A design for an automated Optical Coherence Tomography analysis system

Thesis submitted in partial fulfillment
of the requirements for the degree of

MS in Electronics and Communication Engineering by Research

by

Karthik Gopinath
201450857
`karthik.g@research.iiit.ac.in`

IIIT, HYDERABAD

International Institute of Information Technology
Hyderabad - 500 032, INDIA
July 2017

International Institute of Information Technology

Hyderabad, India

CERTIFICATE OF AUTHORSHIP

I, Karthik Gopinath, declare that the thesis, titled "A design for an automated Optical Coherence Tomography analysis system", and the work presented herein are my own. I confirm that this work was done wholly or mainly while in candidature for a research degree at IIIT-Hyderabad.

17 July 2017 Karthik Gopinath

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "A design for an automated Optical Coherence Tomography analysis system" by Karthik Gopinath, has been carried out under my supervision and is not submitted elsewhere for a degree.

17 July 2017   Adviser: Prof. Jayanthi Sivaswamy

To

Amma and Appa for everything

Acknowledgements

After almost completing three years as a research student under Prof. Jayanthi Sivaswamy, the one question that still puzzles me is " What made her accept me as her student?" . From being a student from Visvesvaraya Technological University, I had very little exposure to research. Undergoing rigorous training under her, I am now addicted to academic research. Interaction with her over the years has made me develop into a better researcher. She has directed me onto the right path by forgiving all my stupid mistakes. It is, only due to her encouragement, that I have become a better thinker from a 'headless chicken'. It is hard to express in words how grateful I am to her. I would also like to thank Dr. Taranum mansuri, Dr. Pavani Reddy and Dr. Satish for providing medical data required for my thesis work.

My initial days at IIIT-Hyderabad was moulded by the valuable advice I received from Arunava sir. He has inspired me in many ways and my experience here would have been totally different without his constant support. For any suggestions, the 'to go' friends for me have been Tabish, Ujjwal, Palash and Nisith. I am lucky and thankful to be in touch with them even at odd hours. Spending time with 'Special friends' Samrudhdhi and Raghav have always been sheer joy. I thank Sam for bearing me for over two years. I thank Raghav for sharing his practical knowledge at all situations . I will always cherish the fun and enjoyable outings with these people.

Studying and regular meetings with the fellow CVIT friends have been a unique learning experience. I wish to have the same wonderful environment CVIT provided to be with me wherever I go next. I thank all my lab mates Aabhas, Jahnavi, Pujitha, Chetan, Sukesh and Alphin for making the group meetings memorable.

My hostel stay would be incomplete without FIFA nights, movie nights, food orders, cricket matches and long talks near water cooler. I thank Rahul, Soumyajit, Sayantan, Pramod, Aditya, Soham, Suman, Govinda, Bharat and Sagar for making my stay joyful. Last but not the least, I thank all my friends in IIIT Hyderabad.

Abstract

Age related macula degeneration (AMD), Cystoid Macular Edema (CME) and glaucoma are retinal diseases that affects vision and often lead to irreversible blindness. Many imaging modalities have been used for screening of retinal diseases. Optical coherence tomography (OCT) is one such imaging modality that provides structural information of retina. Optical coherence tomography angiography (OCTA) also provides vascular information of retina in addition to structural information. With advancement in OCT and OCTA technology, the number of patients being scanned using these modalities is increasing exponentially. Manual analysis of vast data often makes human experts feel fatigue. Also, this extends the time to treat any patient and creates a demand to develop a fast and accurate automated OCT image analysis system. The system proposed in this thesis aims at analysing retinal anatomy and its diseases.

We approach the problem of automatically analysing the retinal anatomy by segmenting its layers using a deep learning framework. In literature these algorithms usually requires pre-processing steps, such as denoising, image flatenning and edge detection, all of which involve separate parameter tunings. We propose a deep learning technique to automate all these steps and handle the presence/absence of pathologies. This model consists of a combination of Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM). The CNN is used to extract layers of interest image and extract the edges, while the LSTM is used to trace the layer boundary. This model is trained on a mixture of normal and AMD cases using minimal data. Validation results on three public datasets show that the pixel-wise mean absolute error obtained with our system is 1.30 ± 0.48 which is lower than inter-marker error of 1.79 ± 0.76. The performance of the proposed module is also on par with the existing methods.

We next propose three modules for disease analysis, first for diagnosing them at volume level, second at slice level and finally at pixel level by segmenting the abnormality. Firstly, we propose a novel method for glaucoma assessment at volume level using data from OCTA modality. The goal of this module is to gain insights about glaucoma indicators from both structural information (OCT volume) and vascular information (angioflow images). The proposed method achieves a mean sensitivity of 94% specificity of 91% and accuracy of 92% on 49 normal and 18 glaucomatous volumes. Secondly, for slice level disease(AMD and CME) classification from OCT volumes, we learn a decision boundary using a CNN in the new extremal representation space of

an image. Evaluating on four publically available datasets with training set consisting of 3500 OCT slices and test set having 1995 OCT slices, this module achieves a mean sensitivity of 96%, specificity of 97% and accuracy of 96%. Finally, we propose an automated cyst segmentation algorithm from OCT volumes. We propose a biologically inspired method based on selective enhancement of the cysts, by inducing motion to a given OCT slice. A CNN is designed to learn a mapping function that combines the result of multiple such motions to produce a probability map for cyst locations in a given slice. The final segmentation of cysts is obtained via a simple clustering of the detected cyst locations. The proposed method is trained on OPTIMA Cyst segmentation challenge (OCSC) train test and achieves a mean Dice Coefficient (DC) of 0.71, 0.69 and 0.79 on the OCSC test set, DME dataset and AEI dataset respectively.

Thus, the complete system can be employed in screening scenarios to aid retinal experts in accurate and faster analysis of retinal anatomy and disease.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Medical image processing is a research area focused on developing computational methods and algorithms to analyze and quantify medical images. With the advancements in making machines understand visual data, an entire field has opened up in pursuit of early prediction of diseases. Imaging modalities have aided medical experts to know the human anatomy better. Medical data in the digital version has allowed researchers across the world to build algorithms for various complex medical image processing tasks like disease classification, identification of abnormality and quantifying anatomical structures. This thesis focuses on analysing images of retina acquired using 3D imaging modality.

1.1 RETINA AND ITS DISEASES

Retina is the light sensitive tissue at the back of the eye. It contains cones which produce visual images and rods which are responsible for peripheral and night vision. Macula forms the central part of retina. A photoreceptor cell is a specialized type of neuron found in the retina. Photoreceptors convert light into signals that can stimulate biological processes. Rods and cones are the two classic photoreceptor cells, each contributing information used by the visual system. The Retinal pigment epethilium (RPE) is composed of a densely packed pigment granules. The main function of RPE is to absorb light. Bruch's membrane (BM) is a strategically located structure between RPE and the choroidal capillaries of the eye. They regulate the exchange of biomolecules, nutrients, oxygen and fluids between the retina and the general circulation [2]. A pictorial depection of a normal eye is shown in Figure 1.1 (a). Retinal diseases can affect any part of the retina and can cause total blindness. Retina of the eye is affected by diseases like Age related macular degeneration(AMD), Cystoidal macular edema (CME), Glaucoma and Diabetic retinopathy. Here in this thesis the focus is on diseases like AMD, CME and Glaucoma. A diagramatic representation of the affected retina is shown in the Fig. 1.1 (b).
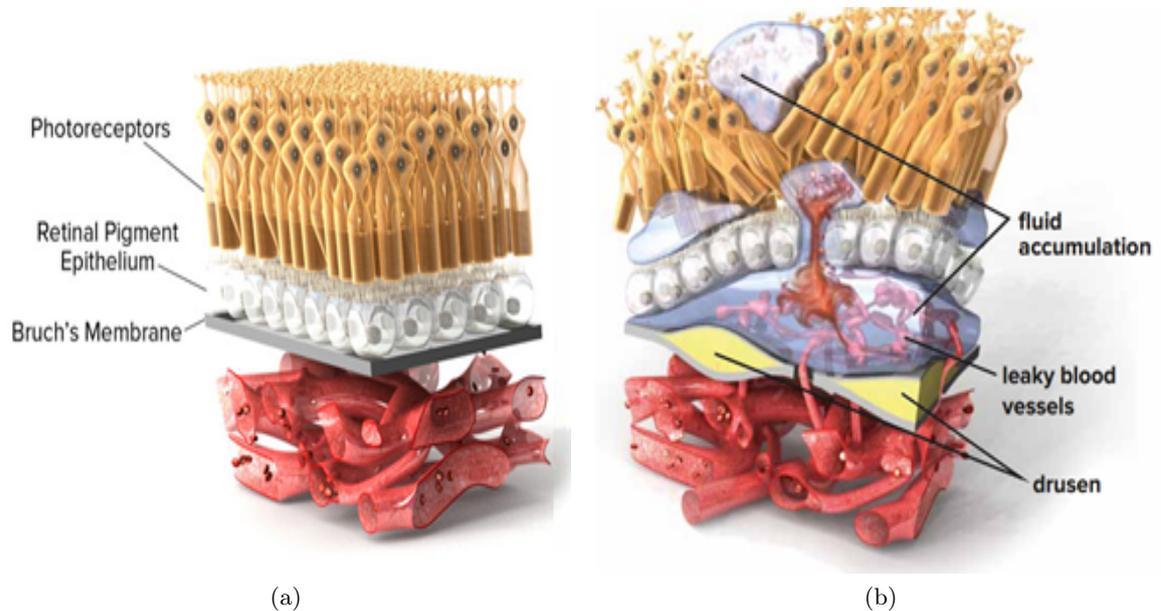
Figure 1.1: 3D representation of retina . (a) Normal retina (b) Retina affected by AMD and CME. (source:https://www.scienceofamd.org/)

1.1.1 Age related macular degeration (AMD)

AMD is the leading cause of irreversible blindness in the world. The cell death in macula affects the central vision of the patient. The difference in vision for a normal subject vs AMD subject is shown in Fig. 1.2. The main characteristics of this disease are Pigment Changes, Drusen, and Geographical atrophy. Pigment changes are visible as coloration difference in the retinal pigment epithelium(RPE) layer. Drusen are yellow deposits in retina resulting from collection of waste products. Geographical atrophy is seen in the final stages of AMD where the RPE layer is absent due to the death of choroidal vessels.

1.1.2 Cystoidal macular edima (CME)

CME is a retinal condition in which swelling develops in the macula. These swollen areas are often filled with fluid, causing the macula to thicken. The cause of the disease is not known. However, it is observed that CME occours usually after cataract surgery. The fluid filled areas are called cystoidal structures or cysts. The difference in vision for a normal subject vs CME subject is shown in Fig. 1.3.

(a)   (b)

Figure 1.2: An illustration of the difference in the vision of a normal person (a) and a person with AMD (b) (source:http://www.optos.com/en-US/Patient/Eye-conditions/Age-related-macular-degeneration-AMD/).

(a)   (b)

Figure 1.3: An illustration of the difference in the vision of a normal person (a) and a person with CME (b) (source:http://eye.md/cystoid-macular-edema/).

Figure 1.4: An illustration of the difference in the vision of a normal person (a) and a person with glaucoma (b).

1.1.3 Glaucoma

Glaucoma is a disease that damages the optic nerve of the eye. It is the second leading cause of blindness after cataracts. Glaucoma is a chronic and irreversible disease in which the optic nerve is progressively damaged, leading to deterioration in vision. Early detection of glaucomatous changes is crucial for timely treatment before the onset of permanent functional visual loss. A key challenge in early detection is that many glaucoma patients are not aware of their condition until vision loss. Glaucoma is detected from visual field loss and manual assessment of the optic nerve head. The difference in vision for a normal subject vs glaucoma subject is shown in Fig. 1.4.

1.2 IMAGING MODALITIES

Understanding the cause and effect of any disease requires visualization of the human anatomy that is being targeted. A wide variety of imaging modalities exist to image human eye. The use of light has played an important role in revealing structural and functional information from the human retina in a non-invasive manner. Multiple imaging modalities exist to image the retina of the eye.

1.2.1 Fundus imaging

Fundus imaging projects the 3D retinal tissue into a 2D image. Each pixel in a color fundus image represents the reflected red, blue, and green waveband. This is a popular imaging technique

(a)   (b)
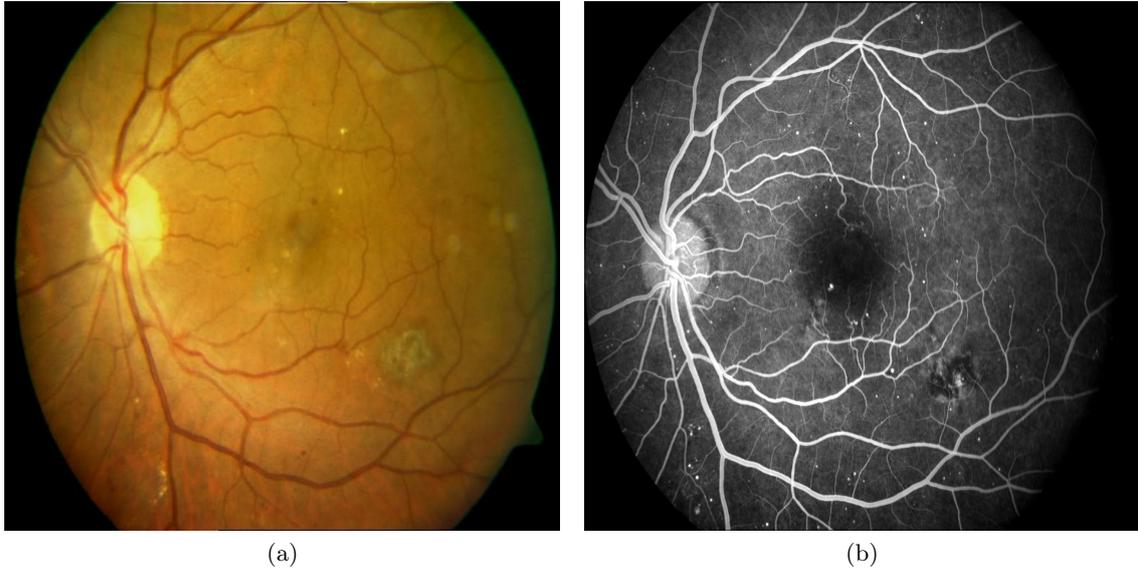
Figure 1.5: Fundus imaging . (a) Color fundus image (b) Fluorescein angiography

used for screening purpose. Since only visible light is used to illuminate the eye, the modality is very cost effective. It can document the appearance of the optic nerve and blood flow in the eye. A technique for imaging vasculature is fluorescein angiography. Here an image intensity depicts the amount of emitted photons from the fluorescein dye injected for circulation. This technique is used to assess the damaged retinal vessels. But, the method is limited because of its invasive nature. An illustration of Color fundus image and fluorescein angiography images is shown in Fig. 1.5

1.2.2   Optical Coherence Tomography Imaging (OCT)

The origin of OCT is traceable to optical ranging used in the telecommunications industry for locating faults or defects in optical fibers. Faults in the fiber produce a partial or complete reflection when optical pulses are incident. The time delay between the original and reflected pulse is measured to localize these faults. Analogous to optical ranging in fibers, optical ranging can also be performed on biological tissue.

OCT imaging works by estimating depth at which a specific backscatter originated by measuring the time taken by light to travel in the eye. The cause of backscatters are typically due to the transitions in refractive index from one tissue to another. The time taken for backscatter is used to differentiate the deeper tissues from the superficial ones. Low-coherent light interferometer is employed by OCT to measure the thickness of retina. A low coherent light is split optically, one reference beam is sent to a mirror at a specific distance to reflect and the other beam is made to
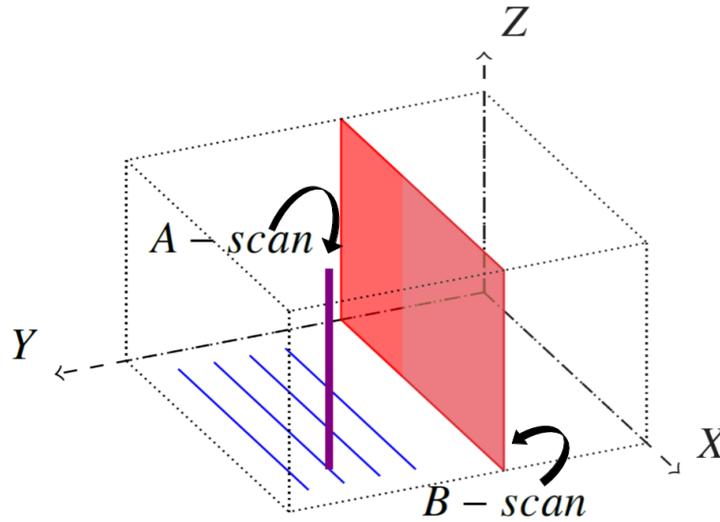
Figure 1.6: Axial scans (A-scan) in purple (Column of the XZ plane) are combined to form a 2D cross sectional slice (B-scan) in red by scanning through the volume in a raster scan pattern in blue (Row in of the XY plane). Multiple B-scans are combined to form a complete OCT volume.

reflect from the retinal tissue. The energy of the interference between the reflected two beams is encoded as intensity in the OCT image. Different depths observed from backscatters is coded with different intensities making it a depth scan also known as A-scan. Aiming to increase the number of A-scans, different approaches have been developed.

1. **Time domain OCT:** Time domain OCT is the initial imaging method in which the reference mirror is moved mechanically to different positions. This mechanical operation restricts the acquisition time of A-scan. The other limitations include poor image resolution and subject inconvenience.

2. **Spectral domain OCT:** Spectral domain OCT differs from its predecessor TD-OCT with an inclusion of a spectrometer in the receiver. This spectrometer according to the Fourier principle assesses the spectrum of reflected light on the retina and transforms it into information about the depth of the structures. The need of mechanically moving reference mirror is eliminated in this approach. This consequently increases the speed at which A-scans are captured and also the axial resolution. A 2-D slice or B-scan is obtained from a single sweep of the scanning beam over the retina in a linear or circular fashion. The increase in the rate of A-scanning created the possibility of capturing multiple B-scans for an examination. Thus, this resulted in imaging retina in 2-D and 3-D with resolution depending on the number of A-scans and B-scans acquired. Figure 1.6 show the acquisition and the coordinate system of the OCT.

(a)   (b)
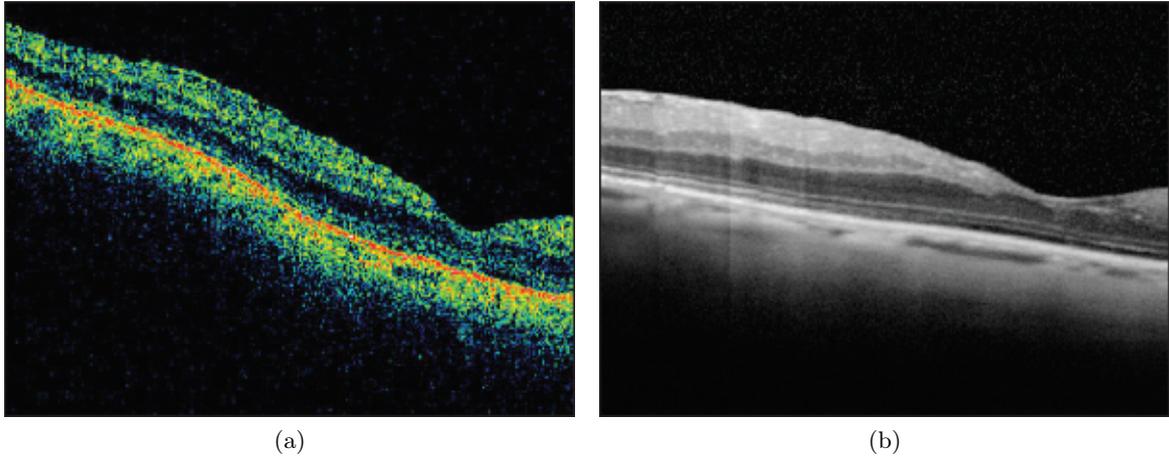
Figure 1.7: Optical Coherence Tomography Imaging . (a) Time domain OCT (b) Spectral domain OCT

Sample Time domain OCT and Spectral domain OCT images are shown in Fig. 1.7. OCT scanning results in cross-sectional images of the retina unlike angiography or fundus photography. Hence, anatomy of the retina greatly facilitates interpretation of normal and pathological conditions.

Optically transparent vitreous is invisible and appears as the black (non-reflective) region in the upper part of an OCT image (see Figure 1.8). This region is not visible in any other imaging modality (Color fundus image or Fundus angiography). Vitreous detachment causes blurry shadows that obscure the vision. The retinal nerve fiber layer (RNFL) appears next as a hyper-reflective band. The thickness of RNFL layers is often measured to assess Glaucoma. RPE is a highly reflective layer visible as a thick line at the base of alternating bright and dark bands. This marks the outer boundary of the retina and is used as a landmark for calculation of retinal thickness. The layers in between together with RNFL and RPE are of great prognostic importance in a range of retinal diseases. The 3D volumetric information from OCT helps in profiling each layer and abnormality across the surface of the retina.

1.2.3 OCT angiography (OCTA)

OCTA is a new imaging modality providing blood flow profiles at various layers of the retina. It is based on a rapid OCT scanning of the eye at the same location in the retina, over a time interval, to look for changes in the scan. All the structures in the retina are static except for the blood flow through vasculature or by movement of the eye itself. Mapping these areas of blood flow is done using the Split-Spectrum Amplitude-Decorrelation Angiography algorithm [3] to derive highly detailed maps of the vasculature in a noninvasive manner as opposed to fluroscein
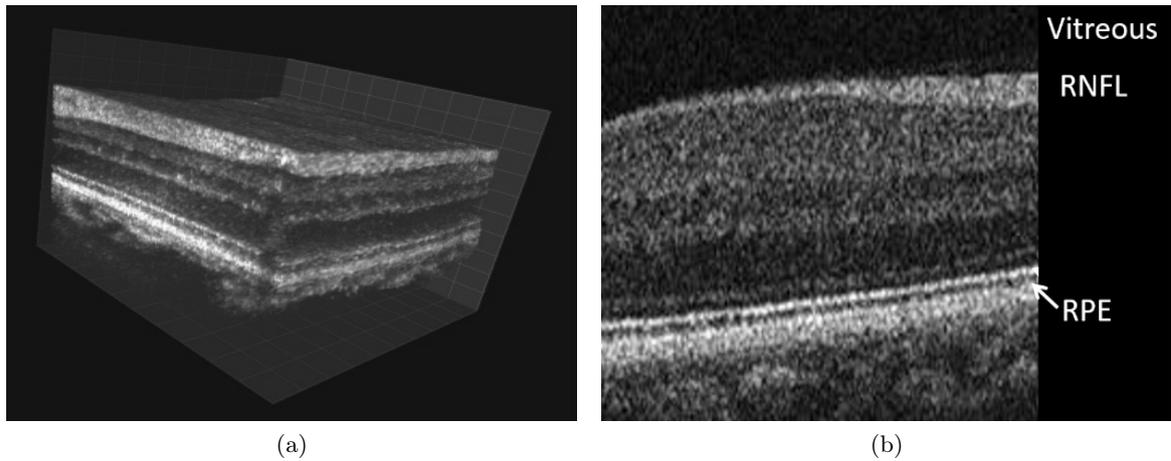
Figure 1.8: 3D visulaization of Retina using OCT (a) and different retinal layers (b).
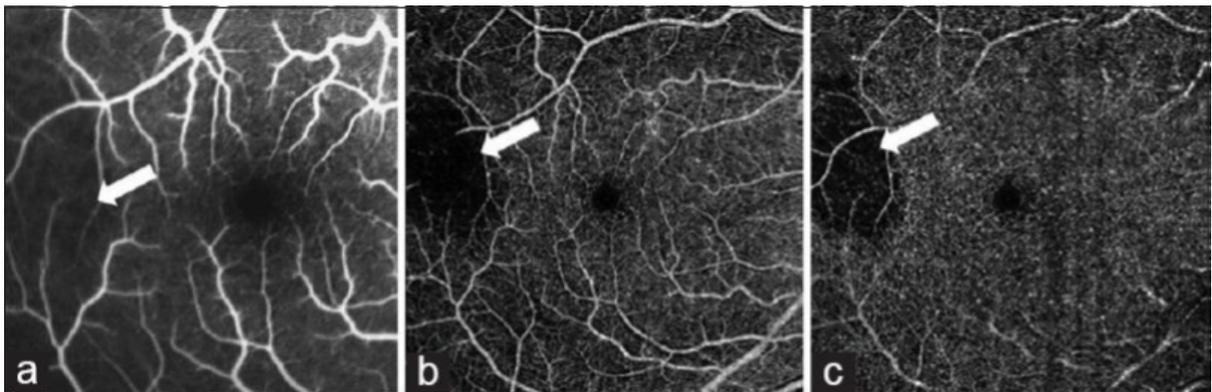
Figure 1.9: Comparison of images of Fluroscein angiography to OCTA enface image. (a) Conventional fluroscein angiography (b) Superficial plexus enface image (c) Deep plexus enface image. (source:[1])

angiography. A comparison of fluroscein angiography in imaging the peripapillary or the deep capillary networks with the OCTA is studied [1] and shown in 1.9

Specifically, in addition to structural information about the retinal layers, OCTA also provides angioflow (enface or top view) images which capture blood flow information at different layers namely, the superficial capillary plexus, deep capillary plexus, outer retina and choriocapillaris area which aid a detailed study of the capillary network. The figure 1.10 illustrates enfaces at different depth of eye and figure 1.10 is a sample slice from an OCTA volume.
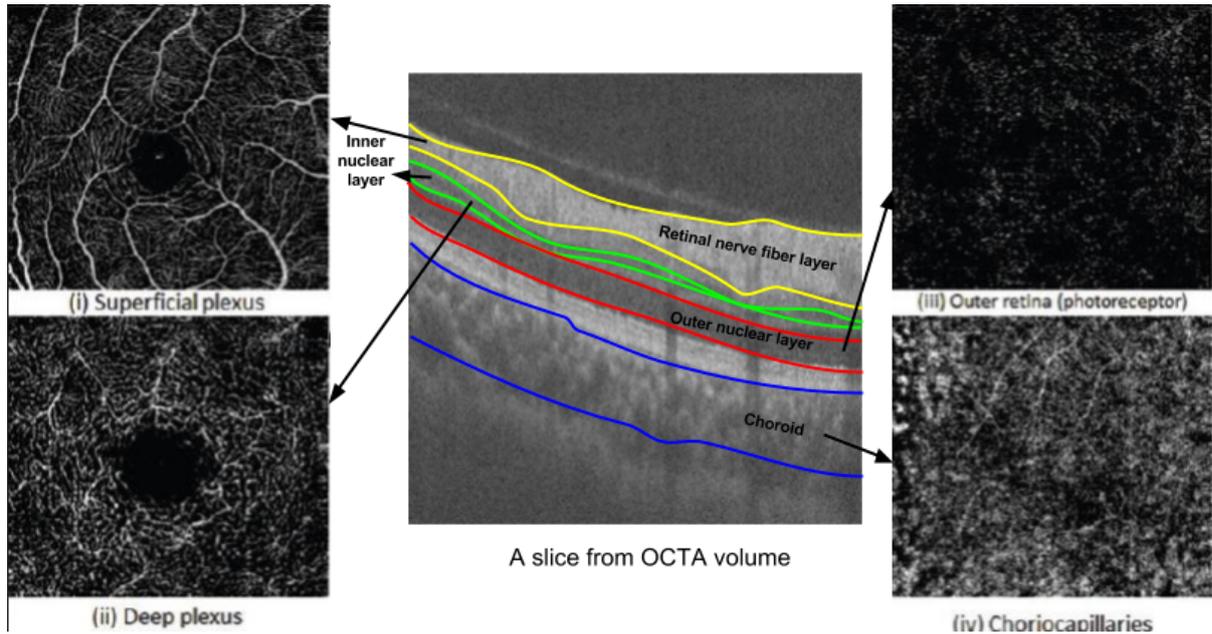
Figure 1.10: The different enface images in relation to a cross-sectional OCT slice. (source:[1])

1.3 MOTIVATION AND THESIS FOCUS

Retina can be examined by visually assessing the images obtained by different imaging modalities. Manifestation of different diseases is clear or invisible according to the imaging modality. Clinical screening of retinal disease is usually done using imaging modalities such as Color fundus image or luroscein angiography or OCT. Color fundus image gives a gross 2D understanding of the health of the eye. Likewise, luroscein angiography captures useful information about retinal vessel integrity. These two imaging modalities are however projection based and lack structural information of the retina. On the other hand, OCT provides morphological information about the examined retinal tissue. OCT provides 3D information which is helpful in estimating the thickness of individual retinal layers and quantitative characterization of abnormality, including area, volume, height. Hence, OCT is widely used in practice,

Cystoidal macular edima results in decreased visual acuity [4]. But, Age related macular degeneration causes irreversible blindness [5]. RNFL thickness is an indicator for assessing glaucoma diagnosis. Early detection of the retinal disease is particularly important to prevent permanent vision loss. Understanding the severity of the disease is crucial for planning image guided therapy. Advancement in OCT imaging modality and its usability for multiple disease assessment is resulting in using it at large screening scenarios. Manually going through each volume for a human grader is laborious. Also, obtaining quantitative information for anatomy and abnormality via manual delineation is often strenuous and time consuming. Automated computer assisted systems are developed with an aim to reduce the human fatigue experienced in assessing
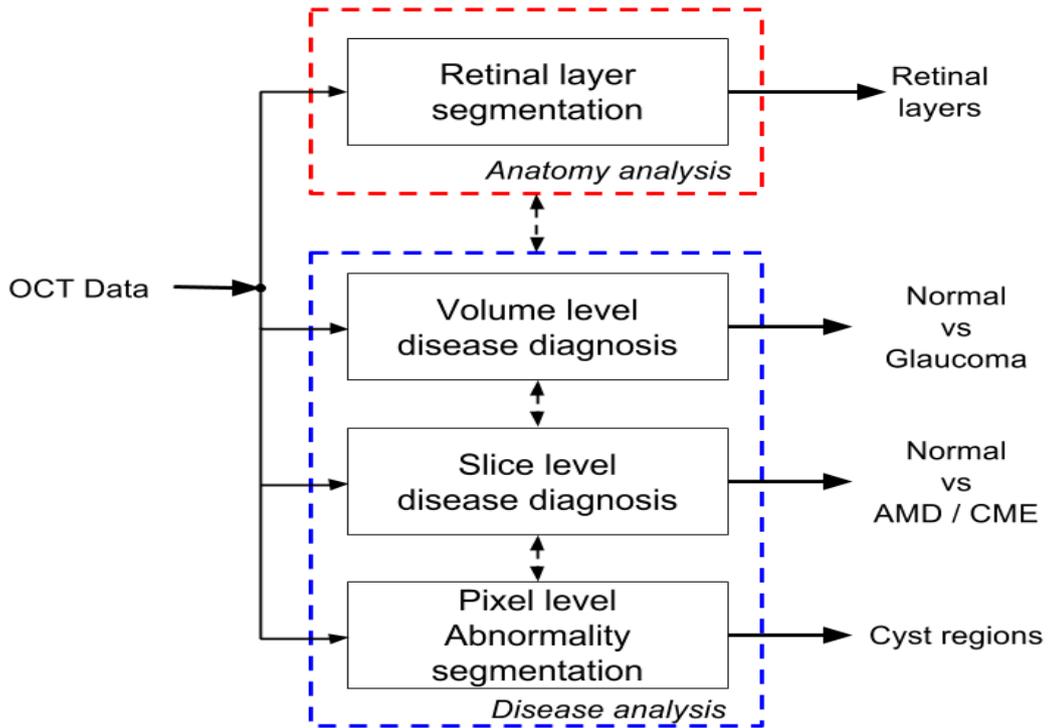
Figure 1.11: Optical Coherence Tomography image analysis system.

medical images. Along with reducing the burden on medical experts, these systems also helps in early detection and providing quantitative information about the disease. This motivates us to develop an automated OCT image analysis system in this thesis.

The proposed system is capable of diagnosing disease and analysing anatomy. The block diagram of the OCT image analysis system is shown in Fig. 1.11. The entire system is developed in four parts. In the first part, we focus on anatomy analysis by developing an automated retinal layer segmentation algorithm. Next, we turn to analysis of the diseases. Initially, we concentrate on developing an algorithm for glaucoma detection at OCT volume level. Next, we develop an algorithm for diagnosis of the retinal disease (Normal vs AMD vs CME) at slice level from an OCT volume. Later, we focus on solving the problem of quantifying abnormality by formulating an automated retinal cyst segmention algorithm. In the Fig. 1.11 there are bi-directional arrow connecting each parts of the system indicating the possibility of information flow within the parts of the system. Combining the four parts, a complete automated OCT image analysis system is formed to assist experts in anatomy and disease analysis. In this thesis, however, we have not made any connections between the parts so as to make it customizable according to the need.

1.4   CONTRIBUTIONS

The major contributions in developing an automated OCT image analysis system described in the thesis are:

1. A novel deep learning based framework for segmenting retinal layers.

2. A novel framework for glaucoma detection from angio-OCT volumes.

3. A Convolutional Neural Network based classifier for slice level diagnosis of retinal disease (Normal vs AMD vs CME).

4. A selective enhancement technique for retinal cysts segmentation using Convolutional Neural Network.

1.5   OUTLINE OF THE THESIS

The organization of the thesis is as follows. In Chapter 2, a deep learning based retinal layer segmentation framework is explained. The novel deep learning architecture has three stages performing layer of interest extraction, edge detection and boundary tracing. The goal of the work is to build a layer segmentation module that can segment retinal layer in presence and absence of pathology. A detailed description of the training procedure and the preparation of the data required for training is given in this chapter. Results of the module are validated against two experts who had marked the retinal layer boundary. Chapter 3 describes the use of angio-OCT data for diagnosis of glaucoma at volume level. A detailed description about angio-OCT and the bio marker used for detecting glaucoma is explained in this chapter. Chapter 4 discusses the disease classification from a given OCT data. The concept of extremal representation from motion pattern for an image is introduced and the use of it for classifying the image is explained. The proposed novel Convolutional Neural Network classifier architecture is also introduced in the chapter. The corresponding input and output for the architecture is discussed in brief and the need for extremal representation is justified. The performance of the classifier for each class is validated by the decisions given by an expert. Chapter 5 discusses the proposed system for localizing and segmenting retinal cysts in detail. The method utilizes a biologically inspired selective enhancement technique for obtaining probable retinal cyst regions. The proposed pipeline along with the architecture of the Convolution neural network used is explained in detail. The results of the work are compared against the available results from state of the art methods. Finally, in the last chapter a general discussion, conclusion and the possible future work is described.

Chapter 2

RETINAL LAYER SEGMENTATION

Optical Coherence Tomography (OCT) is an imaging modality capable of capturing structural composition of retinal tissues at micrometer resolutions. The OCT data are captured with different orientation and scanning protocol. Quantifying the retinal layer thickness is essential for human expert in assessing retinal disease like Glaucoma, Cystoidal Macular Edima and Age related Macular Degeneration. Manually marking accurate layer boundaries on large dataset is laborious and time comsuming. Hence, there is a need to develop automated algorithms to segment visible retinal layers for assisting the experts.

2.1 INTRODUCTION

OCT imaging of a retina aids in visualizing seven structural layers. They are (from top to bottom): Retinal Nerve Fiber Layer (RNFL), Ganglion Cell Layer and Inner Plexiform Layer (GCL+IPL), Inner Nuclear Layer (INL), Outer Plexiform Layer (OPL), Outer Nuclear Layer (ONL), Inner Segment (IS), Outer Segment (OS), Retinal Pigment Epithelium (RPE). An illustration of the same is shown in Fig. 2.1 (a). Accurate segmentation of these layers is necessary to quantify the morphological changes in the retinal tissue that aid in the detection of ocular diseases. The layer information is also further used a position prior for localizing abnormalities. In AMD, the drusen deposits in the RPE layer lead to irregularities and undulations in the RPE_{in} boundary as depicted in Fig. 2.1 (b). In absence of pathologies, assessing each layer in terms of its thickness helps in understanding the progressive health of the retina.

2.1.1 Long Short Term Memory networks

Recurrent neural network (RNN) are neural networks with loops in them. This allows them to retain the information by remembering the past. RNN's in practice fail to learn when the
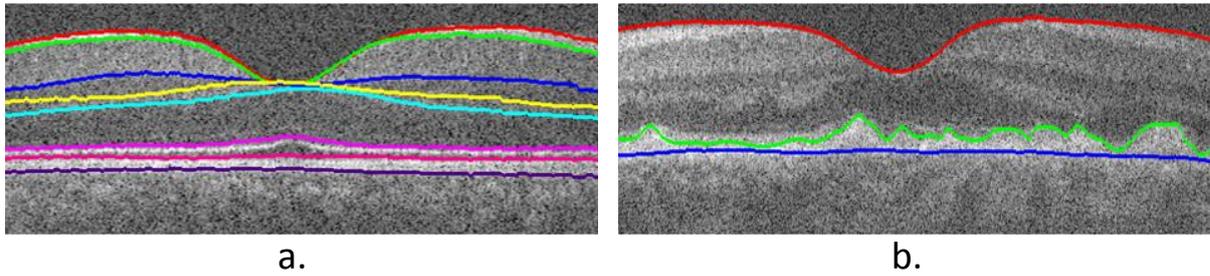
Figure 2.1: Retinal layer boundaries in OCT B-scans of a) Healthy retina, listed from top to bottom: ILM(Red), NFL/GCL(Green), IPL/INL(blue), INL/OPL(yellow), OPL/ONL(cyan), IS/OS(magenta), RPE_{in}(pink) and RPE_{out} (purple); b) Retina with AMD: ILM(Red), RPE_{in}(Green) and RPE_{out}(Blue).

gap between past and present become very large. This is refered to as long-term dependency. Long Short Term Memory networks (LSTM) [6] are a special kind of RNN, capable of learning long-term dependencies. A detailed walk-through on working of LSTM is explained in the blog [7].

2.2 RELATED WORK

Automating the layer segmentation task is challenging due to the presence of speckle noise, vessel shadows and varying layer orientation. These were generally handled using a set of modules as preprocessing. Denoising relied on methods such as median filtering, block-matching and 3D filtering [8] and diffusion filtering [9] [10] [11]. Vessel shadows were removed by explicitly masking out such regions found via vertical projections [12]. Variable orientation of the layers across the dataset was addressed by flattening the structures with respect to one of the roughly estimated layers [13]. All these steps are data dependent and hence require tuning. In layer segmentation, the uppermost and the lowermost boundaries (Vitreous-RNFL and RPE-Choroid) are marked by intensity gradients and hence this information has been used to extract them[14] [15]. Gradient and intensity information along with an active contour approach [16] has also been proposed. By far the most popular approach is based on graph search. In this class there are techniques which use intensity, gradient and 3D context based cost function for optimization [17], shortest path computation with Dijkstra's algorithm [18] and graph based tracing of the layer boundary [13]. These methods detect layers in a sequential fashion by constraining the ROI after each detected layer. Most of these algorithms were proposed for segmenting retinal layers in a normal case.

Presence of pathologies alter the layer morphology locally and thus increases the complexity of the problem. Automatic segmentation of three layers relevant to age related macular generation

and geographic atrophy was proposed [19] by adapting the edge weights used in graph theory and dynamic programming based framework [13]. More recently, information such as slope similarity and non-associativity of layers as edge weight have been explored to handle pathologies [20].

Thus, existing techniques employ *many* (pre)processing blocks all of which require separate tuning of parameters and modify approaches designed for normal cases to handle pathologies. This limits the robustness of the methods. Deep neural networks offer a way to learn the main segmentation task in addition to these early processes. In this chapter, we propose a novel supervised method for layer segmentation applicable to both normal and pathology cases. It is based on a combination of Convolutional Neural Network (CNN) and Bidirectional Long Short-term Memory (BLSTM). The major strengths of the proposed method are (i) no requirement for any preprocessing (ii) one-shot solution multi-layer segmentation (iii) robustness to presence of pathologies (iv) robustness to imaging systems and image quality.

2.3   PROPOSED ARCHITECTURE

OCT images are corrupted due to presence of speckle noise whose characteristics vary across scanners. In order to extract layers of interest which is agnostic to the source of data (scanner), we use a CNN-based first stage. The presence of vessel-shadows and pathologies cause the boundaries between retinal layers to be discontinuous. Naive edge detection algorithm fails to extract eight *continuous* boundaries shared by seven layers. Hence, layer segmentation from the output of first stage is achieved using a cascade of stages: edge detection followed by boundary-tracing. A CNN-based strategy is adopted for the former while for the latter, a specific type of Recurrent Neural Network, namely LSTM is adopted. The LSTM stage learns to trace eight continuous boundaries by following the detected edges, with continuity ensured by using a bidirectional LSTM (referred as BLSTM).

The custom-designed architecture for the proposed system is shown in Fig.2.2 and details are provided in Table 2.1 with CONV being Convolution layer and ACTN as Activation. We describe each stage in the system next.

Table 2.1: Description of the proposed architecture

LOI extraction				Edge detection				Boundary-tracing			
Module	Layer	Filter size	ACTN	Module	Layer	Filter size	ACTN	Module	Layer	Output Nodes	ACTN.
HM1	CONV	(20x30x2)x32	Relu [21]	HM1	CONV	(15x20x9)x16	Relu	fwd-LSTM1	LSTM	64	-
HM2	CONV	(20x30x32)x32	Relu	HM2	CONV	(15x20x16)x16	Relu	bck-LSTM1	LSTM	64	-
VM1	CONV	(30x20x2)x32	Relu	VM1	CONV	(20x15x9)x16	Relu	fwd-LSTM2	LSTM	32	-
VM1	CONV	(30x20x32)x32	Relu	VM1	CONV	(20x15x16)x16	Relu	bck-LSTM2	LSTM	32	-
SM1	CONV	(10x10x64)x32	Relu	SM1	CONV	(10x10x32)x16	Relu	TD	Fully connected	8	Sigmoid
SM2	CONV	(5x5x32)x9	Sigmoid	SM2	CONV	(5x5x16)x1	Sigmoid				
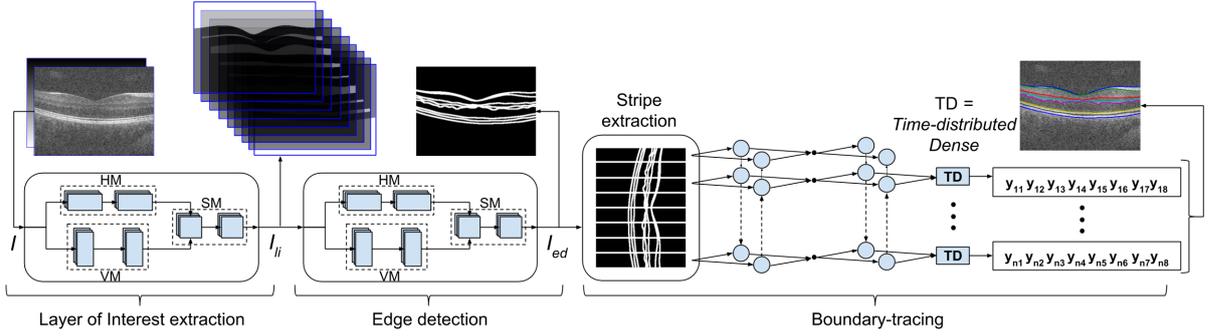
Figure 2.2: Deep learning architecture for layer segmentation.

Stage 1: Layer of interest (LOI) extraction. The input to the CNN is a stack of OCT image I in addition to a position cue image I_{pc} which is defined as $I_{pc}(x,y) = y$. The input passes through 2 parallel modules: Horizontal-filter Module (HM) and Vertical-filter Module (VM). Resultant activations are stacked and merged using a Square-filter Module (SM) to generate a LOI image I_{li} with 9 channels, each depicting one region(7 layers plus the vitreous cavity and choroid). Ideally, each channel of I_{li} is expected to be a Boolean image with 1 denoting inside the respective regions and 0 denoting exterior points. Along with learning interested layers, HM and VM learn intra- and inter-layer characteristics respectively. HM should ideally learn to inpaint pathologies and vessel-shadows, based on the horizontal neighborhood belonging to the same layer. VM should learn to differentiate two neighboring layers. SM ensures equal contribution from both horizontal and vertical neighbors. I_{li} is passed to next stage as well as taken as an independent side-output.

Stage 2: Edge detection. This stage is implemented with a second CNN. LOI image I_{li} is passed through HM, VM and SM similar to stage 1. Here HM and VM learn edges with horizontal and vertical orientations respectively. Both edge-maps are combined using SM to generate a single map I_{ed} capturing edges of all orientations. I_{ed} is passed to next stage as well as taken out as a side-output.

Stage 3: Boundary-tracing. This is implemented using a BLSTM. For each column i of an image, eight boundary-coordinates ($L_i(j), j \in \{1, 2, ..., 8\}$) depend on the edge passing through the neighboring columns. Hence, a stripe representing information of neighboring columns is extracted (*online*) from I_{ed} as follows. I_{ed} is shifted left and right twice ($I_{ed}(x - k, y); k \in \{0, \pm 1, \pm 2\}$) and stacked such that each column is aligned with its neighbors in the 3rd dimension. Each column of this stack is termed as 'stripe'. Stripe extraction from the edge image is illustrated in the Fig. 2.3. Columns of the edge map are separated for the visualization purpose.

Extracted stripes are sequentially passed to the two-staged BLSTM. A BLSTM has two LSTMs, each learns to generate boundary-coordinates for the current stripe in continuation with coor-
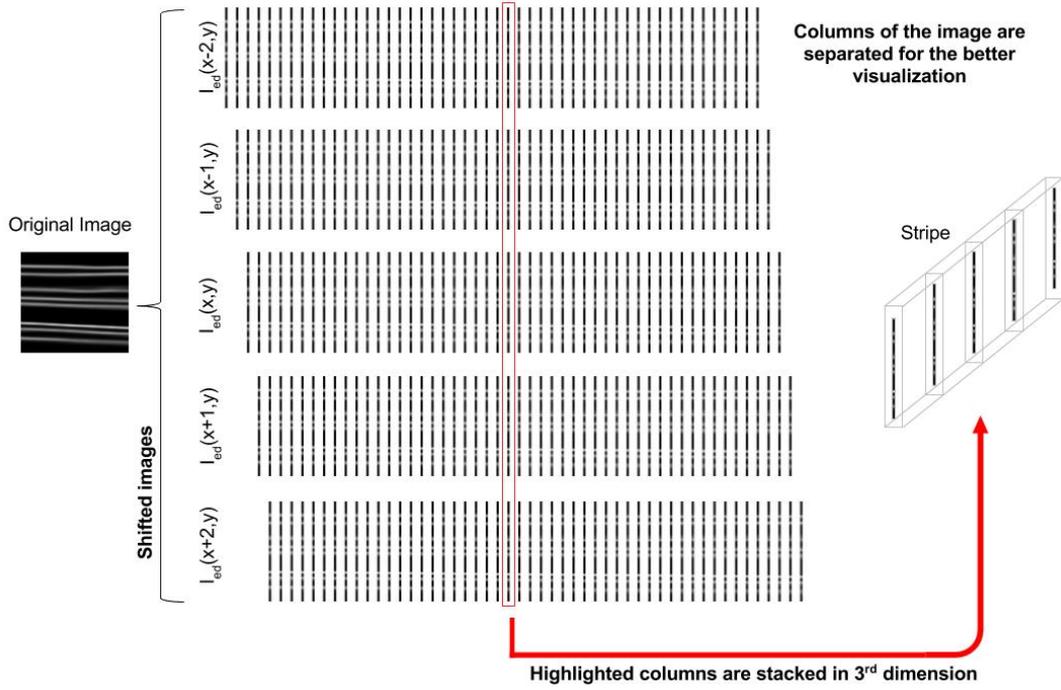
Figure 2.3: Stripe extraction by column stacking

dinates from the right and the left neighboring stripes respectively. Estimated $L_i, i \in \{1, 2, ...,$ No. of columns$\}$ traces the desired 8 layer boundaries simultaneously across the image.

2.4   DATA AND TRAINING METHOD

The proposed system was trained and evaluated using publicly available datasets.

2.4.1   Dataset description

Dataset with normal cases: Two publicly available datasets were considered. The first (we refer to as $Chiu_{norm}$) is made available by Chiu *et al.* [13] and contains 110 B-scans from 10 healthy subjects (11 B-scans per subject) along with *partial* manual markings from two experts. The authors state that, "each expert grader exhibited a bias when tracing layer boundaries" and "manual segmentation tended to be smooth". Thus, manual segmentation does not follow edges tightly. The second dataset (called OCTRIMA3D) is made available by Tian *et al.* [18] and contains 100 B-scans from 10 healthy subjects(10 B-scans per subject); manual markings by two observers are also provided. The authors emphasize that in contrast to the smooth manual labellings in [13], the delineated boundaries in OCTRIMA3D trace small bumps.

(a) Original image

(b) Image after column rolling
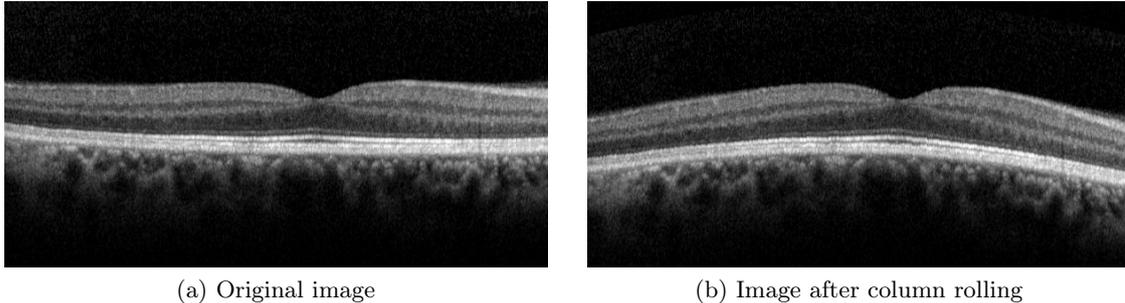
Figure 2.4: Data augmentation using column rolling.

Dataset with pathology cases [19]: This is a dataset made available by Chiu *et al.* [19] and it consists of a total of 220 B-scans from 20 volumes. This dataset is characterized by the presence of pathologies such as drusen and geographic atrophy and hence we refer to it as $Chiu_{path}$. Manual segmentation from two experts are available for only 3 layer boundaries (Vitreous-RNFL, OS-RPE and RPE-Choroid). The dataset includes scans with varying image quality.

2.4.2 Preparation of Training data

Input image I. The datasets have been acquired with varying scanning protocols and hence vary in resolution and layer orientation. This is addressed by standardizing the images to 300×800 pixel resolution as follows. Columns of each image were summed to obtain a 1D projection and a Gaussian function was fitted on the same. The mean value of the Gaussian represents the y-coordinate of the center ($C_{ROI}(x, y)$) of a region containing the layers. Images were shifted/cropped/padded vertically such that resultant height is 300px and y-coordinates of C_{ROI} and the center of the image are aligned. Next, images were padded horizontally with trailing zeros to achieve a standardized width of 800px. The position cue image I_{pc} was generated with pixel value proportional to the y-coordinate of the pixel location. Finally, I and I_{pc} were stacked and the pixel intensities were normalized to have zero mean and unit variance.

Layer boundary coordinates $L_i(j)$. Since the $Chiu_{norm}$ dataset provides partial/discontinuous manual markings from two experts whereas our network requires unique and continuous GT for each image, partial markings from the 1st expert was given to a local expert who completed them. It is to be noted that subjective bias and smoothness of marking maintained by the 1st expert is impossible to be reproduced by the local expert. Also, the markers for all three datasets used for training are different and exhibit different subjective biases and smoothness of marking. In summary, GT used for training is noisy given that they are derived from 3 different markers.

17

LOI image I_{li}. Each I can be divided in 9 regions demarcated by 8 boundaries. 9 binary images were defined to represent each region (see Fig 2.2) which were stacked to generate I_{li}.

Edge-map I_{ed}. For all Is, I_{ed} is defined as a binary image indicating 1 at the boundary pixel and 0 elsewhere (see Fig 2.2).

As mentioned earlier, GT for pathological cases have markings for only 3 boundaries. Hence, $L_i(j)$, I_{li} and I_{ed} for the $Chiu_{path}$ dataset were modified to have only 3 boundaries.

2.4.3   Training

Two copies of the proposed system were trained, one for only normal cases and one for both normal and pathological cases. The first copy, referred to as M_{norm}, was trained for 8 layer-boundaries using two normal datasets. The second copy, referred to as M_{mixed} was trained for detecting the 3 boundaries for which manual markings were available. M_{mixed} was trained using the GT for 3 boundaries in $Chiu_{path}$ dataset and the same ones in $Chiu_{norm}$ and OCTRIMA3D.

For both M_{norm} and M_{mixed}, datasets were divided into training and testing set using a split of 8:2. Equal portion of pathological and normal cases were taken for training and testing M_{mixed}. Online data augmentation was done by applying random rotation, scaling, horizontal and vertical flips, shift and column rolling wherein neighboring columns are rolled up/down systematically (see Fig.2.4(b)).Training of the entire network was done in an end-to-end fashion. All slices(11 B-scans) of a single volume constituted a batch while training. Binary cross-entropy and Mean Square Error loss functions were used at stages-1,2 and stage-3 respectively. ADADELTA optimizer [22] was used with sample *emphasizing scheme* [23]. *Emphasizing scheme* chooses training example with large errors more frequently. This ensures that network sees informative data more often.

Result: L_i
initialize L_0;
for $i \leftarrow 0$ ***to*** *nb_epoch* **do**
$\quad$ $L_a = L_0/(i+1)$;
$\quad$ **for** $b \in batches$ **do**
$\quad\quad$ **for** $j \leftarrow 0$ ***to*** *9* **do**
$\quad\quad\quad$ $[L_{li}, L_{ed}, L_L] = \text{NN}(X[b], Y[b])$;
$\quad\quad\quad$ **if** $L_L < L_a$ **then**
$\quad\quad\quad\quad$ break;
$\quad\quad\quad$ **end**
$\quad\quad$ **end**
$\quad$ **end**
end

Algorithm 1: Homogeneous Training
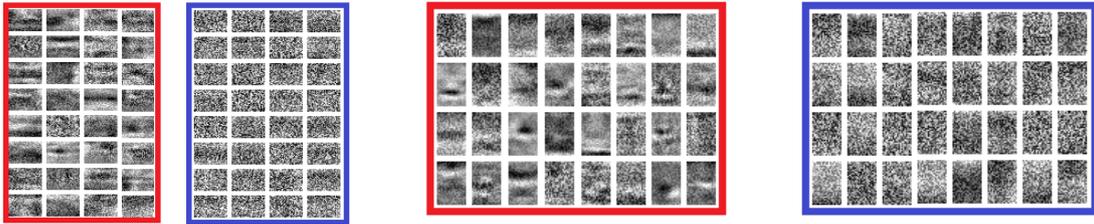
Figure 2.5: Visualization of learnt filters

The proposed pipeline was implemented on a NVIDIA GTX TITAN GPU, with 12GB of GPU RAM on a core i7 processor. The entire architecture was implemented in Theano using the Keras library. Training for 250 epochs took about a week.

2.5 RESULTS

Proposed model has two side outputs (in Fig. 2.6) at stages 1 and 2 respectively. Side output at 1st stage is a 9-channel image. Each channel represents one region (Vitreous, RNFL, GCL+IPL, INL, OPL, ONL+IS, OS, RPE, Choroid). Side output at 2nd stage is a single channel image representing the edges between the two consecutive layers. Both predicted and GT are shown for each stage. Middle panel shows the proposed deep learning model. Top panel depicts side output from M_{norm} network. Bottom panel depicts side output from M_{mixed} network. Learnt filters at layer of interest extraction stage Fig. 2.5 HM1(left) and VM1(right) is shown. Each colored box represents one filter and the subimages inside the box shows channels. Red filters are applied to image I and blue filters are applied to position cue image I_{pc}. It can be observed that both red filters have learnt to retain edges and blobs from noisy input. Blue filters of VM1 has learnt useful vertical gradient in I_{pc}. Blue filters of HM1 could not learn anything useful as there is no horizontal gradient in I_{pc}.

The networks M_{norm} and M_{mixed} were evaluated separately using the mean absolute pixel error (MAE) and layer thickness. M_{norm} was trained for predicting eight layer boundaries on $Chiu_{norm}$ and $OCTRIMA3D$ datasets. M_{mixed} was trained with normal and pathological data ($Chiu_{path}$, $Chiu_{norm}$ and $OCTRIMA3D$) for estimating three layer boundaries. Benchmarking for each dataset was done against the output provided by the authors of the datasets.

Qualitative results comparing manual marking to predicted layer boundaries are shown in Fig. 2.7 and Fig. 2.8. Tables 2.2 and 2.3 list the obtained MAE for the M_{norm} and M_{mixed} networks, respectively. Likewise, Tables 2.4 and 2.5 show the mean layer thickness for M_{norm} and M_{mixed} network. The overall MAE obtained by our system is lower than the inter marker error on all the datasets and our results are comparable to the benchmarks at a subpixel level. It is to be noted that MAE for normal cases is less with M_{mixed} than that achieved for the same cases

Table 2.2: Pixel-level Mean Absolute Error(MAE) for predictions with M_{norm} network. Values indicate mean±std.

Boundary	$Chiu_{norm}$			OCTRIMA3D		
	Inter marker	[13]	ours	Inter marker	[18]	ours
Vitreous - RNFL	2.37±0.79	1.38±0.37	1.11±0.30	1.00±0.24	0.68±0.20	1.49±0.47
RNFL - GCL&IPL	1.73±0.85	1.67±0.77	1.38±0.42	1.70±0.76	1.16±0.34	1.56±0.38
GCL&IPL - INL	1.81±1.44	1.48±0.58	1.42±0.58	1.79±0.47	1.01±0.15	1.24±0.30
INL - OPL	3.02±0.87	1.48±0.46	1.60±0.32	1.44±0.33	1.11±0.41	1.39±0.51
OPL - ONL&IS	2.18±0.97	1.74±0.65	1.88±0.65	1.83±0.60	1.50±0.77	1.78±0.77
ONL&IS - OS	2.85±0.93	1.00±0.30	0.92±0.34	0.76±0.22	0.54±0.10	0.91±0.25
OS - RPE	1.88±1.08	1.14±0.40	1.01±0.23	1.81±0.87	1.22±0.53	1.09±0.32
RPE - Choroid	2.18±1.69	1.26±0.35	1.43±0.68	1.22±0.22	0.76±0.17	0.98±0.27
Overall	2.25±1.08	1.39±0.48	1.34±0.44	1.44±0.47	1.00±0.33	1.30±0.41

Table 2.3: Pixel-level Mean Absolute Error for predictions with M_{mixed} network. Values indicate mean±std.

Boundary	$Chiu_{path}$			$Chiu_{norm}$			OCTRIMA3D		
	Inter marker	[19]	ours	Inter marker	[13]	ours	Inter marker	[18]	ours
Vitreous - RNFL	1.25±0.39	1.14±0.339	0.95±0.28	2.37±0.79	1.15±0.32	1.02±0.28	1.00±0.24	0.70±0.25	1.03±0.35
OS - RPE	2.56±0.75	2.53±0.83	2.41±0.77	1.88±1.08	0.99±0.20	1.12±0.44	1.81±0.87	1.24±0.58	0.87±0.20
RPE - Choroid	1.55±0.71	1.46±1.13	1.97±1.23	2.18±1.69	1.41±0.65	1.32±0.67	1.22±0.22	0.73±0.18	0.92±0.33
Overall	1.79±0.62	1.71±0.76	1.78±0.78	2.14±1.19	1.18±0.39	1.16±0.46	1.34±0.45	0.89±0.348	0.92±0.31

Table 2.4: Mean layer thickness predictions for seven retinal layers with M_{norm} network. Values indicate mean±std.

Layers	$Chiu_{norm}$				OCTRIMA3D			
	Ours	Chiu	man1	man2	Ours	Chiu	Grader 1	Grader 2
Layer 1	7.8417±0.8470	7.8252±2.1241	8.3478±2.5513	9.8632±3.2022	8.7981±1.4777	9.30912.9554	9.5904±3.4771	10.2503±3.2559
Layer 2	20.7719±2.4274	20.1676±3.0416	19.8303±2.7098	18.8198±4.3704	16.4130±2.7645	17.9798±3.5001	17.1814±3.6353	17.3523±3.7262
Layer 3	10.4254±1.5564	11.1408±1.4041	10.8679±1.3723	8.6956±1.9429	8.3506±1.1843	7.2715±1.1823	7.1477±1.1417	6.3490±1.0733
Layer 4	8.54240.5255	8.4742±2.2051	9.4232±1.7948	9.7015±2.4013	7.9476±0.6634	8.0451±1.2718	7.8069±1.2631	8.3618±0.9848
Layer 5	24.5140±3.6324	25.2039±4.3675	24.5995±3.6692	23.7206±6.5841	19.5170±2.3513	19.7355±2.3904	20.7318±2.9901	19.5959±2.9799
Layer 6	7.7199±0.5948	8.1589±0.9218	8.5652±1.0619	9.6131±1.1763	6.8007±0.5259	6.8594±0.9631	6.9885±0.8454	8.4996±0.9128
Layer 7	9.7254±0.4684	8.6572±1.4035	8.1717±1.6276	10.6325±1.9408	10.4403±0.2151	10.5987±1.0266	10.1139±0.8619	8.9593±0.9167

Table 2.5: Mean layer thickness predictions for two retinal layers with M_{mixed} network. Values indicate mean±std.

	$Chiu_{path}$		$Chiu_{norm}$		OCTRIMA3D	
	Total Retina	RPEDC-BM	Total Retina	RPEDC-BM	Total Retina	RPEDC-BM
Ours	81.7147±7.6370	10.1044±0.4008	81.7783±5.6924	9.8518±0.2378	70.7439±6.7810	10.4095±0.3476
Chiu	81.8651±7.4941	9.9326±3.4077	81.1351±5.0261	8.5383±1.3996	70.2256±5.3949	10.8160±1.1683
Mar1	79.8103±6.8810	10.4770±3.3228	81.8147±4.9042	8.2025±1.5518	70.5499±5.5113	10.1879±0.9305
Mar2	79.9374±6.6825	11.1122±4.2274	80.9319±7.4912	0.5320±1.7887	71.4497±5.7757	9.1182±1.0223
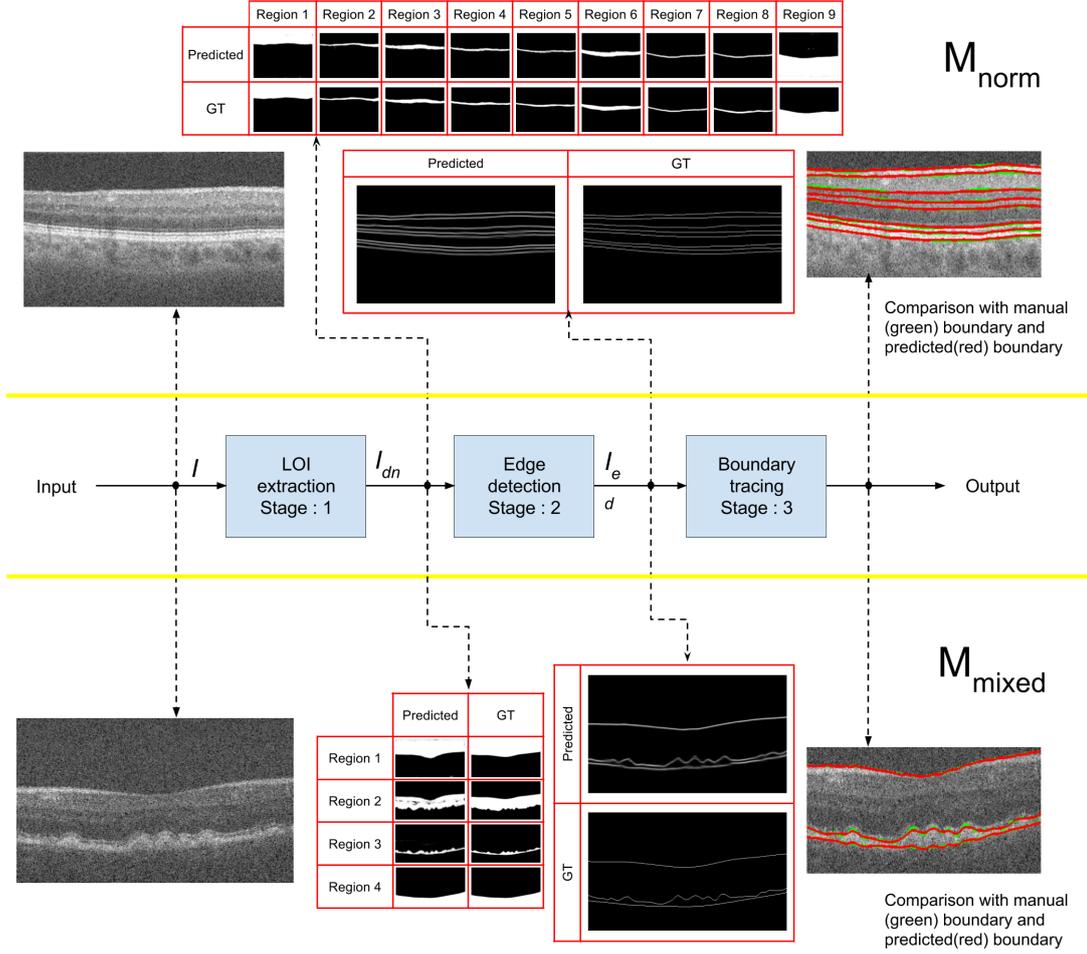
Figure 2.6: Visualization of side output

with M_{norm}. This is due to the larger training set for M_{mixed} compared to M_{norm}. However, the improvement appears to be less for the OS-RPE boundary as the variability to be learnt has also increased. Although training data has increased, it is not enough to learn this increased variability.

2.6 CONCLUDING REMARKS

We proposed a solution for OCT layer segmentation problem using deep networks. It is in the form of a single system which integrates all the processing steps and segments all the layers in parallel. Results show the system to be robust to change in scanner and image quality. These are the major strengths of our method. The performance of this system, even with the limited training, is comparable to the existing benchmarks. Increase in the number of training cases has
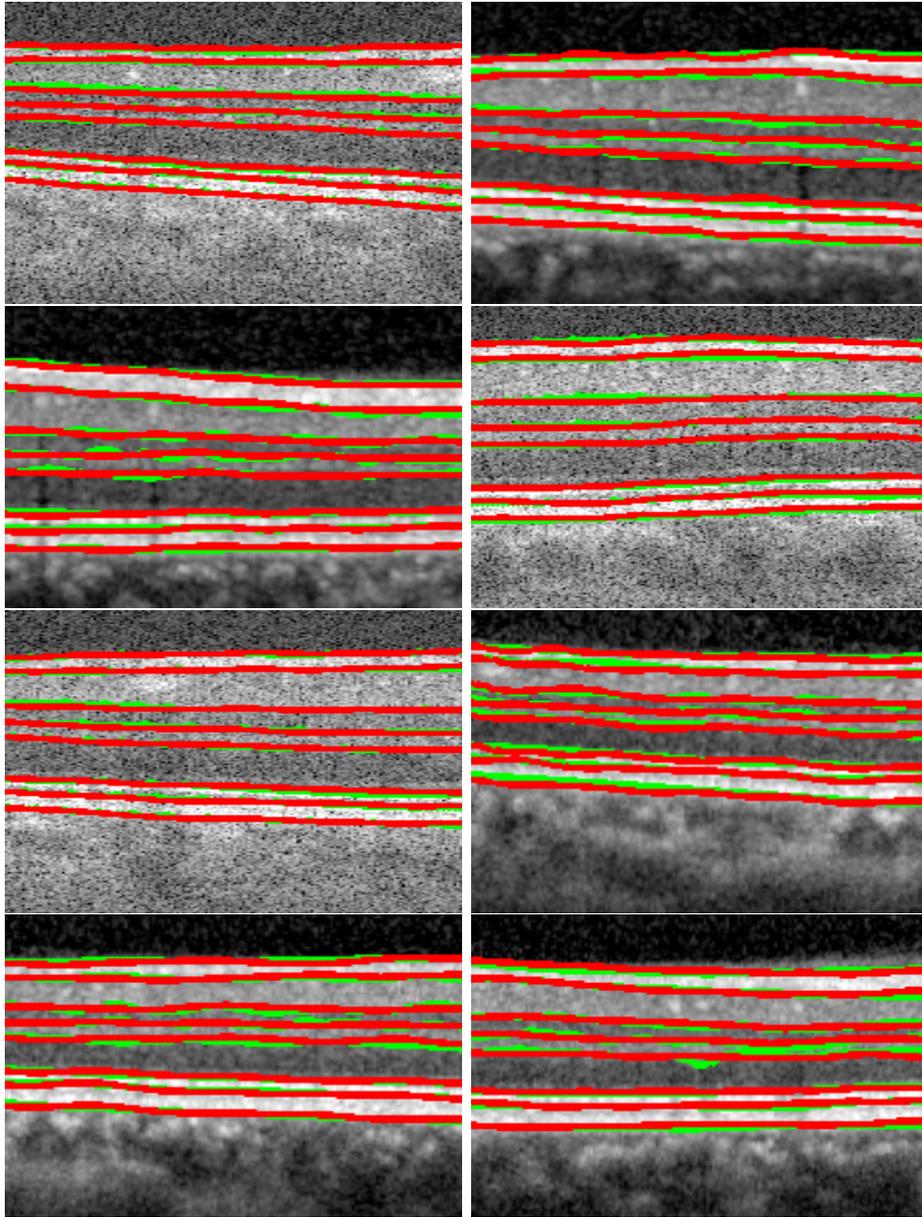
Figure 2.7: Comparison with manual (green) segmentation. Results (in red) for a pathological case predicted by M_{norm}.
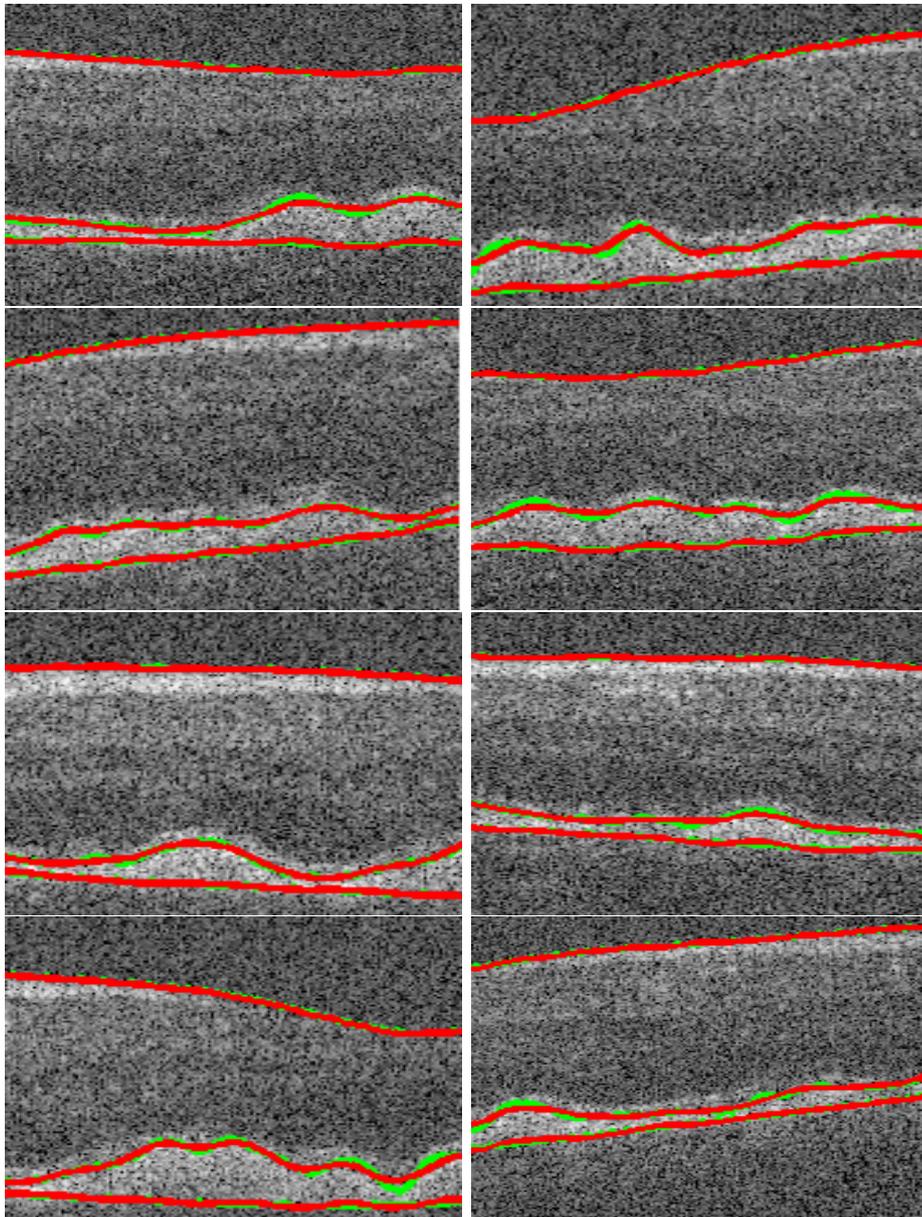
Figure 2.8: Comparison with manual (green) segmentation. Results (in red) for a pathological case predicted by M_{mixed}.

allowed the system to learn more accurate segmentation boundaries. This shows potential for further improvement with larger dataset. Our model also handles both normal and pathological cases without any data dependent tuning. There are two limitations of the proposed work. First is the fact that the ground-truth used for training was sourced from multiple, biased markers. This created ambiguities in the training process. Second is the mismatch in the precision of the output of the edge-detection stage and that required by the input of BLSTM. The former is at most 1 pixel whereas the latter requires sub-pixel precision. The 1 pixel wide edge causes a confusion field for the BLSTM while tracing precise boundaries. Nevertheless, the proposed system for layer segmentation is robust, fast, automatic with potential for further improvements.

Chapter 3

VOLUME LEVEL DISEASE DIAGNOSIS

Glaucoma is a common cause of irreversible blindness. Prevention of glaucoma requires early detection of the disease which in turn depends on the ability to recognize early clinical manifestations. Clinically, retinal experts asses glaucoma by estimating the reduction in retinal nerve fiber layer thickness using structural images. Studies [24] [25] have shown that the death of functional blood capillaries result in the thinning of the retinal nerve fiber layer. Hence, using information (structural and functional) from OCTA modality will help in accurate early prediction of glaucoma.

3.1 INTRODUCTION

The optic disk (OD) is the location where ganglion cell axons exit the eye to form the optic nerve through which visual information of the photo-receptors is transmitted to the brain. The structures within OD region and in its peripheral regions are in Figure 3.1. In glaucoma, the nerve cells and nerve fibers which form connection between eye and brain, die progressively. The Retinal nerve fiber layer (RNFL) thickness is a biomarker for glaucoma [26] and hence, OCT scanner-generated reports generally provide the *deviation* of RNFL thickness of a current measurement from normative database across various regions around the optic disk. Clinically, glaucoma is characterized by enlargement of the optic cup and peripapillary RNFL loss. The cup enlargement is quantified by cup to disc diameter ratio and can be estimated either from colour fundus image or OCT. Some automated methods for computing the cup to disc diameter ratio have been put forth in [27] and [28]. Visualization of retinal nerve fiber layer is difficult using fundus imaging. RNFL thickness measurement by OCT is possible and useful for early diagnosis of glaucoma [29]. Spectral domain OCT is widely used in ocular disease detection and algorithms have been proposed for segmenting the layers from OCT images [30] [31]. Clinical research has also aimed at understanding the cause of glaucoma from *vascular health* point of view. To study
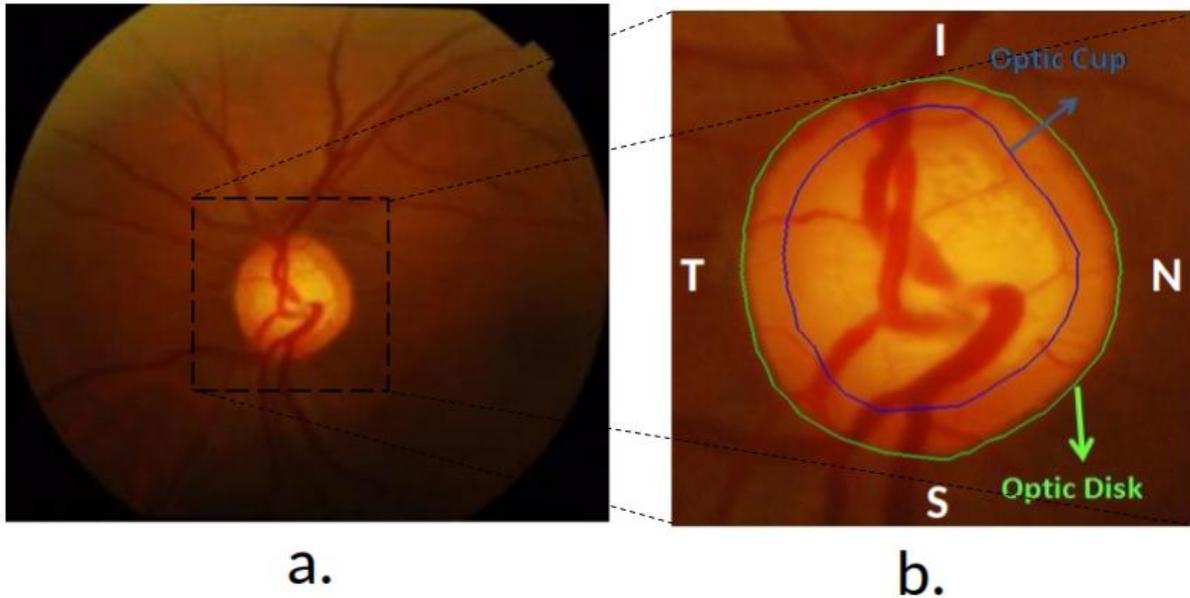
Figure 3.1: A retinal image (a) and a Region of interest (b) around optic disk and related structures.

the correlation between RNFL thickness and capillary volume, postmortem studies on eyes have been done to find selective atrophy of radial peripapillary capillaries in the superficial layers in glaucomatous eyes [24] and in normal eyes [25]. Fluorescein angiography conducted on 75 subjects (50 with glaucoma) has also provided evidence for correlation between glaucomatous disc damage and capillary drop out in the surface layers [32]. OCT based angiograpy (OCTA) provides blood flow information along with the structural information.

3.2 PROPOSED METHOD

OCTA data provides structural information about the retinal layers as a 3D volume and the 2D angioflow images provide information about blood flow in a specific layer. The input for the proposed system is an OCTA volume and four angioflow images of the Choroid disc, Nerve head (NH), Radial Peripapillary Capillaries (RPC) and the Vitreous layer. The proposed system comprises of 3 stages: Extraction of region of interest (ROI) centred around OD; Feature extraction, explained in section 3.2.2 and Classification. The ROI is an annular region around the ONH. The capillary network is extracted using vessel detection and suppression of large vessels from the angioflow images. The RNFL layer thickness is calculated based on the intensity profile (at the slice level) from the OCTA volume. A linear SVM classifier is trained using these information for the classification task.
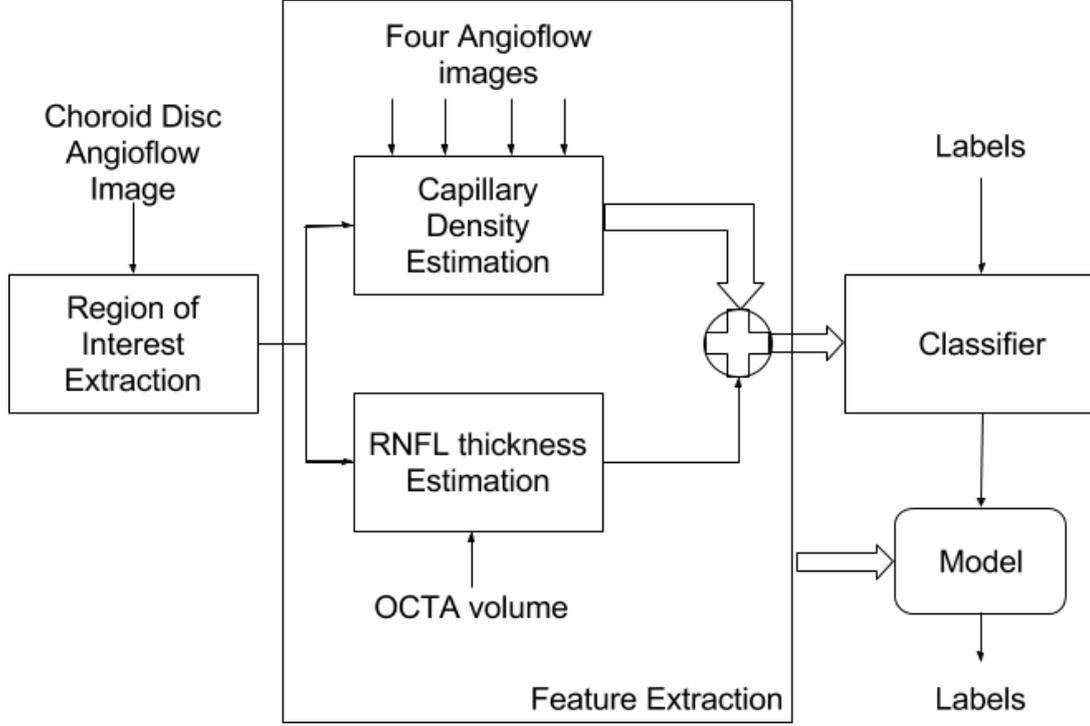
Figure 3.2: Block diagram of the proposed method.

3.2.1 Extraction of Region of interest

Among the 4 OCTA angioflow images, the ONH boundary has the best definition and hence is detectable in the choroid disc angioflow image (I_{cen}). A sample choroid disc angioflow image is shown in Fig. 3.3. The central dark region is the ONH. I_{cen} is used to extract the ROI as follows: The local minima in the image is found using a filter of size 3×3, following which all small objects are removed; The centroid of the large central dark object is found and a circle fitting operation is used to extract the ONH region; Finally, an annular region of 100 pixels width around the detected ONH is extracted as the desired ROI. The ROI is divided into eight sectors with sector angle equal to 45° (Fig. 3.3) to identify the Temporal Upper (TU), Temporal Lower (TL), Superior Temporal (ST), Inferior Temporal (IT), Nasal Upper (NU), Nasal Lower (NL), Superior Nasal (SN), Inferior Nasal (IN) regions. The nasal and temporal sectors are mirrored about the vertical for the right and left eyes.

3.2.2 Feature extraction

For every sector in the ROI, two metrics, namely the capillary density (CD) and RNFL thickness are of interest. The procedure for deriving these metrics is explained next.
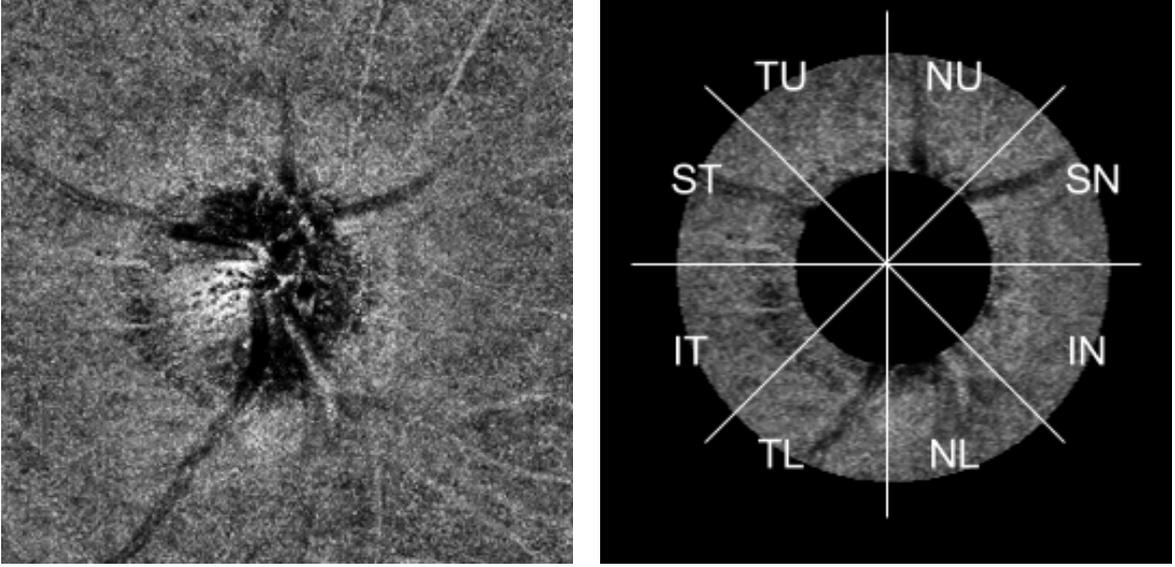
Figure 3.3: Choroid Disc Angioflow image (left) and Eight sectors of ROI (right) for a Right Eye.

3.2.2.1 Capillary Density estimation

An angioflow image is the input for the CD computation. Non-uniform illumination is a problem seen in these images. This is corrected with a modified quotient based approach [33]. Here, the given degraded retinal image I, is modelled as a multiplicative degradation function L applied to an undegraded/original image I_o. L is assumed to be a slowly varying function and thus estimated from the smoothed version (I_s) of the degraded image as:

$$L(x,y) = \begin{cases} \frac{I_s(x,y)}{l_o} & \text{if } I_s(x,y) < lo \\ 1 & \text{if } I_s(x,y) \geq l_o \end{cases} \tag{3.1}$$

where l_o is the desired mean illumination level, chosen to be half the dynamic pixel range of I. Using the above estimate of L, the desired image I_o is found as:

$$I_o(x,y) = \begin{cases} I(x,y) \times \frac{l_o}{I_s(x,y)} & \text{if } I_s(x,y) < l_o \\ I(x,y) & \text{if } I_s(x,y) \geq l_o \end{cases} \tag{3.2}$$

Next, vessels are detected using the Bar-Combination Of Shifted Filter Responses [34]. This method uses a difference of Gaussians (DoG) filter for vessel detection and the choice of parameters of the Gaussian determines the thickness of extracted vessels. The DoG response was thresholded at pixel value of 30 to obtain a binary mask where all thick vessels are set to 0 while others are set to 1. Thus we obtain capillary network from from the angioflow images. Fig. 3.4 (a) and (b) show angioflow images from a sample OCTA and the derived capillary network.

The CD for a given angioflow image is found as

$$CD(m) = \frac{N(m)}{A(m)} \tag{3.3}$$

Here, N is the number of capillary pixels and A is the area of the mth sector with $m = 1, 2, ...8$. CD is computed for each of the 4 angioflow images (per eye), to derive a set of 8-dimensional (8-D) vectors representing the CD at 4 different layers. The capillary loss is inversely proportional to the calculated CD.

(a) Angioflow images.   (b) Detected capillary network.   (c) Capillaries in the ROI.
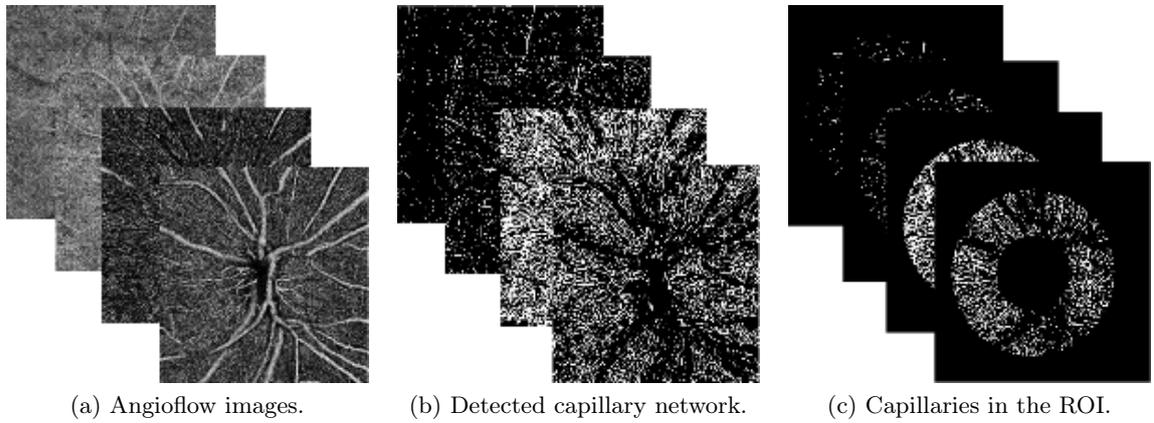
Figure 3.4: Estimation of Capillary Density.

3.2.2.2 RNFL thickness estimation

The RNFL thickness is computed from the slices of the OCTA volume of size $M \times N \times K$. The size of the ONH varies across subjects and as a result the sector area in the ROI is also varies. Hence, we chose to represent the RNFL thickness in a sector by the mean value and not a vector. We propose a computationally efficient method for RNFL layer segmentation. This is based on the analysis of the intensity profiles along the columns of an OCTA slice. Given a $M \times N$ slice, each column is analysed as a M-D intensity profile Fig. 3.5 shows a sample column (vertical line in green) and the corresponding intensity profile. The first major sharp rise in the intensity profile is due to dark to bright transition corresponding to the beginning of the RNFL layer. After this boundary, a large dip in intensity marks the end of RNFL layer. The width of the first major peak in the profile corresponds to the RNFL thickness. We hence locate the first local maxima and fit a second order polynomial to the profile around this peak. The full width at half the maximum value of the first peak is taken as the RNFL thickness.
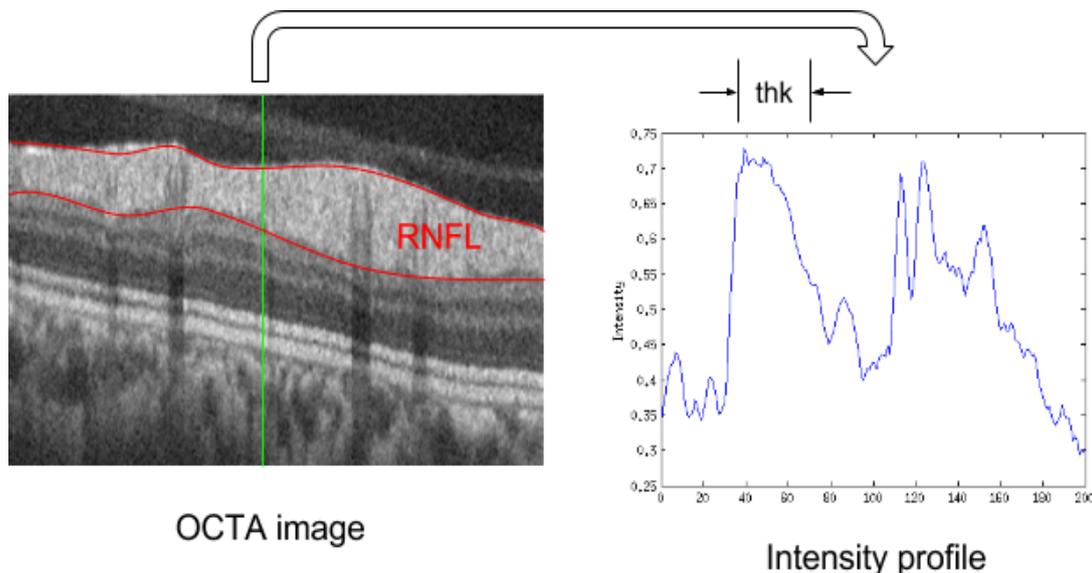
Figure 3.5: Estimation of RNFL thickness from intensity profile.

3.2.3 Classifier

The input to the classifier is a combination of the CD and RNFL thickness features. The CD values for 8 sectors across 4 angioflow images give rise to a 4×8 CD matrix. The RNFL thickness is a 1×8 vector which is added to the CD matrix to form the final 5×8 matrix. This matrix is converted to a 1×40 feature vector which forms the input to a linear SVM.

3.3 EXPERIMENTS AND RESULTS

Dataset: OCTA images from an Optovue scanner (spectral domain RTVue - XR 100 OCT (Avanti edition, Optovue, Inc, CA)) were collected from the Anand Eye Institute hospital, Hyderabad. The dataset consists of images of 67 eyes (49 normal and 18 glaucomatous). The data for each eye consisted of a OCTA volume and angioflow images of the choroidal disc, nerve head, RPC and the vitreous layers. The OCTA volume was of size $640 \times 304 \times 304$) while the angioflow images were of size (304×304). Ground truth for each eye was obtained based on clinical examination by one expert (co-author).

A six-fold validation was done to assess the performance of the system. For each fold of the training set, features were taken from 15 glaucomatous and 40 normal class images. Post training, test samples were given to the learnt model to predict the class. Quantitative evaluation of the method is done by computing sensitivity, specificity and accuracy across the folds. The sensitivity, specificity and accuracy metrics were found using the formulae below. Here, T/FP

Table 3.1: Performance of the proposed system with only RNFL thickness or CD features.

	RNFL alone		CD alone	
	Mean	Std.	Mean	Std.
Sensitivity	0.4444	0.1721	0.8333	0.1826
Specificity	0.8773	0.1370	0.7986	0.1429
Accuracy	0.7601	0.1258	0.8081	0.0971

Table 3.2: Performance of the proposed system with RNFL and CD features from layers. Best results are in bold font.

	RNFL + Choroidal disc		RNFL + NH		RNFL + RPC		RNFL + Vitreous		RNFL + All layers	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Sensitivity	0.7222	0.1361	0.6111	0.2509	0.7778	0.1721	**0.7778**	0.2722	**0.9444**	0.1361
Specificity	0.6944	0.1701	0.7755	0.1658	0.8218	0.1891	**0.9005**	0.1167	**0.9167**	0.1021
Accuracy	0.7008	0.1257	0.7298	0.1429	0.8081	0.0971	**0.8662**	0.0745	**0.9242**	0.0859

denote True/False Positive, T/FN denote True/False Negative.

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

$$Specificity = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The RNFL thickness and CD values across the 8 sectors are presented for one Glaucomatous and one normal eye image in Fig 3.6. Both plots exhibit a common trend in that the values for the glaucomatous eye is lower than that for the normal eye. Based on this trend, we expect the two features to be effective in discriminating between the two classes.

The effect of using either of these features exclusively, on classification was studied. Results are tabulated in Table 3.1. Classification with just RNFL thickness feature is seen to result in low sensitivity and accuracy values, even though the specificity is better compared to that obtained with exclusively CD features computed from all 4 angioflow images. Thus, it appears that the vascular health information is generally more effective in attaining good glaucoma classification.

Next, we report on an experiment in which both CD and RNFL features were used for classification. Results are shown in Table 3.2. The relative importance of each angioflow image can be gauged based on the results reported for CD from each angioflow individually. The top two results are seen to be obtained for the vitreous and RPC angioflow images. Thus, vascular
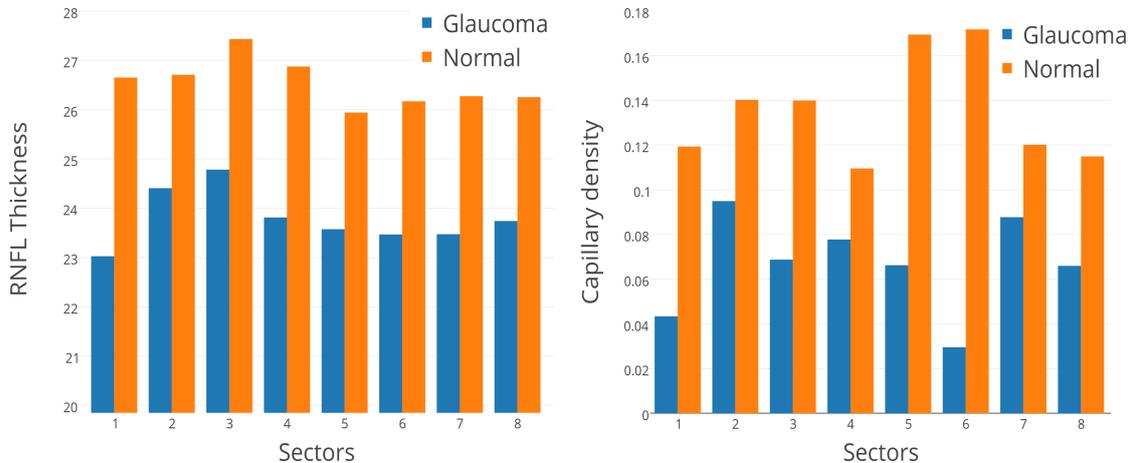
Figure 3.6: RNFL thickness (left) and Capillary density (right) variation for a Glaucomatous and a Normal case across sectors. The latter is averaged over 4 angioflow images(layers).

health information at these top two retinal layers appear to be relatively more important for correct glaucoma assessment. The last column in the Table 3.2 is the result for the full set of features: CD features from all layers and RNFL thickness. The performance is best for this case with all the metrics being above 90%.

The RNFL estimation in our proposed method uses a simple technique. In order to determine if a finer segmentation method can improve the classification performance, we experimented with a graph-based layer segmentation algorithm [30]. However, this yielded only a marginal improvement. This is because the thickness of the RNFL layer in a sector is represented only by its mean value and thus the proposed method is well equipped for the classification task. Since OCTA is a new modality there is no reported work on glaucoma classification from Angio-OCT images and hence bench marking cannot be done.

3.4 CONCLUSIONS

Inspired by the clinical significance of capillary loss and nerve fiber thickness, a novel method for automated glaucoma detection study is proposed using images from a new modality. The method relies on a RNFL thickness and capillary density from OCTA data. While the capillary loss alone appears to be relatively better for glaucoma classification, compared to RNFL thickness, it was also found that the best performance with a comprehensive set of features representing the vascular health (CD) and structural (RNFL) information. Since this is a new imaging modality the proposed work represents the first attempt at its utility for glaucoma assessment application. The obtained results are promising and encourage considering further evaluation on a larger dataset.

Chapter 4

Slice level disease diagnosis

Visualization of retinal internal structures is possible with the availability of OCT imaging modality. In screening scenarios, there is a need to develop automated algorithms to assist medical experts in disease diagnosis. A diagnosis tool for screening of macular disease like Age related macular degeneration (AMD) and Cystoid Macular Edema (CME) is important to public health as it can speed up the diagnosis procedure.

4.1 INTRODUCTION

AMD and CME are the leading causes of blindness [35] [36] in elderly and working-age people across the world. AMD is marked by accumulations of extracellular material (drusen) between Bruch's membrane(BM) and the retinal pigment epithelium (RPE). This is seen as aberration in the BM-RPE layer. CME is marked by leakage from the perifoveal retinal capillaries and accumulation of fluid within the intra cellular spaces of the retina. These are not restricted to any particular retinal layer.

(a) A sample slice of a normal patient with no abnormality.

(b) A sample slice of an AMD patient. A red rectangle is drawn to indicate drusen.

(c) A sample slice of a DME patient. A Yellow rectangle is drawn to cyst.
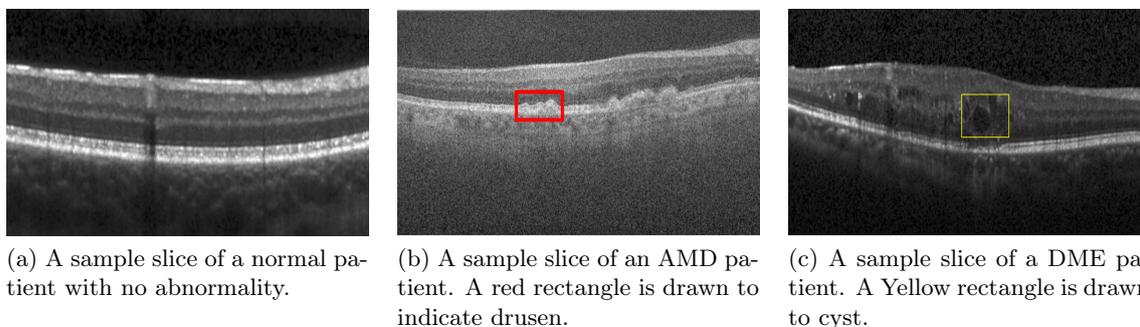
Figure 4.1: A sample slice depicting different retinal diseases

For an AMD patient, detecting drusen (aberration in bright layer) is of interest and likewise retinal cyst (dark regions) need to be detected to DME. These bright and dark abnormalities can be observed in the Fig. 4.1 (b) and Fig.4.1 (c). Both these are major indicators of the diseases AMD and CME in a screening scenario.

4.2 RELATED WORK

Segmentation of retinal layers and assessing the layer thickness have been studied for disease diagnosis. Statistical analysis of layer thickness, with respect to the normal cases, was used to identify the retinal diseases [37] [38]. Segmenting of the abnormalities such as drusen and Cystoid structures for disease classification has also been attempted. Morphological and textural features from the choroidal region were classified with a decision tree in [39] to differentiate between AMD, DME and fibrovascular tissue. A kernel principal component analysis model [40] was proposed into differentiate between AMD and normal eyes using a Bayesian network classifier. Multiscale Local Binary Pattern (LBP) features were used to perform multi class classification of retinal OCT images in [41] for the detection of macular pathologies. The reported performance is excellent but a manual selection of a single high resolution OCT image is required to detect the disease. A recent fully automated method [42] detects retinal diseases using Histogram of Oriented Gradients (HOG) descriptors and SVMs to classify each OCT volume as normal, containing dry AMD, or containing DME. These methods [39] [42] have denoising and image flattening or alignment module as a pre-processing step that are highly data dependent.

In this chapter, we present a fully automated method that detects the presence of retinal pathologies (AMD, DME) in a slice of an OCT volume using a Convolution Neural Network (CNN). The slice-based analysis aids in localization of abnormality in a volumetric data. The input to the CNN is a unique representation of the OCT data, based on motion patterns. This classification approach is presented next in detail.

4.3 PROPOSED METHOD

The proposed work is a cascade of three stages: Region of interest extraction followed by generating extremal representation from motion pattern and finally classification with a CNN. Section 4.3.1 explains how the search region is reduced. The detail about how extremal pattern are obtained is explained in sub-section 4.3.2. The final stage explained in sub-section 4.3.3 describes the proposed architecture of the convolution neural network. The CNN here is used to extract distinct features that are useful for classification of retinal diseases.

4.3.1 Region of interest extraction

The variability observed in the scanning protocol, resolution and layer orientation is handled by resizing all given slices $f(x, y)$ to 512×256 pixels. A rough region of interest (ROI) is derived to reduce the search region. A slice image is projected along columns to obtain a 1D profile and a Gaussian is fit to this profile. The mean value of the Gaussian profile (with row index x_0) is used to extract an ROI of size 250×256, centered at x_0. This ROI is used for next module.

4.3.2 Generating extremal representation

When capturing a high speed object in a natural scene, depending on the camera setting, a streak effect or motion blur is visible in the image. This visual phenomenon is synthesized artificially by inducing translation motion to an image. This representation space of an image has been used previously in discriminating between *images* of normal and abnormal cases [43] [44]. The classification of multiple diseases based on the abnormalities using a neural network has not been addressed before.

An image can be represented in extremal representation space by inducing motion (such as translation, rotation) to that image. Let an image be denoted by f. Applying motion (transformation T) to f results in a stack S of images

$$S = \{f_j\}_{j=-N}^{N} \tag{4.1}$$

where,

$$f_j = f(jT(\delta)); j \in [-N, N] \tag{4.2}$$

where δ denotes the transformation parameter. In our work, T was chosen to be a translation which is applied in the direction θ, in steps of δ resulting in an image stack of size $2N + 1$. This stack of $2N + 1$ transformed images are combined using 'extrema' (max) and (min) as a coalescing function $\phi(.)$ which is given by,

$$F_\theta(f) = \phi\big\{(f(jT(\delta)))\big\}_{j=-N}^{N} \tag{4.3}$$

Figure 4.2 shows the image in (max) and (min) representation space in vertical $(\theta = 90°)$ direction. Two extremal representations at K different θ's are stacked together and fed as the input to the CNN as described in the section 4.3.3. The stack extremal representation is given by,

$$EX_\phi = \big\{F_\theta(f)\big\}_{\theta=1}^{K} \tag{4.4}$$

(a) The original image

(b) The original image in extremal space after translation motion at $90°$, $N = 15$, $\delta = 5$ and $\phi = max$

(c) The original image in extremal space after translation motion at $90°$, $N = 15$, $\delta = 5$ and $\phi = min$
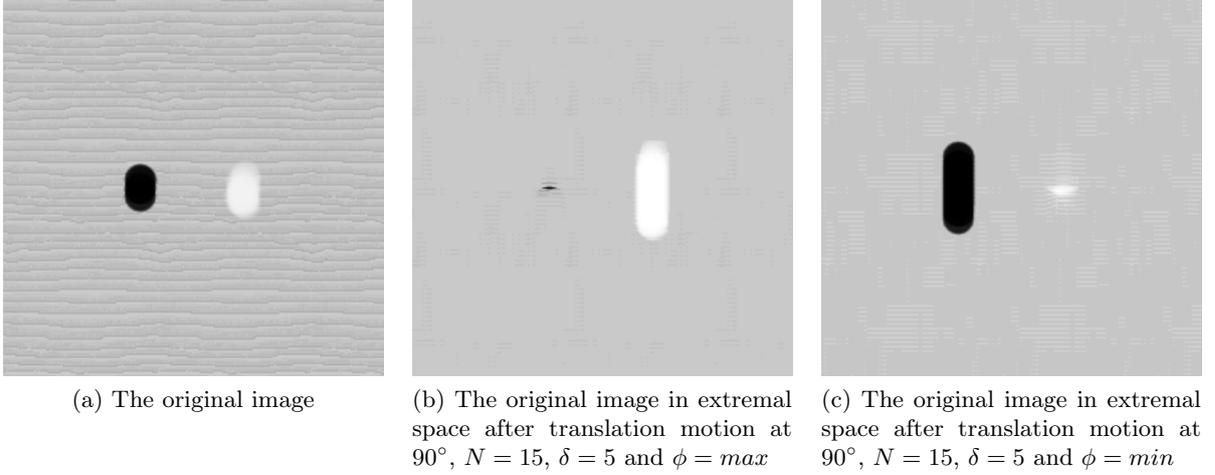
Figure 4.2: An illustration of extremal representation space obtained using translation motion in vertical direction.

4.3.3 Convolution neural network classifier

CNNs are biologically-inspired variants of feed forward multi layer perceptron. A CNN architecture is formed by a stack of layers that transform the input to an output. The building blocks in our CNN are: (i) convolutional layer with filters that learn specific type of feature in the input at some spatial position (ii) maxpooling which downsamples the image by retaining the maximum in a local neighborhood (iii) Fully connected layer have all of its neurons fully connected to all activations in the previous layer, similar to regular Neural Networks. (iv) softmax function that gives probabilistic value of the input belonging to one class.

The CNN architecture shown in Fig. 4.3 is designed to utilize contextual information. The two parallel pipeline in the architecture are fed with separate extremes stack EX_{max} and EX_{min} respectively. In the first pipeline, convolutional layer assesses the input at different field of views and collapses the EX_{max} into a 1-D vector. Similarly, the second pipeline produces a 1-D vector collapsing the EX_{min}. The presence of drusen will produce a high response in the first pipeline and the presence of cyst will produce a response in the second pipeline. Hence, for a normal macula slice, there is an absence of response from either of the pipeline. The two 1-D vectors are merged and fully connected before the softmax activation function. A detailed description for a single pipeline of the architecture is given in the Table 4.1. The same is fixed for the second pipeline.

The CNN architecture combines the extremes representation stack to produce a probabilistic value predicting the disease associated with the slice. The main aim during CNN training, is solving the unconstrained optimization problem by minimizing a loss function that compares the ground truth with the system output. The categorical cross entropy loss function is used
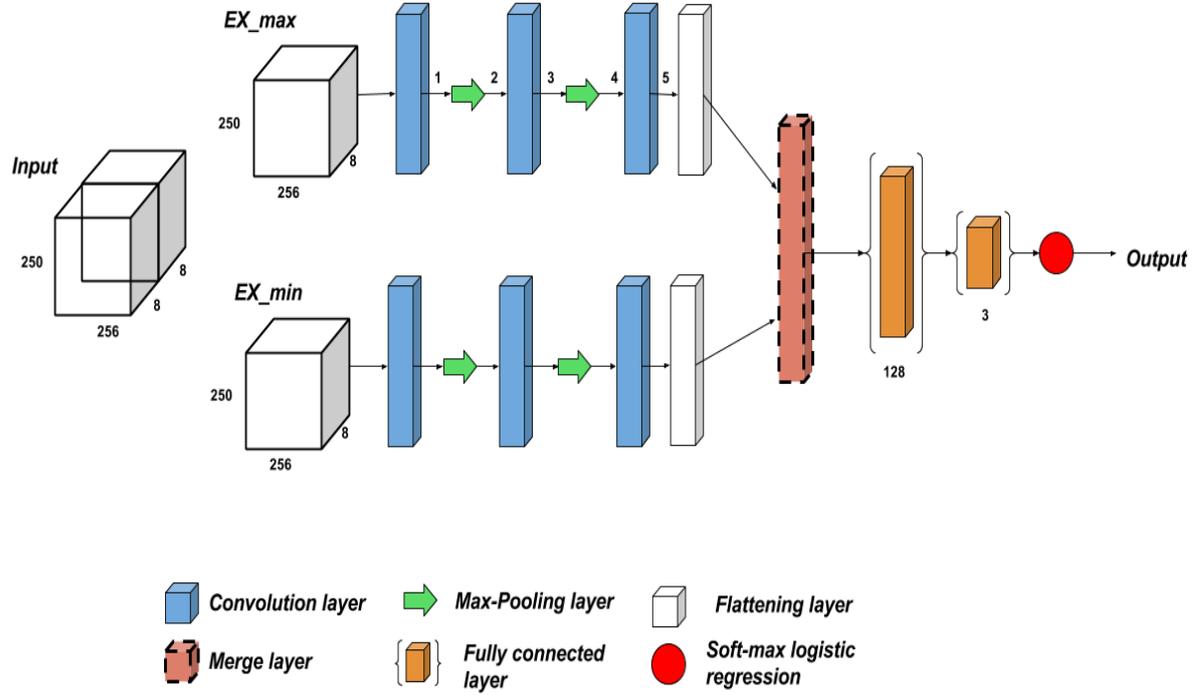
Figure 4.3: The proposed CNN architecture.

Table 4.1: Description of the CNN architecture

Input	C_max	Output1	Output2	Output3	Output4
Stage	Convolution	Max pooling	Convolution	Max pooling	Convolution
# channels	8	8	8	8	8
Filtersize	10x10	-	25x25	-	10x10
Pooling	size	-	2x2	-	2x2
Spatial Input size	250x256x8	250x256x8	125x128x8	125x128x8	62x64x8
Spatial Output size	250x256x8	125x128x8	125x128x8	62x64x8	62x64
Output	Output1	Output2	Output3	Output4	Output5

for a three class classification task. The CNN network was trained for 50 epochs. Training loss as a function of training epochs is shown in the Fig. 4.4.
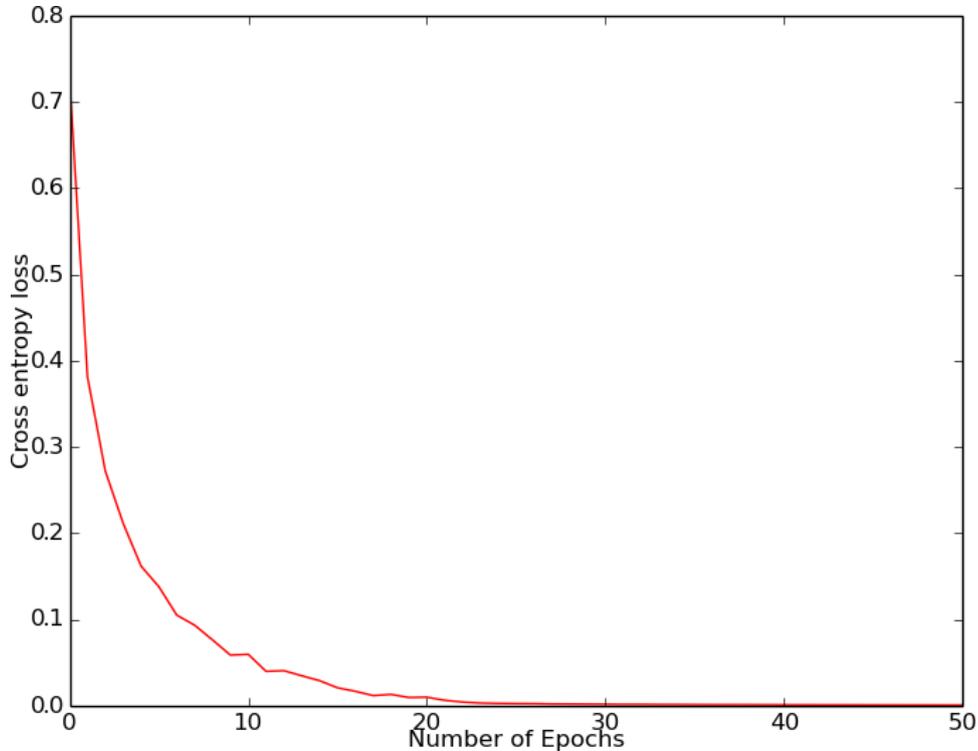
Figure 4.4: Loss function across epochs.

4.4 EXPERIMENTS AND RESULTS

The proposed pipeline was implemented in Theano using Keras library on a NVIDIA GTX Titan X GPU, with 12GB of GPU RAM on a core i7 processor.

Dataset : Four public datasets available online were combined and annotated to create the training and testing split. The OPTIMA cyst segmentation dataset [45] was released in 2015 as a challenge dataset in MICCAI conference. This dataset has SD-OCT volumes captured from four different vendors(Cirrus, Nidek,Spectralis,Topcon). The data consisted of slice level manual cyst markings. The slice that contained cysts were labeled to the CME class for training and testing. The second dataset used for layer segmentation in presence of CME [46] contains manual markings for both layer boundaries and cysts or fluid filled regions. These slices with the cyst markings belong to CME class. The third dataset consists of volumetric scans for 45 patients: 15 normal, 15 AMD, and 15 CME patients [42]. In a volume, not all slices contain abnormality. Hence, a subset of this dataset is annotated to get the training and testing data for each class. The last dataset is the world's largest dataset [47] containing 269 AMD patients

and 115 normal patients volumes. A subset of slices from these volumes were asked to label as Normal/AMD to prove the working of the proposed method. The exact split of the training and testing for each dataset and class is shown in Table 4.2.

Table 4.2: Description of the dataset

Dataset	Train (No. Of slices)			Test(No. Of slices)		
	Normal	DME	AMD	Normal	DME	AMD
[45]	-	600	-	-	175	-
[46]	-	49	-	-	42	-
[42]	623	351	351	377	275	203
[47]	877	-	659	530	-	393
Total	1500	1000	1000	907	492	596

Parameter setting for extremal representation: Extremal representation computation requires choosing a few parameters such as the step size δ, range or extent of translation N and the number of translation directions or number of K in the extremal representation stack EX_ϕ. N and K directly affect the information content and size of cake EX_{max} and EX_{min}. An image in representation space was created by applying translation motion on a ROI slice with $N = 5$, $\delta = 1$ and $\theta = 0°$ to $180°$ (in steps of $22.5°$) over a stack of $K = 8$; EX_{min} was derived by using a min coalescing function and similarly EX_{max} with max coalescing function.

Results : Quantitative evaluation of the method is done by computing sensitivity, specificity and accuracy. The sensitivity, specificity and accuracy metrics were found using the formulae below. Here, T/FP denote True/False Positive, T/FN denote True/False Negative.

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.6}$$

$$Accuracy = \frac{TN + TN}{TP + TN + FP + FN} \tag{4.7}$$

The output of the CNN gives a probability score belonging to one of the class. To compute the sensitivity and specificity for an individual class, we compare it with rest of the classes combined. The ROC plot for individual class is shown in Fig. 4.5 . The Table 4.3 shows the quantitative results for individual classes.

In the dataset [45] and [46], some slices which contain region level cyst markings also has drusen abnormality. Depending on the dominant abnormality in the slice, the proposed system tends to favor one of the class. Due to this, sensitivity of the DME class is reduced a bit. This is evident from the quantitative results listed in the table 4.3.
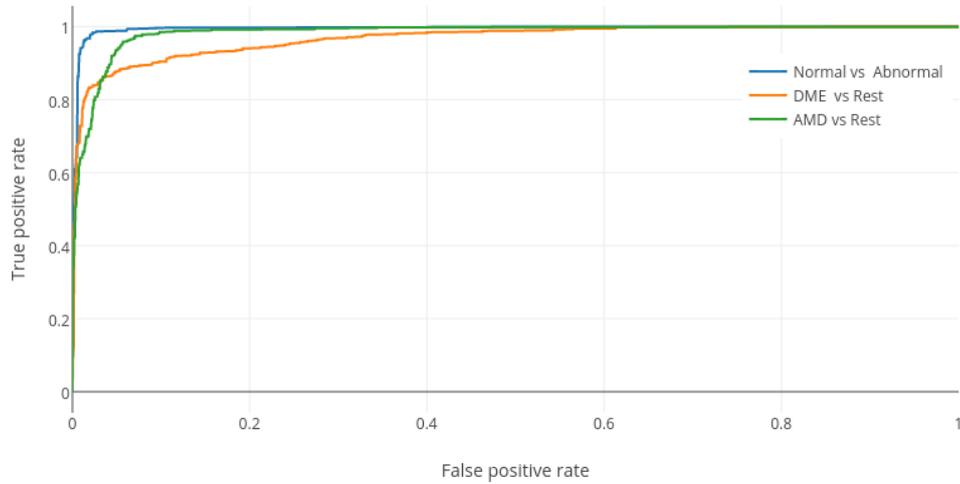
Figure 4.5: ROC plot for individual disease class.

Table 4.3: Performance of the proposed system.

	Sensitivity	Specificity	Accuracy
Normal vs Abnormal	0.98	0.98	0.98
DME vs Rest	0.86	0.98	0.95
AMD vs Rest	0.96	0.96	0.95
Mean Result	0.93	0.97	0.96

4.5 CONCLUSIONS

Moticated by the clinical significance of disease classification and localization of the abnormalities in a SD-OCT volume, a novel method is proposed for automated detection of diseases using slices of an OCT volume. The method relies on a unique way of representing an image/slice using motion pattern. A unified approach for multiple disease classification is learned by a CNN with extrema representation as its input. The performance of the proposed system on multiple dataset shows the robustness of the method. For this work, the method assumes only one type of abnormality present. In future, the model can be altered to learn multiple abnormalities in a single slice. The obtained results are promising and encourages us in developing a system for quantifying the abnormalities by accurately localizing them.

Chapter 5

Pixel level abnormality segmentation

Optical Coherence Tomography (OCT) is an imaging modality useful for assessing pathological changes. Multiple vendors with varying resolution and scanning protocol provide OCT data. Inter scanner variability is reflected in terms of intensity variation and SNR. Retinal experts analyze the abnormality to plan a treatment for the subject. Abnormalities can occur at different regions around macula. Localizing and quantifying them will aid in accurate surgical interventions. This calls for the need for an automated abnormality localization/segmentation algorithm.

5.1 INTRODUCTION AND RELATED WORK

Optical Coherence Tomography (OCT) is used by human experts to assess morphological changes in sub-retinal layers. Low image quality and high acquisition time hindered the use of OCT in the early days. The newer Spectral-Domain OCT (SD-OCT) overcomes these problems.

Cystoid Macular Edema (CME) is a condition in which fluid filled regions (cysts) are formed in the macula leading to swelling [48] [49]. If CME is left untreated for more than 6 to 9 months it can lead to chronic macular changes with permanent impairment of central vision. Therefore, the automated detection of CME from SD-OCT images is of interest. Automated detection however has a few challenges: a)Inter-scanner and intra-scanner variability in images (voxel intensities), b) noisy images, c) ambiguous demarcation between sub-retinal layers and finally d) variable shape, size, appearance and arbitrary locations of cysts in the images.

Early work on SD-OCT image analysis focused on the intra-retinal tissue layer segmentation [50] [17] [13] whereas attention to abnormality (such as cyst) detection is very recent, with cyst segmentation being posed as a challenge problem in MICCAI 2015.

An early approach proposed to detect cysts was semi-automatic, based on GVF-snake to delineate the intra-retinal and sub-retinal fluid regions in 2D OCT B-scans [51]. Later methods were fully automatic and based on 2D analysis. A denoising step was typically followed by a variety of approaches including thresholding and boundary tracing [52] and texture-based classification of regions of interest derived with a watershed algorithm [53]. 3D methods have been reported more recently. A graph-search/graph-cut based approach that simultaneously segments upper retinal surface, lower retinal surface and fluid-filled regions in [54] reports good segmentation only for large cysts. Since smaller fluid regions are also of interest, voxel classification based approaches have been attempted in [55] [56][57]. Among the recent approaches are layer-dependent stratified sampling strategy to address the class-imbalance (far fewer cyst vs healthy tissue voxels) [55], geodesic graph cut method [56] and a kernel regression-based classification [57]. Some of these can segment both the layers and fluid-filled abnormalities while others focus only on the abnormalities.

The MICCAI 2015 OPTIMA cyst segmentation challenge was attempted by participants with a variety of approaches. These range from machine learning with handcrafted features such as intensity and layer thickness, CNN-based voxel labeling [58] and graph-based method with layer-dependent information [59]. The latter method has also been evaluated on the public access DME dataset [60]. The preceding methods are vendor independent. Vendor-specific methods include 3D curvelet transform based dictionary learning [61] and a marker controlled topographical watershed algorithm [62]. Our simple approach [63] to extract candidates with centre-surround filters and classify with an ensemble random forest failed to perform well on the test set. The proceedings of the challenge can be found at Optima website [64]. In this chapter, we propose an entirely different (from existing methods, including ours) approach based on biologically inspired motion patterns and deep learning.

To summarise the current scenario, graph based segmentation such as [54] is inadequate to segment smaller fluid filled regions which is addressed by existing voxel classification based methods which rely on accurate layer segmentation to derive features. However, segmentation of layers in the presence of abnormalities, is still an open problem. The CNN based method [58] employs three separate networks at different scales to handle the size variability of cysts, which is a major drawback. The parametric approaches such as [61] and [62] are highly vendor dependent. An alternate approach could be to directly detect the cystoid regions and then segment them. We take such an approach and propose a novel, biologically inspired pipeline for segmentation of cysts. The novel aspects of the proposed method are as follows. (i) Unique representation for the OCT data, based on motion patterns proposed earlier by our group [43][44]. (ii) Selective enhancement of the objects of interest (cysts) with a Convolutional Neural Network (CNN). (iii) Combining both 2D and 3D information in CNN for detecting the cyst regions. and (iv) Vendor independent and robust system for detecting and localizing cysts. We show that this

novel strategy leads to superior segmentation performance even with a simple, clustering-based post processing step. The next section provides details of the proposed system.

5.2 PROPOSED METHOD

The proposed pipeline consists of three main stages, namely, preprocessing, detection and clustering. The preprocessing stage is described in sub-section 5.2.1; the detection of cyst regions via a selective enhancement operation is explained in sub-section 5.2.2 and segmentation via clustering is described in sub-section 5.2.3. A graphical representation of the work is shown in Fig. 5.1.
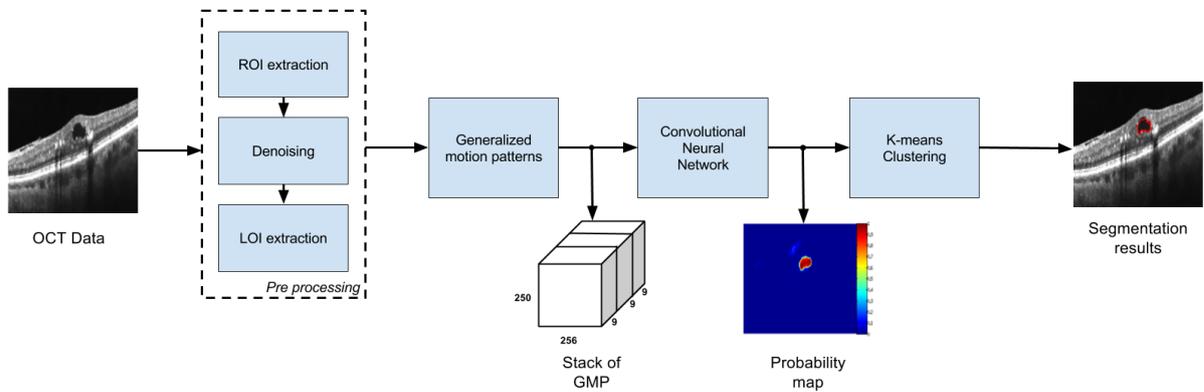
Figure 5.1: Graphical representation of the proposed pipeline.

5.2.1 Preprocessing

5.2.1.1 Region of interest extraction

SD-OCT volumes are generally acquired with different operator-defined protocols. Consequently the layer orientation, resolution and number of slices in the acquired volumes vary widely. In order to address this variability, the size of each slice $f(x, y)$ was standardized to a $512{\times}256$ pixels by resizing. Since, the region of interest is only 40% of each slice, for every given volume, a rough region of interest (ROI) was derived as follows: each slice image was projected along columns to obtain a 1D profile and a Gaussian was fit to this profile. The mean value of the Gaussian profile corresponds to a row index x_0 in the slice. An ROI of size $250{\times}256$, centered at x_0 was finally extracted and used for further processing.

5.2.1.2 De-noising

The SD-OCT volumes are degraded by speckle noise. Speckle noise is signal dependent and hence depends on the structure of the tissue in an OCT volume. A total variational denoising method [65] was used to denoise all ROIs. This reduces the texture content and results in a smoother image. A sample ROI extracted from an original slice and the corresponding denoised result are shown in Fig 5.2(a) and Fig 5.2(b).

5.2.1.3 Extraction of layers of interest

Cysts are known to be restricted to the layers between Internal Limiting Membrane (ILM) and Retinal Pigment Epithelium (RPE). Various segmentation algorithms have been proposed for segmenting retinal layers. Some of these include pixel intensity variation-based ILM and RPE segmentation [15], active contour with a two-step kernel-based optimization scheme [17], complex diffusion filtering with combined structure tensor replacing thresholding [50]. In this work, we used an accurate, graph theory based segmentation approach [13] for extracting ILM and RPE which forms the Layers of Interest (LOI). A mask for this region i.e, ILM and RPE layers was extracted from the ROI (see Fig. 5.2(c)). LOI serves as a position prior in the detection stage as explained below.

(a) A sample ROI.   (b) A ROI after denoising   (c) Extracted mask representing layers of interest between ILM and RPE.
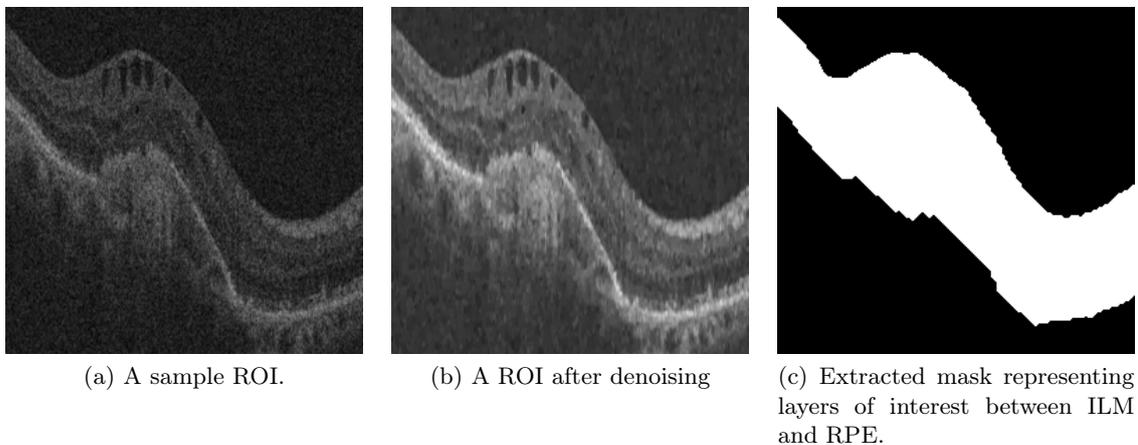
Figure 5.2: Denoising and extraction of layers of interest on a sample ROI.

5.2.2 Detection of cyst regions

Localization of the objects of interest is a vital task in a segmentation problem. We propose to address this task via a selective enhancement process wherein cysts are enhanced without the knowledge of the layer positions. Selective enhancement is done by employing the notion of

Generalized Motion Pattern (GMP). The GMP was introduced in our earlier work [43] and has been shown to be effective in diabetic macular edema detection [44]. In this work, we construct an ensemble of these GMPs and learn a function (using a CNN) to combine them such that only cysts are enhanced. A detailed description is provided next.

(a) A phantom image

(b) Translation GMP at $0°$, $N = 15$, $\delta = 5$

(c) Translation GMP at $45°$, $N = 15$, $\delta = 5$

(d) The sum of $K = 8$ translation GMPs (I_0)
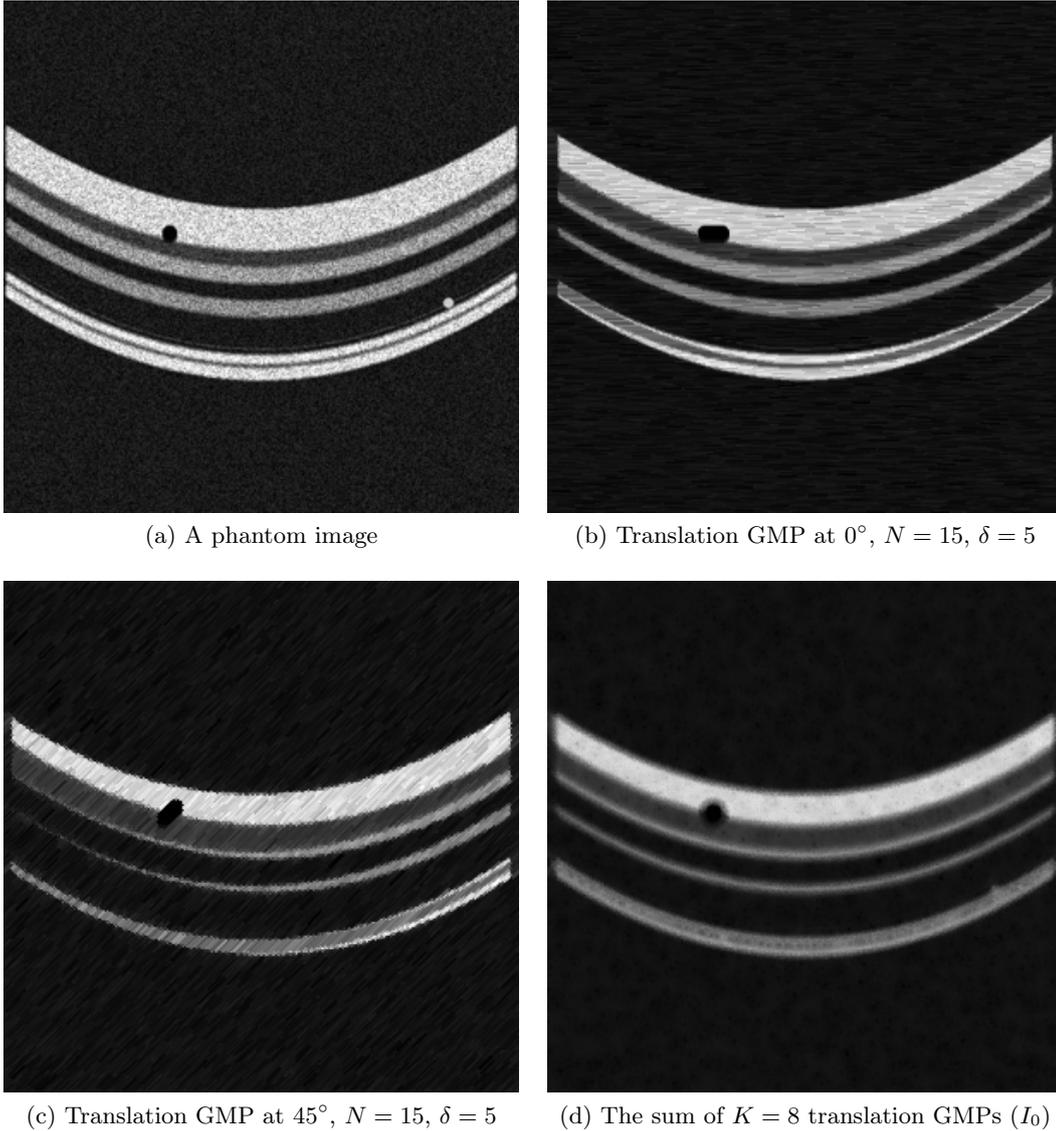
Figure 5.3: Sample generalized motion patterns with a *min* coalescing function and their sum for a phantom image.

5.2.2.1 Generalized motion patterns

When one captures an image of a racing car, depending on the camera setting, a streak effect or motion blur is observed due to the fast moving object. A related phenomenon is seen when a disc with multi color segments is rotated at an appropriate rate: the colors smear and leave an (visual) impression of a white disc. GMP is the synthetic (discrete) equivalent of this phenomenon. It is derived by inducing motion (such as translation, rotation) to an image and can be used to represent an image. Let an image be denoted by f. Applying motion (transformation T) to f results in a stack S of images

$$S = \{f_j\}_{j=-N}^{N} \tag{5.1}$$

where,

$$f_j = f(jT(\delta)); j \in [-N, N] \tag{5.2}$$

where δ denotes the transformation parameter. In our work, T was chosen to be a translation which is applied in the direction θ, in steps of δ resulting in an image stack of size $2N + 1$. Applying a coalescing function $\phi(.)$ to the stack yields a GMP.

$$G_\theta(f) = \phi(S) = \phi\{(f_j)\}_{j=-N}^{N} \tag{5.3}$$

The direction of the translation θ and the choice of coalescing function is dependent on the problem being solved. In the problem at hand, since the object of interest is a cyst which is a dark structure, a minimum (min) operation is an appropriate choice for ϕ.

In our earlier work, GMPs derived using translation and rotation were shown to be useful in discriminating between *images* of normal and abnormal cases [43][44]. The segmentation of abnormalities was not addressed. We propose to solve the cyst segmentation problem by deriving GMPs in multiple directions θ and combining them such that cysts are selectively enhanced. In the problem at hand, since the object of interest is a cyst which is a dark structure, to get a GMP image, a minimum (min) operation is an appropriate choice for ϕ.

A cake of GMPs (C) is constructed by stacking multi-direction GMPs. C is defined as

$$C = \{G_\theta(f)\}_{\theta=1}^{K} = \left\{\phi\{(f_j)\}_{j=-N}^{N}\right\}_{\theta=1}^{K} \tag{5.4}$$

We propose to map the stack to a single image I_o in which the cysts alone are enhanced. Let ψ denote this mapping.

$$I_o = \psi(C) \tag{5.5}$$

The above idea is illustrated with a phantom image designed to model an OCT slice in Fig. 5.3(a). It has a dark blob in the first layer and a bright blob in the last layer representing a cyst and a drusen, respectively. Fig. 5.3(b) and Fig. 5.3(c) show the GMP derived by applying translation motion on the phantom image with $\delta = 10$ in $\theta = 0°$ and $\theta = 45°$ directions

respectively. The GMP was derived by applying a *min* coalescing function on 7 translated images. It can be observed that the dark blob gets smeared or extended in the layer and the relatively dark layers are wider in the direction θ; whereas the bright blob disappears. The final map I_o is shown in Fig.5.3(d) with ψ chosen as a simple sum. Here, both morphological changes as well as blur is evident. With a simple $\sum$ for ψ, no contextual or shape information is exploited. Hence, there is not much differentiation between a blob and layer in terms of the net effect.

This motivates us to *learn* a function that enhances only objects of interest by using relevant information while suppressing other tissues. This type of enhancement is useful when the aim is to localize cysts as it is now possible to interpret the pixel values in the enhanced image as a likelihood (probability) measure for the pixel belonging to a cyst. We propose to use a CNN to learn the function ψ such that cysts are enhanced while other normal anatomical structures are suppressed.
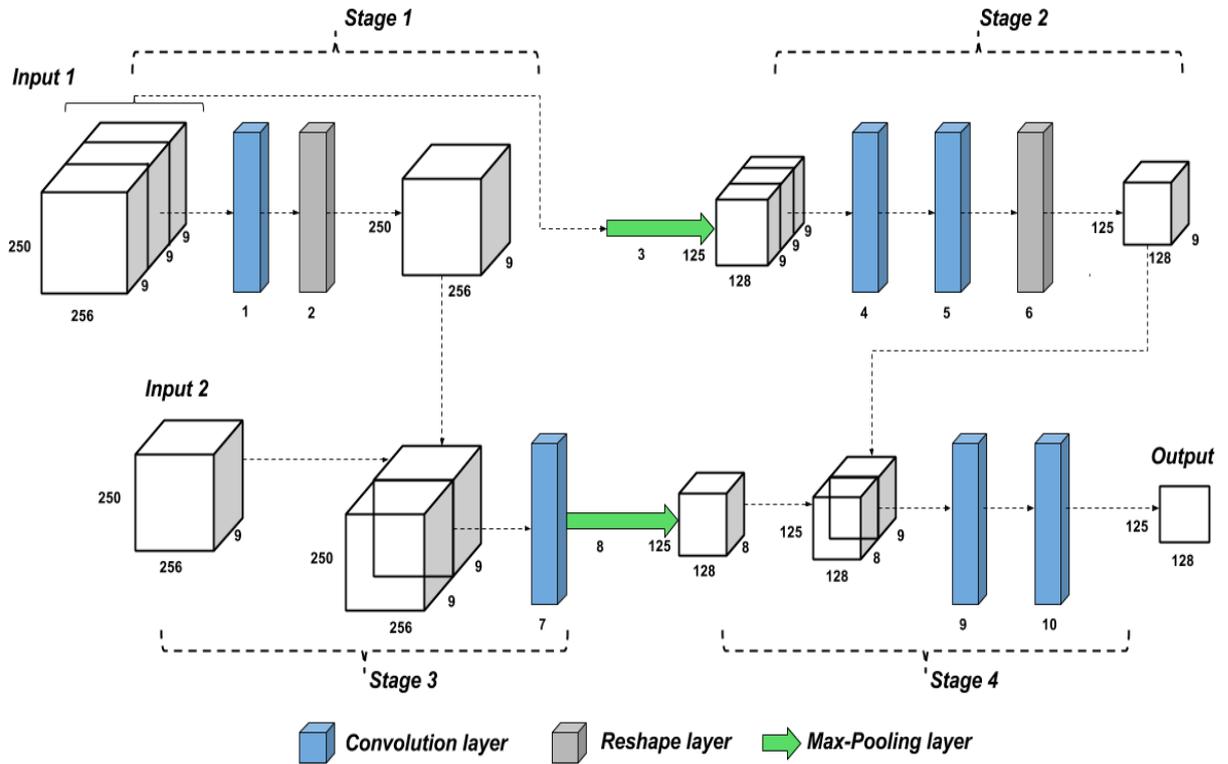

Figure 5.4: Proposed CNN architecture

5.2.2.2 Convolutional Neural Network

CNNs are biologically-inspired variants of feed forward multi layer perceptron. Since convolution filter weights are shared across all spatial positions, the free parameters being learned reduces

and thus reducing the memory requirements to run the network. A CNN architecture is formed by a stack of layers that transform the input to an output. The three building blocks in our CNN are: (i) convolutional layer with filters that learn specific type of feature in the input at some spatial position (ii) maxpooling which downsamples the image by retaining the maximum in a local neighborhood (iii) sigmoid activation function that non-linearly maps the intensity values of an image to an interval [0,1].

The task is to detect cysts from a ROI volume of size $X \times Y \times Z$ where $X = 250$ and $Y = 256$. We utilize the domain knowledge for our work i.e, cysts are 3D structures and thus use consecutive slices as 3D information. The LOI which is mask for the region between ILM and RPE is used as the position prior. In order to detect cysts from the Zth slice, 2 types of inputs (to the CNN) are derived from them. Input1 provides 3D information and hence is a concatenation of the cakes C (of $250 \times 256 \times K$; $K = 8$) for the Zth slice and its 2 neighbors, i.e., $(Z-1)^{th}$ and $(Z+1)^{th}$ slices along with their respective position prior. Thus the size of Input1 is $250 \times 256 \times 3 \times 8+1$) to CNN. Input2 provides just 2D information for the Zth slice and hence is composed of its Cake C along with its position prior. The size of Input2 is hence $250 \times 256 \times 8+1$. All images were normalized to have zero mean value and unit variance. A detailed description of the architecture is given in Table 5.1. Sigmoid activation function is applied on the learned feature map to obtain a probability map.

Table 5.1: Description of the CNN architecture

Input	Stage	# channels	Filter size	# filters	Pooling size	Spatial input size	Spatial output size	Output
Input 1	Convolution	27	1x1x3	9	-	250x256x3x9	250x256x1x9	Output 1
Output 1	Reshape	-	-	-	-	250x256x1x9	250x256x9	Output 2
Input 1	Max-pooling	27	-	-	2x2x1	250x256x3x9	125x128x3x9	Output 3
Output 3	Convolution	27	25x25x3	9	-	125x128x3x9	125x128x3x9	Output 4
Output 4	Convolution	9	1x1x3	9	-	125x128x3x9	125x128x1x9	Output 5
Output 5	Reshape	-	-	-	-	125x128x1x9	125x128x9	Output 6
Input 2;Output 2	Convolution	18	10x10	8	-	250x256x18	250x256x8	Output 7
Output 7	Max-pooling	27	-	-	2x2	250x256x8	125x128x8	Output 8
Output 8;Output 6	Convolution	17	25x25	8	-	125x128x17	125x128x8	Output 9
Output 9	Convolution	8	1x1	1	-	125x128x8	125x128	Probability map

The CNN architecture shown in 5.4 is designed to utilize neighborhood information. Stage 1 in the CNN architecture collapses the 3D information into a 2D map without any contextual information. In stage 2 at a lower scale, a large field of view is assessed to combine 3D information to a 2D map. The output of stage 1 is stacked with the second input(Input2) for Stage 3 which enhances the boundary information at the original scale. For the final stage, output from stage2 and stage 3 are concatenated as to form the input. A large neighborhood in 2D is considered for this stage to strengthen the object of interest and weaken the background information. The size of the receptive field is chosen such that the largest cyst in the dataset is enclosed in the field.

A weighted binary cross entropy loss function was chosen to handle the class imbalance in the data, i.e, cyst versus background pixel. This was minimized using gradient descent. The evolution of the probability maps learned (depicted as heat maps), as epochs increase is shown in Fig. 5.5. It can be seen that cysts gradually take shape in the map.

5.2.3 Segmentation

The output of the trained CNN model is a probability map with cyst pixels having higher probability score than background pixels. This is thresholded to obtain a binary map representing detected cyst regions. This map is multiplied with the ROI image to extract the detected regions in the intensity space. K-means clustering is applied on this product image. Detected regions are clustered into cysts, false positives region and background in the intensity space. Since cysts are relatively darker regions compared to the false positive regions, the clusters with lower mean intensity are retained as desired cyst segments.
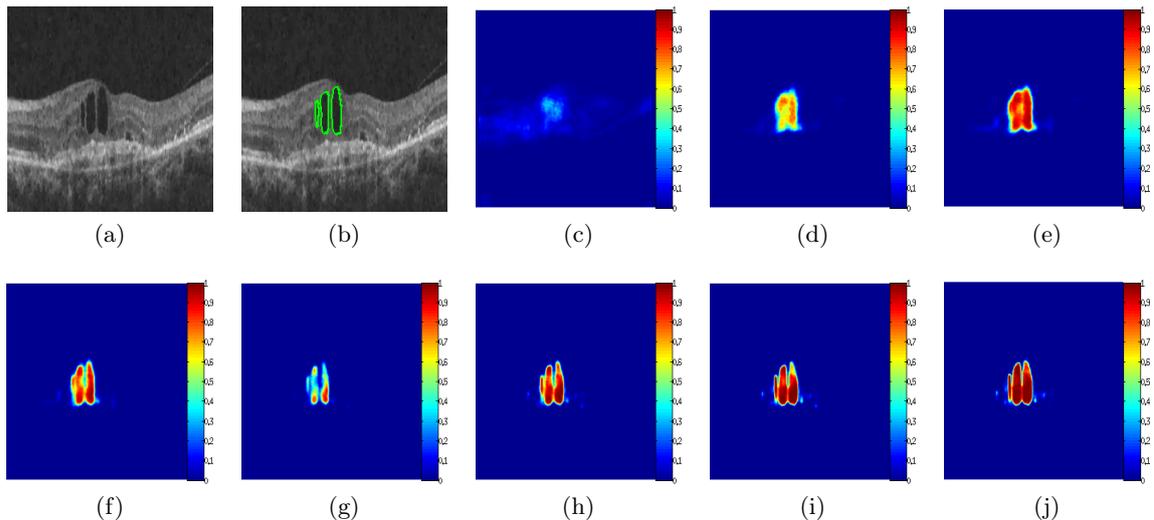
Figure 5.5: Evolution of the mapping function across epochs. Along with (a) ROI input, (b) Ground truth: Grader1 ∩ Grader2, the output of CNN is shown as a probability (heat) map at (c). 10th, (d) 50th, (e) 100th, (f) 200th, (g) 300th, (h) 500th, (i) 900th, (j) 1420th epochs

5.3 EXPERIMENTS AND RESULTS

The proposed pipeline was implemented on a NVIDIA GTX 960 GPU, with 4GB of GPU RAM on a core i3 processor. In order to assess the performance of our method, the experiments were carried out on 3 datasets: the MICCAI 2015's OPTIMA Cyst Segmentation Challenge (OCSC)

dataset [66], the publicly available DME dataset [60] and a locally sourced (from Anand Eye Institute, Hyderabad) dataset referred to as AEI.

The OCSC dataset has 15 training and 15 testing volumes acquired from 4 different scanners; included are manual markings of cysts from two graders. An overview of the dataset is shown in Table 5.2. The resolution and density of the volumes vary from 496×512 to 1024×512 and 5 to 200 B-scans, respectively. The DME dataset has 10 Spectralis scans with 61 B-scans and 496 resolution each. The manual tracings (from two experts) for these data are provided sparsely (11 B-scans per volume) and not for continuous slices. The AEI dataset has 10 volumes from Optovue scanner with varying number of (30-51) B-scans and fixed resolution of 640×304. Cyst boundaries for every fifth B-scan of each volume was collected from an expert.

Table 5.2: An overview of OCSC dataset

Scanner	Cirrus	Nidek	Spectralis	Topcon
Training	4	3	4	4
Testing	4	3	4	4

The proposed system was trained on the OCSC training dataset by taking the intersection of the markings of grader 1(G1) and grader 2(G2) as ground truth (GT). Only slices *with* cysts were used during training. A separate validation set was formed by randomly selecting a cyst containing slice from each volume in the OCSC train dataset. The CNN was implemented in Theano using the Keras library. Training for 1420 epochs took about a day and a half. A stochastic gradient descent optimizer as used to minimize the weighted binary cross entropy loss. The training parameters were: learning rate of 0.001; Nesterov momentum was set to 0.75 and batch size was chosen as 8.

5.3.1 Experiments on system design

Next, we describe a set of experiments that were done to study the impact on system performance. Specifically, these were to study the effect of i) varying the GMP parameters, ii) variants of the proposed system and iii) the choice of threshold value which is used to binarise I_o. Since OCSC is the largest dataset, all experiments were done by training on the OCSC training set and testing on the OCSC testing set. System performance was assessed by computing the Dice Coefficient (DC) defined below.

$$Dice\ Coefficient = 2 \frac{|Detected \cap GT|}{|Detected| + |GT|} \tag{5.6}$$

where |.| denotes the set size. DC$\in$[0,1] can be viewed as a similarity measure over the two sets. DC=1 is the ideal value as it indicates a perfect match between computed result and GT.

(a) ROI of a slice

(b) Ground truth Grader1 ∩ Grader2

(c) Probability map

(d) Segmentation results by thresholding probability map

(e) Segmentation results using K-means clustering

Figure 5.6: Qualitative results for a slice with cysts. Top to bottom rows: Results for a slice from Cirrus, Nidek, Spectralis, and Topcon scanners.

(a) ROI of a slice (b) Ground truth Grader1 ∩ Grader2 (c) Probability map (d) Segmentation results by thresholding probability map (e) Segmentation results using K-means clustering
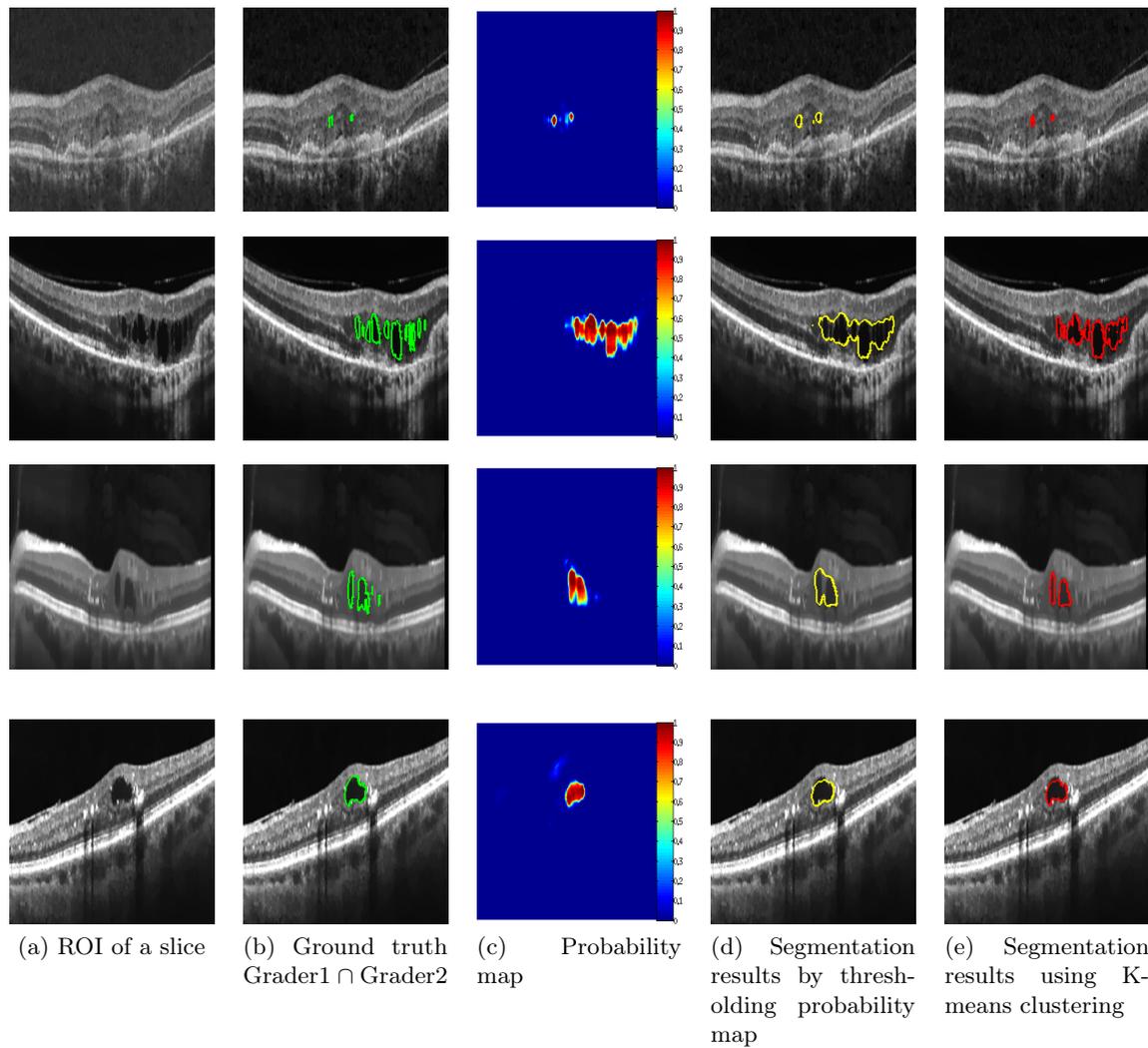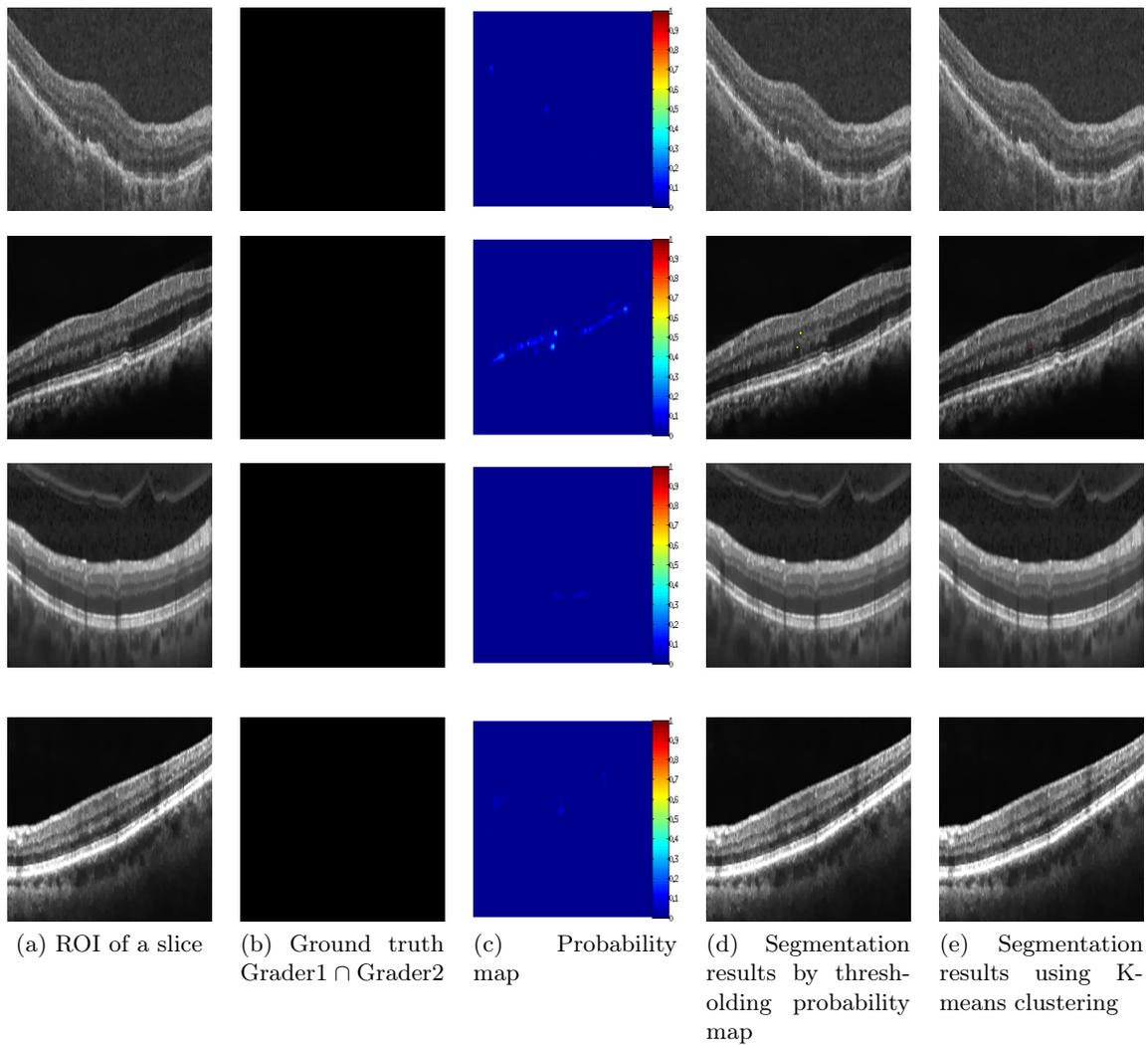
Figure 5.7: Qualitative results for a slice without cyst. Top to bottom rows: Results from a slice from Cirrus, Nidek, Spectralis, and Topcon scanners.

5.3.1.1   Effect of different GMP parameters

GMP computation requires choosing a few parameters such as the step size δ, range or extent of translation N and the number of translation directions or number of GMPs K. The last two directly affect the information content and size of cake C which is mapped by the CNN (by learning a function). K and N were varied while δ was held constant ($= 1$). The output of the CNN (I_o) was thresholded to obtain the detected cyst regions. Since these are highly localized, they were treated as roughly segmented cysts and evaluated using DC. Fig. 5.8 shows the computed DC as a function of K with $N = 5$.
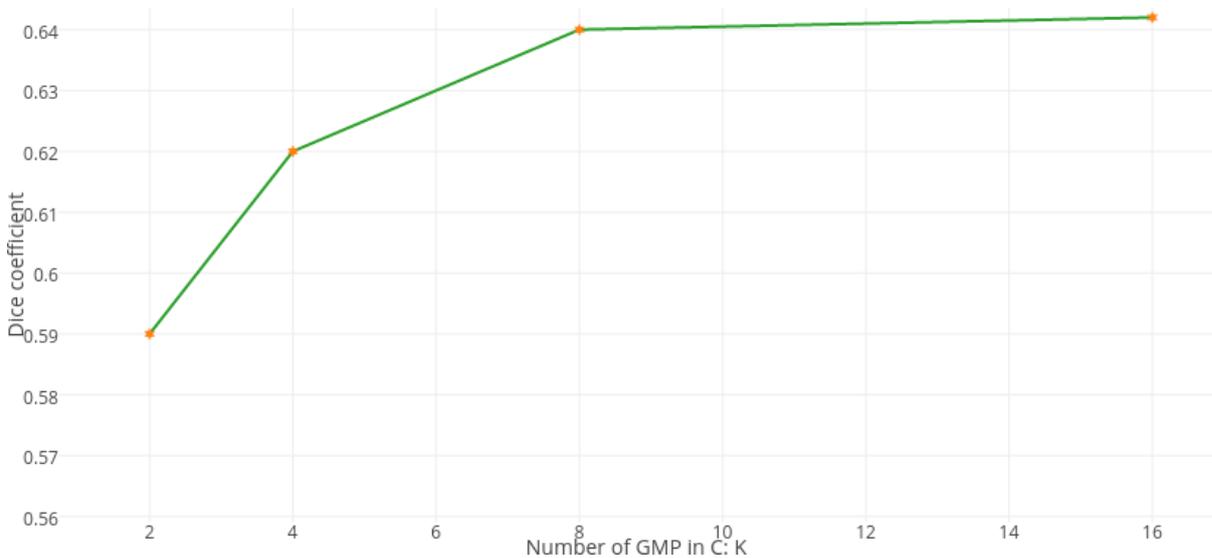
Figure 5.8: Segmentation performance as a function of the number (K) of GMPs.

The DC is seen to increase initially and saturate after $K = 8$. For $K < 8$ the cake is smaller which implies has less information to CNN and hence inadequate learning. Since the training time increases with the cake size, $K = 8$ was chosen for all our further experiments.

Next, the extent of translation N was varied. Specifically, $N = 2, 5, 10$ were considered. Small (large) value for N implies a small (large) range of induced motion. Since motion serves to spatially extend the cyst (along θ), a large N is potentially beneficial for detection of small cysts. However, this can lead to merging of cysts with dark background or give rise to false positives, given the fairly dense retinal layer structure. A small N helps overcome these problems, however at the risk of failure to detect small cysts. Thus the DC versus N plot can be expected to have a mode. This is confirmed in Fig. 5.9.

In experiments described henceforth, each of the GMP was created by applying translation motion on a ROI slice with $N = 5$, $\delta = 1$ and $\theta = 0°$ to $180°$ (in steps of $22.5°$); C was derived by using a *min* coalescing function over a stack of $K = 8$ GMPs.

Figure 5.9: Segmentation performance as a function of the extent (N) of translation in each GMP.

5.3.1.2 Effect of variants of the system

The proposed CNN architecture combines contextual information at 2D as well as 3D levels. To assess their relative effectiveness, a simple CNN architecture which relies only on 2D local information was considered by plugging out stages 1 and 2 in the proposed architecture. A cyst is a 3D structure and hence the 2D CNN model can be expected to perform poorly compared to 2D+3D CNN model (proposed). The DC listed in Table 5.3, is seen to improve with the addition of 3D information by 14%.

5.3.1.3 Effect on threshold at the CNN output

The thresholded result of the probability map (CNN output) represents a rough segmentation. The impact of this threshold on rough and fine segmentation was also assessed. The plot in figure Fig. 5.10 shows the DC and Jaccard index (JI) as a function of threshold level. Both increase initially as the rough segments continue to improve. DC and JI fall with threshold values above 0.5 since more noisy pixels are passed on. When the threshold exceeds 0.7 the DC obtained after clustering drops below the DC obtained prior to clustering. This is because, thresholding at a high probability score leads to fewer false positive (FP) regions and under segmentation, which in turn affects clustering. However, at a threshold of 0.35, clustering is seen to improve the performance by 7.8%. This was chosen as the optimal value for system evaluation.

54

Table 5.3: Mean DC values for different variants of the proposed CNN model on the OCSC dataset

Input	2D	2D+3D	Proposed system
	0.56	0.64	**0.69**

Table 5.4: Mean DC values for different variants of the CNN model

Method	OCSC Test set	DME Dataset	AEI Dataset
U-net with ROI as input	0.58	0.51	0.69
U-net with GMP as input	0.61	0.55	0.71
Proposed system	0.69	0.67	0.79

5.3.2   Evaluation of the system performance

Qualitative results of the proposed method are shown in Fig. 5.6 and Fig. 5.7. The probability map computed by the CNN, thresholded result overlaid on the original image and the final segmentation results are shown for a sample slice from 4 different scanners. It can be observed that the method is able to produce probability maps which highlight the cysts when present and nearly flat maps when cysts are absent. The thresholded outputs (in column d of Fig. 5.6) and final segments (in column e of Fig. 5.6) indicate that segmentation is consistently good for large and small cysts and the false detections are also minimal. The results also appear to be robust to inter scanner variation in contrast/intensity.

Next, we present the segmentation performance of the proposed system. Metrics such as DC, JI, Positive Predictive Value (PPV) and Sensitivity were used by OCSC organizers for evaluating the different segmentation methods. Since DC and JI convey very similar message, we choose to evaluate our work using DC alone. In order to assess the strengths of the proposed solution, we evaluate the work quantitatively against the U-Net [67] a widely used deep network for biomedical image segmentation tasks. It was trained end to end on the OCSC training set twice, once with ROI image as input and next with GMP as input. The obtained mean DC for the OCSC test set and cross validation results on DME and AEI dataset are listed in Table 5.4. The 5.1% improvement seen with GMP as input relative to ROI as input, substantiates the importance of using appropriate representation for the input data.

Finally, we benchmark the proposed system against i) the participants of the challenge on OCSC test set and ii) methods reported on DME dataset. The mean DC values are presented in Table 5.5. The results on the OCSC test set for the participants of the challenge are taken from the OCSC website [64]. The overall DC values obtained for different datasets are listed (in rank order) separately in Table 5.5. Here, G1 and G2 denote the GT from 2 graders; Unmasked (U) denotes the entire volume while Masked (M) denotes a volume with masked central 3 mm

circular region centered at macula as described in [66]; (S) in the row of Esmaeili et al. denotes the results are for only Spectralis scanner. It should be noted that results reported in [61] are for the OCSC training set whereas the DC values listed in the table are for the OCSC test set. The DC values should be nearly equal for M and U volumes if detection is not sensitive to location and this holds only for the proposed method which outperforms all other methods on the OCSC test dataset.

In the DME dataset, the uncertainty of fluid boundaries is reflected by the difference in DC values (0.65/0.9) for the two manual raters. Both [57] [59] have reported results on this dataset though the latter report only the volume similarity error metric and not DC. We consider DC as the standard metric for evaluating the performance of a segmentation task and hence benchmark with [57]. The DC values against the two experts as well as their union on DME dataset are given in Table 5.5. On the AEI dataset, the obtained DC value was **0.79/0.18**. The superior performance of the proposed system on the DME dataset and AEI dataset is noteworthy given that our system was *not* separately trained on these datasets.

Table 5.5: Dice coefficient for the OCSC dataset, DME dataset and AEI dataset

	OCSC Test dataset						DME dataset		
	U			M			Expert 1	Expert 2	Union
	G1	G2	G1∩G2	G1	G2	G1∩G2	(MJA)	(PSM)	(MJA , PSM)
Proposed Work	**0.67/0.17**	**0.68/0.17**	**0.69/0.18**	**0.70/0.17**	**0.70/0.15**	**0.71/0.16**	0.69/0.17	0.67/0.18	**0.67/0.17**
de Sisternes et al.	0.64/0.14	0.63/0.14	0.65/0.15	0.68/0.15	0.67/0.17	0.69/0.15	-	-	-
[58]	0.56/0.2	0.55/0.22	0.54/0.20	0.61/0.19	0.60/0.19	0.59/0.19	-	-	-
[59]	0.48/0.25	0.48/0.22	0.48/0.22	0.60/0.15	0.59/0.15	0.60/0.14	-	-	-
[61]	0.46/0.25	0.45/0.24	0.45/0.25	0.55/0.27	0.55/0.27	0.55/0.28	-	-	-
[63]	0.14/0.08	0.14/0.08	0.14/0.08	0.23/0.15	0.23/0.15	0.23/0.15	-	-	-
[57]	-	-	-	-	-	-	-	-	0.53/0.34
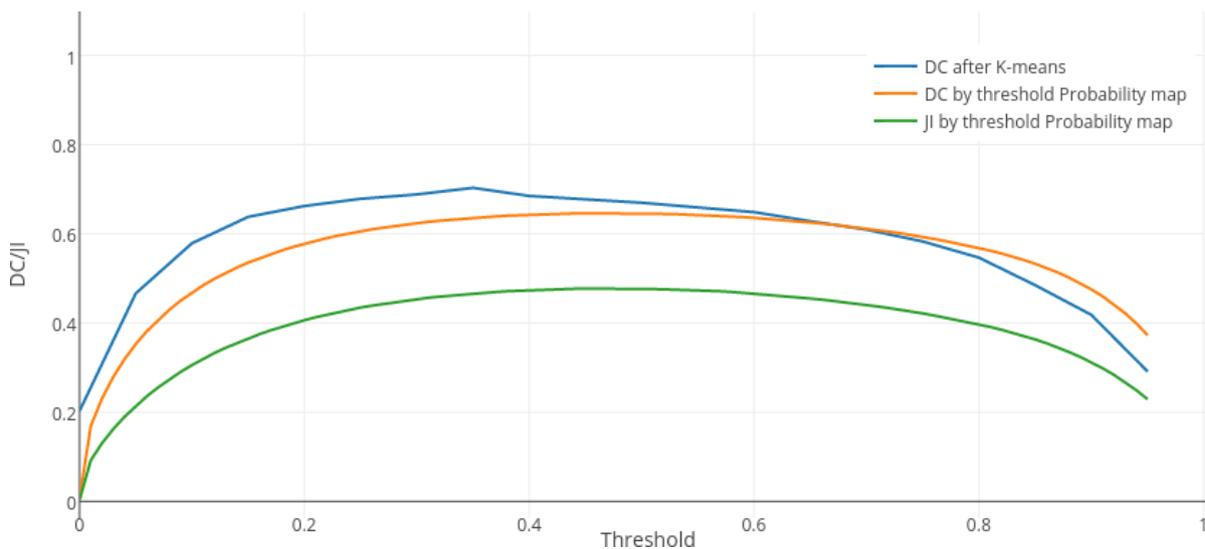
Figure 5.10: DC and JI for various thresholds at different stages

(a) A sub image around cyst

(b) Ground truth (Grader1 ∩ Grader2)
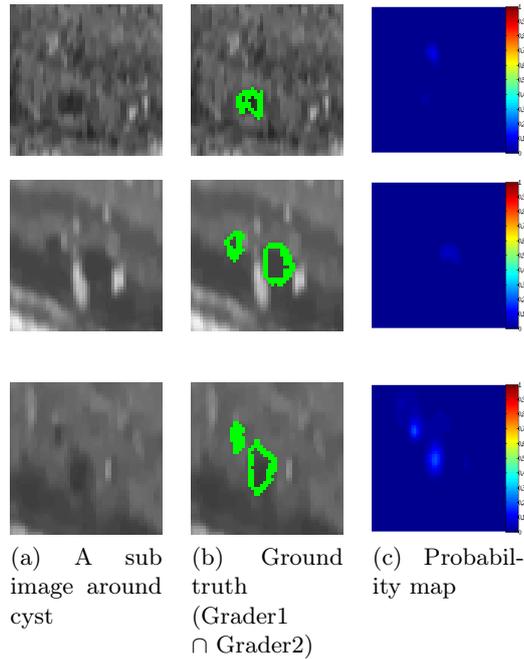
(c) Probability map

Figure 5.11: Sample failure (to detect) cases.

5.4 DISCUSSION AND CONCLUSIONS

An automated retinal cyst segmentation algorithm was presented. It is a generic approach which is independent of scanners and does not require layer segmentation which is in contrast to existing approaches. For instance, layer-dependent statistical information is used to define a cost function in the graph based method in [59] and serves as a feature in machine learning based algorithm in [57]. Our method depends only on extracting an ROI between ILM and RPE for processing. The heart of the proposed method is a novel representation derived by inducing motion. The CNN is used to selectively enhance the object of interest and aid in detection of the object. As only few parameters are to be learnt, CNN training is possible with only 584 sample slices (containing cysts) from the OCSC training dataset. The architecture of the CNN is designed to utilize contextual information in both 2D and 3D and at multiple scales which results in good performance on large cysts and small fluid regions. The effectiveness of the CNN is underscored by the fact that its thresholded output is able to achieve results similar to the other methods proposed during the MICCAI challenge. A simple clustering on this output improved the DC by nearly 7.8%. The cross validation results on the DME dataset and AEI dataset points to the robustness of the proposed method.

The proposed system has some limitations. It fails to detect cyst regions which are very small in a volume especially when they are isolated in depth. These are sometimes treated as noise and rejected. Sample failure cases are shown in Fig. 5.11. It also tends to merge small cysts,

when they are adjacent to a large one and this in turn affects the DC value. One such case is seen in the second row of Fig. 5.6.

Besides cysts which are fluid filled regions, quantifying fatty deposits such as drusen in the RPE layer, which occur in Age related macular degeneration, are also of interest in OCT image analysis. The proposed segmentation method can easily be extended to handle drusen by simply replacing the coalescing function with max. This, we believe, is because our strategy for localization is selective enhancement of the abnormality of interest, agnostic to the location of the abnormality. Further improvement in the work is possible by embedding domain-assisted spatial information in the input to the CNN.

Chapter 6

Conclusions and Future Work

This thesis has taken on the problem of developing an automated OCT image analysis system. While there is an exhaustive literature available for OCT analysis, in this thesis we present four individual modules which are combined to form this system. Firstly, we approach building the system by developing algorithms for OCT anatomy analysis and disease analysis as segmentation and classification tasks from traditional machine learning and image processing perspective. We start in chapter 2 by focusing on extraction of retinal layers by employing deep learning techniques. With advancements in deep learning, we have been able to develop an end to end trainable module for segmenting layer boundaries. Parameter dependent modules like image flatenning, denoising and edge detection employed previously in literature are automated in the proposed pipeline. The segmentation method proposed in chapter 2 is capable of handling OCT data from both normal and AMD affected macula regions. The proposed automatic segmentation algorithm's quantitative results are on par with the existing methods. Shortage of data and boundary markings hinders this deep learning framework to outperform the state of the art methods. However, it is necessary to appreciate the strength of this single model in handling scanner variability and pathological cases. In chapter 3, we have seen the use of additional angio information for glaucoma classification. The RNFL thickness along with capillary density helps in detecting glaucoma better. This establishes the correlation of RNFL and CD for a normal and glaucomatous subject. Chapter 4 proposes an automated classification module for detecting AMD and CME at slice level in an OCT volume. We discuss a pipeline in chapter 4 to analyze macula region of retina from OCT data. We have seen in this chapter, how biologically inspired extremal representation can be used with convolutional neural network (CNN) classifier to diagnose retinal disease. The high performance of the system on data collected from multiple scanner illustrates the capabilities of the module. Apart from identifying disease, the thesis also focuses on segmentation the retinal abnormality like cysts. In chapter 5, we propose an abnormality segmentation algorithm. This algorithm aims at localizing/segmenting retinal cysts from OCT volumes. It is based on learning a mapping function that combines stack of bi-

ologically inspired generalized motion patterns (GMP). Our investigations were primarily cysts and hence the type of motion and coalescing function for GMP is fixed. A CNN in this context learns the required mapping function to produce a an output image with probable cyst regions. Our evaluation of this method has shown,that by a simple thresholding and clustering, it can be used to segment cysts better than the available automated methods. Also, the system performs well with cross validation on multiple datasets without explicit training.

6.1 FUTURE WORK

The following are the possible directions of future work.

- **Anatomy extraction:** In chapter 2, the method segments retinal layers even in presence of macular pathologies. We have not attempted segmentation of retinal layers in presence of pathologies like cysts due to the lack of data required for deep learning. Also, segmentation of retinal layers is useful when Optic disk (OD) is analysed. This problem can be addressed by proposing a unified algorithm for segmenting layer boundaries of OD and macula. This requires large amount of annotated data (layer boundaries marking).

- **Classification of diseases at volume level:** Chapter 3 uses a new imaging modality namely OCTA. The use of this modality for disease diagnosis is established in this chapter. It would be interesting to investigate this imaging modality at macula region and study the effect of vasculature on other retinal disease. We can approach this by obtaining structural and vascular information by mosaicing data obtained at different locations.

- **Classification of diseases at slice level:** The module proposed in chapter 4 can successfully classify a macula OCT as normal or diseased. The module can be enhanced to handle other disease like Diabetic retinopathy. Another possible future direction is in attempting to detect diseases at different stages. We can approach this problem by providing domain knowledge pertaining to the stages of the disease only during training.

- **Abnormality extraction:** Retinal cysts are the abnormalities that is ssegmented in chapter 5. The final stage in the pipeline involves simple clustering based segmentation. This technique proves effective, however, in the future work we can train an end to end conditional random field (CRF) for getting more accurate cyst regions. Methods already exist for learning CRF's with CNN's for segmentation in computer vision applications. We can attempt this problem by combining the CRF module with the proposed CNN architecture.

Related Publications

- *Domain knowledge assisted cyst segmentation in OCT retinal images.*
  Karthik Gopinath and Jayanthi Sivaswamy.
  arXiv preprint arXiv:1612.02675 (2016). (Accepted for oral presnetation in MICCAI 2015 challenge)

- *Automatic glaucoma assessment from angio-OCT images.*
  Karthik Gopinath, Jayanthi Sivaswamy and Tarannum Mansoori.
  IEEE 13th International Symposium on Biomedical Imaging (ISBI) 2016, Prague, Czech Republic.

- Under Review
  Automated segmentation of retinal cysts from Optical Coherence Tomography volumes.
  Karthik Gopinath, Jayanthi Sivaswamy, Andru Pavani Reddy and Satish G Agraharam.
  Medical Image Analysis.

- Under Review
  A deep learning framework for segmentation of retinal layers from OCT images.
  Karthik Gopinath*, Samrudhdhi B*, and Jayanthi Sivaswamy.
  International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer International Publishing, 2017.

---

* equal contribution

Bibliography

[1] R. F. Spaide, J. M. Klancnik, and M. J. Cooney, "Retinal vascular layers imaged by fluorescein angiography and optical coherence tomography angiography," *JAMA ophthalmology*, vol. 133, no. 1, pp. 45–50, 2015.

[2] J. C. Booij, D. C. Baas, J. Beisekeeva, T. G. Gorgels, and A. A. Bergen, "The dynamic nature of bruch's membrane," *Progress in retinal and eye research*, vol. 29, no. 1, pp. 1–18, 2010.

[3] Y. Jia *et al.*, "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography," *Optics express*, vol. 20, no. 4, pp. 4710–4725, 2012.

[4] A. Rothova, M. Suttorp-van Schulten, W. F. Treffers, and A. Kijlstra, "Causes and frequency of blindness in patients with intraocular inflammatory disease.," *British Journal of Ophthalmology*, vol. 80, no. 4, pp. 332–336, 1996.

[5] B. Munoz, S. K. West, G. S. Rubin, O. D. Schein, H. A. Quigley, S. B. Bressler, and K. Bandeen-Roche, "Causes of blindness and visual impairment in a population of older americans: The salisbury eye evaluation study," *Archives of Ophthalmology*, vol. 118, no. 6, pp. 819–825, 2000.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] C. Olah, "Understanding lstm networks," 2015.

[8] K. Dabov *et al.*, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE TIP*, vol. 16, no. 8, pp. 2080–2095, 2007.

[9] G. Gilboa *et al.*, "Image enhancement and denoising by complex diffusion processes," *IEEE TPAMI*, vol. 26, no. 8, pp. 1020–1036, 2004.

[10] D. Fu *et al.*, "Retinal automatic segmentation method based on prior information and optimized boundary tracking algorithm," in *Eighth ICDIP 2016)*, pp. 100331C–100331C, International Society for Optics and Photonics, 2016.

[11] D. C. Fernández, , et al., "Automated detection of retinal layer structures on optical coherence tomography images," *Optics Express*, vol. 13, no. 25, pp. 10200–10216, 2005.

[12] S. Lu et al., "Automated layer segmentation of optical coherence tomography images," *IEEE TBME*, vol. 57, no. 10, pp. 2605–2608, 2010.

[13] S. J. Chiu et al., "Automatic segmentation of seven retinal layers in sdoct images congruent with expert manual segmentation," *Optics express*, vol. 18, no. 18, pp. 19413–19428, 2010.

[14] O. Tan et al., "Mapping of macular substructures with optical coherence tomography for glaucoma diagnosis," *Ophthalmology*, vol. 115, no. 6, pp. 949–956, 2008.

[15] T. Fabritius et al., "Automated segmentation of the macula by optical coherence tomography," *Optics express*, vol. 17, no. 18, pp. 15659–15669, 2009.

[16] A. Mishra et al., "Intra-retinal layer segmentation in optical coherence tomography images," *Optics express*, vol. 17, no. 26, pp. 23719–23728, 2009.

[17] M. K. Garvin et al., "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-d graph search," *IEEE TMI*, vol. 27, no. 10, pp. 1495–1505, 2008.

[18] J. Tian et al., "Real-time automatic segmentation of optical coherence tomography volume data of the macular region," *PloS one*, vol. 10, no. 8, p. e0133908, 2015.

[19] S. J. Chiu et al., "Validated automatic segmentation of amd pathology including drusen and geographic atrophy in sd-oct images," *Investigative ophthalmology & visual science*, vol. 53, no. 1, pp. 53–61, 2012.

[20] M. A. Hussain et al., "Automatic identification of pathology distorted retinal layer boundaries using sd-oct imaging," *IEEE TBME*, 2016.

[21] V. Nair et al., "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th ICML*, pp. 807–814, 2010.

[22] M. D. Zeiler, "Adadelta: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.

[23] Y. A. LeCun et al., "Efficient backprop," in *Neural networks: Tricks of the trade*, pp. 9–48, Springer, 2012.

[24] A. L. Kornzweig et al., "Selective atrophy of the radial peripapillary capillaries in chronic glaucoma," *Archives of ophthalmology*, vol. 80, no. 6, pp. 696–702, 1968.

[25] K. Y. Paula *et al.*, "Correlation between the radial peripapillary capillaries and the retinal nerve fibre layer in the normal human retina," *Experimental eye research*, vol. 129, pp. 83–92, 2014.

[26] M. A. Mayer *et al.*, "Retinal nerve fiber layer segmentation on fd-oct scans of normal subjects and glaucoma patients," *Optical Society of America*, vol. 1, no. 5, pp. 1358–1383, 2010.

[27] J. Cheng *et al.*, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 6, pp. 1019–1032, 2013.

[28] G. D. Joshi *et al.*, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 6, pp. 1192–1205, 2011.

[29] D. Bendschneider *et al.*, "Retinal nerve fiber layer thickness in normals measured by spectral domain oct," *Journal of glaucoma*, vol. 19, no. 7, pp. 475–482, 2010.

[30] S. J. Chiu *et al.*, "Automatic segmentation of seven retinal layers in sdoct images congruent with expert manual segmentation," *Optics express*, vol. 18, no. 18, pp. 19413–19428, 2010.

[31] M. A. Mayer *et al.*, "Fuzzy c-means clustering for retinal layer segmentation on high resolution oct images," *Proceedings of the Biosignal*, vol. 19, 2008.

[32] N. Plange *et al.*, "Fluorescein filling defects and quantitative morphologic analysis ofthe optic nerve head in glaucoma," *Archives of ophthalmology*, vol. 122, no. 2, pp. 195–201, 2004.

[33] S. Garg *et al.*, "Unsupervised curvature-based retinal vessel segmentation," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 344–347, IEEE, 2007.

[34] G. Azzopardi *et al.*, "Trainable cosfire filters for vessel delineation with application to retinal images," *Medical image analysis*, vol. 19, no. 1, pp. 46–57, 2015.

[35] P. T. De Jong, "Age-related macular degeneration," *New England Journal of Medicine*, vol. 355, no. 14, pp. 1474–1485, 2006.

[36] P. Romero-Aroca, "Managing diabetic macular edema: The leading cause of diabetes blindness," *World J Diabetes*, 2011.

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

[38] G. M. Somfai, E. Tátrai, L. Laurik, B. Varga, V. Ölvedy, H. Jiang, J. Wang, W. E. Smiddy, A. Somogyi, and D. C. DeBuc, "Automated classifiers for early detection and diagnosis of retinopathy in diabetic eyes," *BMC bioinformatics*, vol. 15, no. 1, p. 1, 2014.

[39] R. Koprowski, S. Teper, Z. Wróbel, and E. Wylegala, "Automatic analysis of selected choroidal diseases in oct images of the eye fundus," *Biomedical engineering online*, vol. 12, no. 1, p. 1, 2013.

[40] Y. Zhang, B. Zhang, F. Coenen, J. Xiao, and W. Lu, "One-class kernel subspace ensemble for medical image classification," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–13, 2014.

[41] Y.-Y. Liu, H. Ishikawa, M. Chen, G. Wollstein, J. S. Duker, J. G. Fujimoto, J. S. Schuman, and J. M. Rehg, "Computerized macular pathology diagnosis in spectral domain optical coherence tomography scans based on multiscale texture and shape features," *Investigative ophthalmology & visual science*, vol. 52, no. 11, pp. 8316–8322, 2011.

[42] P. P. Srinivasan, L. A. Kim, P. S. Mettu, S. W. Cousins, G. M. Comer, J. A. Izatt, and S. Farsiu, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomedical optics express*, vol. 5, no. 10, pp. 3568–3577, 2014.

[43] K. S. Deepak, N. K. Medathati, and J. Sivaswamy, "Detection and discrimination of disease-related abnormalities based on learning normal cases," *Pattern Recognition*, vol. 45, no. 10, pp. 3707–3716, 2012.

[44] K. S. Deepak and J. Sivaswamy, "Automatic assessment of macular edema from color retinal images," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 766–776, 2012.

[45] J. Wu, A.-M. Philip, D. Podkowinski, B. S. Gerendas, G. Langs, C. Simader, S. M. Waldstein, and U. M. Schmidt-Erfurth, "Multivendor spectral-domain optical coherence tomography dataset, observer annotation performance evaluation, and standardized evaluation framework for intraretinal cystoid fluid segmentation," *Journal of Ophthalmology*, vol. 2016, 2016.

[46] S. J. Chiu, M. J. Allingham, P. S. Mettu, S. W. Cousins, J. A. Izatt, and S. Farsiu, "Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema," *Biomedical optics express*, vol. 6, no. 4, pp. 1172–1194, 2015.

[47] S. Farsiu, S. J. Chiu, R. V. O'Connell, F. A. Folgar, E. Yuan, J. A. Izatt, C. A. Toth, A.-R. E. D. S. . A. S. D. O. C. T. S. Group, *et al.*, "Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography," *Ophthalmology*, vol. 121, no. 1, pp. 162–172, 2014.

[48] T. G. Rotsos and M. M. Moschos, "Cystoid macular edema," *Clinical ophthalmology (Auckland, NZ)*, vol. 2, no. 4, p. 919, 2008.

[49] S. Scholl, J. Kirchhof, and A. J. Augustin, "Pathophysiology of macular edema," *Ophthalmologica*, vol. 224, no. Suppl. 1, pp. 8–15, 2010.

[50] D. Cabrera Fernández, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Optics Express*, vol. 13, no. 25, pp. 10200–10216, 2005.

[51] D. C. Fernandez, "Delineating fluid-filled region boundaries in optical coherence tomography images of the retina," *IEEE transactions on medical imaging*, vol. 24, no. 8, pp. 929–945, 2005.

[52] G. R. Wilkins, O. M. Houghton, and A. L. Oldenburg, "Automated segmentation of intraretinal cystoid fluid in optical coherence tomography," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 1109–1114, 2012.

[53] A. González, B. Remeseiro, M. Ortega, M. G. Penedo, and P. Charlón, "Automatic cyst detection in oct retinal images combining region flooding and texture analysis," in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pp. 397–400, IEEE, 2013.

[54] X. Chen, M. Niemeijer, L. Zhang, K. Lee, M. D. Abràmoff, and M. Sonka, "Three-dimensional segmentation of fluid-associated abnormalities in retinal oct: probability constrained graph-search-graph-cut," *IEEE transactions on medical imaging*, vol. 31, no. 8, pp. 1521–1531, 2012.

[55] X. Xu, K. Lee, L. Zhang, M. Sonka, and M. D. Abramoff, "Stratified sampling voxel classification for segmentation of intraretinal and subretinal fluid in longitudinal clinical oct data," *IEEE transactions on medical imaging*, vol. 34, no. 7, pp. 1616–1623, 2015.

[56] H. Bogunović, M. D. Abràmoff, and M. Sonka, "Geodesic graph cut based retinal fluid segmentation in optical coherence tomography," *Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop, OMIA 2015, Held in Conjunction with MICCAI 2015*, 2015.

[57] S. J. Chiu *et al.*, "Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema," *Biomedical optics express*, vol. 6, no. 4, pp. 1172–1194, 2015.

[58] F. Venhuizen, M. J. van Grinsven, B. van Ginneken, C. C. Hoyng, T. Theelen, and C. I. Sanchez, "Fully automated segmentation of intraretinal cysts in 3d optical coherence tomography," *Investigative Ophthalmology & Visual Science*, vol. 57, no. 12, pp. 5949–5949, 2016.

[59] I. Oguz, L. Zhang, M. D. Abràmoff, and M. Sonka, "Optimal retinal cyst segmentation from oct images," in *SPIE Medical Imaging*, pp. 97841E–97841E, International Society for Optics and Photonics, 2016.

[60] Dataset, "Dme dataset from duke univerity."

[61] R. H. H. F. Mahdad Esmaeili M, Dehnavi AM, "Three-dimensional segmentation of retinal cysts from spectral-domain optical coherence tomography images by the use of three-dimensional curvelet based kâĂŚsvd," *J Med Sign Sence*, vol. 6, pp. 166–171, 2016.

[62] G. Girish, A. R. Kothari, and J. Rajan, "Automated segmentation of intra-retinal cysts from optical coherence tomography scans using marker controlled watershed transform," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pp. 1292–1295, IEEE, 2016.

[63] K. Gopinath and J. Sivaswamy, "Domain knowledge assisted cyst segmentation in oct retinal images," *arXiv preprint arXiv:1612.02675*, 2016.

[64] OPTIMA, "Proceeding of the miccai optima cyst segmentation challenge."

[65] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[66] J. Wu, A.-M. Philip, D. Podkowinski, B. S. Gerendas, G. Langs, C. Simader, S. M. Waldstein, and U. M. Schmidt-Erfurth, "Multivendor spectral-domain optical coherence tomography dataset, observer annotation performance evaluation, and standardized evaluation framework for intraretinal cystoid fluid segmentation," *Journal of Ophthalmology*, vol. 2016, 2016.

[67] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, 2015.