

# **Automatic Classification Of Optical Coherence Tomography Images**

Khaled Abdulhameed Manea Alsaih

Department of LE2I - Laboratoire Electronique, Informatique et Image  
Université Bourgogne Franche-Comté  
University of Technology Malaysia

A Thesis Submitted for the Degree of  
MSc in Vision and Robotics (MSCV)

· 2016 ·

## Abstract

We study in this thesis new methods in order to analyse Optical Coherence Tomography (OCT) volumes in patients that suffer from Diabetes. Early detection of Diabetic Macular Edema (DME) and Age-related Macular Degeneration (AMD) would help patients to get special treatments. Thus, developing a mechanism or algorithm to automatically detecting the retinopathy diabetes is a must to save humans lives and maintain their status healthy.

Basically, the algorithm for automated detection of the retina diseases is going through several steps depends on the task wanted to detect. This project classifies volumes to diseased or normal in the first task, and classifies patches inside the images to cyst or background in the second task. First task is starting by preprocessing, which involves de-noising then flattening then cropping each B-scan of OCT images. After that, Multi-scale histograms of oriented gradient descriptors (HOG) and Local Binary Pattern (LBP) are extracted separately. In other tests, HOG and LBP are combined to create a set of different feature vectors. We also made some tests of the features vectors when projected into a lower-dimensional space through Principal Component Analysis (PCA). Eventually, a Bag of Words (BoW) is also performed prior to feed the data to different classifiers like Linear SVM, kernel RBF SVM and Random Forest (RF). Experimental results show a promising performance in terms of sensitivity (SE) and specificity (SP) of 87.5% and 87.5%, respectively, on a challenging dataset. Another approach for classifying on OCT images is done by extracting Maximally Stable Extremal Regions (MSER) and label the regions extracted using the STAPLE ground truth to fit the labelled data to the autoencoder for learning and extracting the features. After that, the output data of autoencoder are assigned to the softmax layer classifier to classify the cyst patches and the background patches. Experimental results also show a promising performance in terms of sensitivity (SE) and specificity (SP) of 95.0% and 79.0%, respectively, on another challenging dataset provided by OPTIMA.

# Contents

<b>List of Abbreviations</b>	<b>vi</b>
<b>Acknowledgments</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Objectives . . . . .	2
1.2 Thesis Overview . . . . .	2
<b>2 Background</b>	<b>3</b>
2.1 Introduction . . . . .	3
2.1.1 The Eye . . . . .	3
2.1.2 Eyes Abnormalities . . . . .	6
2.1.3 Imaging Techniques . . . . .	10
2.2 Retina Analysis in OCT images . . . . .	14
2.2.1 OCT Voxels Classification . . . . .	15
2.2.2 OCT Patch or Cyst Classification . . . . .	18
<b>3 Methodology</b>	<b>21</b>
3.1 Introduction . . . . .	21
3.2 Automatic classification Of OCT Volumes . . . . .	22
3.2.1 Preprocessing . . . . .	23

3.2.2	Feature Extraction . . . . .	25
3.2.3	Feature Representation . . . . .	27
3.2.4	Classification . . . . .	29
3.2.5	Evaluation Of Classification . . . . .	31
3.3	Automatic Classification Of Potential Regions in OCT Volumes . . . . .	32
3.3.1	Segmentation Process . . . . .	32
3.3.2	Auto-encoder Training And Feature Extraction . . . . .	35
3.3.3	Classification . . . . .	36
3.3.4	Evaluation And Validation . . . . .	37
<b>4</b>	<b>Results And Discussion</b>	<b>38</b>
4.1	Classification Experiments . . . . .	38
4.1.1	Experiments . . . . .	38
4.1.2	Discussion . . . . .	39
4.2	Segmentation Experiments . . . . .	44
4.2.1	Experiments . . . . .	44
4.2.2	Discussion . . . . .	45
<b>5</b>	<b>Conclusion</b>	<b>47</b>
<b>Bibliography</b>		<b>54</b>

# List of Figures

2.1	Anatomy of the Eye [5]	4
2.2	Retina Anatomy [6]	5
2.3	Loss of Central Vision [2]	7
2.4	Cataracts [1]	8
2.5	CMV and HIV [12]	8
2.6	Diabetic Retinopathy (DR) [3]	9
2.7	Glaucoma [8]	10
2.8	Retinal Detachment [14]	10
2.9	Scheimpflug and ultrasonic [16]	11
2.10	Scanning laser ophthalmoscopy [16]	11
2.11	Ocular Ultrasonography [11]	12
2.12	Magnetic Resonance Imaging [13]	13
2.13	Optical Coherence Tomography Angiography [15]	14
2.14	Fundus Photography [7]	14
2.15	Examples of DME cases in SERI dataset.	16
2.16	Examples of DME and AMD cases in OPTIMA dataset.	19
3.1	Classification Pipeline	22
3.2	Original B-scan slice	23
3.3	De-noising B-scan Slice	24

3.4	Flattening B-scan Slice . . . . .	25
3.5	Cropping B-scan Slice . . . . .	25
3.6	Linear-SVM [10] . . . . .	29
3.7	Kernal-SVM [9] . . . . .	30
3.8	Confusion Matrix for evaluation [4] . . . . .	31
3.9	Cyst Classification Pipeline . . . . .	33
3.10	Maximally stable extremal regions . . . . .	34
3.11	Simultaneous truth and performance level estimation . . . . .	35
3.12	Autoencoder Layout . . . . .	36
3.13	Weights of Auto-encoder . . . . .	37

# List of Tables

2.1	Summary of the state-of-the-art methods.	18
4.1	Experiments based Linear-SVM.	40
4.2	Experiments based Linear-SVM-RBF.	41
4.3	Experiments based Random Forest.	42
4.4	Classification results using Histogram+Principal Component Analysis (PCA)+Bag of Words (BoW) representation.	43
4.5	Summary of the classification performance in terms of se and sp in(%)	44
4.6	Auto-encoder Results.	46

## Abbreviations

**ACC** accuracy

**AMD** Age-related Macular Degeneration

**AUC** Area Under the Curve

**BoW** Bag of Words

**DME** Diabetic Macular Edema

**GMM** Gaussian Mixture Models

**HoG** Histogram of Oriented Gradients

**LBP** Local Binary Pattern

**LBP-TOP** Three Orthogonal Planes

**PCA** Principal Component Analysis

**PRE** precision

**RBF** Radial Basis Function

**RF** Random Forest

**SE** sensitivity

**SERI** Singapore Eye Research Institute

**SP** specificity

**SVM** Support Vector Machine

# Acknowledgments

First of all, I would like to thank my amazing supervisor Prof. Fabrice Meriaudeau for his consistent support and clear directions. He never felt annoyed from my questions along with Dr.Desire Sidibe as well. I would also love to thank my colleagues Guillaume Lemaître and Mojdeh Rastgoo for their practical help through the duration of my master project.

I would like to send this thesis work and efforts to My country YEMEN and then to my father Prof. Abdulhameed Alsaih and My Mom for their permanent support through my life. Also I would like to thank my Sister Taif,my brothers Tariq, Mojahed,Dr.Mohammed,Dr.Sakher, Alhalali Ali Almansour and Sabahi.

In addition, I would like to thank my friends that they Always support me starting by Hossam Hassan, Awss,Hamed alboori, Hashem Alsaih, Hamodi, Yousef, Zezo,Sufian, Seper, Paola, Jose, Dennis, Sandeep, Cansen, Raed, Kiril, and others.

As usual, I would like to thank my Seniors Dr.Fares, Dr.Taha and Dr.Abdulaziz for their motivation and help for any inquiries.

Finally, thanks to University of Burgundy and University of Technology Malaysia for the great chance to be part of this course which taught me a lot and looking forward to apply it with the sense of serving humanity in a good manner.

# Chapter 1

## Introduction

Sugar level instability is what determine the existence of Diabetes in human body. Diabetes has two types: type 1 which is insulin-dependent and type 2 which insulin-independent. This critical disease is affecting many parts of the body like the Eye, Heart, Pancreas, body power and others. The scope of this project is to focus in the Eye disease specifically Retina. The sight for human is so important and with the Optical Coherence Tomography (OCT) images, the doctor can decide the status of the patient either it has disease or not. Some statistics are showing from [59] that 29.1 million of Americans in the United States are affected with diabetes, which contributes to 9.3% of the overall population. Diabetes added to 231,404 deaths in the US in 2007. \$245 billion is the total cost of diagnosing the disease in the United States in 2012. With this big number of affected people and the money spent for Research and development (R&D) in this area many algorithms are designed to detect diabetes in the early stage.

Eye diseases such as Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) are the most common causes of irreversible vision loss in individuals with diabetes. DME is defined as the increase in retinal thickness within one disk diameter of the fovea center with or without hard exudate and sometimes associated with cysts. Spectral Domain OCT (SD-OCT) [26] images the depth of the retina with a high resolution and fast image acquisition is an adequate tool, compared to fundus images for DME identification. Automated diagnosis on OCT imaging is rather new and most of the pioneer works on OCT image analysis have focused on the problem of retinal layers or specific lesions (e.g. cysts) segmentation.

In this thesis we developed some methods to early detecting the diabetes with the care of some important things: time processing, cost and accuracy of the method in order to classify the disease. We aimed to do many tests and use the tools provided to enhance the machine learning method we used. The method applied is a great help for the doctors and for the patients if the solution of the sickness also involves any tele-medicine treatment.

## 1.1 Objectives

The project target is to develop an automatic algorithm for retina screening for the sake of detecting the cysts and during the research period the main objectives of this project as follows:

- **Objective 1:** develop an automatic algorithm to diagnose the retinopathy diabetes in the OCT volumes into two approaches:
  - to classify the OCT images using machine learning techniques into diseased volumes or normal volumes.
  - to classify potential regions in OCT images using deep learning techniques into cyst patches or background patches.
- **Objective 2:** evaluate the algorithms based on diabetic retinopathy screening.

After achieving the objectives, some other techniques were added to enhance the results. In addition to the objectives mentioned, different classifiers were tested and compared. An algorithm is developed to detect Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) from the lesions segmented.

## 1.2 Thesis Overview

This project is arranged and sequenced in four chapters. An outline is described as follows:

- **Chapter 2:** explains in details the anatomy of the eye accompanied by the diseases that affect the vision process for humans. In addition, it reviews the previous work in this field with the state of art of this project.
- **Chapter 3:** shows the work done in the period decided to do the project, which has two parts: first, designing an automatic algorithm to classify the OCT images. Second, to classify potential regions in the OCT images based on deep learning techniques.
- **Chapter 4:** discusses the results of the work done and evaluate the performances of the methods.
- **Chapter 5:** sums up the project and recommends work to be achieved in future.

# Chapter 2

## Background

In this chapter the explanation is well-defined in term of the anatomy of the eye and the state of art for classification tasks. This chapter explains the retina layers and the abnormalities that may accompany the eye during the human life to have a better understanding about the project scope and the reasons behind following the algorithm used. After that, a review of previous work in classification of OCT images and the techniques to evaluate and compare the results of the method used. Then, a short summary for the papers is reviewed, followed by making a highlight for the main references is presented for this project

### 2.1 Introduction

#### 2.1.1 The Eye

Eye is one of the most important organs, which is responsible of vision for humans. Saving this organ is critical and taking care of it is a must. Many scientists have written the way the eye is working and they conclude, the eye is like a mirror so whatever it sees it reflects and in the very famous book (**Book of Optics**) [18] written by Ibn Alhaytham, He claims "the eye is affected by light not the opposite". In scientific view, the task of the eye is a sight machine that let the data received from the out atmosphere to be processed by brain for decision making for example: moving right ,left and etc... The camera idea was mainly gotten from the eye in the way it is working in term of light entrance to the lenses used then to the processor (Brain). The camera creates pictures using films or digitally made [17]. Focus, contrast and the brightness also vary from camera to other based on how powerful the processor and other aspects. The camera creates one image (Picture) or sequence of images (Video), it could be also in real time video like the eye with an advantage for the camera that it can record events for later use. After

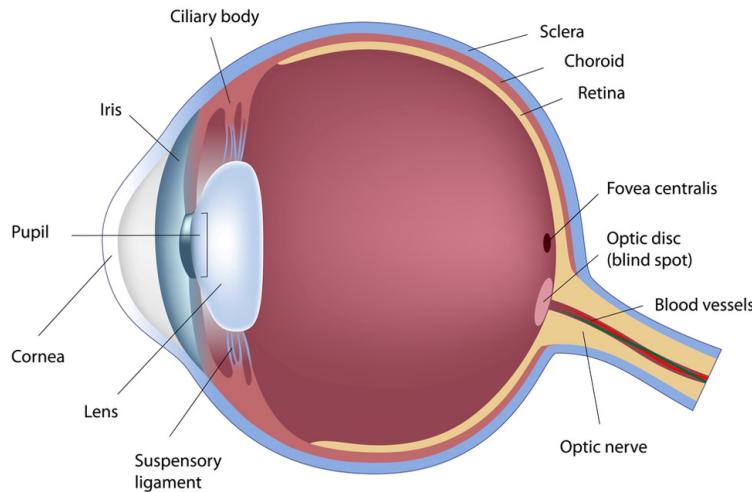


Figure 2.1: Anatomy of the Eye [5]

observing how light travels by Ibn Alhaytham and his book he made the first pinhole camera to be the first design of camera and the development of the camera till our days. The anatomy of the eye is shown in Figure 2.1 with different layers, which is compromised with lenses to process data. Humans may vary in the eye's structure, hence some people has a normal range of vision, others they can not see either far or near. Eyes also can be infected by many diseases, which will be discussed after explaining the process of vision.

Basically, the vision process can be explained as follows:

- **Cornea:** receives light for filtering. Then, it focuses the data received.
- **Aqueous Humour (Anterior Chamber) :** keeps the anterior part of the eye slightly curved and stable and it is made of viscous substance.
- **Pupil:** adapts the eye lenses by relaxing or contracting depending on the amount of light enters the eye.
- **Iris:** does the movement of the muscle based on the pupil order and based on the decision made by the pupil, the light will pass to retina
- **Retina:** converts the light rays into electrical impulses to be delivered to the brain by optic nerve.

In our project we mainly focus in the retina part and it is explained as follows.

## Microscopic Anatomy of the Retina

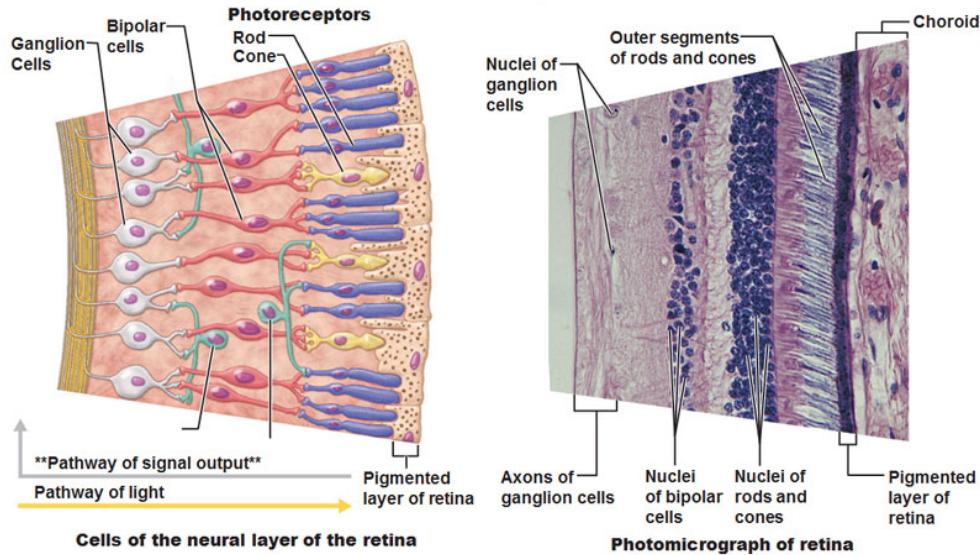


Figure 2.2: Retina Anatomy [6]

### Retina

Retina is a multi-layered tissue located at the posterior of the eye. Figure 2.2 shows the sensory parts in the retina, which they are the photo-receptors that get the rays and convert them into electrical impulses to be able to be read by the brain. Impulses are passing through the optic nerve to the brain for image processing and the result of this process is the vision of the human. There are millions of photo-receptors for sensing and there are two types of them based on their shapes : Rods and Cones.

Rods are located in periphery of the retina and they are imprecise and insensitive to colours so they can only recognize shapes in grey-scale. Whenever there is a change in contrast it is very case-sensitive so it is most effective at night. They are able to detect movement. Cones are responsible for day vision as they are very precise to detect colours. They are located in the Macula [49], [17]. The Fovea is the part of eye which is able to recognize the vision details. It is located in the central of the Macula. Any accidental action to the eye results in losing Fovea will affect the vision in the central area. A dense network of almost 1.2 million nerves is connecting the photo-receptors with the brain from the optic nerve. The blind spot is located also in the retina, which has no photo-receptors.

The inner and outer layers of the retina are encountered to be supplied by their fuel to enrich the retina with its need (oxygen and other important components required to keep the

mechanism of vision) [39]. The number of vessels for the inner layer 35% is smaller compared to the outer layer 65% to supply the retina. Retina has many layers specified in [17] as follows:

- Internal limiting membrane
- Axons of the ganglion cells (Nerve Fibre Layer): the more the age, the thin it gets.
- Ganglion Cell Layer includes the ganglion cells.
- Inner Plexiform Layer have Amacrines and axons of bipolar cells.
- Inner Nuclear Layer have horizontal and bipolar cells.
- Outer Plexiform Layer includes horizontal dendrites and inner segments of photo-receptors cells.
- The photo-receptors are located in the Outer Nuclear Layer.
- External limiting membrane.
- Finally, the layer that generates the colours to retina (Pigment Epithelium).

This complex combination of cells inside the retina is tremendous. The existence of cysts is mysterious, thus it is studied by scholars to get the status of patients with different diseases [37]

### 2.1.2 Eyes Abnormalities

Eyes abnormalities have no early signs, because the pain is not felt usually. The vision also is not interrupted to give a warning that someone is infected by an eye disease to be diagnosed and get a treatment unless the problem is advance. As the body is a complex and compact system, hence any problem faces the eyes might interrupt many organs in the body and could be a sign to diagnose other diseases like heart diseases and diabetics. The main disease, project making analysis on is the Diabetic Macular Edema (DME).

Diabetic sometimes called Diabetes mellitus (DM) where pancreas can not produce enough insulin or the cells can not respond to the insulin produced. As of 2015, an estimated 415 million people have diabetes worldwide [31]. This represents 8.3% of the adult population, with equal rates in both women and men [62]. The number of people with diabetes is expected to rise to 592 million by 2035 [23]. The global economic cost of diabetes in 2014 was estimated to be \$612 billion USD [21]. The best way to protect the eyes is by doing a regular visit to the doctor for eye examination. Failing to treat this disease will yield to other diseases such as cardiovascular disease, stroke, kidney failure and damage to the eyes. In this section, many diseases that might hit the eye will be ordered alphabetically :

1. **Age-related Macular Degeneration (AMD)**: is a disease that will affect the Macula, which is located at the center of the retina. Getting affected by AMD is a sign for diabetes. AMD will lead to blur the vision, a dark area in the center of the vision and straight line distortion as shown in figure 2.3. Activities required focus is affected by this disease like driving, cooking or studying. There are two types of AMD: Dry and Wet. The wet AMD is happening due to generating new blood vessels in the Macula, which will lead to rapid lose of central vision. In the other hand, dry AMD is happening due to the existance of drusen, which they are small yellow particles that they are deposited under the Macula. Slowly, this particles will lead to lose the central vision and it is the most popular type [24].

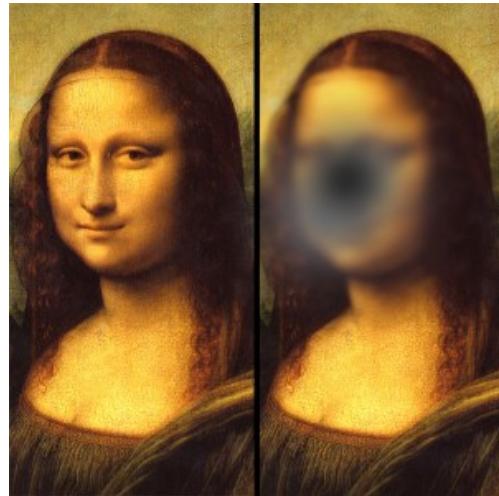


Figure 2.3: Loss of Central Vision [2]

2. **Cataracts**: is an abnormality of the eye, which will affect the crystalline lens. Crystalline lens is responsible in recognizing on stuff and people at different distances. The typical effected patients in this syndrome is old people when the lens get stiffens due to ageing and they call it presbyopia. In other ways, some people might face Cataracts earlier due to some reasons for example: diabetes, birth defects, steroids and heredity. Cataracts patients will suffer from being sensitive to light, bad night vision, colour vision dimming, sometimes double vision in one eye and blurred vision [42].
3. **Cytomegalovirus Retinitis (CMV)**: is a disease that strikes the retina specifically the light sensing cells. A quick treatment is required as this disease is so dangerous and sometimes might lead to blindness. Cytomegalovirus Retinitis does not produce symptoms and usually the immune system is able to fight it. Sometimes, this virus will be accompanied

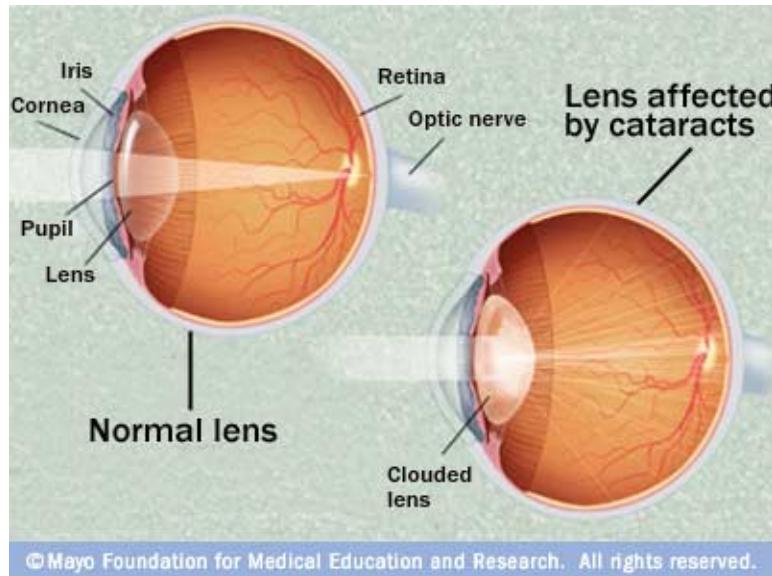


Figure 2.4: Cataracts [1]

by HIV virus. This disease will make the patient suffer from couple of flashes for eyes, blind spots, end of the peripheral vision and blindness [65].

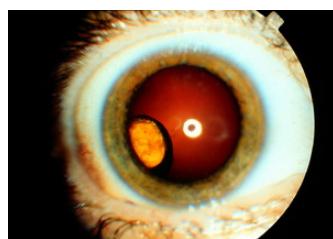


Figure 2.5: CMV and HIV [12]

4. **Diabetic Retinopathy (DR)** is disease that attacks the retina and it is a vascular complication. Diabetic retinopathy (DR) is a vascular complication of DME that leads to vision loss in case of late treatment [64]. Having Diabetes will increase the chances of getting blind than normal people.
5. **Diabetic Macular Edema (DME)** is a retinal disease and caused by fluid accumulation in the Macula dImaging of the eye is critical to improving the diagnosis, assessment of severity and progression, and evaluation of management of eye disease. It is also, a

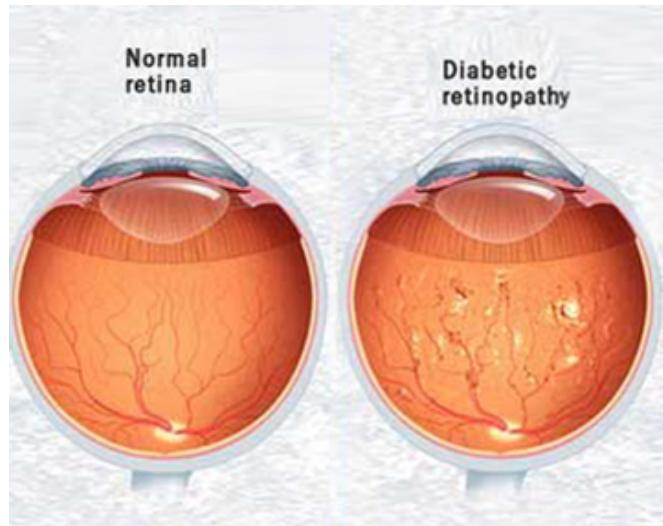


Figure 2.6: Diabetic Retinopathy (DR) [3]

combination of complication of Diabetes Retinopathy that allows the fluid to escape. The cones fill the Macula. When the fluid starts to flow through the Macula to escape, blurred vision is the result of the inability of cones to sense the light. Up to 30% of Diabetes patients are suffering from DME. Two types of DME based on the way the fluid flow, which they are Focal DME and Diffuse DME. To diagnose DME, retina thickness is checked if it is increased or exudate by 500 um distance from the Macula centre [22]. DME only can be diagnosed by 3-D images because the lack of depth information when capturing the image due to the retina thickness increase. Mainly, this project is focusing in detecting DME cysts and classifying the volumes based of how normal the volume is and DME shall look like the table in section 2.2.1

6. **Glaucoma:** this disease will damage the optic nerve, which is responsible for transferring data from the eye to the brain for image processing. Thus, it will lead to vision loss or in the worst case scenario, blindness. As other eye diseases, it is hard to diagnose the disease as it developed with no symptoms. There are four types of Glaucoma: Chronic Open Angle Glaucoma, Acute Closed Angle Glaucoma, Secondary Glaucoma and Normal-Tension Glaucoma. Glaucoma in case of acute angle will happen suddenly and the patient will vomit and feel dizzy. Also, it is accompanied with headache and vision blurring [34].
  
7. **Retinal Detachment:** a sudden damage of the retina allowing the fluid to flow inside it. This action will make the retina to move away from the supportive tissue under it

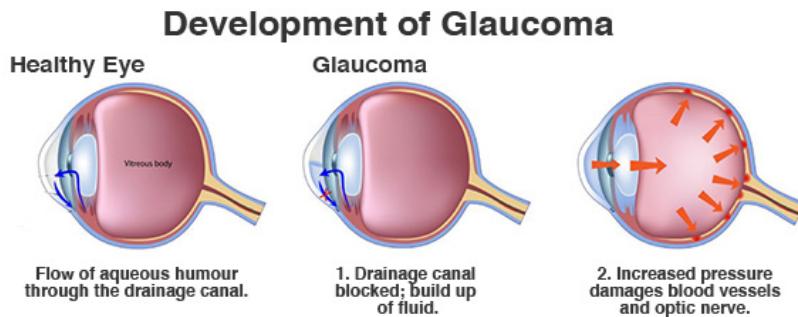


Figure 2.7: Glaucoma [8]

causing the detachment. With this detachment, the retina is no longer able to correctly send the signals of light to the brain. It can be treated but in a very good care [27].

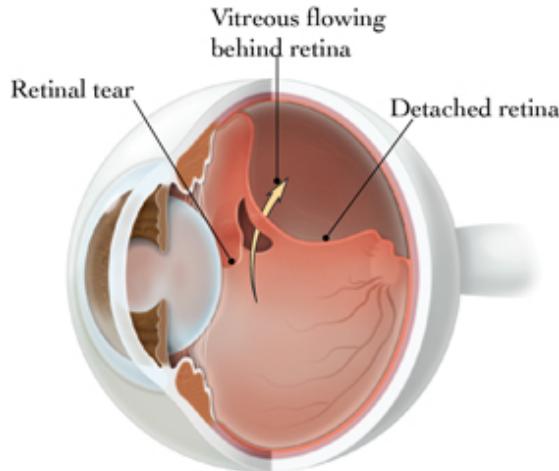


Figure 2.8: Retinal Detachment [14]

### 2.1.3 Imaging Techniques

The eye imaging is critical in developing the diagnosing, analysing the progression and managing the evaluation of the disease. The improvement in hardware such as chips and light sources and software such as image analysis is helpful in diagnosing the disease. Based on the needed part of the eye and the disease to be imaged, the suitable imaging techniques are used by the ophthalmologists. In this project, our data used are SERI and OPTIMA OCT images. This section explains briefly ocular imaging techniques in the eye:

- **Scheimpflug principle:** this technique is named after its inventor the army captain Theodor Scheimpflug from Austria. He has used it in creating a device for correcting the distortion in aerial photographs in a very systematic method. This method has a rule that explains the orientation of the coordinates of optical system when the lenses are not in a parallel plane with the image plane. It can be used in corneal pachymetry, the corneal topography mapping, also prior to refractive eye surgery such as LASIK, and can be used in detection of keratoconus [35].

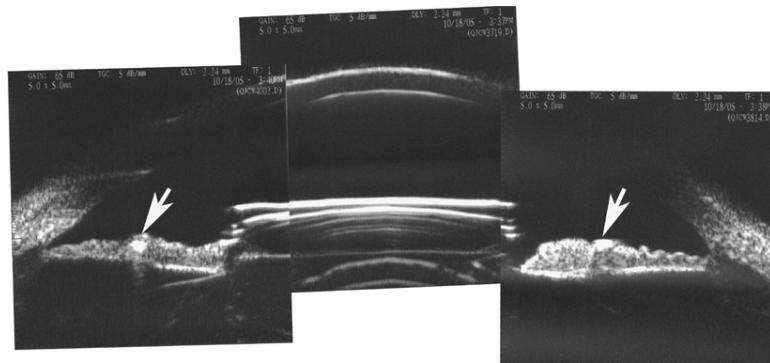


Figure 2.9: Scheimpflug and ultrasonic [16]

- **Scanning laser ophthalmoscopy (SLO):** this method is using confocal laser scanning microscopy for the aim of diagnosing the cornea and the retina at the eye of the human. It employs scanning mirrors in horizontal and vertical coordinates for the aim of scanning a particular region in the retina and create a viewable images at the monitor. It is also able to image the retina in real time and it might face a problem with reflections generated from the cornea. It is helpful in term of diagnosing retinal disorders like DR, AMD, glaucoma and DME due to the high degree of spatial sensitivity [33]. It also has been developed to generate sharper images of the retina when it is combined with adaptive optics technology.

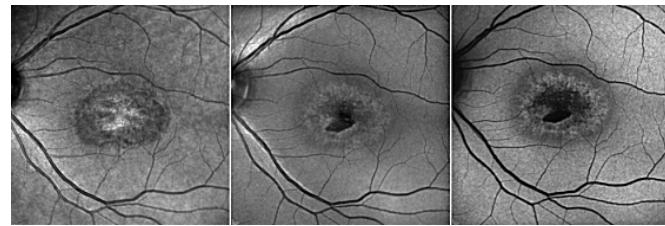


Figure 2.10: Scanning laser ophthalmoscopy [16]

- **Ultrasonography:** this method uses mainly the time taken for high radio frequency pulses measurement, which is generated by piezoelectric devices. The pulses are reflected by the interface of the ocular tissue(A-scan). Another way of scanning (B-scan), which is inserting the probe across the eye to generate a section or moved in a rectangular pattern to create 3D form of structure [28]. As the emphasize of this method to use the high frequency pulses rather than light waves, it has the ability to penetrate opaque corneas to better image the ciliary body and surrounding parts. Depth and resolution of ultrasonography penetration are affected by frequency's frequency.

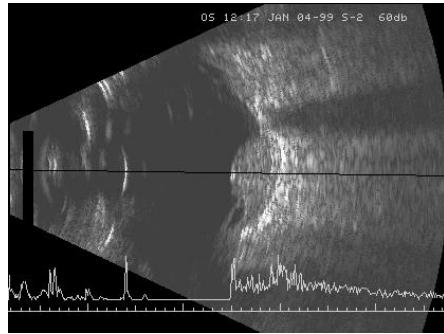


Figure 2.11: Ocular Ultrasonography [11]

- **Magnetic Resonance Imaging(MRI):** is a well-known technique in imaging world, which results in producing a very detailed images of the internal frame of the body. The principle of the MRI is based on nuclear magnetic resonance. Spinning of the nuclei at certain radio frequency to obtain a stored signal that contains information about the physical and chemical frame of molecules that has hydrogen. Also, in the case of placing the object in an uniform magnetic field, the nuclear spins within the targeted structure. Hence, they will align together either in parallel or anti-parallel to the direction of the magnetic field flux. Radio-frequency pulses will radiate echoes of the same frequencies which is detected. The density of the nuclei and the internal structure will affect the magnitude and the decay of the signal. Due to the high price of this device, so it has limited usage, but they use it in the ageing ciliary muscle diameter in phakic eyes [57], checking the 3D structure of the myopic, viewing gained ocular anomalies and some animal models in ocular drug-delivery systems [41].
- **Optical Coherence Tomography (OCT):** the main method of imaging in this project which is a light reflection based. An interface is used instead of calculating the time between emission and detection of light-pulse owing to the fast speed of light in this

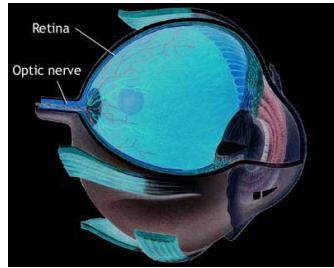


Figure 2.12: Magnetic Resonance Imaging [13]

technique, which is faster than sound speed in almost 872,000 times (344 m/s speed of sound and OCT light speed is 299,792,458m/s). Axial resolution is equal to light source coherence length and lateral resolution is the optics of the device function. While, the laser (Conventional interferometry) has a long coherence length in meters, broadband of light sources are in use to make the laser frequencies shorten to micrometers in order to view high resolution eye image. Any light outside the broadband of short frequencies length will not be used as data in the coherence pattern. There are two types of OCT images, (A and B-scan). A cross-sectional tomography (B-scan) is achieved by aligning a series of axial depths (A-scan). (A-scan) is built by physically scanning the coherence length using the reference mirror in time domain OCT. This type is limiting the resolution of the image and the images shall look like the table presented in section 2.2.1. (B-scan) is used to diagnose DME,AMD and Glaucoma with better precision than fundus images.

- **Optical Coherence Tomography Angiography (OCTA)** is another type of normal OCT, yet modern and from the name angiography, which means a technique to visualize the internal structure of organs and blood vessels. OCTA uses motion contrast imaging to high-resolution volumetric blood flow data producing angiographic images in seconds. OCTA matches the decorrelation signal (differences in the backscattered OCT signal intensity) between sequential OCT b-scans taken at accurately the same cross-section in order to create a map of blood flow. Axial bulk motion from patient movement is removed so sites of motion between repeated OCT b-scans signify strictly erythrocyte movement in retinal blood vessels. OCTA demands higher imaging speeds than most presently accessible OCT systems can deliver in order to attain a densely sampled volume. Conventional OCT device scanning speeds would outcome in too much trade-off between reduced field of view, lower image quality, and greatly increased scanning time [55].
- **Fundus Photography:** is a method to screen the retina directly by using the pupil as the entrance and the exit of light rays for fundus camera. The screening is achieved by

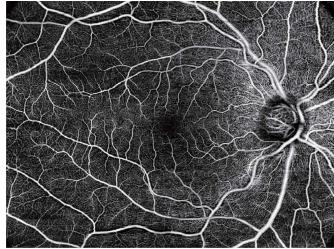


Figure 2.13: Optical Coherence Tomography Angiography [15]

placing the patient in front of the camera with the chin in a chin rest and the forehead against the bar. Then, the photographer will align the camera and presses the shutter to fire a flash to create the fundus image [54]. Fundus images can be coloured with including some dyes like indocyanine and fluorescein.

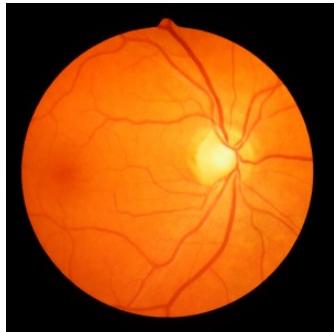


Figure 2.14: Fundus Photography [7]

## 2.2 Retina Analysis in OCT images

In this section, previous work in the classification of the OCT images is presented and compared to have a better understanding of the field chosen. Annotated data or dataset is required to diagnose the retina condition by analysing the retina morphology. The design, implementation and testing of various previous work is reviewed. There are many data for testing purposes and to be used with the permission of patients for scientific purposes. Many datasets are available online for different parts or organs of the human body and also including the retina, where these datasets have different goals to study based on the case wanted to analyse. The target in retina is different based on sickness needed to be diagnosed such as DME, AMD, DR, vessel segmentation and some localization tasks, hence many algorithms and datasets are done for

each sickness to be diagnosed. Since, the datasets are somehow useless if they are presented alone with no ground-truths aside the image or voxel to make the validation of the results and how accurate the designed algorithm. By using the suitable method of image enhancing and extraction of data, ground-truth is giving a hand in training and testing data as the key to evaluate any algorithm's advantages and disadvantages. This project used two types of data of OCT volumes, which they are SERI data and OPTIMA dataset.

### 2.2.1 OCT Voxels Classification

Over the past decades, the improvement in the applications of computer vision and image processing to OCT volumes interpretation have been paid a big focus on improving an automated retinal layer classification and segmentation methods. This section discusses the recent state-of-the-art methods for classification using supervised and semi-supervised methods of SD-OCT volumes. The datasets used in this study were acquired by the Singapore Eye Research Institute (SERI), using CIRRUS TM (Carl Zeiss Meditec, Inc., Dublin, CA) SD-OCT device [25]. The datasets consist of 32 OCT volumes (16 DME and 16 normal cases). Each volume contains 128 B-scan with resolution of 512\* 1024 pixels. All SD-OCT volumes are read and assessed by trained graders and identified as normal or DME cases based on evaluation of retinal thickening, hard exudates, intraretinal cystoid space formation and subretinal fluid as presented in the table below. Within the DME dataset, a large number of lesions were selected to create a rather complete DME dataset. The volumes and sample numbers for DME volumes are presented in figure 2.4:

Before, heading straight away to the algorithm, we shall know what is Machine learning. Inventors have long dreamed and though of creating machines that are able to think just like humans. The difficulties might face any systems depending on hard-coded knowledge said that Artificial Intelligence systems acquire the ability to have their own knowledge, by extracting features in patterns from raw data. This capability is known as machine learning, essentially a form of applied statistics. To have a precise definition we define what learning is first. Learning is provided by [50] A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E: And a human can imagine how computer can integrate couple of experiences, tasks and performances based on how powerful the computer is used for the particular task. Tasks in machine learning can be classification, classification with missing inputs, regression, transcription, anomaly detection and many others.

Srinivasan et al. propose a classification method to distinguish normal, DME, and Age-related Macular Degeneration (AMD) OCT volumes [56]. The SD-OCT volumes are enhanced by:

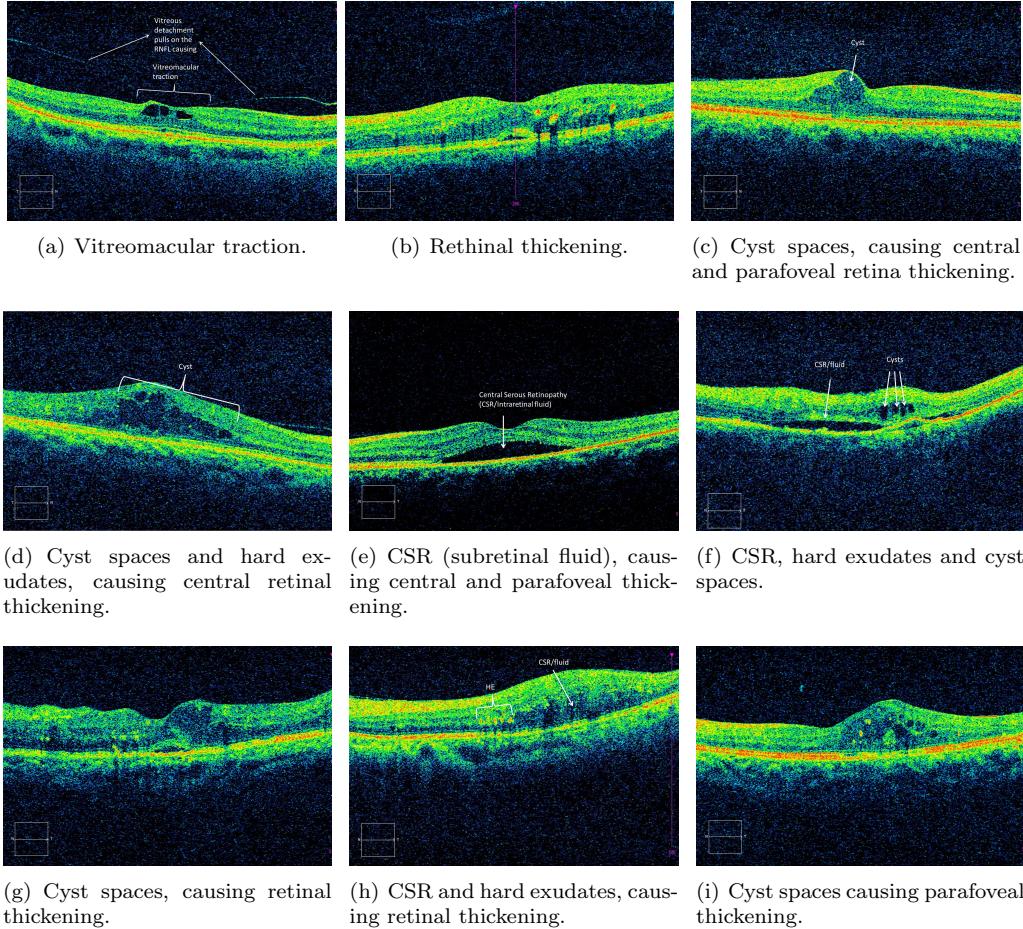


Figure 2.15: Examples of DME cases in SERI dataset.

- reducing the speckle noise through a de-noising method, which enforces the sparsity in a specific transform-domain.
- flattening the retinal curvature.

After that, edge information is extracted using Histogram of Oriented Gradients (HoG) descriptor for each B-scan of a volume and later used to train a linear Support Vector Machine (SVM). This method is evaluated on a dataset of 45 patients equally subdivided into the three mentioned classes and led to a correct classification rate of 100%, 100% and 86.7% for normal, DME and AMD patients, respectively. The dataset used by [56] is publicly available but is already pre-processed (i.e., denoised, flattened, and cropped). Furthermore, this dataset does not offer

a huge variability in terms of DME lesions, have different sizes for the OCT volumes, and some of them, without specifying which, have been excluded during the training. All these reasons, prevent us from using this dataset to benchmark our work.

Venhuizen et al. recently proposed a method to classify AMD and normal OCT volumes in [60] using Bag of Words (BoW) models. In the proposed method, the features are extracted from a set of key-points detected from each individual B-scan. As a feature descriptor, a 9 px × 9 px texton is extracted around each selected key-point and its dimension is reduced, from 81 to 9 using Principal Component Analysis (PCA). A dictionary or codebook is created by clustering the features extracted and each volume is represented in terms of a histogram, which captures the codebook occurrences. These histograms are used as a final feature vectors to train a Random Forest (RF classifier. This classifier is evaluated on a dataset composed of 384 volumes leading to an Area Under the Curve (AUC) of 0.984.

Liu et al. propose a methodology for detecting macular pathology in OCT images using Local Binary Pattern (LBP) and gradient information as attributes [45]. Each B-scans is aligned and flattened and a 3-level multi-scale spatial pyramid is created. Additionally, edges are detected using Canny detector on the same pyramid. Subsequently, an LBP histogram is extracted for each of the layer of the pyramid. All the obtained histograms are concatenated into a global descriptor whose dimensions are reduced using PCA. Finally, a SVM with an Radial Basis Function (RBF) kernel is used as classifier. The method achieved good results in detection of OCT scan containing different pathology such as DME or AMD, with an AUC of 0.93 using a dataset of 326 OCT scans.

Lemaître et al. propose another method based on extracted LBP features from OCT images and dictionary learning using BoW models [43]. Contrary to [56], BoW and dictionary learning are used to perform volume classification is performed rather than B-scan. In this method, the OCT images are first pre-processed using Non-Local Means (NLM) filtering to reduce the speckle noise. Then, the volumes are mapped into discrete set of structures namely: local, when these structures correspond to patches; or global, when they correspond to volume slices or the whole volume. According to different mapping, LBP or Three Orthogonal Planes (LBP-TOP) texture features are extracted and represented per volume using histogram, PCA, or BoW. The final feature descriptors per volumes are classified using RF classifier. Classifying DME versus normal volumes on a balanced dataset of 32 SD-OCT volumes, the classification performance in terms of sensitivity (SE) and specificity (SP) of 87.50% and 75%, respectively, is achieved, while using LBP-TOP features and global mapping.

On the same dataset, Sankar et al. proposed a rather different approach, based on semi-supervised learning, to address the issue of an anomaly detection [53]. In their method, the authors proposed a technique that does not only allow the classification of the OCT volume, but also enables the identification of the abnormal B-scans inside the volume. This approach

Table 2.1: Summary of the state-of-the-art methods.

References	Diseases			Data size	Pre-processing				Features	Representation	Classifier	Evaluation	Results
	AMD	DME	Normal		De-noise	Flatten	Aligning	Cropping					
Srinivasan <i>et al.</i> [56]	✓	✓	✓	45	✓	✓		✓	HoG		linear-SVM	ACC	86.7%, 100%, 100%
Venhuizen <i>et al.</i> [60]	✓		✓	384					Texton	BoW, PCA	RF	AUC	0.984
Liu <i>et al.</i> [45]	✓	✓	✓	326		✓	✓		Edge, LBP	PCA	SVM-RBF	AUC	0.93
Lemaitre <i>et al.</i> [43]		✓	✓	32	✓				LBP, LBP-TOP	PCA, BoW Histogram	RF	SE, SP	87.5%, 75%
Sankar <i>et al.</i> [53]		✓	✓	32	✓	✓		✓	Pixel intensities	PCA	Mahalanobis -distance to GMM	SE, SP	80%, 93%

is based on modelling the appearance of normal OCT images with a Gaussian Mixture Models (GMM) and detecting abnormal OCT images as outliers. The classification of an OCT volume is based on the number of detected outliers. Testing on 32 OCT volumes, their proposed method achieved SE and SP of 93% and 80%, respectively. A conclusion of this techniques is presented on table 2.1

## 2.2.2 OCT Patch or Cyst Classification

This section aims to explain the state of arts of work done in segmenting OCT voxels. Inspired from the research done in segmenting OCT images, we did the classification on potential regions. The SD-OCT data used in this project is provided by the OPTIMA laboratory (Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University of Vienna) for the Cyst segmentation challenge hosted at MICCAI 2015. These data consisted of 15 SD-OCT volumes containing a wide variety of retinal cysts with accompanying clinical ground truth annotation manually drawn by two different experts (Two Ground-Truths). The SD-OCT voxels have 4 different vendors at different resolutions and scanning patterns: four volumes from Cirrus (Carl Zeiss Meditec, Dublin, CA, USA), three volumes from Nidek (NIDEK Co., Hiroishi, Gamagori, Japan), four volumes from Spectralis (Heidelberg Engineering, Heidelberg, Germany) and four volumes from Topcon (Topcon medical Systems, Santa Clara, CA, USA). Results validation in this challenge were compared with the first and second reading by experts and the intersection between them compared to the segmented method done by the researcher. Some examples of the Data used by Optima as shown in figure 2.16:

Many researchers went through this challenge made by Optima laboratory. The method proposed by Luis et al [46] is a machine learning based, where a model is trained using manual markings (to establish the ground truth) and then tested using another 15 volumes of testing data also provided by OPTIMA laboratory. In the preprocessing stage, the SD-OCT data

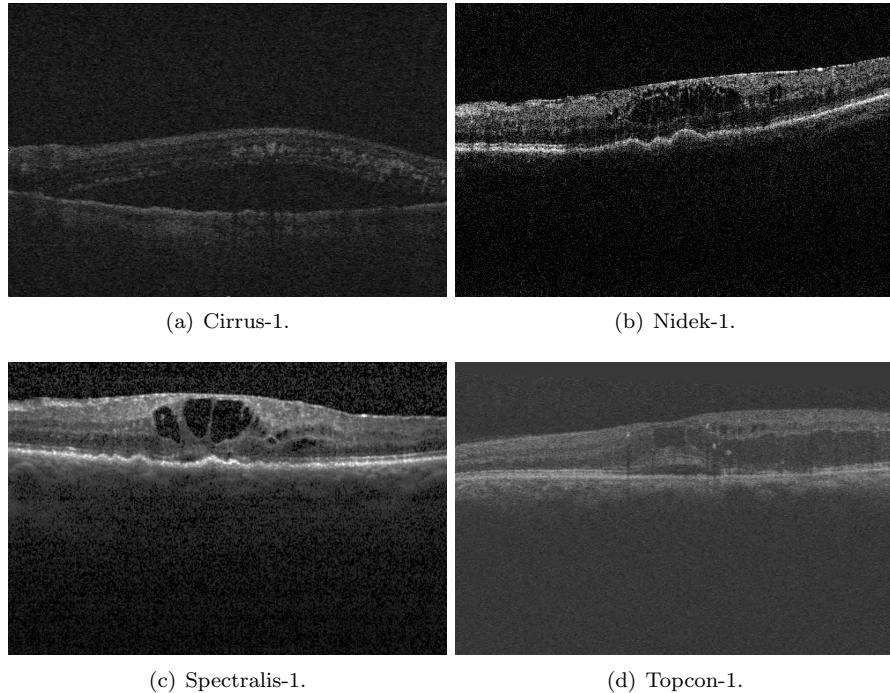


Figure 2.16: Examples of DME and AMD cases in OPTIMA dataset.

normalized and de-noised using non-local means filtering in the axial and horizontal direction. A defined number of boundaries defining the axial location of defined intra-retinal layers are then automatically outlined using a developed segmentation algorithm done by Luis et al in [30] (SOARS: Stanford OCT Automated Retinal Segmentation). A number of quantitative features (34 features) are extracted to characterize each volume located between the segmented internal limiting membrane (ILM) and inner segment junction (IS), where the possible cysts are located in. These features expanded to have four possible resolutions using multi-resolution approach to have set of predictors. After that, they calculate the risk score for each voxel. The final segmentation output is generated automatically by detecting an adaptive threshold to stratify the output scores in those belonging to a cyst or background. The accuracy using this method has achieved a good results using dice coefficient evaluation with 80% of correctly segmenting the B-scan slices.

Another proposed method by Ipek et al., [51], which is using cost function (opposed to machine-learned) that generalizes well to a variety of images. This cost function takes into account the general characteristics of the input image as well as the well-known characteristics of fluid-associated abnormalities. As the background and the cyst color in OCT images is black,

so all methods care about certain area of the B-scan slice which lies between the layers. They create a reliable mask even with the presence of fluid associated with cysts, followed by using a method to correct the Bruch membrane (BM). Then they segment using a cost function and they compare with the experts results. This method fully cover all black holes lie in the layers and they did not target cysts only.

Mahdad et al., [47] used a speckle noise reduction algorithm, which maintains strong edge sharpness and reduce the noise. This can be achieved by applying recursive Gaussian filter to noisy images and to any zero pixel exist in the image. After that, threshold is used to the image to segment fluid space pixels. Then nerve fibre layer (NFL) and retinal pigment epithelium (RPE) layer are extracted in each B- scan. Finally, most of the possible false positives (FPs) are removed based on standard deviation and morphology of extracted candidate pixels.

Another proposed method by Karthik in [40] to segment OCT images. They started by de-noising the image by using Total Variational approach that will reduce the texture content resulting in a smooth piecewise constant images preserving the edges. Then, the make candidate selection as all methods done to specify the process only between NFL and RPE to save time and avoid pixels that they have nothing to do with the assessment of the disease existed in the voxel. After that, Candidate regions are extracted using Maximally stable extremal regions(MSER). This feature computation produces a set of stable region in an image. MSER also used to detect the multi-scale objects without any smoothing. Both small and large structure can be detected based on threshold and other aspects based on the favoured regions wanted to be extracted. Then, based on the texture of pattern calculation, a local descriptor is assigned to each batch after making a bounding box for each region extracted by MSER. Finally , with fifty trees in Random forest classifier is used for results validation and the results were challenging as it gives good results when the cyst is in medium and large size and poor results for small cysts.

A novel method using different way of training and extracting data proposed by Venhuizen [32]. His method mainly using the Convolution Neural Network to segment the images. Two stages define the process, in the first stage, Three convolution neural networks (CNNs) are used to get a segmentation at different image scales. In the second stage, the three CNNs scale segmentations are fused together, redefining the borders of the segmented cysts by combining local information obtained with the lower scale network with contextual information obtained from the higher scale networks.

# **Chapter 3**

## **Methodology**

There are many unpredicted aspects that can result in unwanted images in real world screen environment. For this reason, many algorithms were reviewed in the previous section to give an idea of the project scope. Preprocessing step is a must to have better results, get a noise free image and to be able to specify the disease location clearly, then to develop an algorithm to track or classify it. For some drawbacks in the algorithms reviewed regarding accuracy, speed and sensitivity and etc, we have introduced an automatic algorithm to classify the OCT volumes based on the Data provided SERI and OPTIMA. Another algorithm is introduced in the project, which is classifying potential regions in the OCT images using the auto-encoder. In this chapter, we present the method for both algorithms. A MATLAB code implementation is developed using the cluster provided by University of Burgundy with Xeon 2630 V3 processor, 128 GB of RAM and a GPU Nvidia Tesla K40 which processes the algorithms proposed in an average of 5 hours per tasks.

### **3.1 Introduction**

Many systems for retinopathy automatic detection in OCT images have been improved. The results obtained is promising and discussed in the next chapter. The machine used to train this system proposed is powerful to process a very big data in some hours where normal computers might take days to weeks to achieve the same task. Algorithm used to correct the images light or contract increasing is used in the pre-processing section. However, the image enhanced has limited resolution obtained cause sometimes increasing the threshold will yield to over-blurring , which is not good due to info losing might be useful. Feature extraction in the images is also taking time to process meanwhile the multi-pyramid way is used. In addition to those, feature representation is used to have a clear dictionary to the classifier.

Accuracy quality of results validation is a subjective concept to be determined. It is not a straight task to be determined and this can be noticed with the different experts opinions in term of making a decision of the same aspect. To detect a disease that attacks the eye, same imaging technique is useless. The reason behind that, in Glaucoma we need dark images to better figure out the nerves problem, while dark images are considered low quality if we want to detect Diabetic Retinopathy(DR). Hence, we used confusion Matrix based on how many volumes detected as Normal, DME or wrong for both cases for the first part of the Project.

The second part, is using deep learning technique to train the data after extracting the potential regions using the MSER. The classification is performed by assigning a label to the cyst location then fitting it to the auto-encoder. Then we have a test data with their ground truths to be sent for prediction based on the cluster or network created from the auto-encoder. Accuracy decision is made based on confusion matrix and compared with the two ground truth provided by OPTIMA laboratory and the intersection between them.

Processing speed is another aspect to be taken into account when processing the algorithms, hence a powerful computer is in need for training and classifying the data about. This chapter explains the methodology of the algorithms and it is divided into two parts, first is OCT classification based on machine learning techniques and the second is the OCT classification based on deep learning techniques.

## 3.2 Automatic classification Of OCT Volumes

Inspired by the previous methods and using machine learning tools to obtain an automatic machine able to make a decision. A supervised learning algorithm is designed to map a certain function with a known inputs and outputs. Our classification pipeline is depicted in Fig.3.1. The rest of the section present into details each intermediate step. Our approach is based on the original paper of Srinivasan et al [56]. It has been enhanced with feature representation like PCA , proposed in [60] and BoW. Different classifiers are used and the results presented in the next chapter

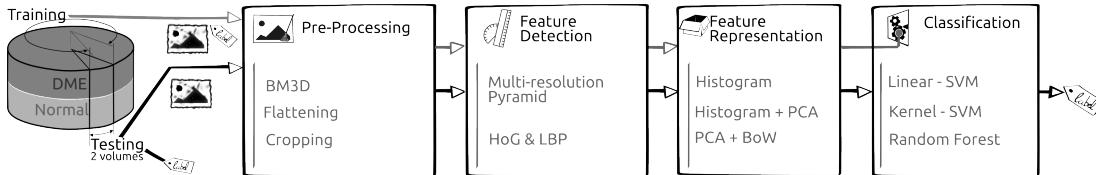


Figure 3.1: Classification Pipeline

### 3.2.1 Preprocessing

Image blurring and sharpening is essential and critical step in machine learning aspects to ensure the high quality of the data presented to be analysed in a process called preprocessing. Fixing the imperfections in the images is important to let the following tools useful. So, this process is the main basis of images processing field. This project introduces the first step as Preprocessing as shown in the figure 3.2 the original image. A sample of the image is shown here:

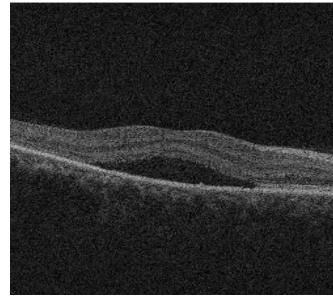


Figure 3.2: Original B-scan slice

Following the work done by Srinivasan et al. [56],and prior to feature extraction, the OCT volumes are pre- processed through de-noising, flattening, and cropping steps, the details as follow:

1. **De-noising:**It is the process of affecting the look of a certain image, by modifying pixels and its neighbours by some math equations applied to the images matrices. Applying some filters to a particular images to enhance the look and remove the noise away. In this part of the project we have used Block Matching 3D filtering to sharpen the image as shown in figure 3.3. In the first step, speckle noise is attenuated through an image de-noising strategy which uses block matching and collaborative filtering in the 3D domain [29], namely Block Matching 3D filtering (BM3D). The core algorithm is composed of three steps:

- **grouping:** consists in grouping similar 2D image patches from different spatial locations, to form 3D blocks.
- **collaborative filtering :** is equivalent to de-noise the 3D blocks by successively applying a 3D transform, a de-noising method, and an inverse 3D transform.
- **5 aggregation:**a de-noised 10 image is reconstructed by making a linear combination of the 2D de-noised patches. The previous algorithm is applied twice in the BM3D frame- work to build:

- (a) **basic estimate:** is computed by grouping 15 noisy 2D patches, de-noising the blocks via hard-thresholding with the value of 190, and aggregating the patches by setting the weights to be inversely proportional to the total sample variance of the blocks.
- (b) **final estimate:** the grouping is built from two distinct blocks by arranging 2D patches from both the 20 noisy image and basic estimate.

The filtering is performed through a Wiener filter driven by the blocks extracted from the basic estimate, considered as the true energy spectrum. The aggregation step is equivalent to the one performed in the basic estimate stage to obtain the final denoised image.

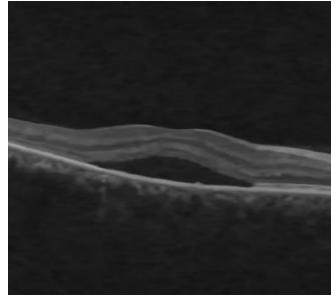


Figure 3.3: De-noising B-scan Slice

2. **Flattening:** is the function used to make the layers of an image almost straight based on the function applied either fitting a polynomial or RANSAC. SD-OCT images of the retina have a natural curvature as noticed in the figure 3.2, which is further distorted due to the common practices in OCT image acquisition and display that varies both between patients and within each SD-OCT volume. Following Srinivasan et al. way we reduce the effects of the perceived retinal curvature when classifying SD-OCT images, we flatten the retinal curvature in each image. To flatten the retinal curvature in each SD-OCT image:

- calculate a pilot estimate of the retinal pigment epithelium layer (RPE).
- calculate the convex hull around the pilot RPE points, and use the lower border of the convex hull as an estimate of the lower boundary of the retina.
- remove outliers by applying a  $[1 * 3]$  median filter (MATLAB notation) to this estimate.
- Unlike Srinivasan et al. in his work for creating the flattened image, he fitted the second order polynomial to make the estimation of retinal lower boundary points. Random Sample Consensus(RANSAC) function is applied to do the same job with

better accuracy of line alignment. An iterative method to achieve parameter estimation of a model from set of data with outliers. With the accuracy achieved in RANSAC but there's a disadvantage of time as it has no limited time to do the job, until it finds the best alignment it will stop, otherwise it will continue calculating.

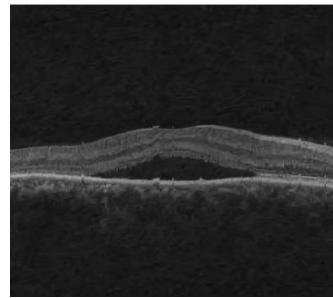


Figure 3.4: Flattening B-scan Slice

3. **Cropping:** this method is used to crop the region interested to save time when the data required for classification might take time. Thus, to differentiate between the volumes with disease or not, limiting the targeted place might help the classifier to set the data clearly. The cropping in our SERI data is bigger than the cropping region done by Srinivasan et al. In the axial dimension, all images are cropped 325 pixels from over the RPE and 30 pixels under the RPE. In the lateral dimension, all images are cropped 340 pixels to the center.

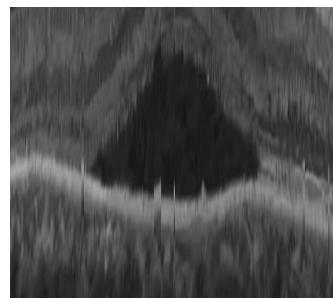


Figure 3.5: Cropping B-scan Slice

### 3.2.2 Feature Extraction

This part of the pipeline is meant to create informative data from measured data and build values (features), which they are non-redundant and lead to better interpretations. It is crazy

to take all features and in our case it reached up to 113215 dimension for 32 volumes, which real data is lost, only 32 points in this great number of dimensions. It describes content of the features in images, algorithms and application and so on, This process finds out the connection between pixels in a particular image. Descriptors describe the main characteristics such as shape, texture and many other aspects in image. Descriptors are divided into two groups:

- **General information descriptors:** give description of colour, shape , texture, basically low level descriptors.
- **Specific domain information descriptors:** give description about events or objects like face recognition.

In the pipeline, implementation of Histogram of Oriented Gradients (HOG) and Local binary patterns (LBP) is applied and the explanation as follow:

1. **Histogram of Oriented Gradients (HOG):** is a feature descriptor for object detection. This descriptors are used based on gradient orientation in well-localized portion in an image. It is computed based on overlapping local contrast normalization on a dense grid on uniform cells to enhance the accuracy. HOG descriptors divides the image into connected cells by counting the strength and orientation of the gradient spatial in every cell and computing the gradients can be done by applying the 1-D centred, point discrete derivative mask in one or both of the horizontal and vertical directions. Applying the mask for filtering the color or intensity data of the image with some filter kernels like: [-1, 0, 1] or its transpose. After that, HOG creates cells histograms in each pixel within the cell. Each pixel has a weight for vote in an orientation-based histogram channel based on the values found in the gradient computation. The cells shape can be radial or rectangular, and the histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, relying on whether the gradient is unsigned or signed. To make the descriptors invariant to shadowing and different illumination, descriptor vectors blocks are normalized over larger overlapping blocks then renormalizing. This results into a feature vector for each image with normalized histograms. This technique is focusing more into edges unlike LBP which focuses in intensity of regions.

In our implementation, as Srinivasan et al. did, the extraction of HOG with [4\*4] cell size, [2\*2] cells per cell, a block overlap of [1\*1] cells, unsigned gradients and 9 orientation histogram bins. Extraction of HOG features at four levels of the multi-scale Gaussian low-pass pyramid to have multiple scale levels structure. Furthermore, concatenate all HOG vectors for each image to obtain the set of histograms to be ready for the feature representation or classification.

2. **Local binary patterns (LBP):** is a visual descriptor and usually combined with HOG to improve the performance of detection on some dataset as our implementation did. It starts with dividing the window of an image to cells (8\*8 pixels for every cell). After that, a comparison of each pixel with its 8, 16 and 24 neighbours with different radius of 1,2 and 3 based on the number of neighbours. Followed by clockwise or counter-clockwise circle along the pixels. The decision making in this process, "0" value is assigned to the pixel if the center pixel value is greater than neighbours value, otherwise assign value "1". Then, the histogram is computed over the cell by the appearing pixels frequency based on the threshold of higher or less than the center. In this project, Non-invariant rotation LBP and invariant rotation LBP are used. Rotation Non-invariant (NRI)LBP is the same as the normal LBP where it results into a histogram of 255 size and this NRI LBP is not grouping the equal patterns into one histogram but it will just add it. Unlike rotation invariant (RI)LBP that will group the equal patterns and put them into single histogram and this will reduce the histograms numbers for example from 255 to 50. Normalizing the vector feature is done. Finally, concatenating all histograms from all cells in all images to make 128 vectors for all volume. this data is ready to be directed for further feature representation or directly to the classifier.

### 3.2.3 Feature Representation

The great number of dimensions resulted from the previous section make the processing time high for classification, hence classification loses the main points in the huge dimensions. Many correlated feature vectors are produced, which reduces the precision and performance of system, so it is very crucial to learn the most important and uncorrelated features. In another words, Feature representation is the way to create a new space for features either by making a new model for features or reducing the real features to more meaningful dimensions.

There many types of feature representation methods, like Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), Sequential Forward Feature Selection (SFFS) and Sequential Backward Feature Selection (SBFS). SFFS and SBFS are used to chose the most correlated dimensions, while PCA and LDA are projecting data into new space by reducing dimensions. Bag of Words (BoW) is a feature representation method that try to find similar patterns based on creating dictionary or words in clusters for features in the space. This section will explain the methods we used in the pipeline:

1. **Principal Component Analysis (PCA):** is feature reduction method for the extracted data, which looks for subspaces where variation is maximized. It is unsupervised, statistical orthogonal linear transformation approach where the subspace transformed data is relevant and uncorrelated . The process of PCA basically, is done by picking up the high-

est variance value in the dimension till it reaches to the number chosen to be reduced. The principal components are orthogonal transformation because they are the eigenvectors of the covariance matrix, which is symmetric. PCA is sensitive to the repetitive scaling of the original variables, hence reducing the dimensions can not be achieved, a kernel of PCA is used to find subspaces.

PCA is the way of fitting some data from a lot of dimensions where each axis means a principle component. Sometimes if the axis is small compare to others so the variance along the axis is small, hence the lose of some useful information is happened by deleting that axis. The math way of PCA is organized as follows

- subtracting the mean from each point in the dataset to gather data around origin center.
- Calculating the covariance matrix of centred data.
- Computing the eigenvectors covariance matrix corresponding eigenvectors.
- Orthogonalizing eigenvectors sets then converting them to unit vectors by normalizing the eigenvectors.

In this project, reduction of HOG and LBP dimensions to 40 and 20 dimensions respectively. The dimensions reduced are concatenated before BoW applied for creating dictionary of the data. Mainly, the reduction was done based on Singular value decomposition (SVD), due to the accuracy produced with acceptable processing speed. The avoidance of using Eigenvalue decomposition (EIG) cause the less accurate results generated compared to SVD, even-though it is faster. There is another method to reduce feature data in PCA using Alternating least squares (ALS) algorithm, which can work well for data sets with a small percentage of missing data at random, but might not perform well on sparse data sets [36].

2. **Bag of Words (BoW):** is a clustering method for the extracted data by modelling them to a new space. The number of repeating of pattern will make the decision how to cluster the data. Clustering belongs to the use of k-means to make the cluster for the feature space so that k represent the number of words, histograms or samples number of minority class. Hence, the centroids of these clusters define the new samples from the majority class. BoW clusters are the set of low-level features using a k-means algorithm to produce a codebook of visual words. Every visual word have a centroid that belongs to a cluster. After generating the words, every centroid is represented by histogram based on the number of the occurrence of the words from extracted features [58]. K-means algorithm is an iterative method, which finds k centroids by alternating assignment and

update steps. Euclidean distance method is used for alternating assignment. In this project, many tests involved using different number of k centroid and vary between 10-60. The results of BoW will be shown in the next chapter.

### 3.2.4 Classification

Classification is a method used to map an input data to its category. This terminology task is to recognize the pattern after giving some data similar to it in a process called training. Training might be attached with labels to be supervised learning or without labels to be unsupervised learning. This part of the project is using supervised learning method with labels for Normal volumes and another label for DME volumes. Three different classifier are used for comparison: Random Forest (RF), linear, and kernel- Support Vector Machine (SVM). Using the feature descriptor provided the first two representations. The classifiers are trained to classify B-scans and volume classification is performed based on the total number of diseased B-scans per volume using a majority vote rule and cross validation. Here in this section we explain briefly about the classifiers used as follow:

1. **linear-Support Vector Machine (SVM):** is a supervised learning algorithm based on training feature vectors to categorize whatever number of categories. SVM creates set of points in space for a category then divided by a reasonable gap and make line to separate points for each category [61]. The new separation of points mapped and predicted to make a decision for the classifier as shown in the SVM figure. In this part, we used the Matlab function `svmtrain` but it did not converge then we used `fitcsvm` function with `predict` function to classify data.

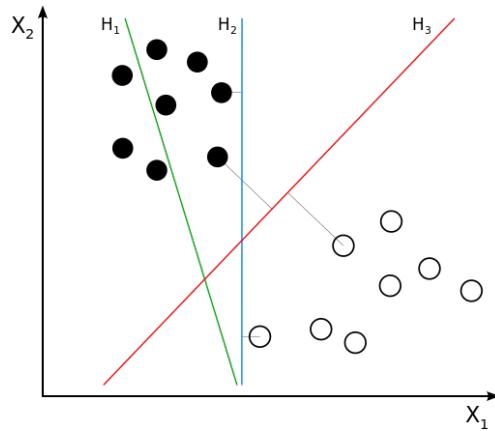


Figure 3.6: Linear-SVM [10]

2. **kernel-Support Vector Machine (SVM):** is exactly the same process of linear SVM but in non-linear way to maximize-margin hyperplanes, as shown in the figure 3.7. This classifier is changing every dot product with the non-linear kernel function. This allows to achieve the maximize-margin hyperplanes in the feature space transformation. The plane might not by linear from input data to be transformed as shown. Working in a higher-dimensional feature space will increase the chances of producing the generalization error of support vector machines, although given enough samples the kernel achieve good performance. Here `fitcsvm` is the Matlab function used to perform this part with kernal RBF.

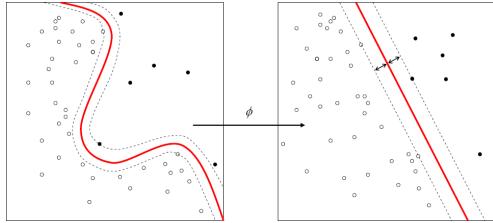


Figure 3.7: Kernal-SVM [9]

3. **Random Forest (RF):** is a classifier that creates trees and then sub-tress based on the targeted number of trees wanted to be involved in the classification task. The more trees to create in this random forest the more time required to compute and the more details that process obtained. This classifier is averaging many trees and sub-trees decision. The average taken is trained on different features in the training set to reduce the variance. This method uses the modified bagging algorithm, which selects random sample with replacement and fits trees into the samples, which differs from the bagging algorithm that targets to decrease the variance which will lead to better performance without even increasing the bias. This means, increasing the number of trees will yield to correlated trees and this will increase the sensitivity to noise, hence bootstrap sampling will de-correlate the trees.

This classifier is enhancing the performance usually, but sometimes with the increasing of trees number the results might stay the same so in here [52] proved that between 64-128 trees are performing good in term of accuracy, memory and time required. In our project we aimed for 80 trees and results shown in next chapter.

### 3.2.5 Evaluation Of Classification

This section briefly explains the way used to validate and evaluate the results obtained from the classifiers mentioned in the previous section. The results were having the assessment based on:

- **Validation:** is the way to generalize the classification results. In this section, we performed leave-two-patient-out-cross-validation, where we take the first voxel from DME images and first voxel from normal images out for testing and the 30 other volumes are processed for training. After that, changing the testing set by substituting the first voxel with second voxel and so on ... for 16 times because we have 32 volumes.
- **Evaluation:** is a method to make results shown and determines the method quality for further comparison. A confusion matrix is done to classify the results and in our case the confusion matrix has two rows and columns, because we classified based on the volume type either diseased or normal.

		Actual Value (as confirmed by experiment)	
		positives	negatives
Predicted Value (predicted by the test)	positives	<b>TP</b> True Positive	<b>FP</b> False Positive
	negatives	<b>FN</b> False Negative	<b>TN</b> True Negative

Figure 3.8: Confusion Matrix for evaluation [4]

The figure shows the evaluation way followed in this project as the True positive (TP) and true negative (TN) are the samples that are correctly predicted to belong to the disease or normal class, respectively. False positive (FP) and false negative (FN) are predicted as the FP in the negative class while it is in the positive class and vice versa. The confusion matrix is the famous method to compute the accuracy (ACC), sensitivity (SE), specificity (SP), and precision. The equations to calculate the accuracy, sensitivity are shown in equation 3.1 and 3.2 respectively. Another aspects used to evaluate the method performance like F-score and Precision.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3.1)$$

and,

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3.2)$$

and,

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3.3)$$

### 3.3 Automatic Classification Of Potential Regions in OCT Volumes

Inspired by deep learning techniques and the segmentation methods discussed in section 2.2.2, this algorithm is introduced to classify the potential regions inside an image. This section is explaining the method used to classify the OPTIMA challenge data as shown in figure 3.9. The rest of the section present into details each intermediate step. This method is based on extracting Maximally stable extremal regions (MSER) "threshold set empirically" and then cropping each region. After that, comparing the cropped images with the Simultaneous truth and performance level estimation (STAPLE) mask images is used for label validation. The total patches extracted from training data and testing data are 354311 and 113149 patches ,respectively. Assigning the patches with labels are fit into auto-encoder for learning and feature extraction, then classification followed by comparing the testing data with the network built in the autoencoder to validate the results.

#### 3.3.1 Segmentation Process

This section explains the steps used to segment an image as many researchers proposed many methods as shown in chapter 2.2.2. and for this project we segment the potential regions as follows:

1. **Maximally stable extremal regions (MSER):** is a method used as blob detection in images. This algorithm proposed by Matas [48], is extracting regions based on the differ of intensity between image elements and create stable regions or images inside each image. This method led to better matching of certain regions with a very fast speed of extracting each image and to compare it with other regions in other images and has the following steps to extract regions:

- Sorting pixels by intensity.

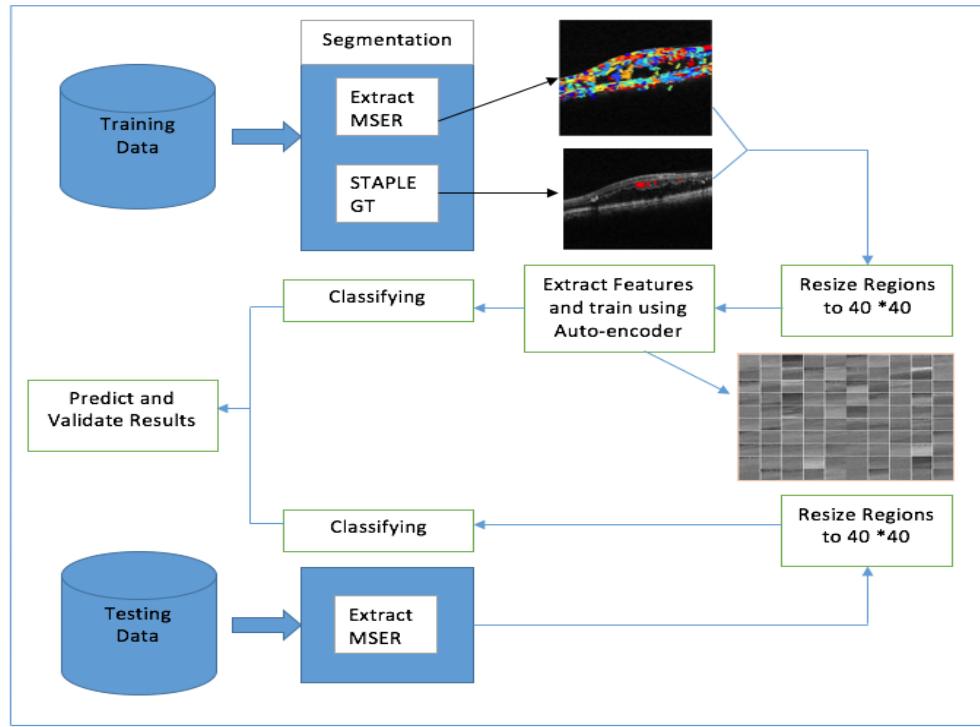


Figure 3.9: Cyst Classification Pipeline

- Marking the set of pixels of each region in different color and the list of merging connected components using union-find algorithm.
- Data structure is produced as a function of intensity of connected pixels.
- two components are merged into larger region if the two groups are smaller than the threshold value set to form one region.
- Regions presented are the the results of the stable regions over large range of threshold.

We extracted MSER regions, which has the cyst regions and non-cyst regions in all images of the training and testing data.

2. **Simultaneous truth and performance level estimation (STAPLE):** is a method used for the validation of segmenting image. The challenge of having optimal method to segment an image is still under research due to the differences in images and resolution. Evaluating the performance of certain algorithm in image segmentation is a difficult process. The reason behind that is back to different opinions of raters or experts in deciding

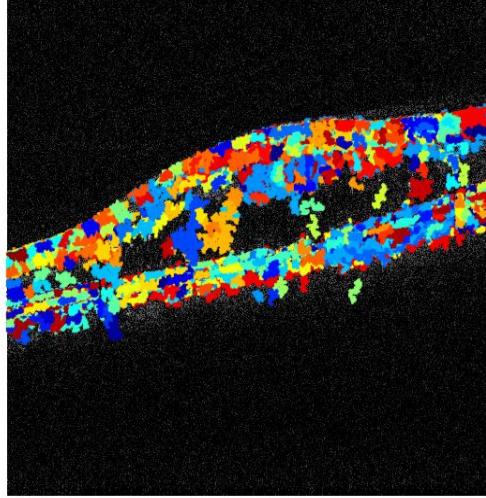


Figure 3.10: Maximally stable extremal regions

how to make ground-truth suiting an image condition, hence the existence of STAPLE to form one ground-truth of all ground-truth presented for one image.

STAPLE is removing the variability introduced by raters to ensure the quality of the proposed method to segment. This method is considering a collection of pixels representing the segmentation and then computing the probabilistic estimate of the true segmentation. The probabilistic estimate of the true segmentation is achieved by predicting the optimal combination of the raters segmentation, weighting each segmentation based on the estimated performance level. Then, incorporating a prior model for the spatial distribution of structures being segmented, as well as spatial homogeneity constraints. STAPLE is a straight method to use and apply in clinical imaging data, it can enable assessment of the performance of an automated image segmentation algorithm, and enables direct comparison of human rater and algorithm performance [63]. OPTIMA data includes two ground truths and using STAPLE algorithm to generate a single ground truth, which contains the data of both ground truths for having a single reference instead of two.

3. **Labelling Patches:**is a process to compare regions extracted from MSER with regions extracted from STAPLE ground truth to assign a label for each patch and then resizing the regions to fit it to the auto-encoder.



Figure 3.11: Simultaneous truth and performance level estimation

### 3.3.2 Auto-encoder Training And Feature Extraction

Auto-encoder is neural network which is copying the inputs to its outputs in the training process. It has different hidden layers internally that represent the input data. The network has two parts which are an-encoder function  $h = f(x)$  and decoder which generate the reconstruction  $r = g(h)$ . Usually auto-encoders are restricted in ways that allow them to learn to copy only input that feeds the training data. Because the model is forced to prioritize, which aspects of the input should be copied, it often learns useful properties of the data and the structure of auto-encoder as shown in figure 3.12. Typically, auto-encoders used for reducing the dimensions or feature extraction.

Auto-encoder is used for training and the algorithm can be summed up as:

- Feed-forward each input  $x$  to calculate the activations in all hidden layers and at the output layers to obtain  $x'$ .
- Using squared error, measure the deviation of  $x'$  from the input  $x$ .
- Backpropagate the error and update the new weight in the net.

Auto-encoder is trained using back-propagation (Fine tuning), which calculates the gradient of the loss function with respect to weights through the network. The weights of this task can be shown in fig 3.13.

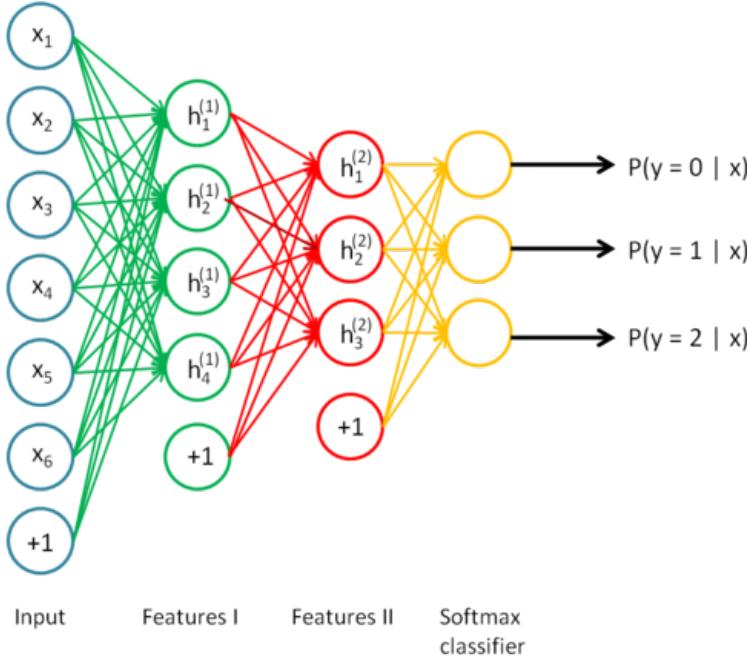


Figure 3.12: Autoencoder Layout

This gradient is used to optimize the algorithm, which is used later to update the weights to minimize the loss function [44]. Backpropagation demands a known output for each input to compute the loss function gradient. This method is supervised and it is used in unsupervised method in auto-encoder. Using backpropagation to train network with many hidden layers will make the first layer when it received the error insignificant, hence conjugate gradient method can solve this problem. Another way to solve this slow process by setting initial weights that estimate the final output. In this project, we train the autoencoder with 100 hidden layers at the first hidden layer then 50 hidden layers at the second layer. then it is classified using softmax layer classifier.

### 3.3.3 Classification

The process of spotting the differences between various classes. This process is explained in 3.2.4. In this part, SVM. or SVM-RBF or RF are not used, but Softmax layer is used instead for classifying the existence of the cyst or not. Softmax layer is a generalization of the logistic function which is implemented at the last layer of the network.

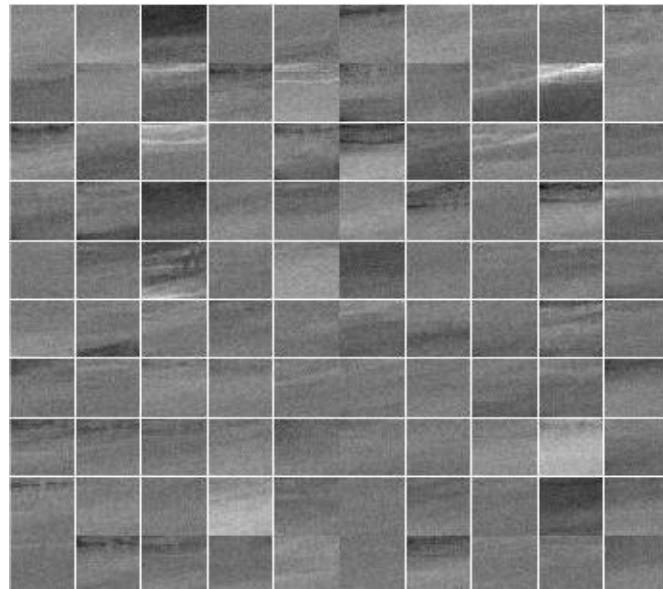


Figure 3.13: Weights of Auto-encoder

#### **3.3.4 Evaluation And Validation**

This process is evaluated using the confusion Matrix. Validation of the results was done based on the same aspects as in section 3.2.5 which they are Precision, Sensitivity and Accuracy.

# Chapter 4

## Results And Discussion

This section presents the results and the discussion of all experiments. Presenting the results based on validation and evaluation process mentioned in the previous chapter. We have divided this section into two parts, which is based on classifying images in the first part and classifying potential regions inside the image for the second part. Inside each section, some subsections presenting the results for each section. In classification section, SERI data is used, and OPTIMA challenge data is used in the second section.

### 4.1 Classification Experiments

#### 4.1.1 Experiments

This section presents the results obtained from different classifiers like linear-SVM, SVM-RBF and RF respectively. First, the 32 volumes are loaded, then two of them are excluded for testing and 30 for training, hence this process is performed 16 times to validate the results based on cross validation. Then, each B-scan image is de-noised using sparsity-based block matching and 3D-filtering (BM3D), which is available online with a harded-threshold , and then flattened the image. Flattening is done by computing the RPE layer, then we compute the convex hull around RPE points and using the lower line of the retina, then applying median filter to remove any outliers. This almost flatten image is then applied to RANSAC to make lines straight. After that, we cropped the targeted location, which consists the disease usual place for having accurate classification and to avoid any not important details.

Instead of dividing images into patches, we extract HOG and LBP at four different scale levels using impyramid then concatenate them into one single vector for each image. After that, for some parts of the test as shown in the tables below, the feature vectors are concatenated from

all volumes and sent directly to the classifiers. Some other cases are representing the features in before sending them to the classifier. Since the great number of dimensions produced from HOG or LBP (rotation invariant (ri) and non- rotation invariant (nri) LBP features with various radius, 8,16,24)or both which are almost 116235 vectors, PCA is used to reduce the dimensions for both descriptors 40 dimensions for HOG and 20 for LBP such that the most discriminative components are kept. Some of experiments sending the PCA of HOG to classifiers directly or lbp or concatenating bot and send them to classifiers.

Another way to represent the feature vectors by creating dictionaries or words then converting the words to histograms based on kmeans and send them to the classifiers and the optimal number of words has been selected heuristically as shown in table 4.4. And finally the classifiers shown the results in table 4.1,4.2,4.3 and 4.4 based on confusion matrix of cross validation.

#### 4.1.2 Discussion

Table 4.1 showed the results of linear-SVM classifier, and the best results was recorded for PCA either for HOG with 0.75 for sensitivity and 0.78 for specificity or LBP 0.88 for both sensitivity and specificity and when mixing both it gave lower results. Table 4.2 showed the lowest results when using the kernel of SVM and some of the experiments did not converge and that in the cases of mixing HOG with LBP. Tabel 4.3 showed the best results and stable classifier using Random Forest and best results were for LBP , and this showed how LBP descriptors performed better in this project. Finally in table 4.4 we picked up the highest results from the previous tables and performed BoW and the results were unexpected as this method shall improve the performance. The results shown was degraded as the image was not divided into patches but it was extracting features directly to the image.

Evaluation of individual features from all tables show that the dimensionality reduction of the features improve the results of B- scan classification. PCA is enhancing the results because it deletes the correlated dimensions and gives limited number of dimensions to make the testing and this will prove the quality of the algorithm made. In addition, comparing individual features, LBP proves to be more discriminative than HoG features as LBP results are giving the best results and this can be explained as HOG is edge detection method and LBP is intensity detection method and SERI data classification depends on intensity to classify the disease. Using only Histogram representation, RF classifier leads to the best performance followed by linear- SVM. RBF-SVM classifier has the lowest performance and over-fits for all the individual features while its performance improves when the number of dimensions are reduced using PCA. Based on Tables results, the combination of LBP and HoG features does not improve the results and decreases the performance of individual features. In this test, RF

Table 4.1: Experiments based Linear-SVM.

Features	Evaluation				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
HoG	0.69	0.94	0.91	0.81	0.78
HoG <sup>PCA</sup>	0.75	0.87	0.85	0.80	0.81
LBP <sub>8-ri</sub>	0.63	0.81	0.77	0.69	0.72
LBP <sub>8-ri</sub> <sup>PCA</sup>	0.88	0.88	0.88	0.88	0.88
LBP <sub>8-nri</sub>	0.63	0.81	0.79	0.73	0.75
LBP <sub>8-nri</sub> <sup>PCA</sup>	0.63	0.88	0.83	0.71	0.75
LBP <sub>16-ri</sub>	0.75	0.75	0.75	0.75	0.75
LBP <sub>16-ri</sub> <sup>PCA</sup>	0.75	0.75	0.75	0.75	0.75
LBP <sub>16-nri</sub>	0.69	0.81	0.79	0.73	0.75
LBP <sub>16-nri</sub> <sup>PCA</sup>	0.63	0.88	0.83	0.71	0.75
LBP <sub>24-ri</sub>	0.69	0.88	0.85	0.76	0.78
LBP <sub>24-ri</sub> <sup>PCA</sup>	0.81	0.81	0.81	0.81	0.81
LBP <sub>24-nri</sub>	0.69	0.88	0.85	0.76	0.78
LBP <sub>24-nri</sub> <sup>PCA</sup>	0.69	0.88	0.85	0.76	0.78
HoG <sup>PCA</sup> + LBP <sub>8-ri</sub> <sup>PCA</sup>	0.69	0.81	0.79	0.73	0.75
HoG <sup>PCA</sup> + LBP <sub>16-ri</sub> <sup>PCA</sup>	0.75	0.88	0.86	0.8	0.81
HoG <sup>PCA</sup> + LBP <sub>24-ri</sub> <sup>PCA</sup>	0.69	0.88	0.85	0.79	0.78
HoG <sup>PCA</sup> + LBP <sub>8-nri</sub> <sup>PCA</sup>	0.69	0.75	0.73	0.71	0.72
HoG <sup>PCA</sup> + LBP <sub>8-nri</sub> <sup>PCA</sup> + LBP <sub>16-nri</sub> <sup>PCA</sup> + LBP <sub>24-nri</sub> <sup>PCA</sup>	0.62	0.75	0.71	0.66	0.68
HoG <sup>PCA</sup> + LBP <sub>8-ri</sub> <sup>PCA</sup> + LBP <sub>16-ri</sub> <sup>PCA</sup> + LBP <sub>24-ri</sub> <sup>PCA</sup>	0.69	0.81	0.78	0.73	0.75

Table 4.2: Experiments based Linear-SVM-RBF.

Features	Evaluation				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
HoG	0.94	0.06	0.50	0.65	0.50
HoG <sup>PCA</sup>	0.13	0.88	0.50	0.20	0.50
LBP <sub>8-ri</sub>	0.94	0.25	0.55	0.70	0.59
LBP <sub>8-ri</sub> <sup>PCA</sup>	0.81	0.81	0.81	0.81	0.81
LBP <sub>8-nri</sub>	1.00	0.00	0.50	0.67	0.50
LBP <sub>8-nri</sub> <sup>PCA</sup>	0.81	0.31	0.54	0.65	0.56
LBP <sub>16-ri</sub>	0.88	0.25	0.54	0.67	0.56
LBP <sub>16-ri</sub> <sup>PCA</sup>	0.81	0.88	0.87	0.84	0.84
LBP <sub>16-nri</sub>	1.00	0.00	0.50	0.67	0.50
LBP <sub>16-nri</sub> <sup>PCA</sup>	0.81	0.50	0.62	0.70	0.66
LBP <sub>24-ri</sub>	0.88	0.50	0.64	0.74	0.69
LBP <sub>24-ri</sub> <sup>PCA</sup>	0.75	0.88	0.86	0.80	0.81
LBP <sub>24-nri</sub>	1.00	0.00	0.50	0.67	0.50
LBP <sub>24-nri</sub> <sup>PCA</sup>	0.69	0.81	0.79	0.73	0.75
HoG <sup>PCA</sup> + LBP <sub>8-ri</sub> <sup>PCA</sup>	0.69	0.81	0.79	0.73	0.75
HoG <sup>PCA</sup> + LBP <sub>8-nri</sub> <sup>PCA</sup>	0.00	1.00	NAN	0.00	0.50
HoG <sup>PCA</sup> + LBP <sub>16-ri</sub> <sup>PCA</sup>	0.19	0.94	0.75	0.30	0.56
HoG <sup>PCA</sup> + LBP <sub>16-nri</sub> <sup>PCA</sup>	0.63	0.94	0.50	0.11	0.50
HoG <sup>PCA</sup> + LBP <sub>24-ri</sub> <sup>PCA</sup>	0.00	1.00	NAN	0.00	0.50
HoG <sup>PCA</sup> + LBP <sub>24-nri</sub> <sup>PCA</sup>	0.13	0.94	0.67	0.21	0.53

Table 4.3: Experiments based Random Forest.

Features	Evaluation				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
HoG	0.63	1.00	1.00	0.77	0.81
HoG <sup>PCA</sup>	0.56	0.94	0.90	0.69	0.75
LBP <sub>8-ri</sub>	0.75	0.81	0.80	0.74	0.78
LBP <sub>8-ri</sub> <sup>PCA</sup>	0.75	0.81	0.80	0.74	0.78
LBP <sub>8-nri</sub>	0.69	0.81	0.79	0.73	0.75
LBP <sub>8-nri</sub> <sup>PCA</sup>	0.75	0.81	0.80	0.77	0.78
LBP <sub>16-ri</sub>	0.81	0.88	0.87	0.84	0.84
LBP <sub>16-ri</sub> <sup>PCA</sup>	0.75	0.94	0.92	0.83	0.84
LBP <sub>16-nri</sub>	0.69	1.00	1.00	0.81	0.84
LBP <sub>16-nri</sub> <sup>PCA</sup>	0.75	0.88	0.86	0.80	0.81
LBP <sub>24-ri</sub>	0.69	0.94	0.92	0.79	0.81
LBP <sub>24-ri</sub> <sup>PCA</sup>	0.75	0.94	0.92	0.83	0.84
LBP <sub>24-nri</sub>	0.69	0.94	0.92	0.79	0.81
LBP <sub>24-nri</sub> <sup>PCA</sup>	0.75	0.88	0.86	0.80	0.81
HoG <sup>PCA</sup> + LBP <sub>8-ri</sub> <sup>PCA</sup>	0.63	0.81	0.77	0.69	0.72
HoG <sup>PCA</sup> + LBP <sub>8-nri</sub> <sup>PCA</sup>	0.63	0.81	0.77	0.69	0.72
HoG <sup>PCA</sup> + LBP <sub>16-ri</sub> <sup>PCA</sup>	0.75	0.88	0.86	0.80	0.81
HoG <sup>PCA</sup> + LBP <sub>16-nri</sub> <sup>PCA</sup>	0.63	0.88	0.83	0.71	0.75
HoG <sup>PCA</sup> + LBP <sub>24-ri</sub> <sup>PCA</sup>	0.63	0.88	0.83	0.71	0.75
HoG <sup>PCA</sup> + LBP <sub>24-nri</sub> <sup>PCA</sup>	0.69	0.88	0.85	0.76	0.78

Table 4.4: Classification results using Histogram+PCA+BoW representation.

Histogram + PCA + BoW					
	Classifier	# Words	Metric		
			SE	SP	PRE
HoG <sup>PCA</sup> +LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	10	0.38	0.69	0.55
HoG <sup>PCA</sup> +LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	20	0.31	0.88	0.71
HoG <sup>PCA</sup> +LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	30	0.31	0.94	0.83
HoG <sup>PCA</sup> +LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	40	0.19	0.94	0.75
HoG <sup>PCA</sup> +LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	50	0.38	0.81	0.67
HoG <sup>PCA</sup> +LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	60	0.13	0.88	0.50
LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	10	0.63	0.75	0.71
LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	20	0.75	0.38	0.55
LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	30	0.44	0.44	0.44
LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	40	0.69	0.56	0.61
LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	50	0.50	0.50	0.50
LBP <sub>8-ri</sub> <sup>PCA</sup>	Linear-SVM	60	0.69	0.31	0.50
LBP <sub>16-ri</sub>	RBF-SVM	10	0.56	0.50	0.53
LBP <sub>16-ri</sub>	RBF-SVM	20	0.56	0.63	0.60
LBP <sub>16-ri</sub>	RBF-SVM	30	0.56	0.63	0.60
LBP <sub>16-ri</sub>	RBF-SVM	40	0.69	0.56	0.61
LBP <sub>16-ri</sub>	RBF-SVM	50	0.50	0.50	0.50
LBP <sub>16-ri</sub>	RBF-SVM	60	0.75	0.38	0.55
LBP <sub>16-ri</sub>	RF	10	0.38	0.63	0.50
LBP <sub>16-ri</sub>	RF	20	0.50	0.50	0.50
LBP <sub>16-ri</sub>	RF	30	0.44	0.63	0.54
LBP <sub>16-ri</sub>	RF	40	0.56	0.50	0.53
LBP <sub>16-ri</sub>	RF	50	0.38	0.31	0.35
LBP <sub>16-ri</sub>	RF	60	0.44	0.38	0.41
LBP <sub>16-ri</sub> <sup>PCA</sup>	RF	10	0.69	0.38	0.52
LBP <sub>16-ri</sub> <sup>PCA</sup>	RF	20	0.56	0.75	0.69
LBP <sub>16-ri</sub> <sup>PCA</sup>	RF	30	0.56	0.38	0.47
LBP <sub>16-ri</sub> <sup>PCA</sup>	RF	40	0.50	0.63	0.57
LBP <sub>16-ri</sub> <sup>PCA</sup>	RF	50	0.69	0.50	0.58
LBP <sub>16-ri</sub> <sup>PCA</sup>	RF	60	0.56	0.63	0.60

Table 4.5: Summary of the classification performance in terms of se and sp in(%)

	Evaluation	
	Sensitivity	Specificity
Lemaître <i>et al.</i> [43]	87.5	75.0
Alsaih <i>et al.</i> [20] [19]	87.5	87.5
Srimivansan <i>et al.</i> [56]	68.8	93.8
Venhuizen <i>et al.</i> [60]	61.5	58.8
Sankar <i>et al.</i> [53]	81.3	62.5
Liu <i>et al.</i> [45]	68.8	93.8

and linear-SVM have similar performance while RBF-SVM overfits.

To conclude with Exp1, tables 4.1,4.2 and 4.3 are producing the highest classification performance in this: LBP-PCA and linear-SVM, LBP-PCA and RBF-SVM, LBP16ri and RF, and LBP16ri with RF classifier. These configurations are later tested in using BoW representation. The results obtained from this experiment show that Histogram+PCA+BoW representation decreases the results. In fact, this approach represents each volume in terms of visual-B-scans rather than visual-patches or visual-subvolumes, which could be a reason why BoW fails. The results obtained are compared with different methods applied by other researchers into SERI data as shown in table 4.5 [38].

## 4.2 Segmentation Experiments

### 4.2.1 Experiments

This section explains the results generated by auto-encoder and softmax layer classifier. This experiment starts by extracting the MSER regions from each image and roughly there are 50-200 regions are extracted. The MSER is extracted to cover almost all cysts appeared in an image. Meanwhile, the STAPLE algorithm is applied to the two-ground-truths to make one reference to test the quality of the algorithm built for segmenting the B-scan images to find the cysts. Based on STAPLE, mask is created full of zeros with equal size of B-scan images and only ones are assigned to the STAPLE ground-truth values. After that, images are cropped around the regions extracted by MSER and with the same cropped images size in B-scan images, mask is cropped.

To label the regions, a method is used based on the number of pixels appearing in the mask cropped images and the threshold of the sum of pixels is decided based on the average of pixels appearing and then tried many thresholds and the results shown in table 4.6. This labelling decision is applied to both sets, training and testing data. The training data after cropping

and labelling, each cropped image is resized to 40\*40 to fit into autoencoder.

As auto-encoder accepts only exact size of images to be trained. Auto-encoder trained two hidden layers with 100 and 50 hidden layers for first and second hidden layer respectively. After that, patches output of second autoencoder is classified using softmax layer. With the full deep network formed as a stacked network, test images are reshaped into a matrix by changing the columns into vectors then forming the matrix. Then, assign the test data to the network deepnetto evaluate the test patches, then compare the results obtained from the network with the actual label and analyse the results. To enhance the results, fine tuning (back-propagation)is applied on the whole multi-layer network. To do this, training data shall be reshaped into vectors then concatenate them to form one matrix, like the testing data. Finally, results are validated using confusion matrix and evaluated based on sensitivity, specificity and precision.

#### **4.2.2 Discussion**

As we crop the regions extracted from MSER, we created a mask, which has ones assigned to the empty image based on the STAPLE groundtruth with equal size of images in the original image. Then we crop the mask regions with exact size of MSER regions cropped in original image. To make the label or the threshold in this experiment, we decided to try couple of number of pixels appearing in the mask images, such as if bigger than 300 pixels means any patch has 300 pixels or more label it as cyst otherwise it is background. The results obtained are promising and the threshold of 200 pixels or more is giving a good results. When the 400 pixels or 900 pixels are used as threshold, the results were bad and does not converge as the cysts size is relatively small and can vary from 50 pixels to 450. For 150 pixels, 200 pixels and 250 pixels are giving a very good results as this number of pixels is around the average of sum of pixels for many cysts. Future work for this method is by applying this deep learning network to SERI data and compare it with table 4.5. In addition to that, CNN can be used for classification.

Table 4.6: Auto-encoder Results.

Threshold	Evaluation		
	Sensitivity	Specificity	Precision
Bigger than 0	0.81	0.24	0.46
Bigger than 0 + Fine tuning	0.76	0.41	0.51
Bigger than 100	0.82	0.19	0.48
Bigger than 100 + Fine tuning	0.78	0.37	0.53
Bigger than 150	0.93	0.53	0.68
Bigger than 150 + Fine tuning	0.93	0.82	0.85
Bigger than 190	0.94	0.51	0.67
Bigger than 190 + Fine tuning	0.95	0.75	0.80
Bigger than 200	0.94	0.54	0.69
Bigger than 200 + Fine tuning	0.95	0.79	0.83
Bigger than 210	0.89	0.12	0.51
Bigger than 210 + Fine tuning	0.82	0.51	0.63
Bigger than 250	0.91	0.56	0.68
Bigger than 250 + Fine tuning	0.95	0.76	0.81
Bigger than 300	0.97	0.12	0.54
Bigger than 300 + Fine tuning	0.87	0.57	0.69
Bigger than 400	1.00	0.00	0.55
Bigger than 400 + Fine tuning	0.72	0.26	0.54
Bigger than 900	1.00	0.00	0.70
Bigger than 900 + Fine tuning	1.00	0.00	0.70

# Chapter 5

## Conclusion

Eye diseases such as Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) are the most common causes of irreversible vision loss in individuals with diabetes. DME is defined as the increase in retinal thickness within one disk diameter of the fovea center with or without hard exudate and sometimes associated with cysts. The status of the patient either it has disease or not will define the classification task, and the cyst location will define the segmentation task. Some statistics are showing from [59] that 29.1 million of Americans in the United States are suffering from diabetes, which contributes to 9.3% of the overall population. Diabetes added to 231,404 deaths in the US in 2007. \$245 billion is the total cost of diagnosing the disease in the United States in 2012. With this big number of affected people and the money spent for Research and development (R&D) in this area many algorithms are designed to detect diabetes in the early stage.

This project tracked algorithms made for the analysis of OCT images with focus to classify DME in the first part and classifying the potential regions of the OCT images in the second part. The introduction to the eye anatomy and the entire structure was given, followed by presenting the diseases might attack the eye structure. After that, the work showed of many screening and imaging techniques to give data about the eye. An automatic algorithm was developed to be able to recognize the volumes and classify them either to DME or normal.

This method was first preprocessed by de-nosing images using BM3D and flattening using RANSAC and the cropping was done to avoid wasting time in un-targeted location of an image and to be specific of job. Then, this followed by extracting features using HOG and LBP and the features resulted are represented in reducing the dimensions using PCA and create visual words and dictionary using BoW. After that, the feature vectors are assigned to classifiers like SVM, SVM-RBF and RF. The validation based on the confusion matrix was done to validate the results aspects in term of precision, sensitivity and accuracy.

The classification of potential regions of OCT images was based on extracting MSER and then compare it with the ground-truth given by the raters. Each volume has two ground-truths to be used for referencing of the cyst location, hence the appearance of STAPLE algorithm to create another reference of ground truth based on the two ground-truths. After that, it was assigned to the auto-encoder for training and feature extraction before sending it to softmax layer for further classification of cyst appearance in image.

The results limitation shown in the previous chapter might be for some reasons like the limited number of collection data from hospitals to have general aspect of the method. In addition to that, there is no standard for cyst existence and this can be shown from the appearance of many ground-truths for one image, which is affecting the algorithm accuracy and performance. The hardware also affects the speed of performing the algorithm and thus will limit the number of experiments in the case of deep learning algorithm due to the huge number of data inserted to the neural network.

A recommended way to improve the thesis in the first part is to have many volumes from OCT for training to have a better performance. In addition to that, the reason of failing the bag of words representation in the results might be happened because of the way data extracted using HOG, which is extracted based on images not patches. For the second part, a great hardware aspects shall be provided to ensure the speed of training and make couple of experiments to know which method can perform better. The cysts are not fully covered when we extracted the MSER regions specially if the cyst size is small, hence it is suggested to denoise the image to have a better extraction regions that cover all cysts. Another way can be used instead of auto-encoder, Convolution neural network is a good method to test the OCT classification in next work.

# Bibliography

- [1] Cataracts. <http://marklatina.com/services/cataracts/>.
- [2] Central loss of central vision. <http://goo.gl/TiJ1jI>.
- [3] Diabetic retinopathy. <http://goo.gl/4hFWia>.
- [4] Discriminatory-power-analysis-by-receiver-operatin. <http://goo.gl/qmp9cj>.
- [5] Eye anatomy. <http://goo.gl/GV2Hts>.
- [6] The eye and vision. <http://goo.gl/qHbf5b>.
- [7] Fundus photography. <http://goo.gl/uDdjz6>.
- [8] Glaucoma. <http://goo.gl/BYmJAe>.
- [9] Kernel svm. <https:// goo.g1/ZTGG4G>.
- [10] Linear-svm. <https:// goo.g1/TZD44P>.
- [11] Ocular ultrasonography. <http://goo.gl/TpC9hl>.
- [12] Ophthalmic manifestations of hiv infection. <http://goo.gl/BuAYKn>.
- [13] Progress on retinal degenerative diseases. <http://goo.gl/9xvr4Z>.
- [14] retinal detachment. <http://goo.g1/V35N40>.
- [15] A review of optical coherence tomography angiography (octa). <https:// goo.g1/S08UDk>.
- [16] Sf and ultralsonic. <http://goo.g1/IyVYok>.
- [17] Michael D Abràmoff, Mona K Garvin, and Milan Sonka. Retinal imaging and image analysis. *Biomedical Engineering, IEEE Reviews in*, 3:169–208, 2010.

- [18] I Alhazen. Book of optics. *The optics of Ibn-Haytham*, 1:965–1040, 1989.
- [19] Khaled Alsaih, Guillaume Lemaître, Joan Massich Vall, Mojdeh Rastgoo, Désiré Sidibé, and Fabrice Mériadeau. Machine learning techniques for dme classification on sd-oct images. In *23rd International Conference on Pattern Recognition (ICPR) (Submitted)*, 2016.
- [20] Khaled Alsaih, Guillaume Lemaitre, Joan Massich Vall, Mojdeh Rastgoo, Desire Sidibe, Tien Y Wong, Ecosse Lamoureux, Dan Milea, Carol Y Cheung, and Fabrice Meriaudeau. Classification of sd-oct volumes with multi pyramids, lbp and hog descriptors: application to dme detections. In *38th IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016.
- [21] Diabetes Atlas. International diabetes federation, 2009. *Online version of Diabetes Atlas: www. eatlas. idf. org*, 2013.
- [22] F Bandello, M Battaglia Parodi, P Lanzetta, A Loewenstein, P Massin, F Menchini, and D Veritti. Diabetic macular edema. Karger Publishers, 2010.
- [23] Jessica Beagley, Leonor Guariguata, Clara Weil, and Ayesha A Motala. Global estimates of undiagnosed diabetes in adults. *Diabetes research and clinical practice*, 103(2):150–160, 2014.
- [24] Neil M Bressler, Susan B Bressler, and Stuart L Fine. Age-related macular degeneration. *Survey of ophthalmology*, 32(6):375–413, 1988.
- [25] B Cense, TC Chen, N Nassif, MC Pierce, SH Yun, BH Park, BE Bouma, GJ Tearney, and JF De Boer. Ultra-high speed and ultra-high resolution spectral-domain optical coherence tomography and optical doppler tomography in ophthalmology. *BULLETIN-SOCIETE BELGE D OPHTALMOLOGIE*, 302:123, 2006.
- [26] Barry Cense, Nader A Nassif, Teresa C Chen, Mark C Pierce, Seok-Hyun Yun, B Hyle Park, Brett E Bouma, Guillermo J Tearney, and Johannes F de Boer. Ultrahigh-resolution high-speed retinal imaging using spectral-domain optical coherence tomography. *Optics Express*, 12(11):2435–2447, 2004.
- [27] Retina Society Terminology Committee et al. The classification of retinal detachment with proliferative vitreoretinopathy. *Ophthalmology*, 90(2):121–125, 1983.
- [28] Andrea Cusumano, D Jackson Coleman, Ronald H Silverman, Dan Z Reinstein, Mark J Rondeau, Roxana Ursea, Suzanne M Daly, and Harriet O Lloyd. Three-dimensional ultrasound imaging: clinical applications. *Ophthalmology*, 105(2):300–306, 1998.

- [29] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, 16(8):2080–2095, 2007.
- [30] Luis de Sisternes, Julia Hu, Daniel L Rubin, and Michael F Marmor. Localization of damage in progressive hydroxychloroquine retinopathy on and off the drug: Inner versus outer retina, parafovea versus peripheral fovearetinal layers in progressive hcq retinopathy. *Investigative ophthalmology & visual science*, 56(5):3415–3426, 2015.
- [31] Ana M Fernández-Alonso, José L Cuadros, Peter Chedraui, Marcela Mendoza, Ángela M Cuadros, and Faustino R Pérez-López. Obesity is related to increased menopausal symptoms among spanish women. *Menopause international*, 16(3):105–110, 2010.
- [32] Carel B. Hoyng Thomas Theelen Bram van Ginneken Freerk G. Venhuizen, Mark J.J.P. van Grinsven and Clara I. Sanchez. Vendor independent cyst segmentation in retinal sd-oct volumes using a combination of multiple scale convolutional neural networks. In *SPIE Medical Imaging*, pages 94141I–94141I. International Society for Optics and Photonics, 2015.
- [33] Daniel C Gray, Robert Wolfe, Bernard P Gee, Drew Scoles, Ying Geng, Benjamin D Masella, Alfredo Dubra, Sergio Luque, David R Williams, and William H Merigan. In vivo imaging of the fine structure of rhodamine-labeled macaque retinal ganglion cells. *Investigative ophthalmology & visual science*, 49(1):467–473, 2008.
- [34] Collaborative Normal-Tension Glaucoma Study Group et al. Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures. *American journal of ophthalmology*, 126(4):487–497, 1998.
- [35] Otto Hockwin, Viorel Dragomirescu, Heike Laser, Alfred Wegener, and Ursula Eckerskorn. Measuring lens transparency by scheimpflug photography of the anterior eye segment: principle, instrumentation, and application to clinical and experimental ophthalmology. *Journal of Toxicology: Cutaneous and Ocular Toxicology*, 6(4):251–271, 1987.
- [36] J Edward Jackson. *A user’s guide to principal components*, volume 587. John Wiley & Sons, 2005.
- [37] Tatjana C Jakobs, Richard T Libby, Yixin Ben, Simon WM John, and Richard H Masland. Retinal ganglion cell degeneration is topological but not cell type specific in dba/2j mice. *The Journal of cell biology*, 171(2):313–325, 2005.

- [38] Guillaume Lematre Carol Y. Cheung Tien Y. Wong Desire Sidibe Fabrice Meriaudeau Joan Massich, Mojdeh Rastgoo. Classifying dme vs normal sd-oct volumes: A review. In *23rd International Conference on Pattern Recognition (ICPR)(Submitted)*, 2016.
- [39] Jost B Jonas, Ulrike Schneider, and Gottfried OH Naumann. Count and density of human retinal photoreceptors. *Graefe's archive for clinical and experimental ophthalmology*, 230(6):505–510, 1992.
- [40] Jayanthi Sivaswamy Karthik Gopinath. Domain knowledge assisted cyst segmentation in oct retinal images. In *Optima Ophthalmic Image Analysis Challenge, Medical Image Computing and Computer Assisted Interventions (MICCAI) 2015*, 2015.
- [41] Stephanie H Kim, Craig J Galbán, Robert J Lutz, Robert L Dedrick, Karl G Csaky, Martin J Lizak, Nam Sun Wang, Ginger Tansey, and Michael R Robinson. Assessment of subconjunctival and intrascleral drug delivery to the posterior segment using dynamic contrast-enhanced magnetic resonance imaging. *Investigative ophthalmology & visual science*, 48(2):808–814, 2007.
- [42] Barbara E Klein and Ronald Klein. Cataracts and macular degeneration in older americans. *Archives of Ophthalmology*, 100(4):571, 1982.
- [43] Guillaume Lemaître, Mojdeh Rastgoo, Joan Massich, Shrinivasan Sankar, Fabrice Mériadeau, and Désiré Sidibé. Classification of sd-oct volumes with lbp: Application to dme detection. 2015.
- [44] Cheng-Yuan Liou, Wei-Chen Cheng, Jiun-Wei Liou, and Daw-Ran Liou. Autoencoder for words. *Neurocomputing*, 139:84–96, 2014.
- [45] Yu-Ying Liu, Mei Chen, Hiroshi Ishikawa, Gadi Wollstein, Joel S Schuman, and James M Rehg. Automated macular pathology diagnosis in retinal oct images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. *Medical image analysis*, 15(5):748–759, 2011.
- [46] Theodore Leng Daniel L. Rubin Luis de Sisternes, Jerry Hong. A machine learning aproach for device-independent automated segmentation of retinal cysts in spectral domain optical cohorence tomography images. In *Optima Ophthalmic Image Analysis Challenge, Medical Image Computing and Computer Assisted Interventions (MICCAI) 2015*, 2015.
- [47] Hossein Rabbani Fedra Hajizadeh Mahdad Esmaeili, Alireza Mehri Dehnavi. 3d segmen-tation of retinal cysts from sd-oct images by the use of three dimensional curvelet based k-svd. In *Optima Ophthalmic Image Analysis Challenge, Medical Image Computing and Computer Assisted Interventions (MICCAI) 2015*, 2015.

- [48] Jiri Matas, Ondrej Chum, Martin Urban, and Tomás Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10):761–767, 2004.
- [49] Connie S McCaa. The eye and visual nervous system: anatomy, physiology and toxicology. *Environmental health perspectives*, 44:1, 1982.
- [50] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [51] Ipek Oguz, Li Zhang, Michael D Abràmoff, and Milan Sonka. Optimal retinal cyst segmentation from oct images. In *Optima Ophthalmic Image Analysis Challenge, Medical Image Computing and Computer Assisted Interventions (MICCAI) 2015*, 2016.
- [52] Thais Mayumi Oshiro, Pedro Santoro Perez, and José Augusto Baranauskas. How many trees in a random forest? In *MLDM*, pages 154–168. Springer, 2012.
- [53] S Sankar, D Sidibé, Y Cheung, TY Wong, E Lamoureux, D Milea, and F Meriaudeau. Classification of sd-oct volumes for dme detection: an anomaly detection approach. In *SPIE Medical Imaging*, pages 97852O–97852O. International Society for Optics and Photonics, 2016.
- [54] Chanjira Sinthanayothin, James F Boyce, Helen L Cook, and Thomas H Williamson. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *British Journal of Ophthalmology*, 83(8):902–910, 1999.
- [55] Richard F Spaide, James M Klancnik, and Michael J Cooney. Retinal vascular layers imaged by fluorescein angiography and optical coherence tomography angiography. *JAMA ophthalmology*, 133(1):45–50, 2015.
- [56] Pratul P Srinivasan, Leo A Kim, Priyatham S Mettu, Scott W Cousins, Grant M Comer, Joseph A Izatt, and Sina Farsiu. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomedical optics express*, 5(10):3568–3577, 2014.
- [57] Susan A Strenk, Lawrence M Strenk, John L Semmlow, and J Kevin DeMarco. Magnetic resonance imaging study of the effects of age and accommodation on the human lens cross-sectional area. *Investigative ophthalmology & visual science*, 45(2):539–545, 2004.
- [58] Hao Sun, Xian Sun, Hongqi Wang, Yu Li, and Xiangjuan Li. Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model. *Geoscience and Remote Sensing Letters, IEEE*, 9(1):109–113, 2012.

- [59] National Diabetes Data Group (US), National Institute of Diabetes, Digestive, Kidney Diseases (US), and National Institutes of Health (US). *Diabetes in America*. Number 95. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 1995.
- [60] Freerk G Venhuizen, Bram van Ginneken, Bart Bloemen, Mark JJP van Grinsven, Rick Philipsen, Carel Hoyng, Thomas Theelen, and Clarisa I Sánchez. Automated age-related macular degeneration classification in oct using unsupervised feature learning. In *SPIE Medical Imaging*, pages 94141I–94141I. International Society for Optics and Photonics, 2015.
- [61] KA Vermeer, J Van der Schoot, HG Lemij, and JF De Boer. Automated segmentation by pixel classification of retinal layers in ophthalmic oct images. *Biomedical optics express*, 2(6):1743–1756, 2011.
- [62] Theo Vos, Abraham D Flaxman, Mohsen Naghavi, Rafael Lozano, Catherine Michaud, Majid Ezzati, Kenji Shibuya, Joshua A Salomon, Safa Abdalla, Victor Aboyans, et al. Years lived with disability (ylds) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the global burden of disease study 2010. *The Lancet*, 380(9859):2163–2196, 2013.
- [63] Simon K Warfield, Kelly H Zou, and William M Wells. Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. *Medical Imaging, IEEE Transactions on*, 23(7):903–921, 2004.
- [64] CP Wilkinson, Frederick L Ferris, Ronald E Klein, Paul P Lee, Carl David Agardh, Matthew Davis, Diana Dills, Anselm Kampik, R Pararajasegaram, Juan T Verdague, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, 110(9):1677–1682, 2003.
- [65] Anne S Yeager, F Carl Grumet, Elizabeth B Hafleigh, Ann M Arvin, John S Bradley, and Charles G Prober. Prevention of transfusion-acquired cytomegalovirus infections in newborn infants. *The Journal of pediatrics*, 98(2):281–287, 1981.