Exploration of multiple imaging modalities for Glaucoma detection

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

Master of Science
in
Electronics and Communication engineering by Research

by

Gamalapati sai jahnavi
201334006
jahnavi.gamalapati@research.iiit.ac.in

International Institute of Information Technology
Hyderabad - 500 032, INDIA
July 2018

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "Exploration of multiple imaging modalities for Glaucoma detection" by G Sai Jahnavi, has been carried out under my supervision and is not submitted elsewhere for a degree.

_____ _____

Date Adviser: Prof. Jayanthi Sivaswamy

To

my Sister

Acknowledgments

With immense gratitude, I would like to acknowledge Professor. Jayanthi Sivaswamy. This thesis would not have been possible without her support. She is my guide, mentor, and inspiration. She constantly encouraged me to be better at my work and life by providing her valuable suggestions. I would also like to thank Dr. Tarannum mansoori for providing medical data required for my thesis work and also for providing relevant clinical knowledge which formed the basis for my thesis.

I thank Ujjwal for helping me learn the basics when I joined the lab and he has been a wonderful mentor to my work. Special thanks to Raghav Mehta whos knowledge and valuable suggestions provided a kick start to my thesis work.

My heartfelt thanks to Karthik Gopinath for being with me with all ups and downs throughout the completion of course. My special thanks to AK Pujitha for being my wonderful teammate during my initial days. Your presence made things simple from the very beginning. Our trips to hospitals, lab trips, and Melbourne are unforgettable. I would like to acknowledge Sukesh Adiga, Chetan for their moral support during the placements and for the completion of my thesis work. Your presence made research lively and interesting. Endless thanks to Mounika, Deepya, Aditya for making all these years memorable and finally, I would like to dedicate my thesis to my parents who believed in me throughout the course completion. Their Patience and constant support helped to deliver my best.

Abstract

Glaucoma is an eye disease characterized by weakening of nerve cells often resulting in a permanent loss of vision. Glaucoma progression can occur without any physical indication to patients. Hence, early diagnosis of Glaucoma is recommended for preventing the permanent damage to vision. Early Glaucoma is often characterized by thinning of the Retinal Nerve Fiber Layer which is commonly called as RNFL defect (RNFLD). Computer-aided diagnosis (CAD) of eye diseases is popular and is based on automated analysis of fundus images. CAD solutions for diabetic retinopathy have reached more maturity than for glaucoma as the latter is more difficult to detect from fundus images. This is due to the fact that nerve damage appears in the form of subtle change in the background around optic disc.

SD-OCT (Spectral Domain - Optical Coherence Tomography), a recently introduced modality, helps to capture 3D information of retina. Hence, it is more reliable for detecting nerve damage in retina compared with fundus imaging. However, a wide usage of OCT is limited due to cost per scan, time and ease of acquisition. This thesis focuses on integrating information from multiple modalities (OCT and fundus) for improving retinal nerve fibre layer defect or detection of RNFLD from fundus images. We examine two key problems in the context of CAD development: i) spatial alignment or registration of two modalities of imaging, namely, 2D fundus and 3D OCT volume images. This can pave way to integrate information across the modalities. Multimodal registration is challenging because of the varied Field of View and noise levels across the modalities. We propose a computationally efficient registration algorithm which is capable of handling complementary nature of modalities. Extensive qualitative and quantitative evaluations are performed to show the robustness of proposed method. ii) Detection of RNFLD from fundus images with good accuracy. We propose a novel CAD solution which utilises information from the 2 modalities for learning a model and uses it to predict the presence of RNFLD from fundus images. The problem is posed as learning from 2 modalities (fundus, OCT images) and predicting from only one (fundus images) with the other (OCT) as missing data. Our solution consists of a deep neural network architecture which learn modality independent representations.

In the final part of the thesis we explore the scope of a new imaging modality angiography-Optical Coherence Tomography (A-OCT) in diagnosing Glaucoma. Two case studies are reported which help in understanding the progression of Retinal Nerve Fiber Layer thickness, Capillary Density in normal

and glaucoma effected patients. The experiments on new modality has shown potential for considering it as a reliable biomarker along with existing modalities.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Glaucoma is the second leading cause of blindness all over the world. It causes damage to the optic nerve of the eye due to increase in Intraocular pressure. This damage results in permanent blindness due to poor transmission of signals between the brain and the eye. The number of people with glaucoma around the world was estimated to be 64.3 million in 2013 and is expected to increase to 76 million by 2020 [4][19]. Glaucoma is termed as silent blinding disease because it neither gives warnings nor obvious symptoms to the patient. If left untreated, most types of glaucoma progress towards permanent damage of the eye. There are two main kinds of Glaucoma: Open-angle glaucoma and Angle-closure glaucoma. Open-angle glaucoma is the most common form of glaucoma. It is caused by the clogging of drain canals, leading to an increased pressure in the retina. It has a wide angle between between the iris and the cornea. The later (Angle-closure glaucoma) is caused by blocked drainage canals leading to a sudden increase in intraocular pressure. It has a narrow angle between the iris and the cornea. (See figure:1.1) Most of the above diseases worsen without displaying any kind of observable symptoms. Thus, screening for glaucoma is important for their early detection and dispersion of effective treatment to prevent permanent blindness.

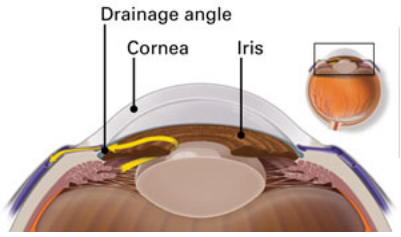


Figure 1.1: Location of Cornea, Iris in eye

For identifying Glaucoma, two kinds of clinical assessments[20] are performed: i) Intra-papillary and ii)Peri-papillary. Intra-papillary is related to geometric changes of Optic Disk (OD). Particularly the presence of glaucoma leads to increase in cup/disk ratio. Whereas later one is related to loss of retinal nerve fiber layer (RNFL defect) occurring in the periphery of the OD region (see figure:1.2 (right)). Parapapillary assessment is mostly done to detect diseases at during early stages of retina.

1.1 IDENTIFICATION OF GLAUCOMA

To understand the progression of Glaucoma, many imaging modalities exist. These imaging modalities unveil the structural and functional information of the retina. Evaluation of multiple modalities further enhances the accurate diagnosis of the patient. In most of the imaging modalities light plays a major role in capturing information. Some of these modalities include Fundus-(Optical imaging), Optical Coherence Tomography-(cross sectional imaging), and Angiography Optical Coherence Tomography-(Flow imaging).

1.1.1 Fundus imaging

It is a non invasive modality and consists of low power microscope with an attached camera. The color fundus camera helps to capture the 3D retina as a 2D image as shown in figure:1.2 (left). This projection captures the presence of disorders and aids to monitor their change over time. Most of the screening systems existing today use fundus imaging as it is very cost effective compared to other techniques. There are several modes of examination using fundus imaging, namely color fundus photography, red free fundus photography, and angiography. In color fundus imaging natural light is used for imaging the retina, where as in red free fundus imaging a specified wavelength of light is used. Angiography is an invasive procedure and less frequent in usage. In this protocol, fluorescent dye is injected and its movement is observed in the vasculature structure.

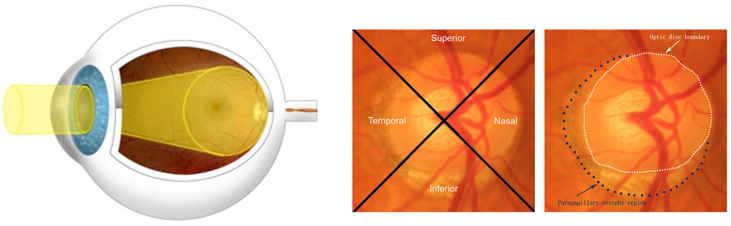
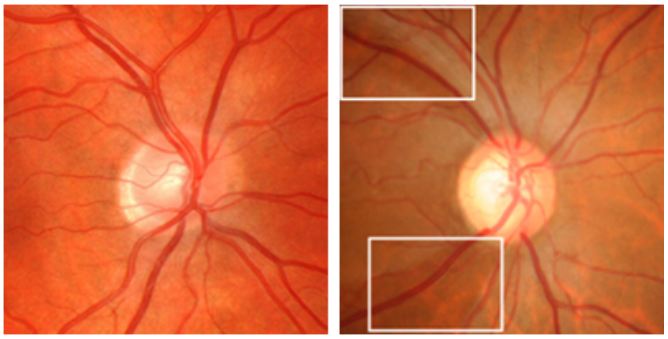
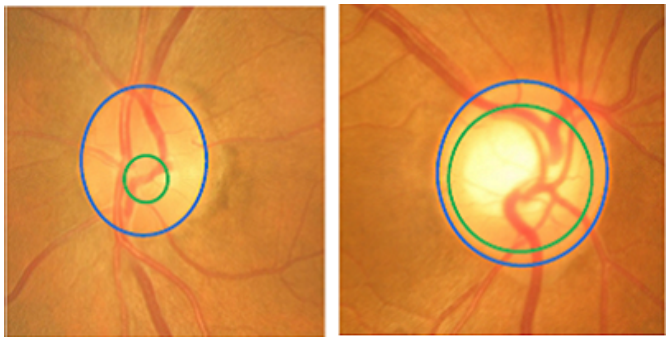


Figure 1.2: Fundus capture of retina (a),Sector division(b,l), Parapapillary region (b,r)

For glaucoma, fundus imaging is the most common modality used for screening systems. In fundus, the presence of glaucoma is indicated by the presence of a bright striped radiation pattern from Optic Disc, enlargement of cup, as shown in figure:1.3a, 1.3b. The radiation pattern changes its visual appearance according to variations in the Retinal Nerve Fiber thickness. The thickness variations can be caused by glaucoma process, or even in the healthy retina, the RNFL thickness varies anatomically depending on angular position around the ONH. In general RNFL measurement is performed in sectors as shown in figure:1.2 (b,r), where S:Superior, T:Temporal, N: Nasal, I: Inferior.



(a) Typical RNFL pattern observed (Region inside white ROI)



(b) Increased cup size

Figure 1.3: Identification of Glaucoma from fundus photography.

1.1.2 Optical Coherence Tomography imaging

It is a non-invasive imaging modality. Optical coherence tomography is a technique for obtaining depth images (cross-section) of retina at a resolution equal to a low-power microscope. It is effectively an optical ultrasound imaging. As it is cross-sectional imaging, we can see each of the distinctive layers of the retina. This allows to map and measure their thickness as each layer is associated with its own significance. For example: Capillary vessels are observed in rpe layer while abnormalities, like drusen, are found between the isos-rpe layer. Glaucoma is caused by thinning of RNFL layer. The qualitative measurements of these layers help with diagnosis. Figure:1.4 (right) shows layers in an OCT volume.

A low coherent light is split optically into two beams. One is sent to a mirror at a specific distance to reflect and the other is made to reflect from the retinal tissue as shown in figure:1.4 (left). The energy of the interference between the two reflected beams is encoded as intensity in the OCT image.There are two types of imaging 1) Time domain OCT, 2) Spectral domain OCT. Time domain OCT is an old imaging method in which the reference mirror is moved mechanically to different positions. This results in poor image resolution, subject to inconvenience, and takes time, whereas Spectral domain OCT includes 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 these structures. The need for mechanically moving the reference mirror is eliminated in this approach. This consequently increases the speed and resolution of scanning.

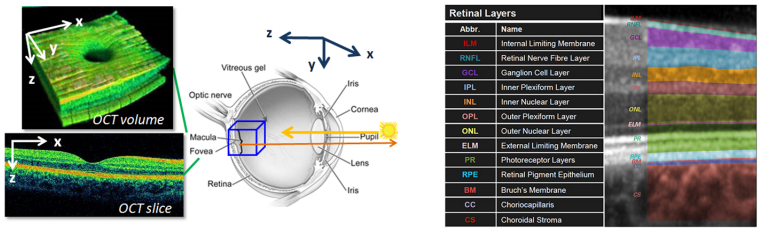


Figure 1.4: Acquisition of OCT (left), OCT layer information (Right)

The use of SD-OCT for glaucoma diagnosis has become a common clinical practice. Numerous studies have demonstrated that RNFL and macular thickness parameters are reproducible, with high sensitivity and specificity in discriminating between healthy and glaucomatous eyes. Figure:1.5 shows the RNFL thickness report of Glaucoma patient.

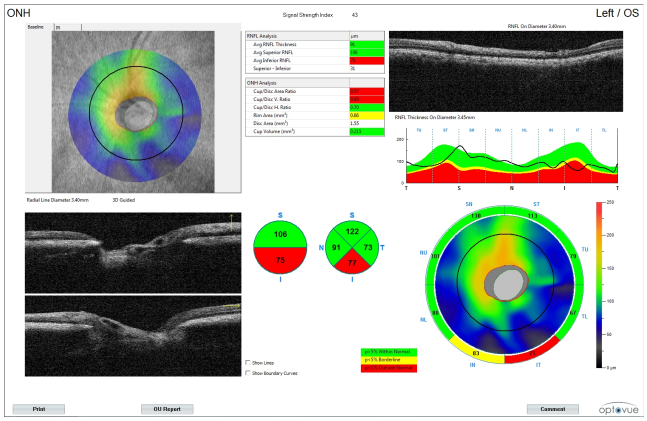


Figure 1.5: Sample OCT report of Glaucoma patient

1.1.3 Angiography - Optical Coherence Tomography imaging

Optical coherence tomography angiography (A-OCT) is a recent non-invasive modality for imaging microvasculature of the retina. This modality uses laser light as a source and maps reflectance from retina vessels into different layers. Unlike fundus angiography, this modality does not require injection of a dye. The A-OCT scan of patients consists of multiple B-Scans to provide cross-sectional structural information. The same tissue is repeatedly captured and the differences between successive scans are analyzed. These differences are detected as flow and no flow regions. For differentiating successive scans, two types of motion detection methods are employed: amplitude decorrelation and phase variance. The former detects differences in amplitude between two different OCT B-scans, where as phase variance calculates the variation in phases of moving objects. To improve visualization and reduce background noise from normal flickering of the eye, two averaging methods were developed, viz. split spectrum amplitude decorrelation technique and volume averaging - .[14] OCT-A is also gaining

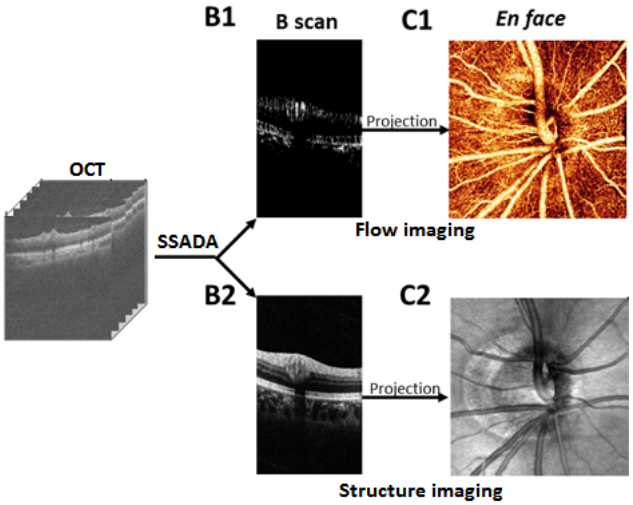


Figure 1.6: Difference between OCT, OCT-A

increasing popularity for optic nerve disorders assessment, such as glaucoma. It has been reported as a useful tool for evaluating optic disc perfusion in glaucomatous eyes, since attenuated peripapillary and macular vessel density is detectable in pre-perimetric glaucoma patients as shown in figure:1.7.

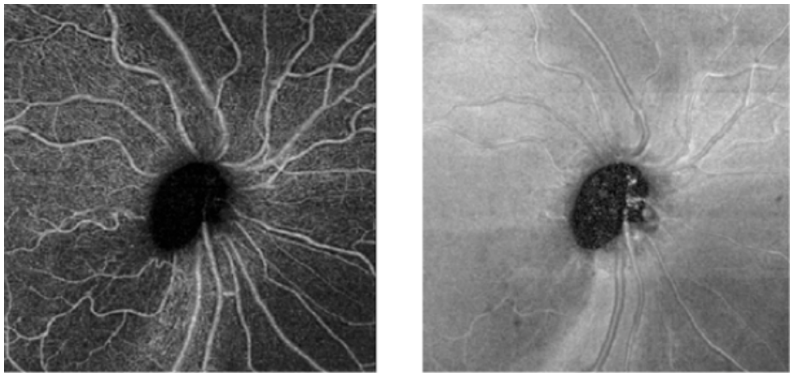


Figure 1.7: OCT-A (left) and en face structural imaging (right) of a glaucoma patient, revealing inferior temporal loss of retinal fiber layer and the anatomical correspondence with decreased vessel density in the OCT-A image.

1.2 SCREENING SYSTEMS - MOTIVATION OF THE THESIS

Screening systems are prevalent in many countries. The patient is classified as Glaucomatous based on measurements such as Visual fields, intraocular pressure, Optic Disc/Cup diameters etc. In almost all screening systems color fundus imaging is used as standard for screening because of its low cost, non-invasiveness and ease of use. Recently several studies revealed that diagnosis of patients are not correlated with usage of fundus and OCT images. In diagnosis of Glaucoma optic cup to disk ratio

is also clearly detected from OCT than fundus image. These studies concluded fundus imaging alone may not be sufficient for screening. This calls for the requirement of multiple modalities for improving diagnosis of patient.

1.3 SCOPE OF THE THESIS

The extent of visibility of diseases in retina depends on the type of imaging modality. The use of other reliable modalities for screening systems such as OCT has its own limitations in terms of cost and ease of acquisition. The work in this thesis focuses on improving disease detection rate on fundus by combining information from OCT and fundus modality(figure:1.8). Prior to integration of modalities, multimodal registration is performed to ensure proper learning. We have also investigated the scope of angiography OCT new imaging flow modality in understanding the progression of disease Glaucoma. Clinical studies prove that this new modality will serve as a reliable biomarker in detection of diseases.

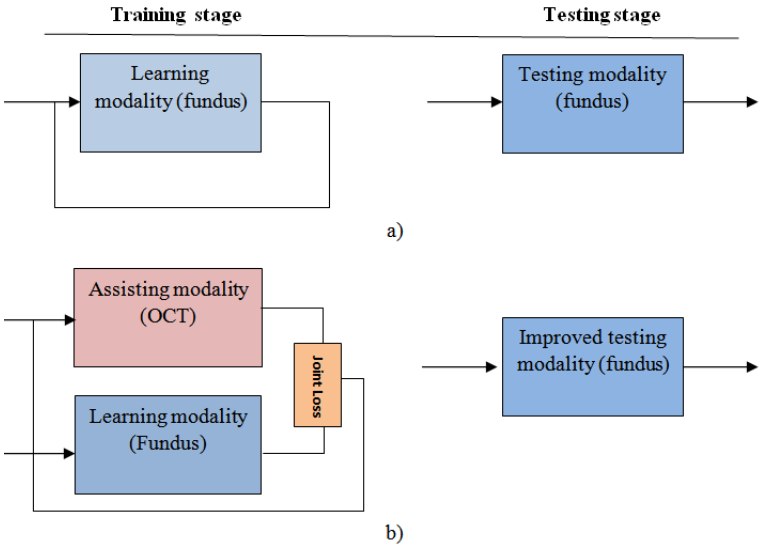


Figure 1.8: a) Existing screening system, b) Proposed idea of improved screening systems

1.4 KEY CONTRIBUTIONS

The major contributions in developing detection algorithms for identifying and understanding Glaucoma from various modalities described in this thesis are:

1. A robust, computationally efficient multimodal registration algorithm for registering fundus and OCT image pairs.

2. A multimodal deep neural network framework for improving the detection of Glaucoma from fundus images.

3. Understanding the progression of Glaucoma through clinical case studies and discovering the scope of new imaging modality angiography OCT for detection of Glaucoma.

1.5 ORGANIZATION OF THE THESIS

The organization of this thesis is as follows. In chapter 2, a robust and computation efficient multimodal registration algorithm is proposed. The algorithm tries to find landmarks and register images in an iterative manner in low Field of View (FOV) images. In chapter 3, a multimodal deep learning framework is proposed in order improve the Glaucoma detection rate on fundus from the model learnt from OCT and Fundus. The training and loss function proposed aims to combine the adversarial loss in learning joint information between fundus and OCT samples and retain the distinctiveness of the two modalities. The description of the proposed architecture is explained in detail in chapter 3. Chapter 4 discusses clinical case studies about new modality in understanding the progression of Glaucoma and further investigates the scope of new modality in detection. Finally, in the last chapter a general discussion, conclusion and the possible future work is described.

Chapter 2

MULTIMODAL REGISTRATION OF RETINAL IMAGES

Registration of multimodal retinal images such as fundus and Optical Coherence Tomography (OCT) images is important as the two structural imaging modalities provide complementary views of the retina. This enables a more accurate assessment of the health of the retina. However, registration is a challenging task because fundus image (2D) is obtained via optical projection whereas the OCT image (3D) is derived via optical coherence and is very noisy. Furthermore, the field of view of imaging possible in the two modalities is very different resulting in low overlap (5-20%) between the obtained images. Existing methods for this task rely on either key-point (junction/corner) detection or accurate segmentation of vessels which is difficult due to noise. We propose a *registration* algorithm for finding efficient landmarks under noisy conditions. The method requires neither accurate structure segmentation nor key-point detection. The Modality Independent Neighborhood Descriptor (MIND) features are used to represent landmarks to achieve insensitivity to noise, contrast. Similarity transformation is used to register images. Evaluation of the proposed method on 142 fundus-OCT pairs results in an RMSE of 2.61 pixels. The proposed method outperforms the existing algorithm in terms of robustness, accuracy, and computational efficiency.

2.1 INTRODUCTION

Fundus photography is based on projective imaging and has been widely used for several years to detect retinal abnormalities occurring in various diseases such as Diabetic Retinopathy (DR), Age related Macular Degeneration (AMD), etc. Optical Coherence Tomography (OCT) on the other hand is a newer modality which enables cross sectional imaging and is used to detect and localize pathologic changes within the retinal layers. A recent study[35] has shown that using fundus images to diagnose diabetic macular edema (DME), a serious condition requiring intervention, led to both false positive and negative results whereas OCT images provided the true condition. The study concluded that fundus images alone may not be a sufficient for retinal screening. OCT also provides unambiguous evidence for AMD which is marked by the appearance of Drusen in between the RPE and ISOS layers of OCT whereas they may or may not be visible in fundus image. Diagnosis of Glaucoma using the optic cup

to disk ratio is also clearly detected from OCT than fundus image. In the proliferative stages of DR and Retinopathy of Prematurity (ROP) where vessels are affected, fundus images provide clearer evidences as against the OCT. Thus, integrating information by registration of OCT-3D and fundus is of interest.

Multimodal retinal image registration is a challenging task[1][7] due to the many factors. 2D (color) Fundus images can be acquired with a wide field of view (FOV) with good clarity for vessels. 3D (greyscale) OCT in contrast can capture a much smaller FOV, is quite noisy due to the coherence phenomenon that underlies image capture and has poor clarity for vessels. Thus, extraction of reliable landmarks for registration from both modalities is difficult. Yet, existing methods rely on establishing correspondence between corners or junctions in the 2 images.

We propose an iterative solution for registering fundus images to an OCT volume. The key contributions of the proposed solution are:

- No requirement for accurate structure (vessel/OD/macula) segmentation or landmark such as corner detection.
- An adaptive and efficient landmark detection method which is capable of handling low FOV images.
- Feature-based registration which overcomes the noisy nature of OCT data and aids handling cross modal data.

2.2 RELATED WORK

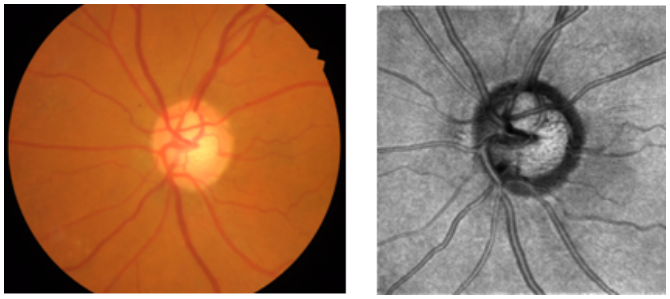
Two approaches are popular in registration literature: area based and feature based. Area based methods operate directly on image intensity values of a specific region or entire image without deriving structural information. This approach follows a typical pipeline of choosing a similarity measure and maximizing it with an optimization method to obtain the desired transformation model. Feature based methods rely on few salient points/features which are prominent in both the images and solving the correspondence between them.

Registration methods have been proposed for both OCT-fundus image pairs and Fluroscene fundus angiography-fundus (FFA) image pairs. FFA provide high contrast view of blood vessels due to the fluroscent dye used in imaging. Both blood vessel ridges [23] and vasculature network [12] [28] have been used as features for registration of OCT-fundus image pairs. Similarity metrics based on pixel distance [23] or similarity between vessels [12] and local similarity [28] have been proposed. A coarse registration by brute force search [23][12] is followed by refinement using ICP algorithm[23] or higher order transform modelling [12][28].

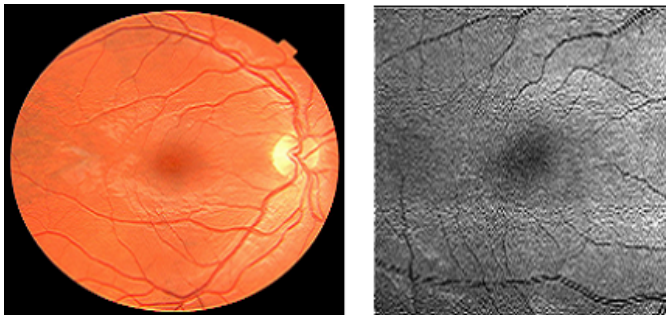
Reported methods [11][7][1] for FFA and fundus registration are mostly feature based. Popular landmarks are Harris corners [11] [7] or junctions [1]. Features include SIFT [11], orientation information [7] and a series of step patterns [1] around each obtained landmark. The extracted features are

matched using bilateral matching [11][7] or nearest neighbour technique [1] and the final transformation is estimated using an affine model [11][7][1].

The limitation of methods in [23][12][28] comes from their dependence on vessel structure as a key feature. While this is relatively easy to extract in fundus images, it is not so in OCT due to the noisy nature of OCT. While Harris[11] [7] and junction [1] detectors perform well on FFA and fundus images, they are unsuitable for the task at hand given the noisy OCT images. Our framework tries to solve these problems by finding landmarks without the need of accurate structure localization. MIND features are chosen so as to nullify the problem of noisy nature of OCT and complementary information provided by both modalities.



(a) Image taken from LD dataset



(b) Image taken from PD1 dataset

Figure 2.1: Sample images of both macula and OD centric.

2.3 METHODOLOGY

The proposed method consists of the following modules: pre-processing and registration module. Pre-processing includes generating 2D OCT en-face images from 3D OCT volumes and illumination correction for OCT and fundus images. The registration module finds landmarks and registers images in an iterative manner. The MIND descriptor is extracted around each landmark and features are matched using Nearest neighbour technique. The obtained matches are finaly used to derive the transformation parameters.

2.3.1 Preprocessing of images

2.3.1.1 OCT-enface creation:

The 3D OCT volume is converted to a 2D OCT enface (see fig:2.2). The variance of the voxel values along columns (Y-axis) of all slices is computed for this purpose. The variance measure is chosen for better visualization of vessels given that there is increased variability of vessels and background across the Y-axis. The final obtained 2D OCT enface is used for registration.

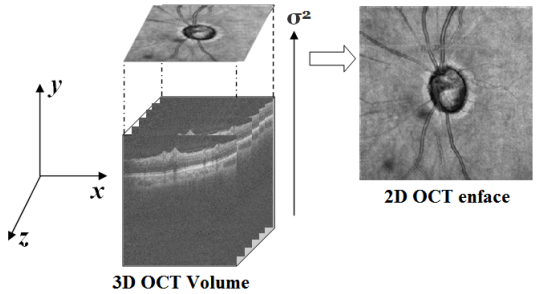


Figure 2.2: 2D OCT en-face creation

2.3.1.2 Illumination correction:

Images suffer from non-uniform illumination and hence this is corrected using the luminosity and contrast normalization method proposed in [18]. Here, the image is modeled as:

$$I(x,y) = T(x,y) * L(x,y) + C(x,y) \tag{2.1}$$

where $T(x,y)$ is the true image, $I(x,y)$ is the observed image, $L(x,y)$ and $C(x,y)$ are multiplicative and additive components. The recovery of the true image is based on estimation of the $L(x,y)$ and $C(x,y)$ components. Hence, the correction of observed image $I(x,y)$ can be obtained by:

$$T(x,y) = \frac{I(x,y) - L(x,y)}{C(x,y)} \tag{2.2}$$

2.3.2 Registration Algorithm

The schematic of proposed registration method is shown in the fig:2.3. This is an iterative approach wherein the registration error is minimized with each iteration. Initial registration is performed with roughly initialized landmarks which are updated in each iteration. Unlike existing methods, the landmarks do not require any accurate structure localization/segmentation. The details of method is pre-

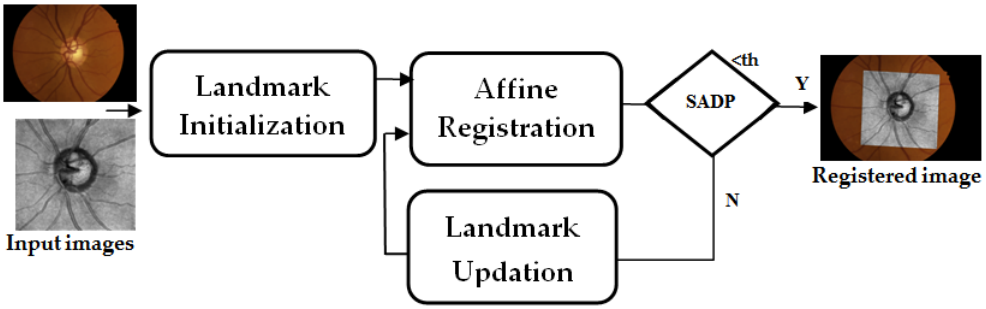


Figure 2.3: Flow diagram for proposed registration framework

sented next. In our work, the enface OCT is used as floating image and registered to the reference fundus image.

2.3.2.1 Landmark Initialization:

Rough initializations of both OD and macula are shown in fig:2.4. A circular Region of interest (cROI-red) is detected and a circle slightly larger than this cROI-$green$ is initialized. The points of intersection of this circle with vessels are considered as initial landmarks for registration. cROI can be the optic disc (OD) or macula depending on the imaging protocol. In order to find cROI from a fundus image, the OD (macula) boundary is detected by finding points of local maxima (minima) of intensity values whereas for an OCT en-face image, OD (macula) boundary is detected by finding points of local minima (minima) of the intensity values. These points are clustered and a best-fit circle is found for the boundary of largest central object. Let C_f (C_o) of radius R_f (R_o) be the circle of fundus(oct). A rough vessel map $I_{vf}(I_{vo})$, is obtained by thresholding the ROI images and thinning it to 1 pixel. The points of intersection of the $I_{vf}(I_{vo})$ with the circle $C_f(C_o)$ are identified as rough landmarks $P_f(P_o)$.

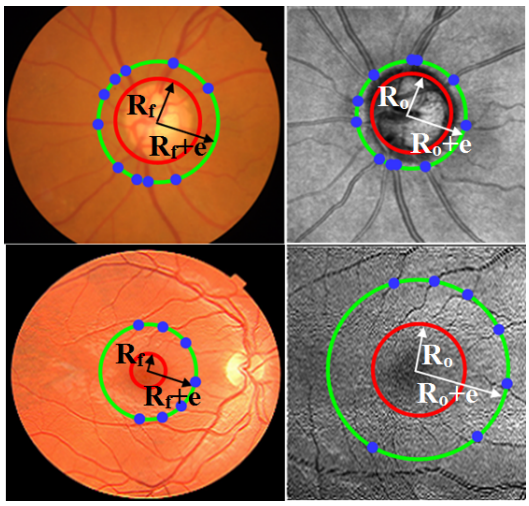


Figure 2.4: Rough landmark initializations (marked in blue)

2.3.2.2 Feature based Affine registration

Feature extraction:

The MIND descriptor[16] is chosen to represent landmarks. MIND is based on the idea of image self-similarity. The assumption is that a local image can be approximated by considering the similarity between all small patches in the image. This approximation aims to extract the distinctive structure in a local neighborhood, which is preserved across modalities. MIND is computed as follows:

$$M(I, x, r) = \frac{1}{n} exp(-\frac{D_p(I, x, x+r)}{V(x, r)}) \tag{2.3}$$

where D_p is the distance measure between two voxels x and $x+r$, r defines the spatial search region, V is variance estimate and n is normalization constant.

A patch of size $W \times W$ is extracted around each landmark from the enhanced image $I_f(I_o)$. Rotational invariance is achieved by rotating the patch to dominant direction found from Histogram of Orientation Gradients(HoG) prior to the the descriptor computation. In our experiments, 12 bins were chosen for HoG computation and for the descriptor was computed over a patch of size $W = 12$. The resultant descriptor was vectorized (144 dimensional) to form a feature vector $M_f(M_o)$.

Feature matching and Outliers rejection:

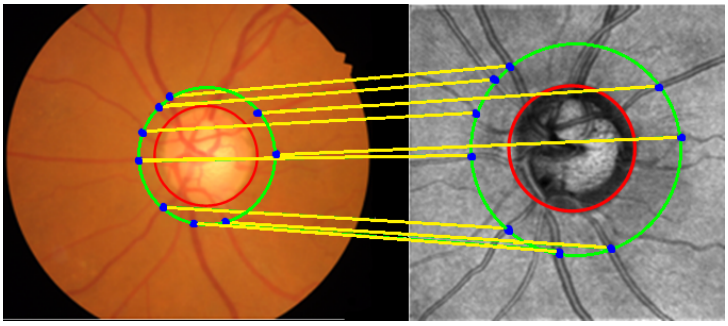


Figure 2.5: Point correspondence after outliers rejection

Given two sets of features M_f, M_o derived from the fundus and OCT images K-nearest neighbor search using normalized Euclidean distance metric was used to match them. Here K is chosen to be 3. So for each feature in set M_f, we obtain $K = 3$ neighbors in set M_o. The normalized Euclidean distance metric between i^{th} feature in set M_f and j^{th} feature in set M_o is given as

$$dist(M_{if}, M_{jo}) = \|M_{if} - M_{jo}\|_2^2 \tag{2.4}$$

The first set of outliers are rejected from the obtained neighbors by matching with dominant orientation of the landmark patch. Random sample consensus (RANSAC) [8] with similarity transformation is applied to all remaining matched pairs (see fig:2.5).

Transformation estimation

The number of matched point pairs obtained from RANSAC is constrained by the number of detectable vessels and FOV of the floating image which is smaller than that of the fundus image. Given the anatomy of the retina, on average, 8-10 matches are obtained for each image. In our problem the retina is a curved surface that can be approximated with higher order model and OCT is a flat projection with very less FOV. Thus, the transformation between two modalities can be assumed to be locally rigid with possible deformations as scaling, translation, rotation. Hence, a similarity model is used for the Transformation function.

2.3.2.3 Landmarks updation

If the registration error from the initialized landmarks is greater than threshold(th) then landmarks are to be updated. This is done in two steps. One is at global level and other is at local level.

Global level update:

At the global level, landmarks are updated by appropriate scaling of R_f and R_o. Let their ratio be Q_R.

$$Q_R = R_f / R_o \tag{2.5}$$

Since the OCT image is taken to be the floating image, the aim is to update R_o such that landmarks extracted from the circle with the updated radius R_{uo} results in a registration error less than th. The registration error is taken to be the sum of absolute distance between points (SADP). SADP is calculated from two thinned vessel sets $I_{vf}(I_{vo})$ by finding matched point pairs in I_{vo} for every point in I_{vf} using Nearest Neighbor search algorithm.

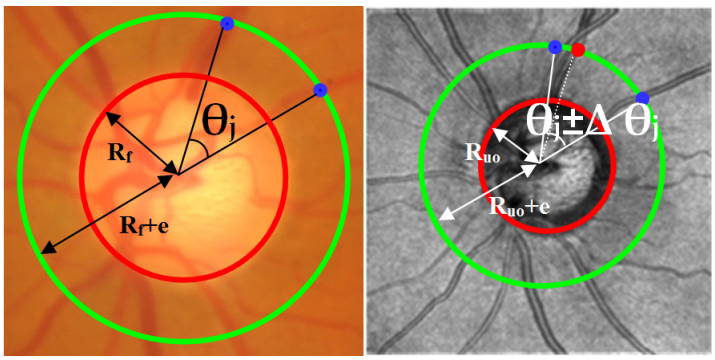


Figure 2.6: Local refinement after updating radii

Local level update:

For every global refinement step, the intersection points P_o are locally refined in the angular direction.

This local refinement is important because skeletonization of vesse]ls can lead to perturbation of landmarks (see fig:2.6). The landmark points are updated by making use of the principle that given two circles C_f and C_o for a fixed angle, the ratio of their radii (Q_R) to the corresponding arc lengths (Q_A) are equal where the arcs are defined to be between every pair of successive points in P_f, P_o subtending angles θ_j, $j = 0, 1, 2..J$. The updation rules for global and local refinement are shown below in fig:2.7. After updating θ_j, the intersection points are also updated with appropriate angular shifts. This

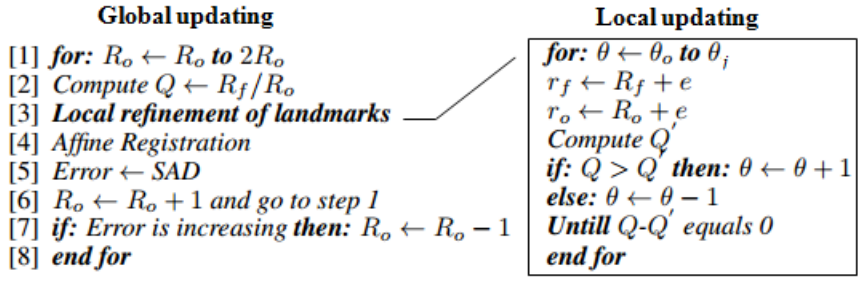


Figure 2.7: Global and Local updation rules

is done until the criterion of $Q_R - Q_A = 0$ is satisfied. The θ_j updated subject to it being in the range $min(\frac{\Delta\theta_j}{2}, \theta_{max})$ where θ_{max} is the upper bound for radial shift of landmarks. This is done to ensure that landmark update is bounded.

2.4 DATASETS

Three datasets were considered for evaluation of proposed registration algorithm.
1) **PD1[12]**: a public dataset, consisting of 22 pairs (Fundus-OCT) of which 17 pairs are of macular region and 5 pairs are of OD region with a variety of retinal diseases. Each OCT volume is of size 650 $\times$ 512 $\times$ 128 voxels and the fundus image is of size 1200 $\times$ 1143 pixels.
2) **PD2[25]**: a public dataset, consisting of 100 pairs (Fundus-OCT) of macular region, with no abnormalities. Each OCT volume is of size 650 $\times$ 512 $\times$ 128 voxels and the fundus image is of size 1200 $\times$ 1143 pixels.
3)**LD:** a locally sourced dataset, consisting of 42 pairs (Fundus and OCT) of OD region obtained. 20 pairs are of glaucoma cases and remaining are normal. Each OCT volume is of size 650 $\times$ 304 $\times$ 304 and the fundus image is of size 4288 $\times$ 2848 pixels.

2.5 EXPERIMENTS AND RESULTS

The MIND descriptor was computed with the following parameters $\{\sigma = 2, N = 6\}$ where σ is the variance estimate which controls the response of feature and N is the neighborhood of spatial search. For evaluating registration in every iteration the threshold th was chosen to be 5 pixels. The number of

iterations depends on the error (SADP). If the SADP converges (less than 5 pixels) the least registration error (SADP) among next 5 iterations is considered.

Code was written in MATLAB 2016a and executed on a 2.5GHz Intel Core i5 desktop with 8GB memory. The average execution time for an image pair was 10.12 seconds, of which 5.8 seconds was for en-face creation and the remaining time was for registration.

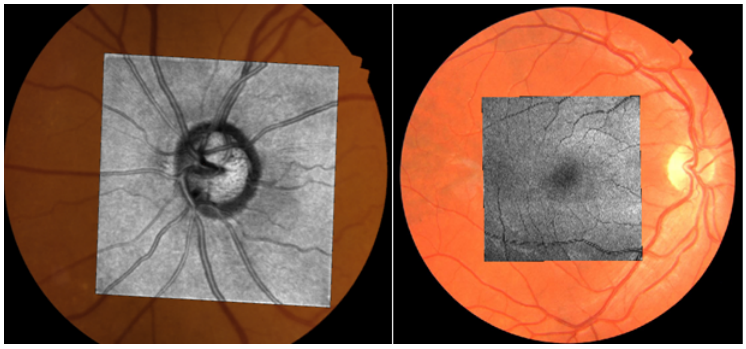


Figure 2.8: Sample registration results

Accuracy of registration is evaluated by Root mean square error (RMSE), Mean median error (MME[7]), Mean maximum error (MAE[7]) and Sum of Absolute Distance between points (SADP[12]). Ground truth control points were manually marked mostly at reliable positions such as junctions of thin and thick vessels for measuring RMSE, MME, MAE. SADP was calculated from two sets of thinned vessel maps $I_{vf}(I_{vo})$ by finding the matched point pairs using Nearest Neighbour search Algorithm.

$$SADP = \sum_{i=1}^{n} |d(x_i, y_i)| / N \tag{2.6}$$

where x and y denote vessel pixels extracted from fundus image and OCT enface images, respectively; d is the distance between ith observation of x and y_i which is the corresponding closest point in y and N is the length of vector x. Table 2.1 shows the average SAD, RMSE, MME, MAE over all datasets. It is to be noted that for a total image set of 142 pairs 90.22% pairs are successfully registered with a average RMSE error of 2.61 pixels. The performance metrics of other methods are given below table:2.2. Method[1] implementation is done and results are reported on LD dataset and method [12] results are mentioned in paper. SDP metric is only mentioned in the paper. In comparison with the our results it is observed that proposed algorithm performs exceptionally well considering the amount of data that is tested.

Next we report results of test of robustness of the proposed method for rotation invariance and scale sensitivity. This was tested by applying rotation and scaling to the floating images.

Scale change Test:

Floating images were up-sampled by a factor ranging from 1-10 and the results were registered to the fixed image. A registration error of $\leq$5 pixels is considered as perfect registration (or 100%). The

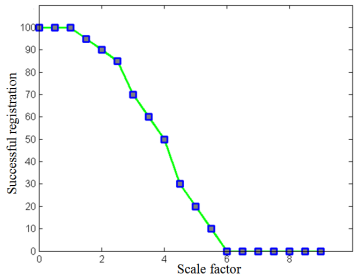


Figure 2.9: performance with scale change

success rate is proportionately reduced by 10% for an increase in error by 5 pixels. The plot in fig:2.9 indicates that the proposed registration is robust to a scale change of up to 1.6.

Rotation invariance Test:

Floating images were rotated by 0^o-180^o degrees in steps of 30^o. The registration was found to be robust for all angles in the tested range. This is to be expected as the patches used for feature computation are pre-rotated to dominant orientation. This makes the features invariant to rotation.

Table 2.1: Registration performance on datasets

Datasets	Success rate %	SDP	RMSE	MME	MAE
PD1	86.66	3.92	3.12	3.83	6.17
PD2	96	4.21	3.70	3.93	5.39
LD	90	1.51	1.01	2.22	2.46
Average	90.22	3.21	2.61	3.32	4.67

Table 2.2: Registration performance of existing methods

Method	Dataset	SDP	RMSE	MEE	MAE
[12]	PD1	6.01$\pm$1.82 (stage-1)	-	-	-
[1]	LD	4.8	3.19	4.32	7.33

2.6 CONCLUSION

Multimodal registration is a challenging task as the images are varied by FOV, noise, contrast etc. These differences affect the registration in terms of finding landmarks and extracting features. These problems are overcome in our proposed registration method by minimizing the registration error by simultaneously refining the landmark positions in an iterative framework. Unlike earlier methods, the proposed method does not require precise ROI/vessel segmentation which is usually challenging. The

choice of MIND descriptor ensures robustness as the extracted features are insensitive to modality, noise and contrast across modalities. The proposed algorithm was found to achieve high registration accuracy and outperform existing algorithms when evaluated over three datasets i.e. 142 image pairs.

Chapter 3

MULTIMODAL LEARNING OF RETINAL IMAGES

Images of fundus are widely used in screening for eye diseases due to their fast acquisition and cost-effective nature, compared to Optical Coherence Tomography (OCT). However, OCT imaging has been found to be more reliable in disease identification as compared to fundus images[35]. The complementary nature of the two modalities, cross sectional imaging (OCT) and optical projection imaging (fundus), provide nearly orthogonal sources of information. Harnessing features from both these modalities can thus help in designing more reliable screening systems for medical diagnosis. However the high setup cost associated with OCT is a deterrent for its inclusion in CAD-based screening systems. Hence, an alternative is to consider OCT data along with fundus images for training but assume only fundus images to be available during screening (i.e. testing/evaluation). The design of a system trained on the multiple modalities but with missing modalities during evaluation time is an active area of research. This problem is often framed as a missing data problem and several deep generative architectures have been proposed as a possible solution. These assume availability of tuples of data which are difficult in medical imaging. Hence the application of such generative models is limited. To resolve this we propose an approach to OCT assisted fundus detection systems. The proposed method improves the performance of Retinal Nerve Fiber Layer Defect (RNFLD) detection rate on fundus with the help of OCT. The main contribution in this work is a novel loss function which combines the adversarial loss in learning joint information between fundus and OCT samples and the distinctiveness of the two modalities.

3.1 INTRODUCTION

Fundus imaging has been widely used to detect retinal abnormalities in diseases such as Diabetic Retinopathy (DR), Diabetic Macular Edema (DME), etc. OCT on the other hand detect and localize pathological changes within the retinal layers as there is depth information associated with every point on retina. Recent studies[35][30] have reported high sensitivities for OCT imaging compared to fundus in diagnosing DME, Glaucoma and retinal layer irregularities. These studies suggested that fundus imaging alone may not be a sufficient for retinal screening studies. On the contrary, in the proliferative stages of DR and Retinopathy of Prematurity (ROP) where vessels are affected, fundus images provide

clearer evidences as against the OCT. This calls for the requirement of multiple modalities for improving medical diagnosis of patient. Thus, integrating information of OCT and fundus is of interest. In this chapter we propose an approach to OCT assisted fundus detection system. In our proposed method we aim to improve the detection of RNFLD in a given fundus image which is one of the key and early symptoms of Glaucoma. The presence of Glaucoma leads to structural changes of optic disc (OD)

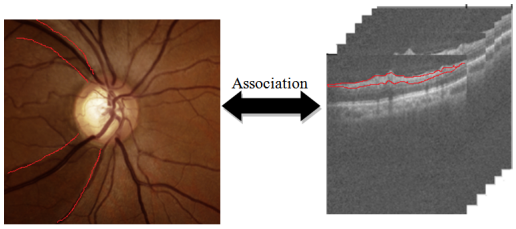


Figure 3.1: Fundus OCT association

and damage to Retinal Nerve Fiber Layer (RNFL). In a fundus image this loss appears wedge shaped gradient radiating from OD. RNFLD is a challenging task because the difference between background and foreground intensities is very subtle. In an OCT, RNFLD is usually observed as a thinning of RNFL (first layer among 7 layers) in a given OCT slice. Thus the thinning of RNFL layer in OCT corresponds to contrast difference in fundus image (see fig:3.1).

3.2 RELATED WORK

Various methods are proposed for RNFLD detection from fundus images. A typical pipeline for RNFLD detection on fundus image consists of rough Optic Disc (OD) boundary detection followed by feature extraction around OD and classification. In [22] rough OD boundary is obtained by morphological processing and features such as intensity[22][20], Local Binary Patterns (LBP)[20][31], Entropy[31], and Directional Differential Energy (DDE)[31] are extracted at different radii around OD. The features that are computed are used either for classification[20][22] or for obtaining finer boundary around OD by passing features to a Recurrent Neural Network (RNN)[31]. The obtained boundary is further analyzed for classification of the region into Normal/RNFLD patch. The limitation of above mentioned methods are their dependence on OD detection and their inability to capture the radial/angular degradation of RNFL within a given patch. The performance reported by above mentioned methods are on very limited dataset which limits the scalability of the algorithms to screening systems.

Recently OCT has shown its potential in accurate detection of abnormalities. [13] showed an attempt for glaucoma assessment in screening systems considering RNFL thickness measurement from OCT. The obtained results are promising compared with fundus. However using fundus for screening systems is prefered over OCT as fundus is cost effective. In [29] fused information obtained from fundus image and RNFL thickness map from OCT is used for predicting RNFL thickness from fundus image patch. This is done by estimating the correlation coefficient using regression models. With the

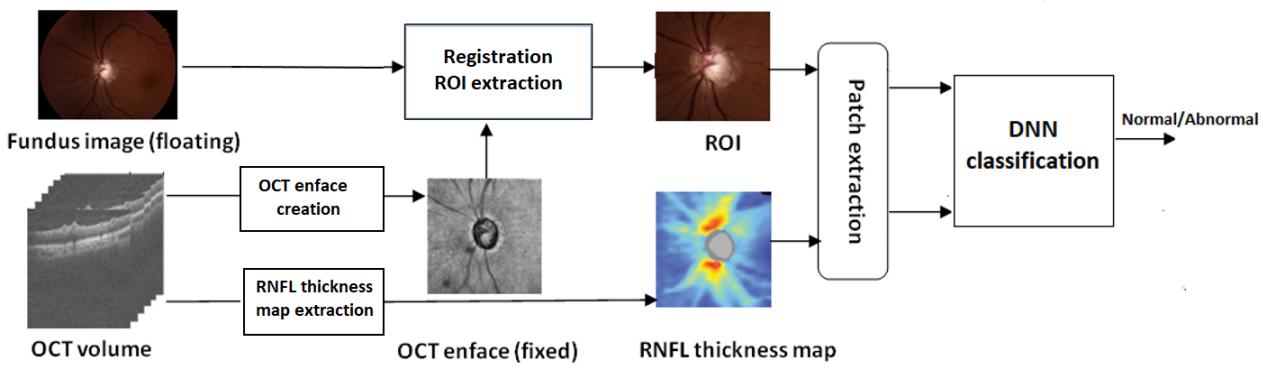


Figure 3.2: Block diagram for the proposed method

trained regression model there is no requirement of OCT for generating RNFL thickness map. We consider the case of detection of RNFLD as a classification problem. We assume the availability of paired OCT and fundus images during modeling (training) while the absence of OCT during validation stage. This is considered challenging because missing of one modality at validation stage has been shown to significant drop in performance [15].

There are several methods proposed to learn joint information for the missing modalities problem. Missing modality is well defined problem is deep learning based systems. [33] used Deep Boltzmann Machine for learning a generative model to generate missing modalities whereas in [15] missing modalities are handled by mean and variance filling to avoid sudden drop in performance of system. The main drawbacks of above methods is requirement of enormous amount of paired data for generating the missing modalities and the certain drop of performance [15] claims that there will always be a drop of performance which cant be avoided though missing modalities are replaced with mean/variance filling.

Recently few Domain Adaptation (DA) based methods have been proposed. A DA based domain independent representation in [34] which uses adversarial loss for adapting source to target samples. [10] [32] proposed an architecture to simultaneously learn a domain-invariant representation and a task-related network in a single Deep Learning (DL) architecture. With this as motivation we propose a deep multimodal architecture for learning joint representation from fundus image and OCT. In the proposed method we learn representation from fundus image and OCT by reducing the variance between the input modalities using adversarial loss function along with retaining modality specific input embeddings.

3.3 METHODOLOGY

The proposed method consists of following stages: (See fig:3.2). Preprocessing of OCT, registration and patch extraction followed by classification using a Deep Convolutional Neural Network (DCNN). Pre-processing includes generating 2D projection of OCT (enface) and extracting RNFL thickness map from 3D OCT volume. 2D projection of OCT (enface) is registered to fundus for image-image learning.

The patches are extracted from the registered images and given as input to deep CNN for classification into Normal / RNFL Defect (RNFLD) patch.

3.3.1 Preprocessing OCT volume

OCT-enface creation

Processing 3D volumes in computationally intensive. Hence volume is converted into 2D enface representation. This is done by computing variance of all voxel values along columns(Y-axis) in all slices. It is observed that within a slice across Y-axis there is maximum variability of vessels and background. Earlier approaches computed sum or mean for enface generation. We chose variance to provide high vessel definition. The definition of vessels is important for accurate registration of fundus image and 2D OCT enface.

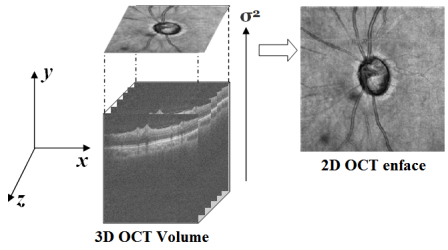


Figure 3.3: 2D OCT enface creation

RNFL thickness map creation:

RNFLD is caused by thinning of the RNFL layer and this layer is of interest for RNFLD hence a 2D RNFL thickness map is generated from 3D OCT volume. This is computed from slices of 3D OCT volume by analyzing the intensity profiles along each column of the OCT slice. The width of the RNFL layer in every column is projected on 2D image as thickness value which is considered as RNFL thickness map. The complete description of the method is described in [13].

Registration and patch extraction

Registration

Our previously developed multimodal retinal registration algorithm [17] was used for registration of fundus and OCT enface images. The registration step is to enable the DCNN model to learn the correspondence between modalities during training. Registration is done by choosing fundus as the floating image and OCT en-face as the fixed image. The estimated transformation parameters are applied to floating image (fundus) and a Region of Interest (ROI) bounding the fixed image(OCT) is extracted.

patch extraction

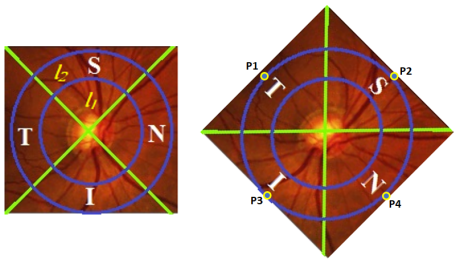


Figure 3.4: Left: shows ROI in the given fundus image along with the ISNT sectors and labels derived from patch l_1, l_2 are overlaid. Right: Result of rotation applied to the image on the left.

Patch type	**Label** (l_1,l_2)
Abnormal	0 0
Abnormal	0 1
Normal	1 1
Normal	1 0

Table 3.1: Ground truth labels (per patch)

Our method works at a patch level. Patches of fixed size 152×152 are extracted from registered pairs of OCT thickness map and fundus images. Clinically RNFLD is reported by dividing the region around OD into sectors S-Superior, I-Inferior, N-Nasal, T-Temporal (see fig:3.4,left). RNFLD typically [27] occurs in either Superior(S)/Inferior(I) or both regions. A square patch is extracted using vertices P1-P4 as in (see fig:3.4,right). This is subdivided into 4 patches according to I,S,N,T sectors. Ground truth markings for each patch are assigned under supervision of expert considering the spatial information i.e. by observing RNFL trend within the patch.

Typical RNFL patterns observed in a given patch:

- RNFLD appears only in S,I regions. In remaining sectors it is absent[27]
- In RNFLD, the loss typically starts from OD and progresses towards the periphery
- In normal cases, RNFL thickness decreases with distance starting from OD[26]

For assigning labels we hypothetically divide the patches into two circular regions as shown in fig:3.6a,right. Let l_1 be label at limited ROI (inner circle) and l_2 be label at extended ROI (outer circle) in a given patch.In normal cases RNFL layer thickness decreases with distance hence we label the patch (1,X) where X∈0,1 where as for RNFLD patch since loss starts from OD to outwards the label assignment for patch is (0,X) where X∈0,1.

3.3.2 Overview of Deep Architecture

Our architecture (see fig:3.5) has three deep networks running parallely 1) Deep convolutional network-fundus (D_f) 2) Deep convolutional network- OCT (D_o) 3) Deep convolutional network-modality (D_m). In the proposed architecture, there are separate identical architectures training on OCT (D_o) and fundus samples(D_f). These architectures aim to classify patches into normal/RNFLD. During training we

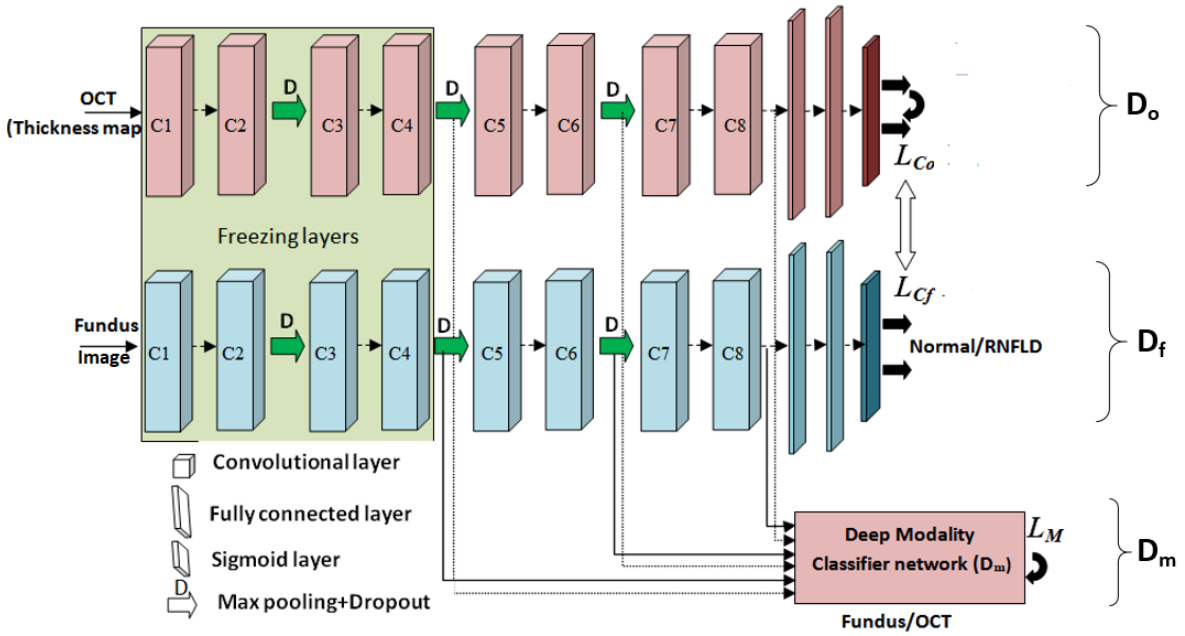


Figure 3.5: Proposed Deep Multimodal Architecture

freeze some part of network parameters and some are shared (joint learning) between two modalities. Joint learning is introduced by adding D_m objective function in the classification objective functions of D_f and D_o networks. D_m architecture aim is to discriminate the input patches of OCT and fundus. Our overall training objective is to learn a model which is robust to missing modalities. The details of each network and how robustness is achieved are explained in later sections.

Deep convolutional network- fundus (OCT) D_f (D_o)

The architecture of D_f (D_o) is deep convolutional neural network (CNN) for classification of fundus (OCT) into Normal/RNFLD patch. In our implementation a variation of VGG architecture model is used for 2D CNN. It has 8 convolutional layers, 2 fully connected layers and 1 sigmoid layer. For training D_f, fundus patches are used along with labels. Similarly for training D_o OCT patches are used along with labels.

Deep Convolutional network- modality (D_m)

The D_m network discriminates the input modality of the sample i.e. fundus/OCT. We train two models D_f and D_o on paired set of samples along with their corresponding labels. Labels in this case are of fundus or OCT. The output activations of convolutional layers of D_f and D_o networks capture inherent representation of modalities. These activations of the convolutional layers (upsampled a_4,a_6,a_8 obtained from output of C_4,C_6,C_8) are given as input to $D_m$network. Our D_m architecture(shown as

box 3.5) has 4 convolutional layers, 1 fully connected layer and 1 sigmoid layer. The complete system is obtained by adding D_m, D_o networks to D_f as shown in 3.5.

3.4 TRAINING DEEP ARCHITECTURE

Offline training

Before Online(actual) training, we pre-train D_f and D_o networks using fundus and OCT samples separately along with ground truth labels. The main idea behind this is to ensure models are well fitted to their respective modalities. This is needed as we are not only trying to improve performance on fundus image but also preserve the modality specific learnings. Let $Lo_{Cf}(\theta)$ (and $Lo_{Co}(\theta)$ be training losses corresponding to D_f and D_o which are to be minimised during offline training. The best fit parameters are obtained by minimising loss functions using Stochastic Gradient Descent (SGD).

Online training

From the obtained pre-trained models we freeze first few convolutional layers of D_f and D_o networks and perform online training over remaining layers with the addition of D_m network as in (see fig:3.5). For training D_m we provide resampled outputs from C_i i.e. a_i, i∈4,6,8 activations as input. We incorporate modality classification loss $L_M(\theta)$ (from D_m network) into RNFLD classification training objective $L_{Cf}(\theta)$, $LCo(\theta)$ (from D_f and D_o networks). The overall training objective to be minimized is given in equation below.

$$L_{Cti}\left(\Theta\right) = L_{Ci}\left(\Theta\right) - L_M\left(\Theta\right) \tag{3.1}$$

During training our aim is to minimize the loss function in eqn: 3.1. These functions aim to minimize RNFLD classification loss $L_{Cf}(L_{Co})$ and increase modality classification loss L_M during process of training. The increase in modality classification loss corresponds to making inputs least discriminative. Since we have freezed initial layers during training process the initial modality specific learnings are preserved. Coming to the training strategy the loss functions L_{Ctf}/L_{Cto} are optimized by alternative training. In alternate training the multiple loss functions are minimized by separate optimisers for each. In our method we performed joint learning by alternately calling each optimiser, which means we are transferring information between modalities by means of discovering the commonality across the modalities.

3.5 DATASETS and IMPLEMENTATION DETAILS

Dataset details

The dataset details are mentioned in table:3.2.

Local dataset (LD):

The OCT-fundus paired dataset is obtained from a Local hospital (Anand eye hospital, Hyderabad). The dataset contains 45 image pairs out of which 25 are Glaucoma with most of them are early, moderate and advanced with focal wedge shaped RNFL loss and the remaining 20 are normal cases. Each image set contains fundus and 3D OCT volume where Fundus image is of size 3000 $\times$ 2000 and OCT volume is of size 304$\times$640$\times$304.

Private dataset (PD): The dataset is obtained from authors of [20] consists of 25 fundus images (for RNFL defect detection). It is mentioned that sector-wise decision on the presence of RNFL defect was collected from a glaucoma expert.

Table 3.2: Datasets used in evaluating proposed method

Dataset	Type	Normal	Glaucoma	Total
LD	Paired	20	25	45
PD	Fundus	0	25	25

Parameter selection and Implementation details of the proposed method

Training-validation-test data is generated from 45 paired samples from LD dataset. Data augmentation is done by applying transformations to patches extracted from training images(30). This includes rotation at angles - (0^0-180^o in steps of 30^0), vertical and horizontal flips, Gaussian noise addition, log transform (Y correction), scaling 2-4. Training-validation-test sets are created with 960-300-312 samples by randomly sampling in 1:1 ratio of normal and abnormal. In our proposed architecture D_f/D_o architecture consists of 8 Conv layers of filter size (5,5) and 2 Fully Connected (FC) layers. The obtained probabilities are converted to class labels by applying sigmoid activation function. The D_m architecture takes input from Conv-4,6,8 layer outputs and it has architecture similar to D_f/D_o succeeding the Conv-4 output. The training losses ($L_{Cti}(\theta)$) are chosen to be binary cross entropy and is minimized using SGD with learning rate chosen 0.0001 for finding network parameters. Before training the network, the patches are normalized between 0 and 1. We used Lasagne library with Theano backend for our implementation and system is trained using an Nvidia TitanX GPU.

3.6 RESULTS

The effectiveness of proposed method is analyzed using the following evaluation metrics: Sensitivity, Specificity, Receiver Operating characteristics (ROC), Area Under Curve (AUC).

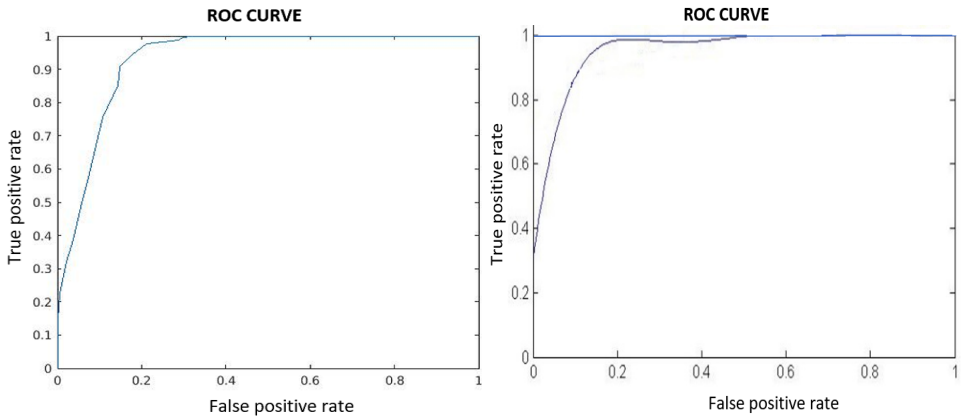
Patch level predictions

A test set of 312 patches from LD dataset are used for validation. ROC is plotted (see fig:3.6a) by varying the threshold on class probabilities obtained before applying sigmoid activation. From the obtained ROC we can infer a high **sensitivity of 91.99%** is achieved at a **specificity of 90.83%** with **AUC 0.92**. We compare our method by performing validation on the dataset obtained from [20]. The specificity is reported for the sensitivity mentioned in [20]. The results are tabulated in table:3.3. The results in [20] are reported only on 10 images as they have used remaining images (15) for training their architecture.

Table 3.3: Patch level performance of proposed method on datasets

Method	Dataset	Data size(Np/Ni)	SN	SP	Acc
Proposed method	LD	312*	0.80	0.96	0.924
Proposed method	PD	375/25	0.80	0.97	0.89
Joshi et al[20]	PD	150/10	0.80	0.93	0.78

*number of patches; N_P denotes the number of patches, N_i denotes number of images.



(a) ROC curve at patch level (b) ROC curve at image level

Figure 3.6: ROC curve for predictions at the (a) patch level and (b) image level

Image level predictions

Image-level RNFLD is analyzed by varying thresholds on the number of abnormal patches detected in a given image (see fig:3.6b). We reported our results and ROC curve on PD(25 images) obtained from

[20]. The results are tabulated in table:3.4. However for comparing image level performance mentioned in [20] we computed Precision and Recall for 25 images as per paper. The precision is 0.92 and recall 0.93. In the paper the reported precision and recall values are found to be 0.91.

Table 3.4: Image level predictions on datasets

Dataset	No of images	SN	SP	Acc
LD	10	0.9398	0.9289	0.9362
PD	25	0.8795	0.8928	0.8868
Average	**NA**	**0.9096**	**0.9108**	**0.9115**

3.7 EXPERIMENTS

Various experiments are performed to understand the effectiveness of the proposed method. The experiments are described in detail in later sections.

Performance improvement on fundus samples

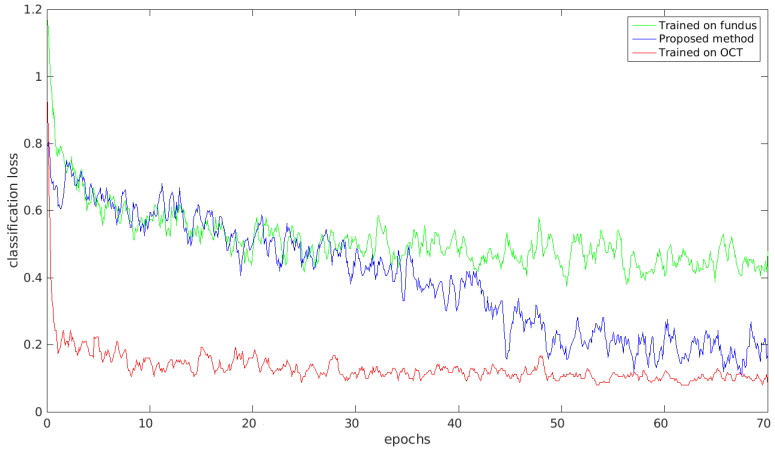


Figure 3.7: Performance improvement on fundus samples with proposed method (Loss vs epochs)

To understand the performance improvement over fundus samples, we plotted validation loss versus epochs (fig:3.7) for three cases viz. 1) (Train & Test) with fundus samples using D_f network 2) (Train & Test) with OCT samples using D_o network 3) Proposed method (as described in section 3.3.2) - test with fundus samples. From fig:3.7 it can be seen the loss is less(blue) in proposed method compared to training on fundus samples alone (green). When compared these curves with the validation curve of OCT(red) the proposed method is observed to perform well.

Multimodal training & testing

Traditional Multimodal learning:
In traditional multimodal learning method [24] two separate modality specific models are learned and the features obtained at FC layers are concatenated. We perform this test on our modalities and results are reported (see table: 3.5). The results are shown to work well if both modalities are present at the test time. When tested only with fundus/OCT the performance is deteriorated.

Proposed method):
We trained the classifier using samples of fundus and OCT for our proposed network. Our method learns model by transferring information across the modalities in a away without losing the modality specific identities. This improves accuracy of fundus samples to 8-10% compared to performance reported on system trained on fundus samples alone (monomodal training). (Mono-modal training of sub-network D_f with fundus samples achieved accuracy is 83.2 $\%$ and training of sub-network D_o using OCT samples achieved accuracy is 95.2$\%$.)

Table 3.5: Patch based results (Accuracy) of Multimodal modal training & testing

Training set/Test set	**Test with OCT**	**Test with fundus**	**Test with OCT and fundus**
Traditional multimodal architecture	0.5671	0.66018	0.9557
Proposed method	0.8941	**0.924**	NA

Choice of input patch sizes

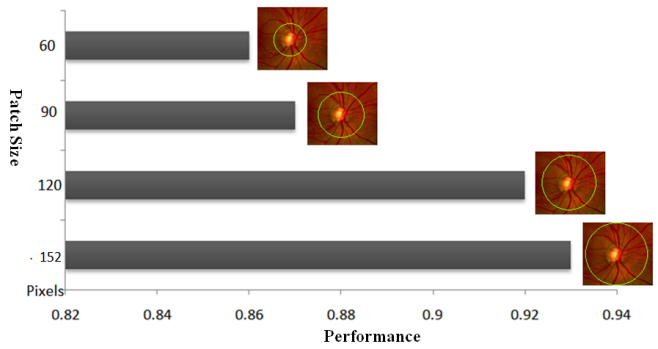
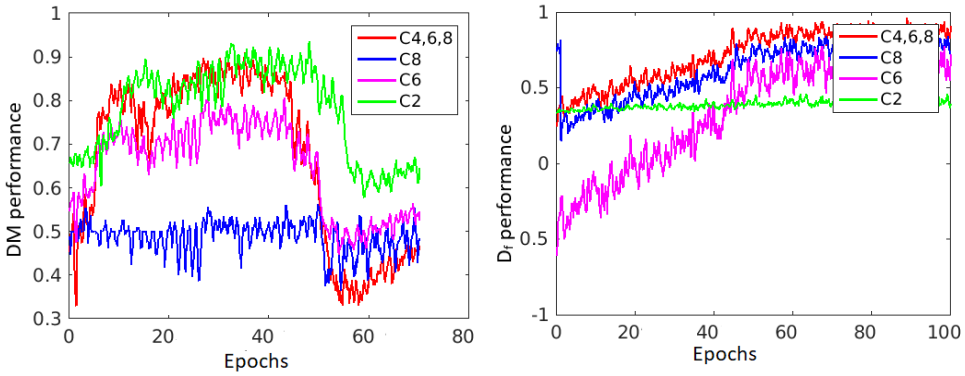


Figure 3.8: patch size vs performance (accuracy) on fundus

Choice of size of input patches plays an important role in capturing information. If the patch is chosen very limited around OD the network fails to capture RNFL pattern. To derive the optimal patch size we trained on different patch sizes i.e. at various radii around OD. Fig:3.8 indicate the performance of our system at various patch sizes. It is observed that there is decrease in classification performance when trained on samples extracted at limited ROI. In these patches context information about RNFL

trend is absent, whereas if we observe in extended ROI's the spatial information is captured well which resulted in high accuracy.

Effect of choice of activations



(a) D_m performance (b) D_f performance

Figure 3.9: Performance (Accuracy) vs epochs (a) D_m performance and (b) D_f performance.

The choice of activations that are input to D_m play a key role in improving the performance of the system. Each activation obtained from the output of convolutional block serves as distinct feature of particular modality. We now investigate how the choice of adapted layers affects the performance of our system. For this we perform our method by changing the inputs to the D_m.

fig:3.9a shows how activations input to D_m effects the overall performance of the system. As we observe from D_m curves, the sudden drop in performance indicates the starting point from where input modality samples are becoming least discriminative. This corresponds to increase in classification performance in D_f (see fig:3.9b).

It is observed that activations sampled from earlier layers (C2) are more discriminative, whereas activations sampled from later layers are less discriminative. This has conversely impacted on classification performance (see fig:3.9b)

3.8 CONCLUSION

Detection of RNFLD from fundus images is a challenging task as pattern is very complex and subtle in nature when compared with background. To address this problem we presented multimodal learning algorithm which allows one modality to be unavailable during validation stage. The proposed approach improved RNFLD classification over fundus than traditional monomodal training. The main novelty of the method lies in learning model between fundus images and OCT without losing the original identity of the fundus images. The adversarial loss introduced in training objective ensure that model is robust to input modality given at validation stage without worrying about missing modality. The proposed

method performs well across datasets on patch and image level which depicts the robustness of obtained model. It is clear from experiments that the proposed method detected RNFLD more accurately than the existing methods considering the size of test set etc. The results have demonstrated that the proposed system can be used in automated detection and grading of RNFLD.

Chapter 4

ANGIO-OCT FOR ASSESSMENT of GlAUCOMA

4.1 INTRODUCTION

The study of progression of disease with respect to particular modality is critical in order to develop efficient detection algorithms. As mentioned in section:1.1.3 A-OCT is a non-invasive flow imaging technique for capturing micro vasculature of retina. It uses Split Spectrum Amplitude De-correlation Algorithm to compute and map the flow across the OCT volume. The typical features that are visualized through A-OCT scan are Capillary Density, Nerve Fiber Layer (NFL) thickness. Several clinical studies [2] [21] reported that loss of capillary vessels, NFL loss during early stages of Glaucoma. In this section the trends of CD (See figure:4.1), NFL thickness in Normal and Glaucoma are explored by studying the following:

1. The possibility of presence of Glaucoma correlating with Capillary density, along with sectors that are affected.

2. Extent to which loss of capillaries are correlated with thinning of RNFL, as Glaucoma is caused by thinning of RNFL.

3. Modelling CD and NFL in normal retina from OD towards periphery and how the model is affected by the presence of Glaucoma.

4.2 CASE STUDIES

4.2.1 Capillary density of abnormal and normal patients

Glaucoma is characterized by loss of NFL as well as capillary density. Evaluation of CD is limited in the past modalities because of difficulty in imaging capillaries, whereas with A-OCT it is made feasible. The purpose of this study is to compare vessel density or Capillary density in normal and Glaucoma subjects using A-OCT scans.

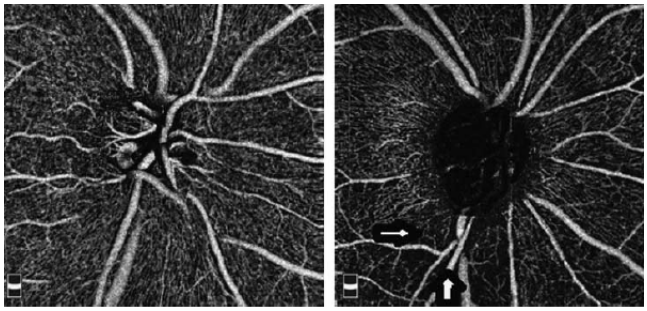


Figure 4.1: Normal capillary network flow (left), Loss of network in infero-temporal sector (right)

Materials and methods

52 healthy patients and 24 early Glaucoma are participated in this study. RNFLT measurements are obtained using scan protocol. Eight RNFL sectors were measured at 3.45mm diameter circle around the optic disc, and defined in clockwise order for the right eye and counterclockwise order for the left eye and designated as shown in figure:4.2 as superior nasal (SN), superior temporal (ST), temporal upper (TU), temporal lower (TL), nasal upper (NU), nasal lower (NL), inferior nasal (IN), and inferior temporal (IT).

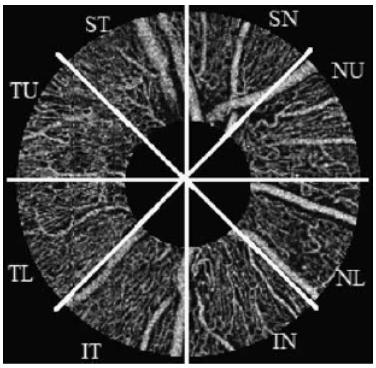


Figure 4.2: Sectors for evaluation of Capillary density (left eye)

Image processing and Analysis

The proposed methodology comprises of 3 stages: (1) extraction of the region of interest (ROI) around ONH, (2) detection and suppression of thick vessels, and (3) estimation of CD.

Stage-1: Extraction of the ROI around ONH:

The angioflow image (A-OCT scan) has the entire vessel network, where only the capillaries are of interest to us. Eight sectors with a sector angle of 45 degrees were selected for quantitative analysis and designated as superior nasal (SN), superior temporal (ST), temporal upper (TU), temporal lower (TL), nasal upper (NU), nasal lower (NL), inferior nasal (IN), and inferior temporal (IT). The annulus-shaped ROI is extracted, with the inner edge marked manually by a single operator as a set of user defined points at the ONH edge and the best-fit circle to these points is found. The outer edge of the annulus

was chosen to be at a fixed value of 3.4-mm circle diameter around the ONH.

Stage 2: Detection and Suppression of Thick Vessels:
Vessels are detected from the Angio flow enface RPC image using the Bar-Selective Combination of Shifted Filter Responses method [3]. The 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 image 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. Multiplying this mask with the RPCs Angio flow image yields the desired image (I_C) figure:4.3 with only the capillary network.

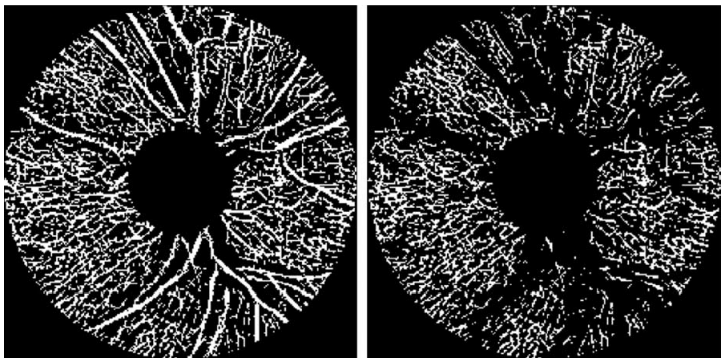


Figure 4.3: Suppression of thick vessels (I_C) (right)

Stage 3: Estimation of CD:
The CD is the ideal measure of capillaries per unit area. I_C is a binary image, where white pixels represent capillaries and black pixels represent nonvessel and thick vessel pixels (as they have been suppressed). The CD was calculated using the following formula:$CD(i) = \frac{N_w(i)}{A(i)}$; where i= 1,2,y,8, is the sector index, $N_w(i)$ the number of white pixels, and A(i) the area of the i_{th} sector of I_C. As both the numerator and the denominator are pixel counts, the CD being reported is a dimensionless value, which lies between 0 and 1. **Observations** The mean CD in Glaucoma patients is lower than that in the normal eyes. There was no significant correlation between CD and corresponding RNFL sector analysis between normal and Glaucoma eyes except for the ST and IT sectors.The ST and IT sectors correlated significantly with RNFL thinning. OCT-A demonstrated reproducible results in patients with early Glaucoma when compared with normal patients. The results of the study suggest that the CD measurements may have a value in the diagnosis and monitoring of glaucoma.

4.2.2 Understanding topography effect by correlating RNFL and CD

The purpose of this study was to analyse the extent of the vessel (capillary) network using wide-field montage images of OCT-A in the normal human eyes. In addition, we measured and correlated capillary density and RNFLT at various distances from the ONH margin in the montage images. **Materials**
In the current study 50 healthy subjects underwent imaging with OCT-A. For montaging scans of Angio

disc (OD centric) and Angio retina (Macula centric) were combined to create a wide-field montage image of the capillary network. Capillary density and RNFLT was calculated at different circle diameters around the ONH, and their correlation was measured. **Image processing and Analysis**

Stage-1: Montage creation:

The enface montage image is obtained using online tool Feature-based retinal image registration by Li Chen[6] which uses vessel bifurcation and crossovers as landmarks, and the transformation type is chosen as rigid. The montage image along with sector markings is shown in figure:4.4.

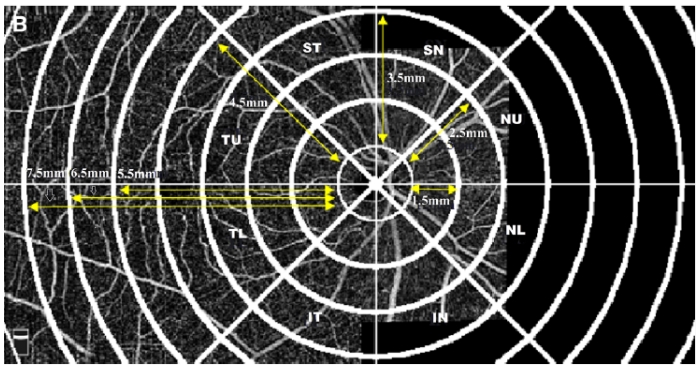


Figure 4.4: Montage image along with sector markings

Stage-2: Measurement of Capillary density:

To measure CD, The peripapillary region was divided into eight equal sectors, designated as superior nasal (SN), superior temporal (ST), temporal upper (TU), temporal lower (TL), nasal upper (NU), nasal lower (NL), inferior nasal (IN) and inferior temporal (IT). RNFLT and RPC density were measured at various distances (1.5 to 8.5 mm) from the ONH margin. An annulus-shaped region of interest was marked by manually selected points at the edge of ONH as the inner circle, and the outer edge of the annulus was taken at 1.5 to 8.5 mm from ONH at different distances. Vessels were detected in the montage image, using Frangi filter method [9] which computes vessels, based on eigenvalues of the Hessian. Vesselness maps were computed at 5 scales with varying sigma (r),r = 3, 5, 7 for extracting the thick vessel and r = 1, 2, 3, 4, 5 for extracting small and thick vessels. The vessel maps obtained was thresholded at 0.001 to get the final capillary map which was used to estimate the capillary density (CD). Capillary image (I_{Cm}) is a binary image where white pixels represent capillaries and black pixels represent non-capillaries. The CD at a distance r was calculated by considering a small sector region between radius r and (r + 1) mm. Here r varies from 1.5 to 8.5 mm from ONH margin (Fig. 2b). The CD was calculated as: CDr(i)= $N_w(i)/A(i)$ where i = 1, 2,,8 is the sector index, $N_w(i)$ denotes the number of white pixels, and A (i) denotes the area of the ith sector region of I_C. CD was normalized by the sector area, which lies between 0 and 1. Colour coding (figure:4.6) was done for every sector region, ranging from dark red to dark blue, indicating maximum and minimum density, respectively, based on the scale bar.

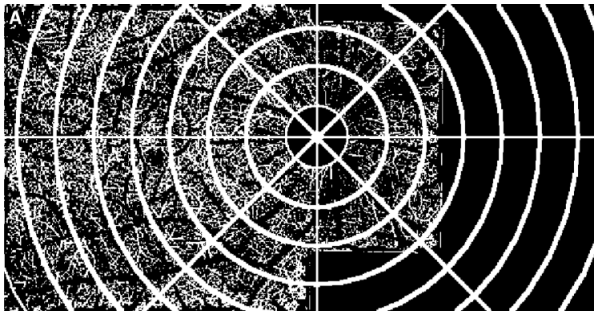


Figure 4.5: Visualization of capillaries after suppression of large vessels I_{Cm}

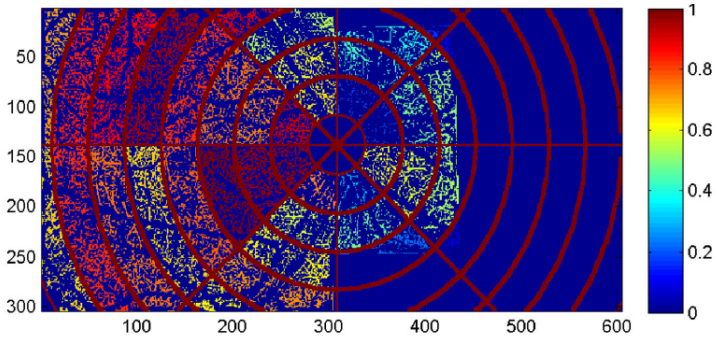


Figure 4.6: Color coded representation of CD variation with distance in normal subjects

Stage-3:Measurement of RNFL thickness:

In the intensity profile, along the columns of OCTA slice, the RNFL boundary was marked by the first abrupt change (from dark to bright intensity) in the ILM layer followed by a large dip in intensity and RNFLT was determined by computing the distance between these two points for the entire volume (see figure:4.7). The RNFL at distance r is calculated by considering sector region between radius r and (r+ 1) mm, where r varies from 1.5 to 8.5 mm from ONH margin. The RNFLT is calculated as: RNFLr(i) = mean(Th(i)) where i = 1,2,,8 is the sector index.

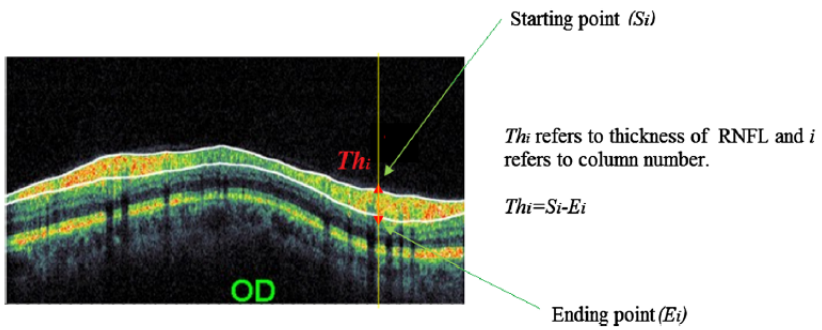


Figure 4.7: Evaluating RNFL from OCT volume

Conclusion

The mean capillary density and RNFLT are highest at 1.5 mm from ONH margin, and there was a trend in its decline, in a distance-dependent manner, with the least density at 8.5 mm . Overall mean CD density correlated significantly with the overall mean RNFLT. The Wide-field montage OCT-A can visualize expansion of the capillary network, which is useful in obtaining information about various retinal disorders. The results obtained support the hypothesis that the vessel network could be responsible for RNFL nourishment.

4.3 EXPERIMENTS

4.3.1 Classification of Glaucoma using CD, NFL as features

The CD and RNFL values that are reported from montage image are considered for classification. Features (CD/RNFL) are reported individually for 8 sectors. For each of the sector distance (d=1.5 to 8.5) varied CD is also noted. This forms 8*8 feature matrix. Two types of evaluations are performed and classificiation results are evaluated by computing Accuracy, Sensitivity(SN), Specificity (SP). The formulae are as mentioned below where TP: True positive, FN: False negative, TN: True negative, FP: False positive.

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.2}$$

$$Accuracy = \frac{TP + TN}{FN + FP + TP + TN} \tag{4.3}$$

Classification considering all sectors:

We performed classification by considering all the eight sectors CD/RNFL values as feature vectors. The 8*8 feature matrix is vectorized to form 64 dimensional vector. The results are reported as shown in table:4.1. We have reported individual performance of RNFL and CD alone.

Table 4.1: Feature selection based performance

Features	Accuracy	SN	SP
CD alone	0.9121	0.8911	0.9356
RNFL alone	0.8269	0.8750	0.8243
CD+RNFL	**0.9213**	**0.9087**	**0.9322**

Classification considering specific sectors:

Through case studies it is shown that Glaucoma is affected only in ST,IT regions. As montage is directed

towards Temporal(Upper/Lower) region we excluded the regions (TL,TU) in performing classification with RNFL and CD as features. The results are reported in Table:4.2.

Table 4.2: Sector wise performance

Region	Accuracy	SN	SP
All sectors	0.9391	0.9195	0.9432
TU+TL	0.9032	0.9199	0.8733
All sector-{TU,TL}	0.9506	0.9471	0.9610

Observations

It is observed that RNFL and CD both contribute in accurate diagnosis of Glaucoma. Coming to sector wise feature selection, as Glaucoma is not affected in TL,TU regions, excluding them in feature selection will result in better results.

Experimental details:

Experimental dataset consists of 50 Normal and 20 Glaucoma patients. K- fold validation with K=4 with 15 abnormal, 15 Normal for training and remaining for testing. Classification is performed by linear SVM classifier[5].

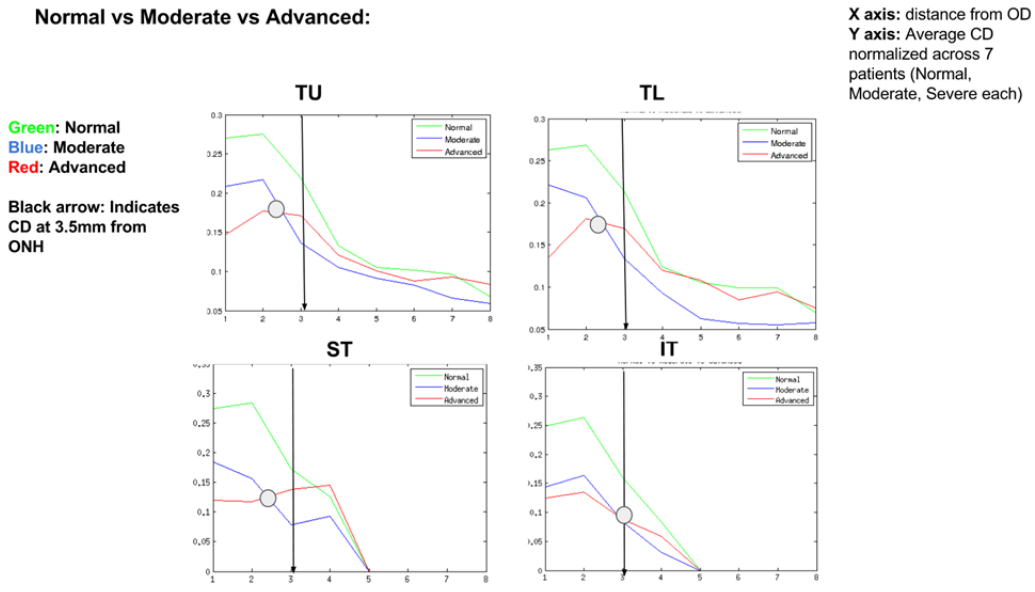


Figure 4.8: Trends of CD in Normal, Early and Advanced Glaucoma

Trends of CD/RNFL in Normal, Early and Advanced Glaucoma

The following are the observations that are inferred by analyzing the CD with distance.

1. Normal and Abnormal (Moderate and Glaucoma) are differentiated in all sectors. We observe a CD drop of 30% for moderate and 60% for advanced.

2. If we consider Moderate vs Advanced, CD of moderate is greater than advanced within 1.5-3mm distance from ONH (with a avg drop of 20,30% in TL and TU sectors and 25-30% drop in ST region and ¡10% drop in IT region). Then after 3mm CD (of abnormal and moderate) started to crossover in between 3-3.5mm distance in TU,TL and ST regions. and after 3.5mm Advanced CD is greater than moderate in all sectors.

4.4 CONCLUSION

We have established two case studies which give better understanding of disease Glaucoma. These findings are evaluated by clinical experts community and further these findings can be employed in developing algorithms. With this as motivation we developed detection algorithms considering Capillary density, nerve fiber layer thickness as features. The results are promising and with reduced setup cost this modality can be considered as reliable biomarker in developing screening systems.

Chapter 5

CONCLUSIONS

This thesis focuses on detection of Glaucoma from multiple modalities. The treatment of Glaucoma is only possible if it is diagnosed in early stages. To ensure timely treatment, screening systems are required for frequent examination of retina. Most of the existing literature for screening systems focus on developing detection algorithms on fundus images. Several studies have shown that the usage of fundus images have resulted in higher false positives as compared with other modalities. Thus combining multiple modalities is required. The use of multiple modalities in a CAD setting improves detection due to availability of multiple sources of information across modalities. However, collecting scans of multiple modalities during the screening stage is an expensive process and normalizing across multiple sources of data is crucial for learning a robust model.

In this thesis, we tried to resolve the above limitations by proposing an alternative for existing fundus based screening solutions. We focused on detection of Retinal Nerve Fiber Layer Defect Detection (RNFLD) which is one of the key and early symptoms of Glaucoma. In the proposed approach we tried to learn from OCT and fundus modalities during training stage. Our method is capable of (chapter-3) handling absence of OCT modality during the screening stage. Using the proposed approach we observed an 8-10$\%$ improvement of RNFLD detection when compared with traditional mono-modal training. For normalizing data across modalities and incorporate image-image learning we proposed the multimodal registration algorithm (chapter-2), especially for low FOV images and noisy images. This method is supported by extensive quantitative and qualitative results. In addition to that, we investigated the scope of new imaging modality (Angiography-Optical Coherence Tomography) A-OCT for detecting Glaucoma. Many research studies are being conducted to reduce the setup cost of A-OCT scanners. With reduced setup cost A-OCT modality has potential for usage in screening systems. The main findings for each problem are chapter specific and were summarized within their respective chapters.

Although this thesis offers promising results and analysis, there are certain areas open to further research. Exploring the following areas will improve the diagnosis of the patient on a screening level stage.

- In chapter-2 A robust multimodal registration for low FOV images is proposed. An extensive experimentation is performed. However, improving accuracy to subpixel is still an open-ended problem which needed a closure.
- In chapter-3 An approach to OCT assisted fundus detection system is proposed. By exploring different architectures and training strategies the detection of RNFLD can further improved. Further, this can be extended to multiple modalities/diseases. Scalability of the proposed method needs to be evaluated by validating on relatively large datasets.
- Chapter-4 An extensive understanding of RNFLD in A-OCT via case studies has gated new potential solutions for diagnosis. The same understanding can be extended to other diseases.

Finally, this thesis concludes by positing that learning from multiple modalities strongly helps in improving diagnosis of the patient. The system is able to capture the strengths of all modalities during the learning process.

Related Publications

1. *Multimodal registration of retinal images*
   jahnavi G.S, Jayanthi Sivaswamy.
   National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, 2017, IIT Mandi (2nd chapter)

2. *Topography and Correlation of Radial Peripapillary Capillary Density Network with Retinal Nerve Fiber Layer Thickness*
   Tarannum Mansoori*, Jayanthi Sivaswamy, **jahnavi G.S** Nagalla Balakrishna.
   Journal of Ophthalmology, 2017 (4th chapter)

3. *Radial peripapillary capillaries density measurement using optical coherence tomography angiography in early glaucoma*
   Tarannum Mansoori*, Jayanthi Sivaswamy, **jahnavi G.S** Nagalla Balakrishna.
   Journal of Glaucoma, 2017 (4th chapter)

4. *Measurement of Radial Peripapillary Capillary Density in the Normal Human Retina Using Optical Coherence Tomography Angiography*
   Tarannum Mansoori*, Jayanthi Sivaswamy, **jahnavi G.S** Nagalla Balakrishna.
   Journal of Glaucoma, 2016 (4th chapter)

5. *Neovascularization detection using semi supervised learning*
   jahnavi G.S*,Pujitha AK*, Jayanthi Sivaswamy.
   IEEE 14th International Symposium on Biomedical Imaging (ISBI) 2017,Melbourne, Australia.

Bibliography

[1] J. Addison Lee, J. Cheng, B. Hai Lee, E. Ping Ong, G. Xu, D. Wing Kee Wong, J. Liu, A. Laude, and T. Han Lim. A low-dimensional step pattern analysis algorithm with application to multimodal retinal image registration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1046–1053, 2015.

[2] M. Alterman and P. Henkind. Radial peripapillary capillaries of the retina. ii. possible role in bjerrum scotoma. *The British journal of ophthalmology*, 52(1):26, 1968.

[3] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov. Trainable cosfire filters for vessel delineation with application to retinal images. *Medical image analysis*, 19(1):46–57, 2015.

[4] R. Bock, J. Meier, L. G. Nyúl, J. Hornegger, and G. Michelson. Glaucoma risk index: automated glaucoma detection from color fundus images. *Medical image analysis*, 14(3):471–481, 2010.

[5] C.-C. Chang and C.-J. Lin. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):27, 2011.

[6] J. Chen, R. T. Smith, J. Tian, and A. F. Laine. A novel registration method for retinal images based on local features. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 2242–2245. IEEE, 2008.

[7] J. Chen, J. Tian, N. Lee, J. Zheng, R. T. Smith, and A. F. Laine. A partial intensity invariant feature descriptor for multimodal retinal image registration. *IEEE Transactions on Biomedical Engineering*, 57(7):1707–1718, 2010.

[8] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[9] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 130–137. Springer, 1998.

[10] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016.

[11] Z. Ghassabi, J. Shanbehzadeh, A. Sedaghat, and E. Fatemizadeh. An efficient approach for robust multimodal retinal image registration based on ur-sift features and piifd descriptors. *EURASIP Journal on Image and Video Processing*, 2013(1):1–16, 2013.

[12] M. Golabbakhsh and H. Rabbani. Vessel-based registration of fundus and optical coherence tomography projection images of retina using a quadratic registration model. *IET Image Processing*, 7(8):768–776, 2013.

[13] K. Gopinath, J. Sivaswamy, and T. Mansoori. Automatic glaucoma assessment from angio-oct images. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, pages 193–196. IEEE, 2016.

[14] I. Gorczynska, J. V. Migacz, R. J. Zawadzki, A. G. Capps, and J. S. Werner. Comparison of amplitude-decorrelation, speckle-variance and phase-variance oct angiography methods for imaging the human retina and choroid. *Biomedical optics express*, 7(3):911–942, 2016.

[15] M. Havaei, N. Guizard, N. Chapados, and Y. Bengio. Hemis: Hetero-modal image segmentation. *CoRR*, abs/1607.05194, 2016.

[16] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, M. Brady, and J. A. Schnabel. Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012.

[17] S. jahnavi and J. Sivaswamy. Multimodal registration of retinal images. In *National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*. IEEE, 2017.

[18] G. D. Joshi and J. Sivaswamy. Colour retinal image enhancement based on domain knowledge. In *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*, pages 591–598. IEEE, 2008.

[19] G. D. Joshi, J. Sivaswamy, and S. Krishnadas. Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment. *IEEE transactions on medical imaging*, 30(6):1192–1205, 2011.

[20] G. D. Joshi, J. Sivaswamy, R. Prashanth, and S. Krishnadas. Detection of peri-papillary atrophy and rnfl defect from retinal images. In *International Conference Image Analysis and Recognition*, pages 400–407. Springer, 2012.

[21] A. L. Kornzweig, I. Eliasoph, and M. Feldstein. Selective atrophy of the radial peripapillary capillaries in chronic glaucoma. *Archives of ophthalmology*, 80(6):696–702, 1968.

[22] S. Lee, K. Kim, J. Seo, D. Kim, H. Chung, K. Park, and H. Kim. Automated quantification of retinal nerve fiber layer atrophy in fundus photograph. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 1, pages 1241–1243. IEEE, 2004.

[23] Y. Li, G. Gregori, R. W. Knighton, B. J. Lujan, and P. J. Rosenfeld. Registration of oct fundus images with color fundus photographs based on blood vessel ridges. *Optics express*, 19(1):7–16, 2011.

[24] L. Ma, Z. Lu, L. Shang, and H. Li. Multimodal convolutional neural networks for matching image and sentence. In *Proceedings of the IEEE international conference on computer vision*, pages 2623–2631, 2015.

[25] T. Mahmudi, R. Kafieh, H. Rabbani, M. Akhlagi, et al. Comparison of macular octs in right and left eyes of normal people. In *SPIE Medical Imaging*, pages 90381W–90381W. International Society for Optics and Photonics, 2014.

[26] T. Mansoori, J. Sivaswamy, J. S. Gamalapati, and N. Balakrishna. Topography and correlation of radial peripapillary capillary density network with retinal nerve fibre layer thickness. *International Ophthalmology*, pages 1–8.

[27] T. Mansoori, J. Sivaswamy, J. S. Gamalapati, and N. Balakrishna. Radial peripapillary capillary density measurement using optical coherence tomography angiography in early glaucoma. *Journal of glaucoma*, 26(5):438–443, 2017.

[28] S. Niu, Q. Chen, H. Shen, L. de Sisternes, and D. L. Rubin. Registration of sd-oct en-face images with color fundus photographs based on local patch matching. 2014.

[29] J. Odstrcilik, R. Kolar, R.-P. Tornow, J. Jan, A. Budai, M. Mayer, M. Vodakova, R. Laemmer, M. Lamos, Z. Kuna, et al. Thickness related textural properties of retinal nerve fiber layer in color fundus images. *Computerized Medical Imaging and Graphics*, 38(6):508–516, 2014.

[30] Y. Ouyang, F. M. Heussen, P. A. Keane, S. R. Sadda, and A. C. Walsh. The retinal disease screening study: Prospective comparison of nonmydriatic fundus photography and optical coherence tomography for detection of retinal irregularitiesnonmydriatic fp versus oct. *Investigative ophthalmology & visual science*, 54(2):1460–1468, 2013.

[31] R. Panda, N. Puhan, A. Rao, D. Padhy, and G. Panda. Recurrent neural network based retinal nerve fiber layer defect detection in early glaucoma. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 692–695. IEEE, 2017.

[32] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[33] N. Srivastava and R. R. Salakhutdinov. Multimodal learning with deep boltzmann machines. In *Advances in neural information processing systems*, pages 2222–2230, 2012.

[34] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell. Adversarial discriminative domain adaptation. *arXiv preprint arXiv:1702.05464*, 2017.

[35] Y. T. Wang, M. Tadarati, Y. Wolfson, S. B. Bressler, and N. M. Bressler. Comparison of prevalence of diabetic macular edema based on monocular fundus photography vs optical coherence tomography. *JAMA ophthalmology*, 134(2):222–228, 2016.