, "Compact akinetic swept source optical coherence tomography angiography at 1060 nm supporting a wide field of view and adaptive optics imaging modes of the posterior eye," *Biomed. Opt. Express*, vol. 9, no. 4, p. 1871, Apr. 2018.

[59] A. Dubra *et al.*, "Noninvasive imaging of the human rod photoreceptor mosaic using a confocal adaptive optics scanning ophthalmoscope," *Biomed. Opt. Express*, vol. 2, no. 7, p. 1864, Jul. 2011.

[60] T. Y. Chui, H. Song, and S. A. Burns, "Adaptive-optics imaging of human cone photoreceptor distribution," *J. Opt. Soc. Am. A*, vol. 25, no. 12, p. 3021, Dec. 2008.

[61] O. P. Kocaoglu *et al.*, "Imaging cone photoreceptors in three dimensions and in time using ultrahigh resolution optical coherence tomography with adaptive optics," *Biomed. Opt. Express*, vol. 2, no. 4, p. 748, Apr. 2011.

[62] M. Pircher, R. J. Zawadzki, J. W. Evans, J. S. Werner, and C. K. Hitzenberger, "Simultaneous imaging of human cone mosaic with adaptive optics enhanced scanning laser ophthalmoscopy and high-speed transversal scanning optical coherence tomography," *Opt. Lett.*, vol. 33, no. 1, p. 22, Jan. 2008.

[63] S. S. Choi *et al.*, “In vivo imaging of the photoreceptor mosaic in retinal dystrophies and correlations with visual function,” *Investig. Ophthalmol. Vis. Sci.*, vol. 47, no. 5, pp. 2080–2092, May 2006.

[64] J. I. Wolfing, M. Chung, J. Carroll, A. Roorda, and D. R. Williams, “High-Resolution Retinal Imaging of Cone-Rod Dystrophy,” *Ophthalmology*, vol. 113, no. 6, 2006.

[65] H. Song *et al.*, “Phenotypic diversity in autosomal-dominant cone-rod dystrophy elucidated by adaptive optics retinal imaging,” *Br. J. Ophthalmol.*, vol. 102, no. 1, pp. 136–141, Jan. 2018.

[66] J. L. Duncan *et al.*, “High-resolution imaging with adaptive optics in patients with inherited retinal degeneration,” *Investig. Ophthalmol. Vis. Sci.*, vol. 48, no. 7, pp. 3283–3291, Jul. 2007.

[67] K. E. Talcott *et al.*, “Longitudinal study of cone photoreceptors during retinal degeneration and in response to ciliary neurotrophic factor treatment,” *Investig. Ophthalmol. Vis. Sci.*, vol. 52, no. 5, pp. 2219–2226, Apr. 2011.

[68] Y. Makiyama *et al.*, “Macular cone abnormalities in retinitis pigmentosa with preserved central vision using adaptive optics scanning laser ophthalmoscopy,” *PLoS One*, vol. 8, no. 11, Nov. 2013.

[69] Y. Kitaguchi, S. Kusaka, T. Yamaguchi, T. Mihashi, and T. Fujikado, “Detection of photoreceptor disruption by adaptive optics fundus imaging and fourier-domain optical coherence tomography in eyes with occult macular dystrophy,” *Clin. Ophthalmol.*, vol. 5, no. 1, pp. 345–351, 2011.

[70] A. Nakanishi *et al.*, “Pathologic changes of cone photoreceptors in eyes with occult macular dystrophy,” *Investig. Ophthalmol. Vis. Sci.*, vol. 56, no. 12, pp. 7243–7249, Nov. 2015.

[71] K. Y. Li and A. Roorda, “Automated identification of cone photoreceptors in adaptive optics retinal images,” *J. Opt. Soc. Am. A*, vol. 24, no. 5, p. 1358, May 2007.

[72] B. Xue, S. S. Choi, N. Doble, and J. S. Werner, “Photoreceptor counting and montaging of en-face retinal images from an adaptive optics fundus camera,” *J. Opt. Soc. Am. A*, vol. 24, no. 5, p. 1364, May 2007.

[73] D. H. Wojtas, B. Wu, P. K. Ahnelt, P. J. Bones, and R. P. Millane, “Automated analysis of differential interference contrast microscopy images of the foveal cone mosaic,” *J. Opt. Soc. Am. A*, vol. 25, no. 5, p. 1181, May 2008.

[74] R. Garrioch, C. Langlo, A. M. Dubis, R. F. Cooper, A. Dubra, and J. Carroll, “Repeatability of in vivo parafoveal cone density and spacing measurements,” *Optom. Vis. Sci.*, vol. 89, no. 5, pp. 632–643, May 2012.

[75] S. J. Chiu *et al.*, “Automatic cone photoreceptor segmentation using graph theory and dynamic programming,” *Biomed. Opt. Express*, vol. 4, no. 6, p. 924, Jun. 2013.

[76] F. Mohammad, R. Ansari, J. Wanek, and M. Shahidi, “Frequency-based local content adaptive filtering algorithm for automated photoreceptor cell density quantification,” in *Proceedings - International Conference on Image Processing, ICIP*, 2012, pp. 2325–2328.

[77] R. F. Cooper, C. S. Langlo, A. Dubra, and J. Carroll, “Automatic detection of modal spacing (Yellott’s ring) in adaptive optics scanning light ophthalmoscope images,” *Ophthalmic Physiol. Opt.*, vol. 33, no. 4, pp. 540–549, Jul. 2013.

[78] D. Cunefare *et al.*, “Automatic detection of cone photoreceptors in split detector adaptive optics scanning light ophthalmoscope images,” *Biomed. Opt. Express*, vol. 7, no. 5, p. 2036, May 2016.

[79] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science*, vol. 349, no. 6245. American Association for the Advancement of Science, pp. 255–260, 17-Jul-2015.

[80] A. Lee, P. Taylor, and J. Kalpathy-Cramer, “Machine Learning Has Arrived!,” *Ophthalmology*, vol. 124, no. 12. Elsevier Inc., pp. 1726–1728, 01-Dec-2017.

[81] P. Liskowski and K. Krawiec, “Segmenting Retinal Blood Vessels with Deep Neural Networks,” *IEEE Trans. Med. Imaging*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.

[82] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, “A cross-modality learning approach for vessel segmentation in retinal images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 109–118, Jan. 2016.

[83] S. P. K. Karri, D. Chakraborty, and J. Chatterjee, “Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration,” *Biomed. Opt. Express*, vol. 8, no. 2, p. 579, Feb. 2017.

[84] L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, “Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search,” *Biomed. Opt. Express*, vol. 8, no. 5, p. 2732, 2017.

[85] D. Cunefare, L. Fang, R. F. Cooper, A. Dubra, and J. Carroll, “Open source software for automatic detection of cone photoreceptors in adaptive optics ophthalmoscopy using convolutional neural networks,” *Sci. Rep.*, no. June, pp. 1–11, 2017.

[86] J. Donahue *et al.*, “DeCAF: A deep convolutional activation feature for generic visual recognition,” in *31st International Conference on Machine Learning, ICML*

2014, 2014.

[87] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” in *Advances in Neural Information Processing Systems*, 2014.

[88] D. S. Kermany *et al.*, “Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning,” *Cell*, vol. 172, no. 5, pp. 1122-1131.e9, Feb. 2018.

[89] D. Cunefare *et al.*, “Deep learning based detection of cone photoreceptors with multimodal adaptive optics scanning light ophthalmoscope images of achromatopsia,” *Biomed. Opt. Express*, vol. 9, no. 8, p. 3740, Aug. 2018.

[90] D. Cunefare *et al.*, “Deep learning based detection of cone photoreceptors with multimodal adaptive optics scanning light ophthalmoscope images of achromatopsia,” *Biomed. Opt. Express*, vol. 9, no. 8, p. 3740, Aug. 2018.

[91] F. Chollet, “Keras Documentation,” *Keras.Io*, 2015. .

[92] S. Ooto *et al.*, “High-resolution imaging of the photoreceptor layer in epiretinal membrane using adaptive optics scanning laser ophthalmoscopy,” *Ophthalmology*, vol. 118, no. 5, pp. 873–881, May 2011.

[93] R. F. Cooper, M. A. Wilk, S. Tarima, and J. Carroll, “Evaluating descriptive metrics of the human cone mosaic,” *Investig. Ophthalmol. Vis. Sci.*, vol. 57, no. 7, pp. 2992–3001, Jun. 2016.

[94] K. Y. Li and A. Roorda, “Automated identification of cone photoreceptors in adaptive optics retinal images,” *J. Opt. Soc. Am. A*, vol. 24, no. 5, p. 1358, May 2007.

[95] R. F. Cooper, M. A. Wilk, S. Tarima, and J. Carroll, “Evaluating descriptive metrics of the human cone mosaic,” *Investig. Ophthalmol. Vis. Sci.*, vol. 57, no. 7, pp. 2992–3001, Jun. 2016.

[96] C. A. Curcio, K. R. Sloan, R. E. Kalina, and A. E. Hendrickson, “Human photoreceptor topography,” *J. Comp. Neurol.*, 1990.

[97] H. Song, T. Y. P. Chui, Z. Zhong, A. E. Elsner, and S. A. Burns, “Variation of Cone Photoreceptor Packing Density with Retinal Eccentricity and Age,” *Invest. Ophthalmol. Vis. Sci.*, 2011.

[98] M. N. Muthiah *et al.*, “Cone photoreceptor definition on adaptive optics retinal imaging,” *Br. J. Ophthalmol.*, 2014.

[99] M. Heisler *et al.*, “Strip-based registration of serially acquired optical coherence tomography angiography,” *J. Biomed. Opt.*, vol. 22, no. 3, p. 036007, Mar. 2017.

[100] B. C. Chauhan and C. F. Burgoyne, “From clinical examination of the optic disc to clinical assessment of the optic nerve head: A paradigm change,” *American Journal of Ophthalmology*. 2013.

[101] A. S. C. Reis *et al.*, “Influence of clinically invisible, but optical coherence tomography detected, optic disc margin anatomy on neuroretinal rim evaluation.,” *Invest. Ophthalmol. Vis. Sci.*, vol. 53, no. 4, pp. 1852–1860, Apr. 2012.

[102] M. Young, S. Lee, M. Rateb, M. F. Beg, M. V. Sarunic, and P. J. Mackenzie, “Comparison of the clinical disc margin seen in stereo disc photographs with neural canal opening seen in optical coherence tomography images,” *Journal of Glaucoma*, vol. 23, no. 6. Lippincott Williams and Wilkins, pp. 360–367, 2014.

[103] P. Zang *et al.*, “Automated boundary detection of the optic disc and layer segmentation of the peripapillary retina in volumetric structural and angiographic optical coherence tomography,” *Biomed. Opt. Express*, vol. 8, no. 3, p. 1306, Mar. 2017.

[104] M. K. Garvin, M. D. Abràmoff, R. Kardon, S. R. Russell, X. Wu, and M. Sonka, “Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search,” *IEEE Trans. Med. Imaging*, vol. 27, no. 10, pp. 1495–1505, Oct. 2008.

[105] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, “Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation,” *Opt. Express*, vol. 18, no. 18, p. 19413, Aug. 2010.

[106] A. Lang *et al.*, “Retinal layer segmentation of macular OCT images using boundary classification,” *Biomed. Opt. Express*, vol. 4, no. 7, p. 1133, Jul. 2013.

[107] S. Lee *et al.*, “Age and glaucoma-related changes in retinal nerve fiber layer and choroid: point-wise analysis and visualization using functional shapes registration,” *Front. Neurosci.*

[108] S. Lee *et al.*, “Comparative analysis of repeatability of manual and automated choroidal thickness measurements in nonneovascular age-related macular degeneration,” *Investig. Ophthalmol. Vis. Sci.*, vol. 54, no. 4, pp. 2864–71, 2013.

[109] S. Lee, M. Heisler, P. J. Mackenzie, M. V Sarunic, and M. Faisal, “Quantifying Variability in Longitudinal Peripapillary RNFL and Choroidal Layer Thickness Using Surface Based Registration of OCT Images,” *Transl. Vis. Sci. Technol.*, vol. 6, no. 1, pp. 1–20, 2017.

[110] S. Lee *et al.*, “Age and glaucoma-related characteristics in retinal nerve fiber layer and choroid: Localized morphometrics and visualization using functional shapes registration,” *Front. Neurosci.*, 2017.

[111] S. Lee *et al.*, “Atlas-based shape analysis and classification of retinal optical coherence tomography images using the functional shape (fshape) framework,”

Med. Image Anal., vol. 35, pp. 570–581, 2017.

[112] S. Lee, S. X. Han, M. Young, M. F. Beg, M. V. Sarunic, and P. J. Mackenzie, “Optic nerve head and peripapillary morphometrics in myopic glaucoma,” *Investig. Ophthalmol. Vis. Sci.*, vol. 55, pp. 4378–4393, 2014.

[113] L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, “Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search,” *Biomed. Opt. Express*, vol. 8, no. 5, p. 2732, May 2017.

[114] A. G. Roy *et al.*, “ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks,” *Biomed. Opt. Express*, 2017.

[115] Y. He *et al.*, “Deep learning based topology guaranteed surface and MME segmentation of multiple sclerosis subjects from retinal OCT,” *Biomed. Opt. Express*, vol. 10, no. 10, 2019.

[116] S. Apostolopoulos, S. De Zanet, C. Ciller, S. Wolf, and R. Sznitman, “Pathological OCT retinal layer segmentation using branch residual U-shape networks,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10435 LNCS, pp. 294–301.

[117] M. Pekala, N. Joshi, T. Y. A. Liu, N. M. Bressler, D. C. DeBuc, and P. Burlina, “Deep learning based retinal OCT segmentation,” *Comput. Biol. Med.*, vol. 114, p. 103445, Nov. 2019.

[118] S. Sedai, B. Antony, D. Mahapatra, and R. Garnavi, “Joint Segmentation and Uncertainty Visualization of Retinal Layers in Optical Coherence Tomography Images Using Bayesian Deep Learning,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11039 LNCS, pp. 219–227.

[119] S. K. Devalla *et al.*, “DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images,” *Biomed. Opt. Express*, 2018.

[120] S. K. Devalla *et al.*, “A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head,” *Investig. Ophthalmol. Vis. Sci.*, 2018.

[121] P. Zang, J. Wang, T. T. Hormel, L. Liu, D. Huang, and Y. Jia, “Automated segmentation of peripapillary retinal boundaries in OCT combining a convolutional neural network and a multi-weights graph search,” *Biomed. Opt. Express*, vol. 10, no. 8, p. 4340, Aug. 2019.

[122] M. Young, S. Lee, M. F. Beg, P. J. Mackenzie, and M. V Sarunic, “High Speed

Morphometric Imaging of the Optic Nerve Head with 1μm OCT," *Invest. Ophthalmol. Vis. Sci.*, vol. 52, no. 14, p. 6253, Apr. 2011.

[123] S. Lee, S. X. Han, M. Young, M. F. Beg, M. V. Sarunic, and P. J. Mackenzie, "Optic nerve head and peripapillary morphometrics in myopic glaucoma," *Investig. Ophthalmol. Vis. Sci.*, 2014.

[124] M. Bhalla *et al.*, "Longitudinal Analysis of Bruch Membrane Opening Morphometry in Myopic Glaucoma," *J. Glaucoma*, vol. 28, no. 10, pp. 889–895, Oct. 2019.

[125] S. Lee, M. F. Beg, and M. V. Sarunic, "Segmentation of the macular choroid in OCT images acquired at 830nm and 1060nm," in *Optical Coherence Tomography and Coherence Techniques VI*, 2013, vol. 8802, p. 88020J.

[126] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 5967–5976, Nov. 2016.

[127] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[128] J. M. D. Gmeiner, W. A. Schrems, C. Y. Mardin, R. Laemmer, F. E. Kruse, and L. M. Schrems-Hoesl, "Comparison of bruch's membrane opening minimum rim width and peripapillary retinal nerve fiber layer thickness in early glaucoma assessment," *Investig. Ophthalmol. Vis. Sci.*, vol. 57, no. 9Special Issue, pp. OCT575–OCT584, Jul. 2016.

[129] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: Learning dense volumetric segmentation from sparse annotation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9901 LNCS, no. 8, pp. 424–432.

[130] N. Cheung, P. Mitchell, and T. Y. Wong, "Diabetic retinopathy.," *Lancet*, vol. 376, no. 9735, pp. 124–36, 2010.

[131] J. FRIEDENWALD and R. DAY, "The vascular lesions of diabetic retinopathy.," *Bull. Johns Hopkins Hosp.*, vol. 86, no. 4, pp. 253–254, Apr. 1950.

[132] J. S. FRIEDENWALD, "Diabetic retinopathy.," *Am. J. Ophthalmol.*, vol. 33, no. 8, pp. 1187–1199, Aug. 1950.

[133] D. G. COGAN, D. TOUSSAINT, and T. KUWABARA, "Retinal vascular patterns. IV. Diabetic retinopathy.," *Arch. Ophthalmol. (Chicago, Ill. 1960)*, vol. 66, pp. 366–378, Sep. 1961.

[134] N. Demirkaya *et al.*, "Effect of age on individual retinal layer thickness in normal

eyes as measured with spectral-domain optical coherence tomography.," *Invest. Ophthalmol. Vis. Sci.*, vol. 54, no. 7, pp. 4934–4940, Jul. 2013.

[135] S. Rajaraman, S. Jaeger, and S. K. Antani, "Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images," *PeerJ*, 2019.

[136] J. J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[137] A. Hoover, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imaging*, 2000.

[138] L. K. Hansen and P. Salamon, "Neural Network Ensembles," *IEEE Trans. Pattern Anal. Mach. Intell.*, 1990.

[139] Z.-H. Zhou, J. Wu, and W. Tang, "Ensembling neural networks: Many could be better than all," *Artif. Intell.*, vol. 137, no. 1, pp. 239–263, 2002.

[140] L. Breiman, "Bagging predictors," *Mach. Learn.*, 1996.

[141] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap. Monographs on Statistics and Applied Probability, No. 57. Chapman and Hall, London, 436 p*. 1993.

[142] R. E. Schapire, "The Strength of Weak Learnability," *Mach. Learn.*, 1990.

[143] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1995.

[144] Y. Freund, "Boosting a weak learning algorithm by majority," *Inf. Comput.*, 1995.

[145] Z. Ji, Q. Chen, S. Niu, T. Leng, and D. L. Rubin, "Beyond Retinal Layers: A Deep Voting Model for Automated Geographic Atrophy Segmentation in SD-OCT Images," *Transl. Vis. Sci. Technol.*, vol. 7, no. 1, p. 1, 2018.

[146] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016.

[147] G. Shi *et al.*, "Automatic Classification of Anterior Chamber Angle Using Ultrasound Biomicroscopy and Deep Learning," *Transl. Vis. Sci. Technol.*, vol. 8, no. 4, p. 25, Aug. 2019.

[148] P. Mehta, A. Lee, C. Lee, M. Balazinska, and A. Rokem, "Multilabel multiclass

classification of OCT images augmented with age, gender and visual acuity data," *bioRxiv*, 2018.

[149] W. Lu, Y. Tong, Y. Yu, Y. Xing, C. Chen, and Y. Shen, "Deep Learning-Based Automated Classification of Multi-Categorical Abnormalities From Optical Coherence Tomography Images," *Transl. Vis. Sci. Technol.*, vol. 7, no. 6, p. 41, 2018.

[150] A. A. Jammal *et al.*, "Detecting Retinal Nerve Fibre Layer Segmentation Errors on Spectral Domain-Optical Coherence Tomography with a Deep Learning Algorithm," *Sci. Rep.*, vol. 9, no. 1, p. 9836, 2019.

[151] S. Maetschke, B. Antony, H. Ishikawa, G. Wollstein, J. Schuman, and R. Garnavi, "A feature agnostic approach for glaucoma detection in OCT volumes," *PLoS One*, 2019.

[152] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks," in *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, 2018.

[153] K. Li, Z. Wu, K. C. Peng, J. Ernst, and Y. Fu, "Tell Me Where to Look: Guided Attention Inference Network," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018.

[154] J. Lo *et al.*, "Microvasculature Segmentation and Inter-capillary Area Quantification using Transfer Learning," *Submitt. to Transl. Vis. Sci. Technol.*, 2019.

[155] M. Heisler *et al.*, "Deep learning vessel segmentation and quantification of the foveal avascular zone using commercial and prototype OCT-A platforms." ArXiv ID: eess.IV/1909.11289, 2019.

[156] P. L. Nesper *et al.*, "Quantifying Microvascular Abnormalities With Increasing Severity of Diabetic Retinopathy Using Optical Coherence Tomography Angiography," *Invest. Ophthalmol. Vis. Sci.*, 2017.

[157] T. E. de Carlo *et al.*, "DETECTION OF MICROVASCULAR CHANGES IN EYES OF PATIENTS WITH DIABETES BUT NOT CLINICAL DIABETIC RETINOPATHY USING OPTICAL COHERENCE TOMOGRAPHY ANGIOGRAPHY.," *Retina*, vol. 35, no. 11, pp. 2364–2370, Nov. 2015.

[158] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for Simplicity: The All Convolutional Net." ArXiv ID: cs.LG/1412.6806, 2014.

[159] S. Rajaraman *et al.*, "A novel stacked generalization of models for improved TB detection in chest radiographs.," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2018, pp. 718–721, Jul. 2018.

[160] D. Bhanushali *et al.*, “Linking Retinal Microvasculature Features With Severity of Diabetic Retinopathy Using Optical Coherence Tomography Angiography,” *Invest. Ophthalmol. Vis. Sci.*, vol. 57, no. 9, pp. OCT519–OCT525, 2016.

[161] S. K. Johannesen, J. N. Viken, A. S. Vergmann, and J. Grauslund, “Optical coherence tomography angiography and microvascular changes in diabetic retinopathy: a systematic review,” *Acta Ophthalmologica*. 2019.

[162] Y. Lu *et al.*, “Evaluation of automatically quantified foveal avascular zone metrics for diagnosis of diabetic retinopathy using optical coherence tomography angiography,” *Investig. Ophthalmol. Vis. Sci.*, 2018.

[163] H. S. Sandhu *et al.*, “Automated diabetic retinopathy detection using optical coherence tomography angiography: a pilot study.,” *Br. J. Ophthalmol.*, vol. 102, no. 11, pp. 1564–1569, Nov. 2018.

[164] X. Li, L. Shen, M. Shen, F. Tan, and C. S. Qiu, “Deep learning based early stage diabetic retinopathy detection using optical coherence tomography,” *Neurocomputing*, Aug. 2019.

[165] X. Liu, M. D. Chan, X. Zhou, and X. Qian, “Transparency guided ensemble convolutional neural networks for stratification of pseudoprogression and true progression of glioblastoma multiform.” ArXiv ID: q-bio.TO/1902.09921, 2019.

[166] J. Hamwood, D. Alonso-Caneiro, D. M. Sampson, M. J. Collins, and F. K. Chen, “Automatic detection of cone photoreceptors with fully convolutional networks,” *Transl. Vis. Sci. Technol.*, vol. 8, no. 6, Nov. 2019.

Appendix A.

Peer Reviewed Journal Publications:

A1. Z. Mammo, C. Balaratnasingam, P. Yu, J. Xu, **M. Heisler**, P. Mackenzie, A. Merkur, A. Kirker, D. Albiani, K. B. Freund, M. V. Sarunic, D.-Y. Y. Yu, "Quantitative Noninvasive Angiography of the Fovea Centralis Using Speckle Variance Optical Coherence Tomography of Macula," Investig. Ophthalmol. Vis. Sci. **56**(9), 5074–5086 (2015).

A2. Z. Mammo, **M. Heisler**, C. Balaratnasingam, S. Lee, D.-Y. Y. Yu, P. Mackenzie, S. Schendel, A. Merkur, A. Kirker, D. Albiani, E. Navajas, M. F. F. Beg, W. Morgan, and M. V. V. Sarunic, "Quantitative Optical Coherence Tomography Angiography of Radial Peripapillary Capillaries in Glaucoma, Glaucoma Suspect, and Normal Eyes," Am. J. Ophthalmol. **170**, 41–49 (2016).

A3. P. Prentasic, **M. Heisler***, S. Lee, Z. Mammo, A. B. Merkur, E. Navajas, M. F. F. Beg, M. V. Sarunic, S. Loncaric, "Segmentation of the Foveal Microvasculature Using Deep Learning Networks," J. Biomed. Opt. **21**(7), 075008 (2016). ***Co-first Author**

A4. C. E. Pang, E. V Navajas, S. J. Warner, **M. Heisler**, and M. V Sarunic, "Acute Macular Neuroretinopathy Associated With Chikungunya Fever," Ophthalmic Surg Lasers Imaging Retin. **47**(6), 596–9 (2016).

A5. Y. Jian, S. Lee, M. J. J. Ju, **M. Heisler**, W. Ding, R. J. J. Zawadzki, S. Bonora, and M. V. V Sarunic, "Lens-based wavefront sensorless adaptive optics swept source OCT," Sci. Rep. **6:27620**(February), 1–10 (2016).

A6. H. R. G. W. Verstraete, **M. Heisler**, M. J. Ju, D. Wahl, L. Bliek, J. Kalkman, S. Bonora, Y. Jian, M. Verhaegen, and M. V Sarunic, "Wavefront sensorless adaptive optics OCT with the DONE algorithm for *in vivo* human retinal imaging," Biomed. Opt. Express **8**(4), 2261–2275 (2017).

A7. **M. Heisler**, W. L. Quong, S. Lee, S. Han, M. F. Beg, M. V Sarunic, and P. J. Mackenzie, "Anterior Segment Optical Coherence Tomography for Targeted Transconjunctival Suture Placement in Overfiltering Trabeculectomy Blebs," J. Glaucoma **00**(00), 1–5 (2017).

A8. S. Lee, **M. Heisler**, P. J. Mackenzie, M. V Sarunic, and M. Faisal, "Quantifying Variability in Longitudinal Peripapillary RNFL and Choroidal Layer Thickness Using Surface Based Registration of OCT Images," Transl. Vis. Sci. Technol. **6**(1), 1–20 (2017).

A9. M. J. Ju, **M. Heisler***, D. Wahl, Y. Jian, and M. V. Sarunic, "Multiscale sensorless adaptive optics OCT angiography system for *in vivo* human retinal imaging," J. Biomed. Opt. **22**(12), 1 (2017). ***Co-first Author**

A10. S. Lee, **M. L. Heisler**, K. Popuri, N. Charon, B. Charlier, A. Trouvé, P. J. Mackenzie, M. V. Sarunic, and M. F. Beg, "Age and glaucoma-related characteristics in retinal nerve fiber layer and choroid: Localized morphometrics and visualization using functional shapes registration," Front. Neurosci. (2017).

A11. **M. Heisler**, S. Lee, Z. Mammo, Y. Jian, M. Ju, A. Merkur, E. Navajas, C. Balaratnasingam, M. F. Beg, and M. V. Sarunic, "Strip-based registration of serially acquired optical coherence tomography angiography Strip-based registration of serially acquired optical," J. Biomed. Opt. **22**(3), 036007 (2017).

A12. **M. Heisler**, M. J. Ju, M. Bhalla, N. Schuck, A. Athwal, E. V Navajas, M. F. Beg, and M. V Sarunic, "Automated identification of cone photoreceptors in adaptive optics optical coherence tomography images using transfer learning.," Biomed. Opt. Express **9**(11), 5353–5367 (2018).

A13. P. Gong, D.-Y. Yu, Q. Wang, P. K. Yu, K. Karnowski, **M. Heisler**, A. Francke, D. An, M. V Sarunic, and D. D. Sampson, "Label-free volumetric imaging of conjunctival collecting lymphatics ex vivo by optical coherence tomography lymphangiography.," J. Biophotonics **11**(8), e201800070 (2018).

A14. M. J. Ju, **M. Heisler**, A. Athwal, M. V Sarunic, and Y. Jian, "Effective bidirectional scanning pattern for optical coherence tomography angiography." Biomed. Opt. Express **9**(5), 2336–2350 (2018).

A15. D. An, C. Balaratnasingam, **M. Heisler**, A. Francke, M. Ju, I. L. McAllister, M. Sarunic, and D.-Y. Yu, "Quantitative comparisons between optical coherence tomography angiography and matched histology in the human eye.," Exp. Eye Res. **170**, 13–19 (2018).

A16. Navajas EV, Schuck N, Govetto A, Akil H, Docherty G, **Heisler M**, Sarunic MV, and Sarraf D., "*EN FACE* OPTICAL COHERENCE TOMOGRAPHY AND OPTICAL COHERENCE TOMOGRAPHY ANGIOGRAPHY OF INNER RETINAL DIMPLES AFTER INTERNAL LIMITING MEMBRANE PEELING FOR FULL-THICKNESS MACULAR HOLES.," Retina (2019)

A17. Lu D, **Heisler M**, Lee S, Ding GW, Navajas E, Sarunic MV, Beg MF., "Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network.," Med Image Anal. 54:100-110 (2019)

A18. Bhalla M, **Heisler M**, Han SX, Sarunic MV, Beg MF, Mackenzie PJ, Lee S., "Longitudinal Analysis of Bruch Membrane Opening Morphometry in Myopic Glaucoma.," J Glaucoma. 28(1):889-895 (2019)

A19. Akil H, Karst S, **Heisler M**, Etminan M, Navajas E, Maberley D., "Application of optical coherence tomography angiography in diabetic retinopathy: a comprehensive review.," Can J Ophthalmol. 54(5):519-528 (2019)

A20. Sonja Karst, **Morgan Heisler**, Julian Lo, Nathan Schuck, Abdollah Safari, Marinko Sarunic, David Maberley, and Eduardo Navajas, "Evaluating signs of microangiopathy secondary to diabetes in different areas of the retina with swept source OCTA.," Invest Ophthal Vis Sci. Submitted September 18th, 2019.

A21. **Morgan Heisler**, Sonja Karst, Julian Lo, Zaid Mammo, Timothy Yu, Simon Warner, David Maberley, Mirza Beg, Eduardo Navajas, and Marinko Sarunic, "Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography," Trans Vis Sci. Submitted September 30th, 2019.

A22. Julian Lo, **Morgan Heisler**, Vinicius Vanzan, Sonja Karst, Eduardo Navajas, Mirza Beg, and Marinko Sarunic, "Microvasculature Segmentation and Inter-capillary Area Quantification of the Deep Vascular Complex using Transfer Learning," Trans Vis Sci. Submitted September 30th, 2019.