JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF BIOMEDICAL ENGINEERING

BIOMEDICAL IMAGING STREAM

MASTER'S THESIS

Automated Detection and Classification of Glaucoma and Macular Degeneration Using Convolutional Neural Network

Author: Muluken Gerbi

A Thesis Report Submitted to School of Graduate Studies of Jimma Institute of Technology, Jimma University, in Partial Fulfillment for the Degree of Master of Science in Biomedical Engineering (Biomedical Imaging).

Jimma, Ethiopia

May, 2020

JIMMA UNIVERSITY

JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF BIOMEDICAL ENGINEERING

BIOMEDICAL IMAGING STREAM

MASTER'S THESIS

Automated Detection and Classification of Glaucoma and Macular Degeneration Using Convolutional Neural Network

Author: Muluken Gerbi

A Thesis Report Submitted to School of Graduate Studies of Jimma Institute of Technology, Jimma University, in Partial Fulfillment for the Degree of Master of Science in Biomedical Engineering (Biomedical Imaging)

Advisor: Gizeaddis L. Simegn (Ph.D.)

Co-Advisor: Kokeb Dese (M.Sc.)

Jimma, Ethiopia

May, 2020

Declaration

This thesis entitled with "**Automated Detection and Classification of Glaucoma and Macular Degeneration Using Convolutional Neural Network"** is my original work and has not been presented for a degree in any other university.

Done By:

Muluken Gerbi _____ _____

Signature Date

We the advisors of this thesis with the title "**Automated Detection and Classification of Glaucoma and Macular Degeneration Using Convolutional Neural Network**" confirm that this research is approved as an MSc thesis for the student.

Gizeaddis Lamesgin (Ph.D.) _____ _____
Advisor Signature Date

Kokeb Dese (M.Sc.) _____ _____
Co-Advisor Signature Date

Chair: Name _____

_____ _____
Signature Date

Internal Examiner Name: _____

_____ _____
Signature Date

External Examiner Name: _____

_____ _____
Signature Date

Acknowledgments

First of all, I would like to give my heartfelt thanks to my almighty God for helping me with all my works. Secondly, I would like to express my deepest gratitude to my advisor Dr. Gizeaddis Lamesgin and co-advisor Mr. Kokeb Dese for helping me in all aspects of the research. Next, I would like to give thanks to Dr.Timothy Kwa, Mr. Hakkins Raj, Mr. Samuel Sisay and Ms.Selamawit Hadush for helping while starting the research and also to all members of the BME staff. Finally, I need to thank Dr. Kume from JUMC (Jimma University Medical Center) Ophthalmology Department staff, Mr.Biruk from the Korean Hospital Ophthalmology Department and to all my classmate's students.

Abstract

Eye is the most complex and sensitive organs of our body which enables us to visualize things. Globally, around 1.3 billion people will live with visual impairment. Most eye diseases may gradually lead to permanent vision loss if not detected and treated early. Glaucoma is a chronic eye disease often called "silent thief of sight" as it has no symptoms. If it's not detected at an early stage, it may cause permanent blindness. Age-related macular degeneration is another chronic eye disease that causes blurred or reduced central vision leading to blindness. Regular eye health checkup is required, especially for the elders, to avoid blindness. Tonometry, ophthalmoscopy, visual field test, and optical coherent tomography are used to diagnose glaucoma and age-related macular degeneration. Most of the modern imaging devices are expensive and unaffordable for health facilities in low resource setting. Also, diagnostic methods are manual, time-consuming and exposed to intra/inter observer variability and require expert supervision.

Most works are focused on binary classification of eye disease. However, grading of the disease is also an important diagnosis factor that is used to identify the disease progression. Segmenting cup and disk areas are one of challenging task in the area. Segmentation of OC and OD using traditional machine learning approach is difficult for extraction of deep feature and learning hidden pattern by itself which reduce computation time, reliability and performance of the system.

This thesis presented automated glaucoma and AMD detection using a convolutional neural networks approach from fundus image. Glaucoma detection and grading have been performed based on the cup to disk ratio calculation after the segmentation of OC and OD features. The experiments shows that Unet++ achieves state-of-the-art OD and OC segmentation results compared to Unet. The detection accuracy was found to be 91%. AMD classification has also been done with 97% accuracy using the VGG16 model.

Keywords: AMD, Glaucoma, OC, OD, Unet, VGG16.

Table of Content

List of Tables

List of Figures

Acronyms

AMD → Related Macular Degeneration

ANN → Artificial Neural Network

CAD → Computer-Aided Diagnosis

CNN → Convolutional Neural Network

CDR → Cup to Disc Ratio

CLAHE → Contrast Limited Adaptive Histogram Equalization

DL → Deep Learning

DMD → Dry Macular Degeneration.

DR → Diabetic Retinopathy

DWT → Discrete Wavelet Transform

EWT → Empirical Wavelet Transform

FCN → Fully Connected Neural Network

FOV → Filed of View

GPU → Graphical Processing Unit

GUI → Graphical User IInterface.

IOP → Intraocular Pressure

JUMC → Jimma University Medical CCenter.

LMIC → Low- And Middle-Income Countries

OCT → Optical Coherence Tomography

OC → Optic Cup

OD → Optic Disk

ONH → Optic Nerve Head

Mask RCNN → Masked Region-based Convolutional Neural Network

MICCAI → Medical Image Computing and Computer-Assisted IIntervention.

PPA → Peripapillary Atrophy.

RELU → Rectified Linear Unit

ROI → Region of Interest

RNFL → Retinal Nerve Fiber Layer

SDC → Sparse Dissimilarity Constrained Coding

SVM → Support Vector Machine

WHO → World Health Organization

WMD → Wet Macular Degeneration

Chapter 1: Introduction

1.1 Background

Eye is the most complicated and sensitive organ of our body which is important to visualize things. Major parts of the eye are cornea, tear duct, iris, pupil, lens, optic nerve and retina. The weakening of one or more of these parts may cause a reduction of vision and visual impairment leading to blindness. Globally, the prevalence of moderate, severe visual impairment and blindness is 285 million, with 65% of visually impaired people and 82% of all blind people being 50 years and older [1]. Age-related vision reduction happens to many people as they grow older.

Eye diseases like age-related macular degeneration, glaucoma, cataract, and diabetic retinopathy are the most common ones. Age-related macular degeneration (AMD) is occurred by gradual damage to the cells of the macula and it is the most common among the elder. Glaucoma is caused by the increasing pressure of fluid inside our eyes. It is the second leading cause of blindness. But, cataract is occurred by clouding of the lens of the eye. AMD and glaucoma occurrence is shown by a structural change like retinal nerve, macula, cup and disk size. Figure 1.1 shows part of the eye where age-related diseases occur prevalently.

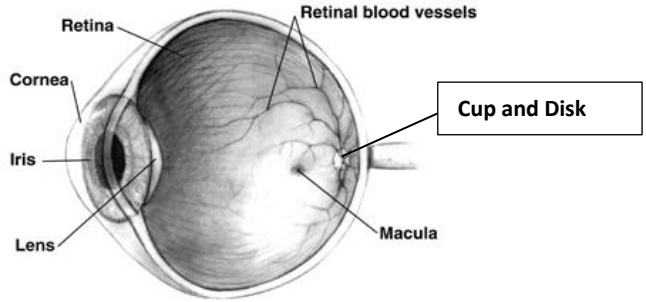

Figure 1.1: Shows part of the eye where age-related disease occurred.

1.2 Age-related eye disease.

Aging sometimes brings changes that weaken our vision and eyes, but there are things you can do to maintain lifelong eye and overall health. The solution may be as simple as using brighter lights around the house to help prevent accidents caused by weak eyesight or seeing your doctor more frequently to screen for age-related diseases, there is a high probability of developing age-related eye disease including age-related macular degeneration, glaucoma, retinopathy, and cataract.

1.2.1 Glaucoma and Types of Glaucoma

Glaucoma is the 2nd leading cause of blindness in the world following to cataract [2] and the third leading cause in Ethiopia [3]. Glaucoma is caused by increased intraocular pressure (IOP) of the eye, IOP caused by blockage of fluidic humor flowing outside channel of the eye. When a blockage occurs in the channel of the eye, intraocular pressure will increase. That leads to damage to the optic nerve. Because of damage to optic nerve the communication between the retina and brain will stop, which means blindness happens to the patient. It's painless and didn't show immediate symptoms, vision loss may happen gradual without noticing the disease.

There are two main kinds of glaucoma open-angle and angle-closure glaucoma. Open-angle glaucoma is the most common type, sometimes called wide-angle glaucoma. The drain structure or trabecular meshwork looks normal, but fluid doesn't flow out as it should. But, angle-closure glaucoma is less common and sometimes called acute angle-closure glaucoma. This occurs when the eye doesn't drain and the space between iris and cornea becomes too narrow.

If glaucoma is detected early there is some curing method to prevent vision loss like opening the channels by surgery [4]. Since the eye screening process is time consuming and tedious process, however, digital fundus image can be used to effectively diagnose glaucoma by using image features like a cup to disk ratio, optic head and cup diameter. Using a fundus camera is a cost-effective and easy method for the diagnosis of glaucoma [5]. As shown in Figure 1.2 diagnosis of glaucoma using the necked eye is very difficult. So, it is very important to extract the best features like a cup to disc ratio (CDR) and retinal nerve fiber layer (RNFL). Figure 1.2 A shows the normal

fundus image and 1.2 B shows cupping of in the retinal fundus image and indicates the presence of glaucoma but, difficult to quantify without segmenting the cup and disk.

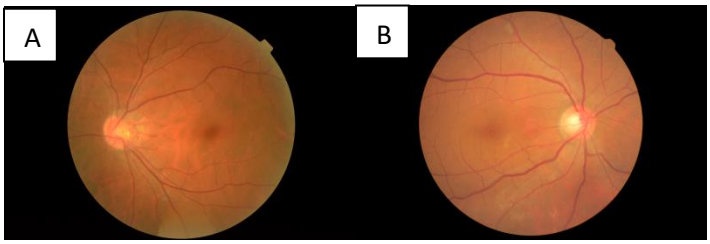

Figure 1.2: A normal fundus image, B is a glaucoma fundus image.

1.2.2 Age-related macular degeneration.

AMD is a medical condition which results in blurred or no vision that occurs in the center of the eye. It is caused due to damage to the macula of retina. AMD is chronic eye disease because, it does't show any symptoms at an early stage. Most of the times AMD may not result in complete blindness, but it can cause central vision loss and makes hard to recognize faces, drive, read or perform other activity of daily life.

The is a progressive accumulation of yellow deposits called drusen (buildup of extracellular proteins and lipids) in the macula (a part of the retina), between the retinal pigment epithelium and the underlying choroid. This accumulation is believed to damage the retina over time. Amyloid-beta builds up in Alzheimer's disease brains, is one of the proteins that accumulated in AMD. This is the reason why AMD is sometimes called "Alzheimer's of the eye" or "Alzheimer's of the retina". There's no treatment for dry macular degeneration. But, if the AMD condition is diagnosed early, there are steps to take to help slow its progression, such as taking vitamin supplements, eating healthy foods and not smoking. So, early detection of AMD is very important because the treatments can delay or reduce the severity of the disease.

Dry AMD type is a common eye disorder that may lead to a gradual loss of visual acuity. But, the wet type which is more severe than dry type. The progression of dry AMD can be slowed in many patients through dietary supplements when detected at an early stage. Wet AMD type is also sometimes called choroidal neovascularization (CNV), qualified by ingrowth of a choroidal vascular structure into the macula followed by increased vascular permeability. The increase in vascular permeability leads to abnormal fluid collection within or below the retina that causes visual dysfunction. Each type requires different approaches in treatment the impact of AMD is precarious without timely treatment. Therefore, patients in rural areas lacking access to ophthalmologists have a higher risk of blindness caused by AMD.

1.3 Glaucoma and Macular Degeneration in Ethiopia.

According to World Health Organization(WHO), glaucoma is the second leading cause of avoidable blindness after cataract, contributing 8% of total blindness worldwide, 15% in are low- and middle-income countries (LMIC), particularly in sub-saharan Africa [6]. Blindness and low vision are major public health issues in Ethiopia. The national prevalence of blindness in Ethiopia is 1.6% (1.1% for urban and 1.6% for rural populations) and that of low vision is 3.7% (2.6% for urban and 3.8% for rural populations) [7]. Ethiopia has low access to eye health care services. Lack of awareness about age-related disease and poor health-seeking behavior observed in people living in rural areas that may result in the late diagnosis of the disease with severe visual impairment. Furthermore, most of the people from the rural area had poor income to afford their diagnostic cost and treatment-related expenses.

1.4 Problem statement

Nowadays, glaucoma and AMD are the most common causes of blindness in the world. Glaucoma is predicted to affect about 80 million people by 2020 [8]. Glaucoma is a chronic and irreversible eye disease that can lead to vision loss by slowly damaging the optic nerve. Although it is incurable, the effect can be slowed down by treatment if detected early. On the other hand, AMD is the 3rd leading cause of blindness that affected approximately 20–25 million people across the

world according to the WHO 2015 survey. Among these, approximately 8 million peoples have lost vision due to AMD only [9].

Currently, eye disease diagnosis methods in developing countries are dependent on manual diagnostic devices like tonometry and ophthalmoscopy. Modern imaging methods like coherent tomography (OCT) are expensive and not available, they usually require highly skilled ophthalmologists to diagnose glaucoma and macular degeneration.

In this thesis, automatic detection of glaucoma and classification of AMD has been proposed using a deep learning approach. It comes up with the idea of using a widely available device like a fundus camera and to build an image processing algorithm to easily detect and manage the progression of such disease with less expert knowledge.

1.5 Significance of the research

This study will help the health facilities to easily screen glaucoma and AMD patient with better accuracy, less duration of time and without using highly experienced ophthalmologists. It is the best method for a country like Ethiopia which has a high number of patient flow. Nowadays, most hospitals in Ethiopia use manual estimation of cup to disc ratio using slit lamp for glaucoma and AMD diagnosis. Manual diagnosis is very difficult, time-consuming, it very hard for follow-up, sensitive and it is directly dependent on expertise skill and knowledge. Because of the lack of automated diagnostic devices, cost of the diagnosis, very poor health awareness and unavailability of the skilled person most peoples suffered for visual impairment.

CAD (computer-aided diagnosis) screening systems can reduce the time and effort wasted on the analysis of glaucoma eye disease. It is also important to develop CAD systems using image analysis for clinical experts to differentiate between normal and glaucoma retinal images because it is difficult for ophthalmologists to make this discrimination and to follow the progress of glaucoma. Rather the patient may suffer for unnecessary surgery or blindness.

1.6 General objective

The general objective of this thesis is to develop an automatic glaucoma and age-related macular degeneration detection and classification system using a convolutional neural network algorithm.

1.7 Specific objectives

- ✓ To grade glaucoma stages using convolutional neural network architectures.
- ✓ To segment cup and disk from retinal fundus image.
- ✓ To classify age-related macular degeneration.

Chapter 2: Diagnosis System and Image Analysis

2.1 Diagnosis system

Clinical diagnosis for glaucoma is done using different techniques depending on the ophthalmologist's need and available equipment. The most common test methods are tonometry, pachymeter, visual field test, ophthalmoscope, gonioscope, and computer-aided diagnostic technology.

Tonometry: Measure the IOP of the eye. It is easy, quick and painless tests but, it is less accurate, manual and indirect measurement method.

Pachymeter: It measures the central corneal thickness of the eye. It uses anesthetics and a small probe to measure the IOP of the eye. This method is also an indirect measurement for glaucoma.

Visual field test: Measures the extent of damage to the optic nerve from elevated IOP. This is to the measurement of peripheral vision loss. This test is an advanced test, but it is very expensive.

Ophthalmoscope: is an instrument which used to see directly inside of the optic nerve through the pupil. To diagnose glaucoma and AMD based on the color and appearance of the retina.

Gonioscopy: together with a slit lamp or operating microscope to view the irido-corneal angle or the anatomical angle formed between the eye's cornea and iris. It is important in diagnosing and monitoring various eye conditions associated with glaucoma.

2.1.1 Computer-aided diagnosis (CAD)

Computer-aided diagnosis (CAD) of glaucoma and AMD is performed by various structural and texture feature extraction techniques of the fundus image using different algorithms. It is a cost-effective method and has less inter/intra-observer variability. When compared to the manual detection method, CAD takes much less time [8, 9]. Digital fundus image is one of the best retinal imaging modalities used to diagnose glaucoma and AMD. The development of glaucoma is manifested by the structural and functional changes of the retinal fundus image. This can be used as an indicator of the glaucoma level through the cup to disc ratio, RNFL, disc diameter, and

peripapillary atrophy (PPA) [12]. The optic cup to disc ratio is one of the principal physiological characteristics which is employed for early detection of glaucoma [13].

2.1.2 Cup to disc ratio-based glaucoma detection

The manual estimation of a cup to disc ratio (CDR) is a widely used parameter to diagnose glaucoma in its early stages. However, this process of diagnosing glaucoma is time consuming and has limited potential in population-based screening for early detection. Moreover, this subjective evaluation is susceptible to intra and inter-observer errors and clinical experiences play an important role in glaucoma detection and grading [14]. To reduce subjective assessment errors, different clinical indicators or metrics are designed to aid in the quantitative grading of glaucoma. Among this metrics, the ratio of the optic cup to disc, known as the cup-to-disc ratio (CDR) [15], is very popular and widely used for detecting and classifying glaucoma early. The appearance and size of the optic cup change rapidly during glaucoma, hence, it is considered as a commonly used indicator for glaucoma identification and detection [8].

2.1.3 Related work for diagnosis of Glaucoma and AMD

In literature, automatic detection and classification of glaucoma through the cup to disc ratio (CDR) and extraction of texture features have been used widely. However, it is very difficult for the CAD system to segment the cup and disk regions in a robust manner. To extract features, there is a need to segment the optic disk (OD) and optic cup (OC) regions that require lots of image processing techniques. After defining image features, there is also need domain-expert knowledge to select the most discriminative features. Therefore, the detection of glaucoma is a challenging task for ophthalmologists and CAD systems [16]. The different techniques of glaucoma and AMD classification and detection are discussed below.

Non-model-based OD and OC segmentation approaches include thresholding techniques, morphological operations, and super-pixel classification-based methods. These non-model-based methods take less time as compared to model-based techniques [17]. The boundary detection mainly uses some algorithms such as thresholding, morphological operations and pixel clustering.

However, the identification process of the optic disc boundary had some problems due to obscuration by bridging retinal vessels. This approach is not widely used for glaucoma detection.

Active contours and deformable models-based segmentation are also widely used for OC and OD segmentation. They segmenting the objects having geometrical shapes. Active contour modeling (ACM) and Gradient vector flow (GVF) models are commonly one [18, 20, 21]. This approach refers to the shape approximation of an object using statistical approaches to learn the boundary shape of the optic disc from a training set.

Machine and deep learning-based approach, recently deep learning-based methods have shown promising results for OD and OC segmentation [21]. Most feature extraction techniques involve simplifying the number of resources required to describe a large data set accurately. Using instead of extracting features manually, deep learning requires only a set of data with minor preprocessing, if necessary, feature extraction is less tedious and saves a lot of time also [28, 29]. Pixel-level feature extraction with the pattern is difficult for the traditional machine learning approach [30, 31]. And also, when the sample size and diversity are increased in number, shallow machine learning models cannot adapt to the increasing complexity [26]. But, deep learning is best approach to extract dominat freature and to discriminate different features in more robust manner. Related works under all approaches stated above are discussed below in chronological order with pros and cons.

In 2015, regional wavelet features of optic nerve head(ONH) were used to discriminate between glaucoma and normal fundus image. Based on regional wavelet features of ONH they have conducted the classification based on wavelet energy of different frequency bands. They used 158 images from RIMONE database. They have achieved 93% accuracy. The research presented that wavelet-based feature extraction can accurately classify glaucoma than another methods. Classification accuracy of normal and glaucoma images using regional that wavelet features of different regions with 93% outperforms all other feature sets. Determination of wavelet features at different decomposition level wavelet analysis at each decomposition level is not performed [27].

Other research in 2015, automatic detection of glaucoma by using features learning through a deep-learning algorithm on retinal fundus images is proposed [28]. The CNN model learn the features with a linear and nonlinear activation function. They used glaucoma and non-glaucoma patterns to differentiate for the training of the CNN model. They performed experiments on the ORIGA and SCES datasets and reported 0.838 and 0.898 of AUC values, respectively. Even though they used the state of art image classification technique, but the performance of the system is not good [28].

In 2016, a novel algorithm was developed to detect glaucoma using a support vector machine (SVM) instead of using an advanced deep-learning algorithm along with the hybrid feature set [2]. The hybrid feature set for glaucoma detection and distinguished color and texture features from retinal fundus images and some other properties of CDR of OD/CUP ratio to determine the severity-level of glaucoma eye disease. The system also designed for glaucoma detection in rural areas where lack of expertise is available. This developed approach was evaluated on 100 patients and they obtained 100% average sensitivity and 87 % specificity, respectively. But, the system might not perform accurately as expected in big data and small OC segmentation [2].

In 2017, glaucoma detection from the retinal fundus image using deep learning CNN architecture for extracting features and deep – belief network is used for selecting discriminative features [29]. This system also used images for training and testing. They conclude that the system can easily diagnose glaucoma in a large-scale environment, but the system can't incorporate another disease that can be detected by using a fundus image like AMD. Furthermore, another paper in 2017 proposed optic disc and cup segmentation methods for glaucoma detection by modifying U-Net convolutional neural networks [30]. They experimented on publicly available databases like DRIONS-DB, RIM-ONE v.3 and DRISHTI-GS. But the system gets difficulty in cup segmentation and no comparison was made with other deep learning segmentation methods.

Later in 2017, the adaptive deformable model used for the automatic segmentation of optic disc and cup to aid glaucoma diagnosis. This research provides robustness in capturing a range of variations and different shapes. The optic cup segmentation results are highly dependent on the

accurate segmentation of the optic disc segmentation. However, optic cup segmentation is difficult include the normal images which have a very small cup size [18].

Other research in 2017, used machine vision for the screening of age-related macular degeneration using fundus images. They have used image processing techniques like media filtering, binary thresholding, and total white pixel ratio to classify AMD. But, this system aims only classifying the type of AMD and grading is not performed [31]. Additionaly, another research in this year developed to automatically detect AMD from fundus images using anovel application of deep learning methods. Deep convolutional neural networks that are explicitly trained for performing automated AMD grading were compared with an alternate deep learning method that used transfer learning and universal features. The deep convolutional neural network method yielded accuracy that ranged between 88.4% and 91.6% with different neural network architecture. The data set class imbalance regarding the number of fundus images and the moderate level have a effect on performance of the system [32].

In 2018, Proposed OD and OC segmentation using M-Net deep learning architecture and polar transformation with the multi-label deep learning concept and achieved AUC of 0.89 and 0.85 respectively for SCES and ORIGA dataset [33]. The work is incompetent of achieving robustness and accuracy enough for evaluating the optic disc and cup in clinical practice. The work considered smooth segmentation of cup and disk but, some shape irregularity may be missed while using polar transform. In other papers in 2018, deep learning-based algorithm identifies glaucoma disks by using a microscopic fundus image [34]. They have used image data from ydney-based clinical studies, RIMONE and HRF. They have achieved 92.6%accuracy of, 89.3% sensitivity and 97.1% specificity. But the cup segmentation is not included in the work.

The gaps reviewed in the literature show most works are focused on binary classification of glaucoma. However, grading of the disease is also an important diagnosis factor that is used to identify the disease progression. The literature also lacks to try the system for diagnosis of multi ocular disease, diagnosis of like AMD and glaucoma from fundus images. Researching on finding a robust fast and accurate system for diagnosing ocular disease is very important.

This thesis introduces a new approach for the detection of glaucoma and AMD using fundus images. A new condition of segmentation and classification of glaucoma fundus image has been proposed by comparing Unet and Unet++ segmentation performance of the CDR feature. Additionally, the implementation of the VGG16 model for detecting AMD.

2.2 Image Analysis

2.2.1 Artificial Neural Networks

Artificial neural networks are famous machine learning algorithms that have been used since the 1950s. The neural network consists of neuron, which is the computational unit of the network. There's an input layer where data enters the network, followed by one or more hidden layers transforming the data as it flows through, before ending at an output layer that produces the neural network's predictions. Figure 2.1 shows the simplest model of a multilayer perception's neural network [26].

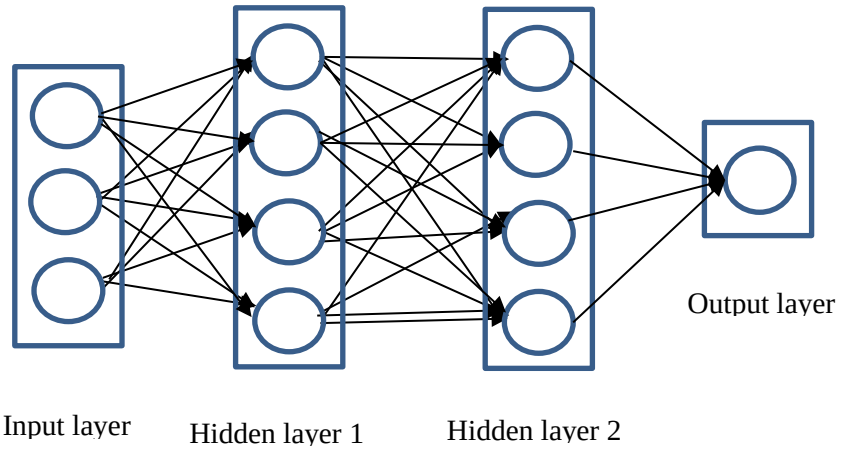

Figure 2.1: Simplest model class of neural network is the multilayer perception [35].

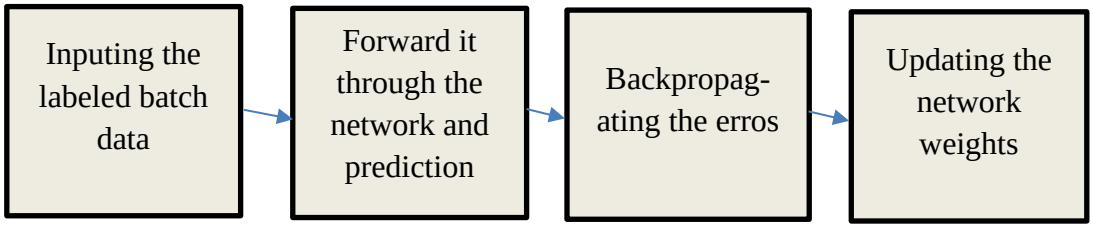

Figure 2.2: Workflow of training a neural network.

The basic ideas behind training neural networks are simple. The training data is fed through the network, compute the gradient of the loss function concerning every weight using the chain rule, and reduce the loss by changing these weights using gradient descent. It is based on simple computational units, called neurons, organized in layers. Figure 2.2 summarize general workflow of the training process of the neural network. The input is then multiplied by some random weight, then add bias and activation will be applied. On the next layer, the neuron will take the output of the previous layer neuron. And it continues like this until it reaches output neuron then, the final value is representing the probability that the initial input belongs to one's class.

2.2.2 Convolutional Neural Network (CNN)

Deep learning is mostly triggered by convolutional neural networks. Convolutional neural networks were inspired by the biological processes of the brain. CNN's are regularized versions of multilayer perception's. They take advantage of the hierarchical pattern in data and assemble more complex patterns using smaller and simpler patterns. Therefore, on the scale of connectedness and complexity, CNN's are on the lower extremity , it takes advantage of little pre-processing, better feature learning performance and speed.

2.2.3 Deep Features Extraction

Deep learning (DL) is an active research topic that learns discriminative representations of data. The DL architectures are formed by the composition of multiple linear and non-linear transformations of the data, to yield more abstract and ultimately more useful representation of data. Convolutional neural networks are deep learning architectures, recently been employed successfully for image segmentation and classification tasks [4]. In this thesis, effectively capturing the deep features of glaucoma and AMD based on deep CNN is a major task.

CNN architecture is used for learning image feature because of speed, flexibility and high learning rate capability. It uses distributed representation and learn in multiple level representation. CNN architecture contains an input layer, hidden layer, and output layer as shown in Figure 2.3. Hidden Layers learn more abstract representations as you head up [36]. Figure 2.3 shows the typical CNN architecture.

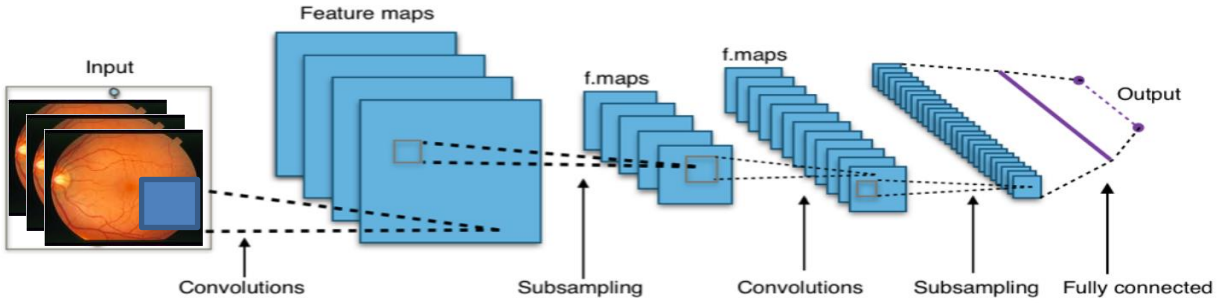

Figure 2.3: Typical CNN architecture [37]

Convolutional layers apply a convolution operation to the input, passing the result to the next layer. The convolution emulates the response of an individual neuron to visual stimuli. Pooling layers used to reduce the dimensions of the image data. Max pooling is most commonly than average pooling because of average pooling can't extract good features sometimes. ReLU applies the non-saturating activation function f (x)= max (0, x). It effectively removes negative values from an activation map by setting them to zero. It's advantage is fast, not expensive and prevents vanishing gradient. Fully connected layer connect every neuron in one layer to every neuron in another layer. Back propagation is the most commonly used gradient descent optimization algorithm to adjust the weight of neurons by calculating the gradient of the loss function.

Regularization refers to a set of different techniques that lower the complexity of a neural network model during training and prevent over-fitting. It is a technique for training deep neural network that standardizes the inputs to layer for each mini-batch. CNN's use of various types of regularization like dropout is used to prevent overfitting [38]. Batch normalization is inserted after fully connected for purpose improves gradient flow through the network. It also allows higher learning rates to reduces the strong dependence on initialization, regularization in a funny way, and reduces the need for dropout.

The final layer of connections in the network is fully connected which connects every neuron from the max-pooling layer to every one of the output neurons. The soft-max classifier is generally used to recognize the classes in deep-learning algorithms. Soft-max classifier is used for differenting between binary images and also multi dimentional classification problems. The vector of weights and the bias are called a filter and represent some features of the input (e.g. particular shape).

Chapter 3: Materials and Methods

3.1 Overview

Automated diagnosis system for retinal disease like glaucoma and AMD from fundus camera image are extremely important as they help to reduce diagnosis time and also plays a critical role to know the progress of the disease [27]. Nowadays, convolutional neural network-based image processing methods is showing remarkable results in the area of image classification and segmentation tasks. This chapter discusses the methods and materials used for the diagnosis of glaucoma and AMD.

3.2 Data collection

Image data for this research were collected from Drishti-GS1 database for cup and disk segmentation [39]. It consists of 101 fundus images, taken with eyes dilated and with 2896 × 1944 pixel size. The Drishti-GS1 dataset was used for glaucoma diagnosis by segmenting cup and disk. The images were originally acquired from the Aravind Eye Hospital, India. The ODIR-5K dataset is collected form Shanggong Medical Technology, China. The data set was used for AMD classification training. Fundus images are taken by different cameras such as Canon, Zeiss and Kowa and with mixed image resolutions. Annotations are categorized by qualified human readers with high-quality control management. Both left and right eye images and patient ages are also considered during image collection [40]. Table 3.1 summarizes the acquired image dataset. In addition, 100 images were collected from Korea hospital, Ethiopia. The image was taken by a Canon camera with eye dilated, with 960 X 960 pixel size and contain both right and left eye.

Table 3.1: Acquired image data.

Database	**Data Size**	**Image Size**	**Image format**	**Used for**
Drishti-GS1	101	2896 × 1944	PNG	Cup and Disk segmentation
ODIR-5K	400	-	JPG	AMD
Korea Hospital	100	960 x 960	JPG	AMD

3.3 The General Method

Figure 3.1 describes the overall steps used for these researches like data acquisition, pre-processing, preparing model, training, segmentation, post-processing and classifications steps are involved in the research for glaucoma and AMD disease detection and classification.

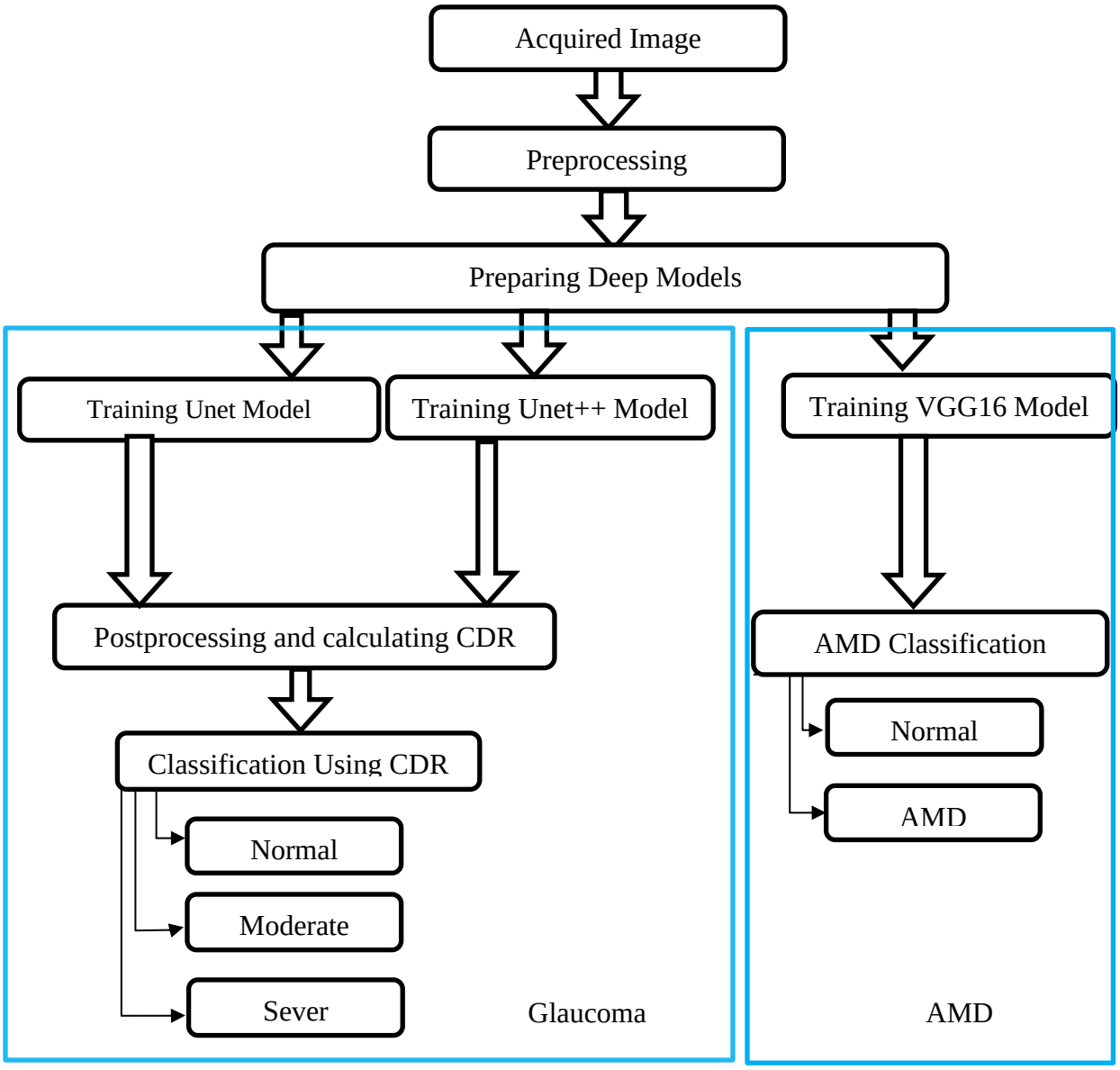

Figure 3.1: Overall Block diagram of the research.

3.4 Pre-processing

Most raw ophthalmology images are not appropriate for training with the model because of their poor quality, due to movement artifacts and iris color [41]. So, it is important to pre-process the images acquired before feeding them for models. These steps help to improve the overall performance of the system and to reduce or remove all unnecessary information from the image data. Resizing, contrast limited adaptive histogram equalization (CLAHE) and gamma correction are very important techniques that are used to pre-process the raw fundus image. Figure 3.2 summarizes image pre-processing steps.

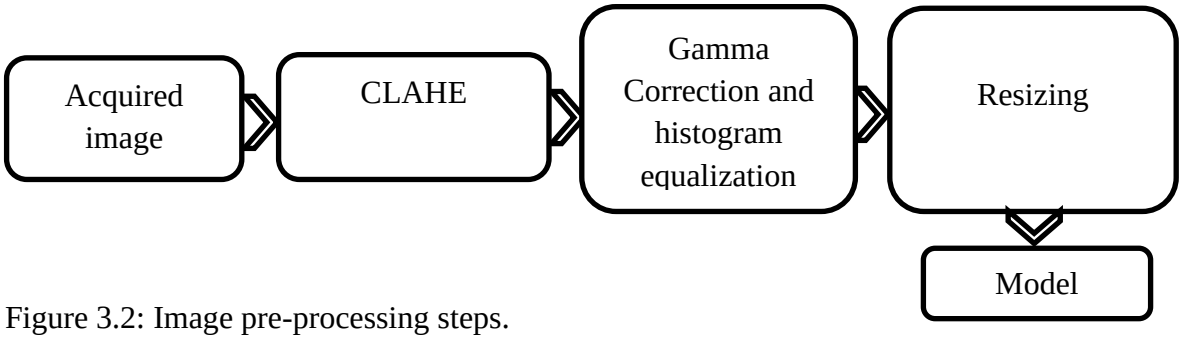

Figure 3.2: Image pre-processing steps.

To reduces the dimensionality of the image pixel matrix by mapping traditional three-dimensional color data to one-dimensional gray data. Figure 3.3 shows the original image and gray-scale image.

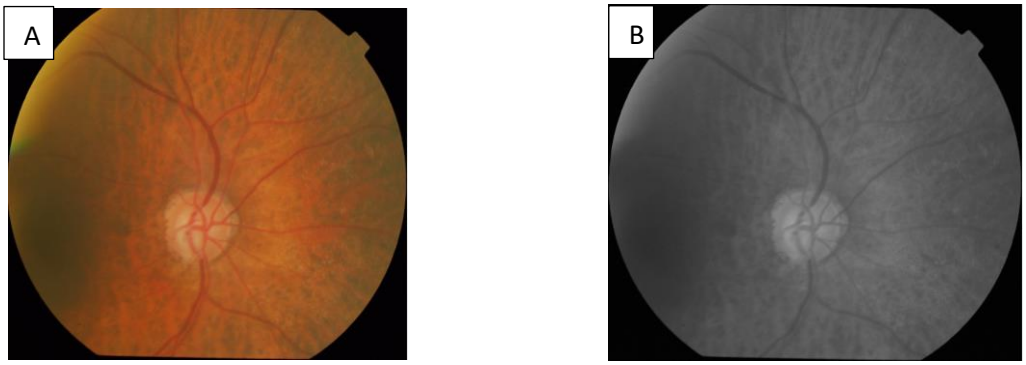

Figure 3.3: A: Fundus color image, B: gray-scale image

For imporving both local contrast and global contrast different image pre-processing algorithm have been applied. CLAHE is applied for improving local contrast problem. It improves the local contrast, prevent the over amplification of noise and enhances the definitions of edges in each region of an image. The slope of the transformation function gives as amplification of contrast in the vicinity of a given pixel value. Figure 3.4 show Image after CLAHE is applied and histogram plot.

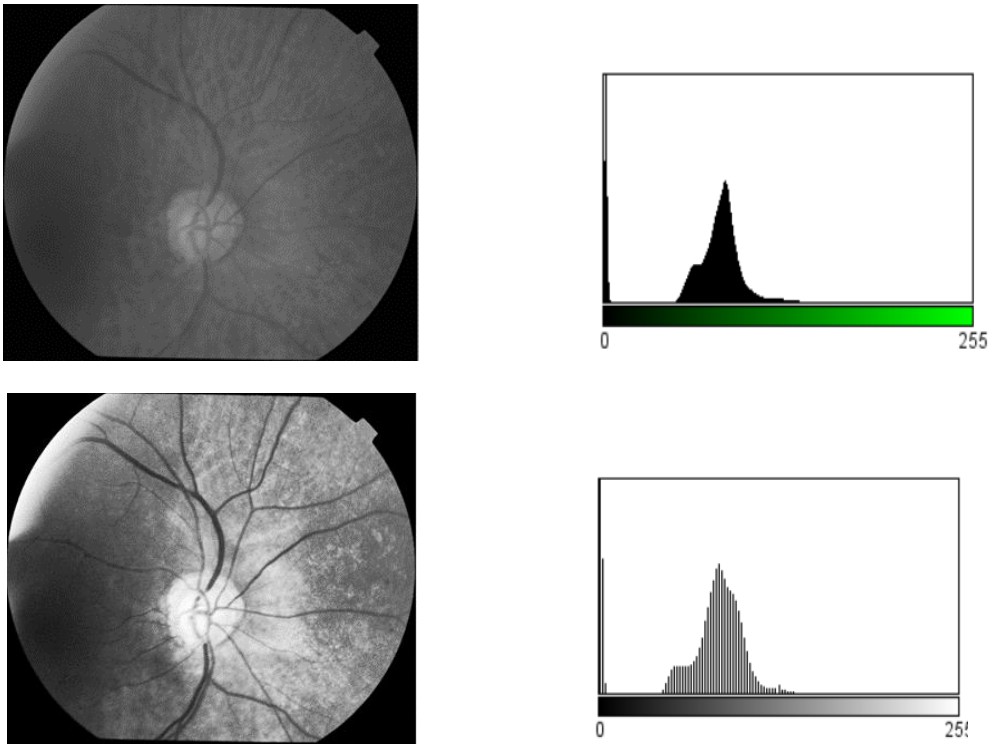

Figure 3.4: Image after CLAHE and histogram plot.

Gamma correction is the most known image processing method, which used to transform luminance non-linearity in digital image. Because of the camera quality limitations, most of the captured images are of low quality. Therefore, the image may not show the main features of the objects in some cases. This drawback is usually referred to as gamma distortion.

It defines the relationship between a pixel's numerical value and its actual luminance. Without gamma, shades captured by digital cameras wouldn't appear as they did to our eyes [42]. It's also referred to as gamma correction, when the transformation curve expresses the normalized pixel value of the luminosity channel, γ is a constant and v is the transformation has a simple point-wise operation to the input u. Equation 3.1 shows transformation of input with gamma value.

$$v = ur^{\gamma} \qquad (3.1)$$

Histogram equalization is also used in pre-processing to equalize contrast by stretching the image regions and interpolating the result across them. It is very important for removing global contrast problem. Finally, it is also important to resize the image in order to fit the training and test image to the input of the model. Figure 3.5 shows the image histogram plot and image before and after histogram equalization and gamma correction steps.

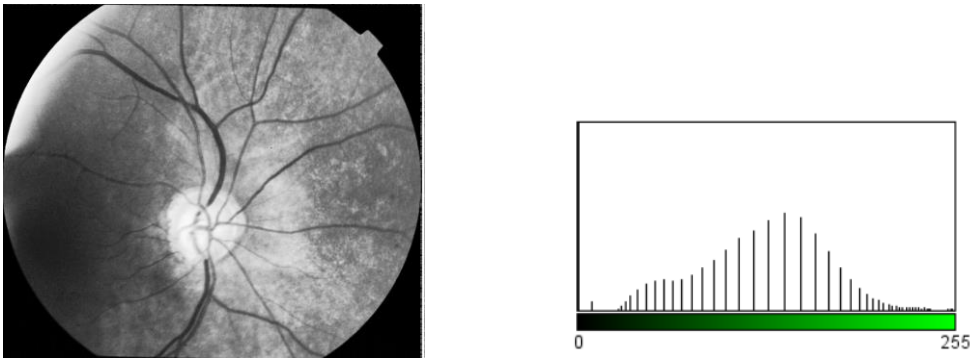

Figure 3.5: Image after Gamma Correction and Histogram equalization with histogram plot.

The pre-processing steps in Figure 3.2 has been followed for glaucoma detection But, for AMD image pre-processing color image is used no gray-scale conversion steps has been applied. This is because of the yellowish color of the macula which is important for differentiating AMD and normal image. Since, deep learning model's performance can be improved by the amount of dataset. Image augmentation has also been used, keras image data generator has been used to generate image while training the model. Image data generator will create an artificial image for training by shifting, rotating and flipping the image data while training.

3.5 Deep Learning Approach for Glaucoma and AMD Classification

3.5.1 Deep learning approach for image segmentation

Deep learning methods has been used for different applications in medical imaging, including image registration, anatomy localization, lesion segmentation, tissue segmentation and optical nerve detection [43]. Generally, there are three levels of image analysis in deep learning approaches are image classification, segmentation and object detection.

Classification is categorize the entire image into a class by categorizing image in to normal image or abnormal image, know a time it is commonly used in medical imaging. Whereas, object detection used to detect objects within an image by drawing a rectangle around them. Recently, object detection has been widely applied for medical image analysis like lesion, cell, nerve and organ detection.

Image segmentation is also used for identifying different parts or subsection of the image. Image segmentation lays the basis for performing object detection and classification. Image segmentation is even deeper than object recognition [44]. In image segmentation process, there are two levels of granularity those are instants segmentation and semantic segmentation. Instance segmentation identifies each instance of each object in an image. But it differs from semantic segmentation in that it doesn't categorize every pixel as an instant object. If there are three class of objects in an image, semantic segmentation classifies all the objects as one instance, while instance segmentation identifies each individual object.

For this research semantic segmentation method has been selected for cup and disk segmentation. State-of-the-art semantic segmentation methods like Fully Convolutional Network (FCN), Mask RCNN, Unet and Unet++ are among the most commonly used image segmentation method. Table 3.2 summarize semantic segmentation methods with their advantage and disadvantage.

Table 3.2: Deep learning image segmentation methods comparison .

Method	Description	Advantage and Disadvantage of the methods
FCN	-FCN uses convolutional layers to classify each pixel in the image. Contain full blocks of convolutional layers followed by max-pooling operation to decompress an image. It then makes a class prediction at this level of granularity.	**Advantage:** there are no limitations on image size and high-level feature map representations. **Disadvantage**: smooths detailed structures and ignores small objects to reduce the feature map resolution.
Mask RCNN	-It is extensions of Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. Mask R-CNN is simple to train and adds only a small overhead to Faster R-CNN.	**Advantage**: It effectively locates different objects in an image with bounding boxes and realigning RoIPool is also more accurate in the case of Mask RCNN. **Disadvantage:** Segmentation masks may overlap and over-segmentation may occur sometimes [45].
Unet	-Its network architecture is made up of a contracting path on the left and an expansive path on the right. -It introduced skip connection between encode and decoder.	**Advantage:** Few training images can yield more precise segmentation. It is much faster to run than FCN or Mask RCNN. Unet is reducing lone objects to create accurate segmentation masks [45]. **Disadvantage**: it may over compress information and increasing memory consumption as well.

Unet++	The encoder and decoder sub-networks are connected through a series of nested, dense skip pathways.	**Advantage:** it is accurate and powerful in medical image segmentation than Unet. And also have improved gradient flow [46]. It will alleviate the vanishing gradient problem, strengthen feature propagation and encourage feature reuse. **Disadvantage:** Relatively use high GPU memory for larger images.

3.6 Unet ++ Approach

3.6.1 Background of Unet

Unet is a more powerful architecture developed for medical image segmentation. First Unet was introduced by Ronneberger in 2015 [47]. Unet contains three basic parts this are contracting (encoder), bottleneck and expanding (decoder) parts. The network gets its name because of the U-shaped architecture of the contracting path and expansive path. The contracting path is a typical convolutional network that consists of repeated application of convolutions, each followed by a rectified linear unit (ReLU) and a max-pooling operation. During the contraction, the spatial information is reduced while feature information is increased. The expansive pathway combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the contracting path. The concatenation helps the vanishing gradient and reduce over-fitting. Figure 3.6 shows normal Unet architecture.

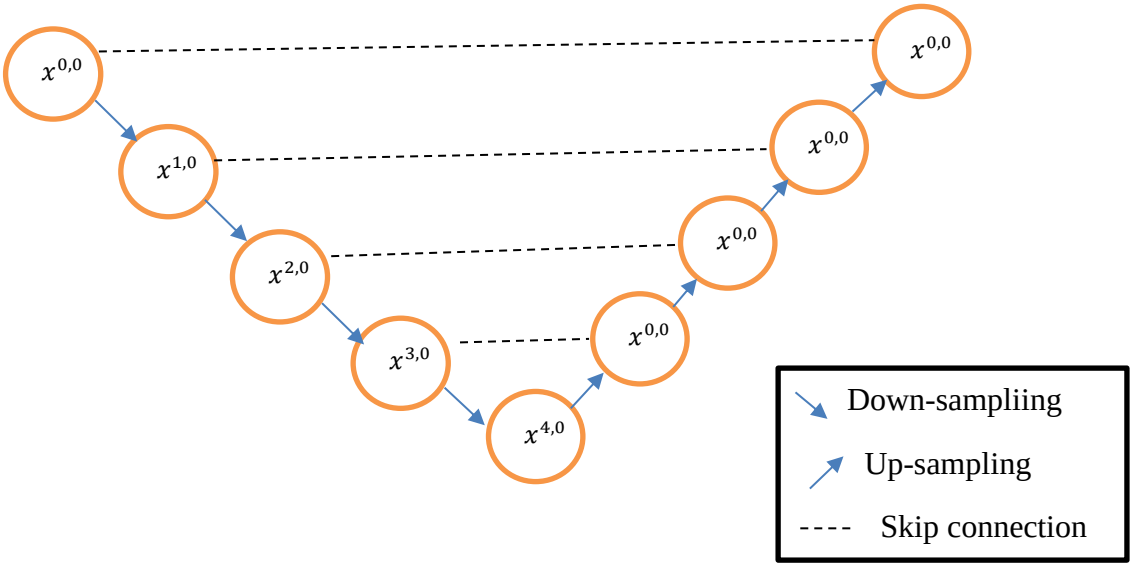

Figure 3.6: Unet network architecture.

The convolution feature extractor is important in feature representation and has fast convergence speed. But also, it has the drawback of feature extractor and overfitt on an easy task. The down-sampling path is robust against small input variance, reduces over-fitting, reduces computation cost and enlarges receptive field area however, It has the drawback of compressing information and the invisibility of small objects. The up-sampling has the advantage of recovering lost resolution in down-sampling and guides encoder to select important information but, sometimes may not recover large object boundaries accurately.

3.6.2 Gradient Descent

The key concept behind nearly all deep learning is to reduce the error by gradually iterating and updating learning parameters, each step moving the parameters in the direction that incrementally lowers the loss function. This process is called gradient descent. It is the most popular algorithm to perform optimization on neural networks. It is an iterative optimization algorithm used to find the minimum value for a function.The loss function then compares these predictions to the targets, producing a loss value measure of how well the network's predictions match what was expected.

The optimizer uses this loss value to update the network's weights. Figure 3.7 summarizes the training process and loss function.

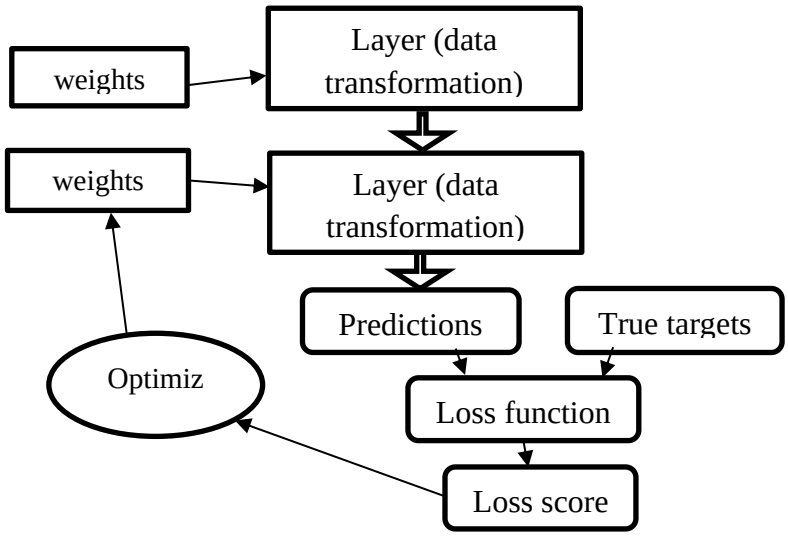

Figure 3.7: Loss function during training

Given a collection of data points **X** and a vector containing the corresponding target values **y**, the weights in the neural network is the most important factor in converting an input to impact the output. To find the weight vector **w** and bias term **b** that associates each data point **xi** with an approximation of Ŷi is expressed in terms of a single data point.

$$\hat{Y} = Xw + b \qquad (3.2)$$

3.6.3 Loss Function

Loss function measure the error between the real value and the predicted value of the price. Usually, the smaller the value, the smaller the error. Squre funcution is used for calculating loss function. For given parameters w and b, we can express the error of our prediction on a given sample by using the following equation 3.3.

$$l^{(i)}(w,b), \frac{1}{2}(\hat{y}^{(i)} - y^{(i)})^2 \qquad (3.3)$$

In summary, Unet encoder-decoder like architecture performs consistent and continuous improvements. However, it still finds a semantic gap along the skip connections. Re-design the skip pathways is important. Figure 3.8 shows Unet ++ architecture design. In Unet, the feature maps of the encoder are directly received in the decoder. In Unet++, they undergo a dense connection whose number of convolution layers depends on the pyramid level. The skip connection is important in fighting vanishing gradient problem, learns pyramid level features and recovers information loss in down-sampling. Figure 3.8 illustrates Unet++ re-designed skip pathways.

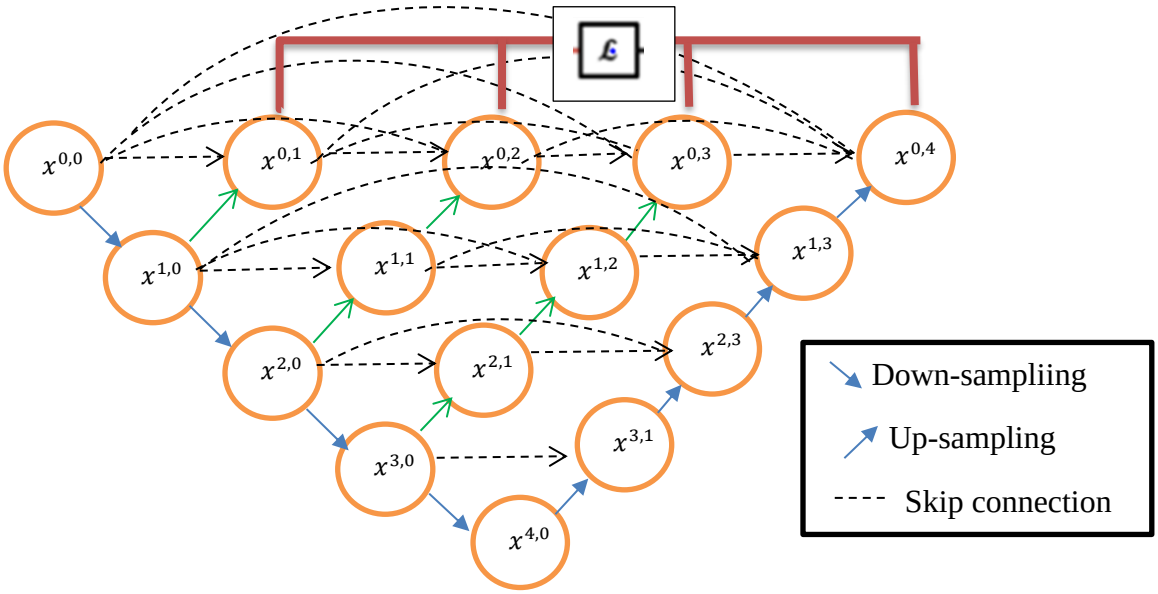

Figure 3. 8: Unet ++ architecture re-designed skip pathways

3.6.4 Re-designed skip pathways

Re-designed skip pathways change the connectivity of the encoder and decoder sub-networks [46]. In Unet, the encoder's feature maps are obtained directly in the decoder, but Unet++ undergoes a complex convolution block between encoder and decoder, the number of convolution layers depends on the pyramid level. In Unet, the feature maps of the encoder are directly received in the

decoder however, in Unet++, they undergo a dense convolution block whose number of convolution layers depends on the pyramid level. Formally, the skip pathway stack of feature maps represented by $x^{i,j}$ is computed as equation 3.4.

$$x^{i,j} = \begin{cases} H\left(x^{(i-1),j}\right), & j = 0 \\ H\left(\left[\left[x^{i,k}\right]_{k=1}^{j-1}, u(x^{i+1,j-1})\right]\right), & j > 0 \end{cases} \quad (3.4)$$

where function H is a convolution operation followed by an activation function, U denotes an up-sampling layer and [] denotes the concatenation layer. By assuming xi,j denote the output of node of Xi,j where i indexes the down-sampling layer along the encoder and j indexes the convolution layer of the dense block along the skip pathway. Basically, nodes at level j = 0 receive only one input from the previous layer of the encoder nodes at level j = 1 receives two inputs, both from the encoder sub-network but, at two consecutive levels and nodes at level j > 1 receive j +1 inputs, of which j inputs are the outputs of the previous j nodes in the same skip pathway and the last input is the up-sampled output from the lower skip pathway. Generally, it shows the feature map at every dense block.

3.6.5 Deep Supervision

Unet++ generates full resolution feature maps at multiple semantic levels of architecture [46]. The deep supervision works in two mode. The first is accurate mode where in the outputs from all segmentation are averaged. And the second is the fast mode where in the final segmentation is selected from one of the segmentation branches and the selection made by model pruning and speed gain.

Unet++ differs from the original Unet in three ways the first is having convolution layers on skip pathways, which bridges the semantic gap between encoder and decoder feature maps, the second is having dense skip connections on skip pathways, which improves gradient flow and the third is having deep supervision.

3.6.6 Hyper Parameters

The hyper parameter tuning process is to achieve a balance between under-fitting and over-fitting. Different hyper parameters available for a DL model approaches used for selecting the right set of hyper parameters for a model. Here are some of the important hyper parameters used.

Learning rate: it defines the length of each step in simple terms, how large the updates to the weights in each iteration can be made. If our learning rate is too small than optimal value, then it would take a much longer time (hundreds or thousands) of epochs to reach the ideal state.

Number of Epochs: is a single training cycle, which contains many steps depending on the data and configuration. After all training samples are finished or seen by the network, one epoch has elapsed, and then, the second set will be started at second epoch. Sometimes, just increasing the number of epochs for model training delivers better results, although this comes at the cost of increased computation and training time.

Batch size: is the number of images that can be processed in one full iteration. Using a moderate batch size always helps achieve a smoother learning process for the model. A batch size of 5 or 10, irrespective of the dataset size and the number of samples, will deliver a smooth learning curve in most cases.

Number of Layers: It is true that just adding a few more layers will generally increase the performance, at least marginally. But the problem is that with an increased number of layers, the training time and computation increase significantly. Moreover, you would need a higher number of epochs to see promising results. In case of Unet++ architecture lower pyramid level is used for training to over come over-fitting.

3.7 Training the Unet Models

The best hyper-parameters were selected for the model to prevent over-fitting and for better training performance of the model. The Unet model contains a contracting path that uses a convolutional neural network. It consists of the repeated application of two 3×3 convolutions, each followed by a batch normalization layer, a rectified linear unit (ReLU) activation and 2×2 max pooling operation with stride 2 for down-sampling. For every step for contracting path the feature map will double. Every step in the expansive path consists of an up-sampling of the feature map followed by a 2×2 deconvolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly feature map from the contracting path and two 3×3 convolutions, each followed by batch normalization and a ReLU activation. The final layer a 1×1 convolution is used to map each component feature vector to the desired number of classes. Table 3.3 shows the Unet++ layer used to build the model. Appendix A and B shows the source code implementation of Unet and Unet++ model architecture.

Table 3. 3: Number and type layers used for Unet++ model.

	Layers used for Training Unet ++	No layers
1	Conv2D	14
2	Max pooling 2D	2
3	Concatenate	3
4	Conv2DTranspose	3

Sometimes, increasing the number of epochs for model training often delivers better performance, although, this may come at the cost of increased computation and training time as well as model may also over-fit. Using a moderate batch size always helps achieve a smoother learning process for the model. Batch size of 10 and 5 is used for training, irrespective of the dataset size and the number of samples will deliver a smooth learning curve in most cases. Sometimes just adding a few more layers will generally increase the performance, at least marginally. But the problem is that with an increased number of layers, the training time and computation increase significantly.

3.8 Training VGG16 architecture

AMD image classification using VGG16 architecture is the second model used. Training contains 13 Conv-2D and 3 dense layers. Convolution layers use a 3x3 filter with a stride 1 and always used padding and max pool layer of 2x2 filter of stride 2. From the collected image 70% of the image data has been used for training and validation and the rest 30% of the image is used for training. And the validation split is 30% for validation and 70% for training the model. While training the model Image data generator is used for augmentation of the image or artificially extending the training data. It is very important to improve the performance of the system by multiplying training data by shifting, rotating and flipping the image data. Table 3. 4 shows VGG16 layers used. And also appendix C, shows the detail architecture of the model.

Table 3.4: Number and Type of layers used for VGG 16 models.

	Layers used for VGG 16 models	Number of layers
1	Conv2D	13
2	MaxPooling2D	5
3	Dense (fully connected)	3

Selecting the best hyper-parameter is very important for achieving a better performance model. So, different hyper-parameters is available for a DL model various approaches used for selecting the right set of hyper-parameters for a model. Here are some of the important hyper-parameters used. Learning rate 0.001, 0.0001 and 0.00001 are used for training the VGG16 model. Then, when the learning rate is too small than optimal value, then it would take a much longer time or hundreds of epochs to reach the ideal state. So, different learning rates are used to check the output prediction accuracy. Also, batch size of 5 and10 is used for these models, irrespective of the dataset size and the number of samples can deliver a smooth learning curve in most cases.

For optimization, Adam optimizer algorithm is used for stochastic gradient descent for training deep learning models. It also helps to provide an optimization algorithm that can handle sparse gradients on noisy problems and also it's easy to configure.

3.9 Post-processing

The model may sometimes suffer for abnormal small object prediction during segmentation because of contrast in the image. Mathematical morphology like erosion and dilution is used to remove unnecessary objects and noise form a predicted image. Binary opening is erosion followed by dilution while keeping the initial region size used for filling unnecessary holes form in the regions while keeping the initial region size. Binary closing is dilution followed by erosion while keeping the initial region size also used to fill the holes inside the segmented images.

The Erosion of binary image ***f*** by structuring element ***s*** produces a new binary image ***g = f-s*** with ones in all locations ***(x,y)***. The dilution of an image ***f*** by a structuring element and ***s*** denoted ***g= f+s*** with ones in all locations ***(x,y)*** of a structuring element origin at which that structuring element ***s*** its input image ***f, g (x,y)=1.***

3.10 Materials

Python programming language version 3.6 is used for this thesis with graphics packages and toolsets like numpy, scipy, tensorflow, keras and matplotlib. Table 3.3 summarizes the material used during the research.

Table 3.5: The material used for this research

Hardware	**Software material**
✓ Fundus camera	✓ Python 3.6
✓ Laptop computer ✓ RAM 8 GB ✓ Hard Disk 1TB ✓ GPU ✓ Speed 1.6 Hz 8th Gen	✓ TensorFlow ✓ Keras ✓ OpenCV ✓ Scipy
	✓ Fiji
	✓ Colab (GPU Nvidia Tesla K80)

Chapter 4: Result and Discussion

4.1 Preprocessing for Glaucoma Images

For this research, fundus camera images have been collected to develop the system. All the images used are preprocessed using different techniques such as resizing, CLAHE, histogram equalization and gamma correction before feeding into the model. Finally, the contrast of the image is well articulated, informative and descriptive after preprocessing. This will help to easily detect the objects from the retinal image and to discriminate images. Figure 4.1 summarizes the result of pre-processing and compares the histogram plots.

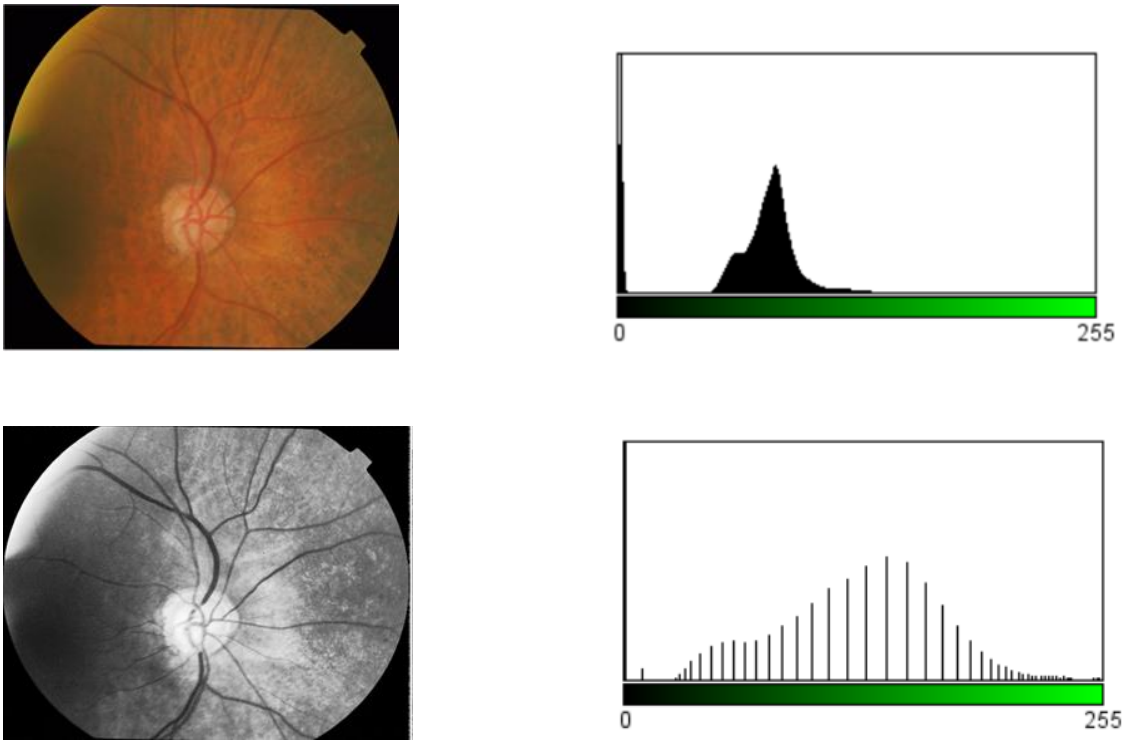

Figure 4. 1: The raw dataset image and Image after pre-processing with its histogram plot

4.1.1 Retinal Cup And Disk Segmentation Training Using Unet

4.1.1.1 Cup segmentation learning curve and prediction using Unet

After pre-processing the models have been trained with different hyper-parameters values. Figure 4.2 shows the learning curve of the model shows validation accuracy of 0.975 and validation loss of 0.26. The training curve converged at an early stage and achived better validation accuracy but, the validation loss is still high. Figure 4.3 shows cup segmentation learning performance of the model. Dice coefficient score for segmentation of cup area achieved for the model is 0.85 dice score.

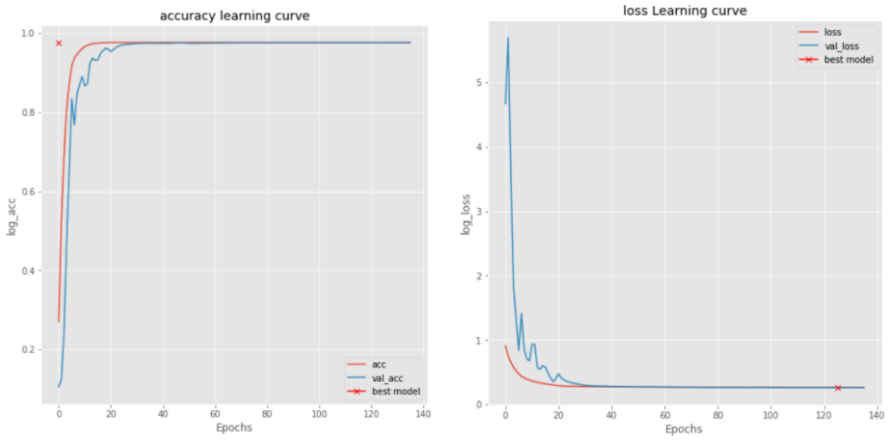

Figure 4.2: Retinal Cup segmentation Unet learning curve

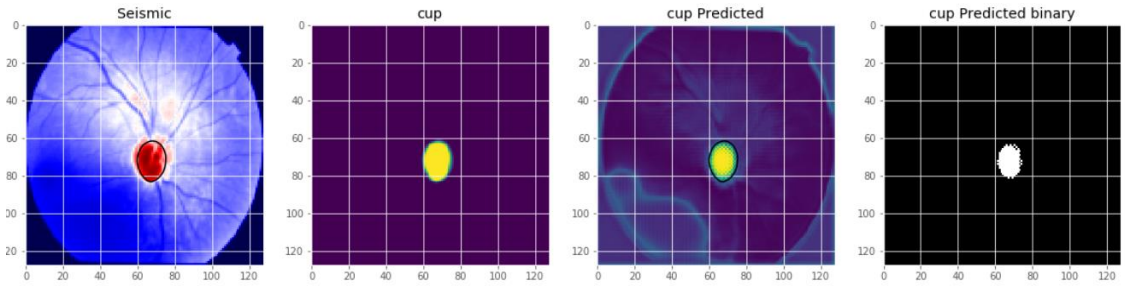

Figure 4.3: Retinal Cup segmentation using Unet trained model.

4.1.1.2 Disk segmentation learning curve and prediction using Unet

It has been shown that the learning curve for retinal disk segmentation is increased exponentially but, Initially, the loss is high because of the random selection of gradient descent. The model tries to find the exact shape and structure of the disk. The training and validation curve shows a slight difference compared to cup segmentation. Figure 4.4 illustrates the learning rate of the model is 0.9754 validation accuracy and 0.25 validation loss. The dice score for segmentation of disk area is 0.86. Figure 4.5 shows manually masked retinal and disk segmentation using Unet model.

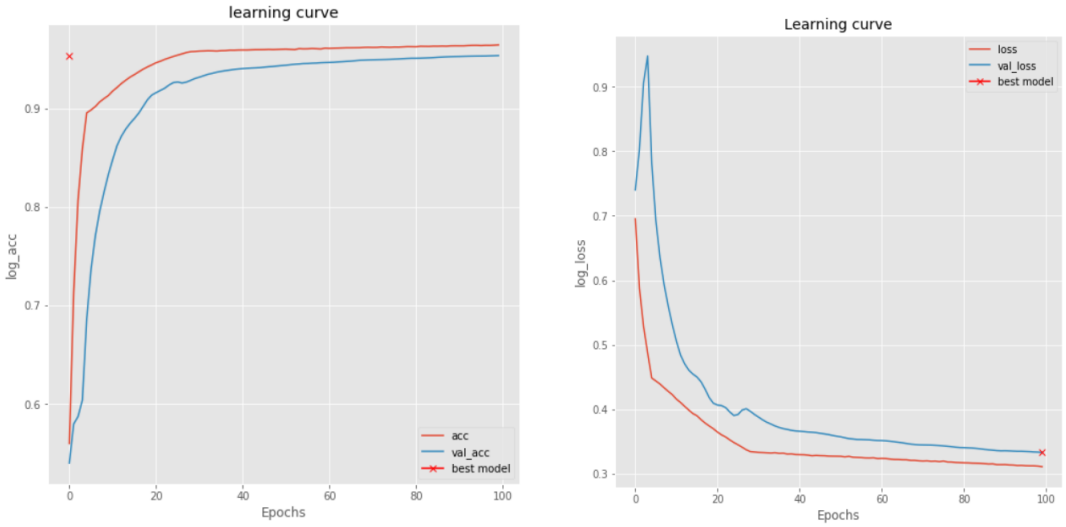

Figure 4. 4: Retinal disk segmentation learning curve using Unet trained model

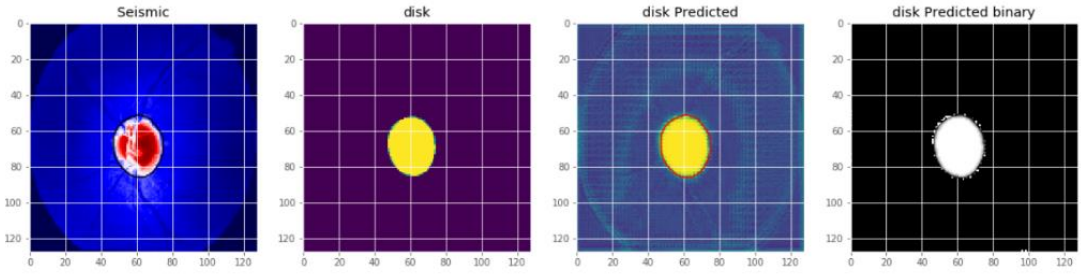

Figure 4. 5: Retinal disk segmentation using Unet trained model

4.1.2 Retinal Cup And Disk Segmentation Training Using Unet ++

4.1.2.1 Cup segmentation learning curve and prediction using Unet++

It has been shown that retinal cup segmentation training using Unet architecture resulted 0.973 validation accuracy and 0.18 validation loss. The curve shows validation loss improvement than Unet. but some differences in the loss curve of cup segmentation are shown. Figure 4.6 illustrates the learning curve. The segmentation prediction also shows clear and precise segmentation however sometimes small dots and objects are detected. While, the dice coefficient score for segmenting disk area found to be 0.87. Figure 4.7 shows the manually masked retinal cup image with Unet predicted cup image.

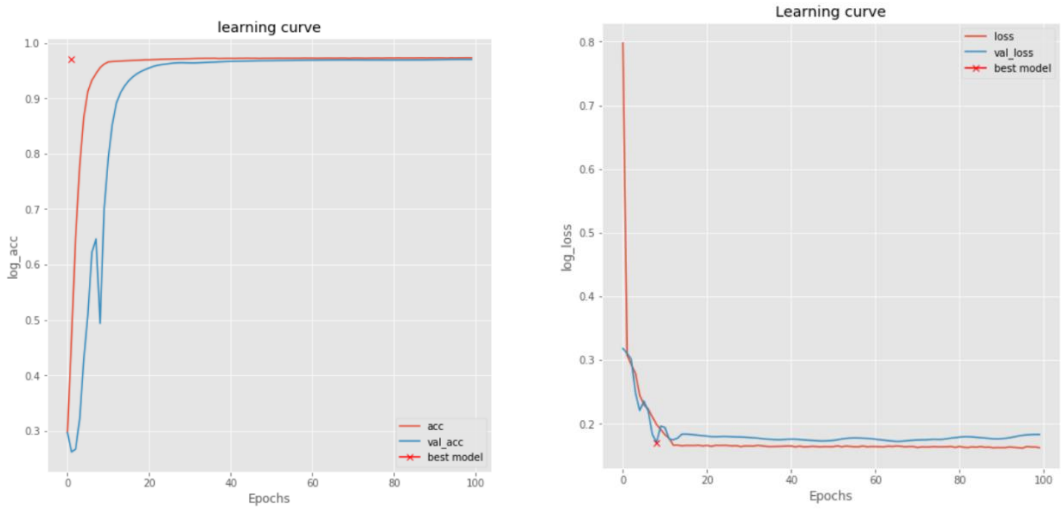

Figure 4.6: Retinal Cup segmentation learning curve using Unet++.

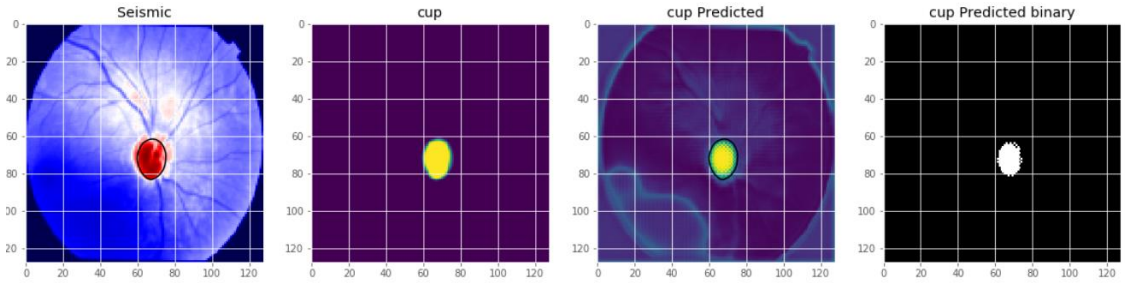

Figure 4.7: Retinal Cup segmentation using Unet++ trained model.

4.1.2.2 Disk segmentation learning curve and prediction using Unet++

The training model for retinal disk segmentation using Unet architecture illustrated a 0.976 validation accuracy and 0.12 validation loss. This training curve shows big improvement when compared to Unet. The model converges fast but, also learning curve is not smoothed well. And also dice coefficient score for segmentation is found to be 0.894. Figure 4.8 summarizes the learning curve and Figure 4.9 illustrates the manually masked retinal disk image with Unet predicted disk image.

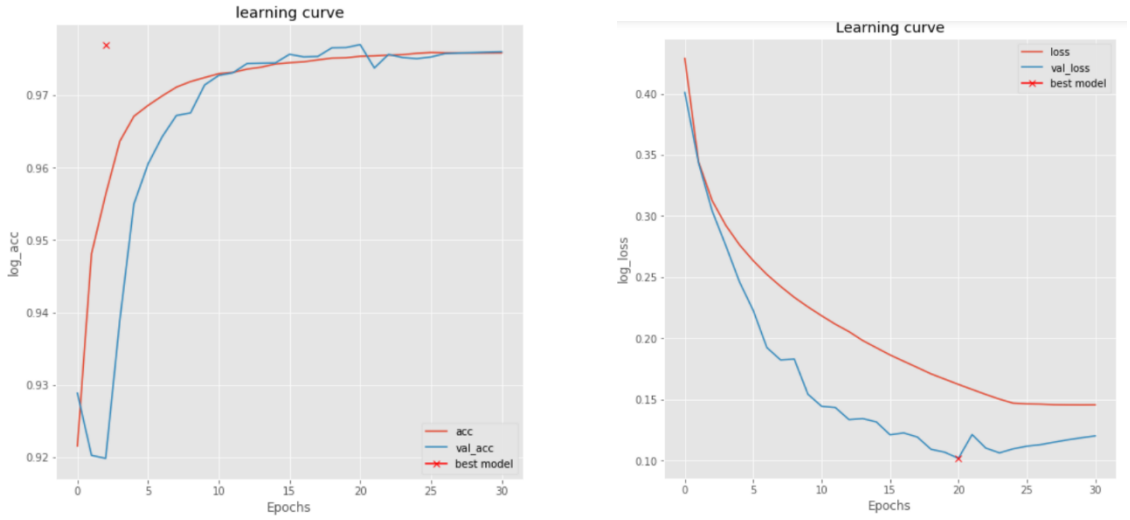

Figure 4.8: Retinal Disk segmentation learning curve using Unet++

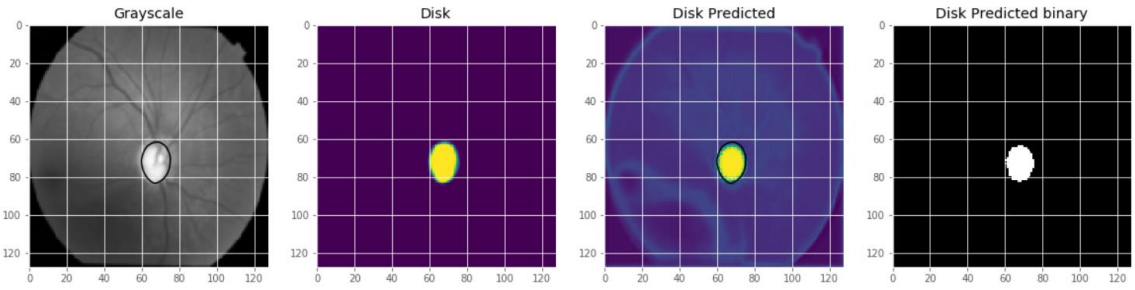

Figure 4.9: Retinal Disk segmentation using Unet++ trained model

4.2 Post-processing

Erosion and dilution is capable of removing small artifacts from the predicted image and provide clear and concise segmentation results. Figure 4.10 shows the complete removal of these artifacts from the predicted image step by step. Small dot predicted by the model removed by eroding the pixel of the image and diluting the disk of the fundus image back. Erosion and dilution applied for this research using binary opening and binary closing principle. The result has shown post-processing is very important for improving the segmentation accuracy acquired from the model and to provide clear and precise segmented object. After clearing all irregularity, areas of cup and disk is calculated by counting the white pixels. Figure 4.10 summarizes the post-processing result from disk and cup segmentation.

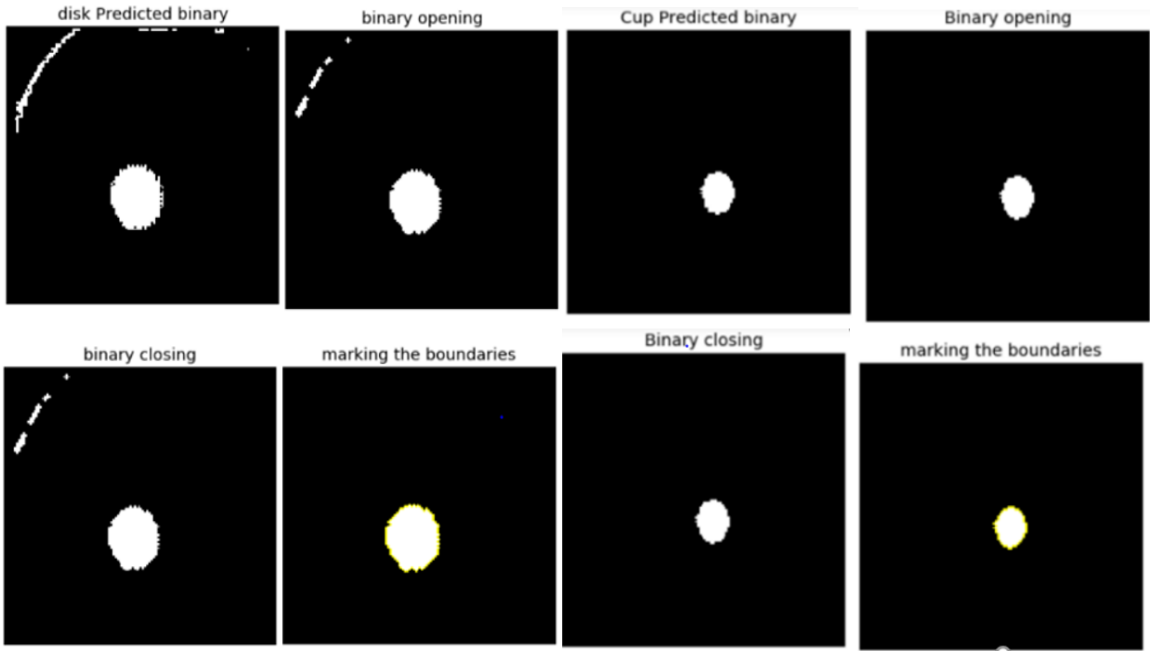

Figure 4.10: Disk and Cup segmented image post-processing.

4.3 Cup to disk ration-based Classification performance using Unet++

The CDR is obtained by dividing the area of the cup to the area of the disk. Both the area is obtained from the segmented image by counting the white pixels. Segmentation of the cup and disk area determines the severity level of glaucoma by CDR value. For the CDR value of less than 0.4 is taken as normal, for the CDR value greater than 0.4 and less than 0.8 is taken as moderate and for CDR value greater than 0.8 is taken as sever glaucoma patient. The test performance metrics are summarized using the confusion metrics in Table 4.1. The Unet++ model test indicated most of the predicted results are classified as moderate and two test data are falsely classified as normal while their true label is moderate and sever. And one data is falsely classified as sever while the true label is moderate.

$$\text{CDR} = (\text{area of Cup})/(\text{area of Disk}) \qquad (4.1)$$

Table 4. 1: Confusion matrix for glaucoma classification using Unet++and Unet models.

Unet++ Model	***Normal***	***Moderate***	***Sever***
Normal	7	1	1
Moderate	0	18	0
Sever	0	1	8

Unet Model	***Normal***	***Moderate***	***Sever***
Normal	6	1	1
Moderate	0	19	1
Sever	1	0	7

The overall classification performance for glaucoma using Unet and Unet++ is calculated from the confusion matrix as follows. Using the value of True Positive (TP), False Negative (FN), True Negative (TN) and False Positive (FP) the performance metrics like accuracy, precision, recall and F1score have been calculated. Equation 4.2, 4.3, 4.4 and 4.5 below shows computational formula used for performance metrics. Tables 4.2 and 4.3 summarizes the performance of Unet++ model and Unet model respectively.

$$\text{Accuracy} = (\text{TP} + \text{FN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \qquad (4.2)$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (4.3)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (4.4)$$

$$\text{F1score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4.5)$$

Table 4. 2: Overall Unet++ model classification performance.

Class	n(truth)	n(Classified)	Accuracy	Precision	Recall	F1 Score
Normal	7	9	100%	0.78	1.0	0.88
Moderate	20	18	90%	1.0	0.90	0.95
Sever	9	9	88.89%	0.89	0.89	0.89

Table 4. 3: Overall Unet model classification performance.

Class	n(truth)	n(Classified)	Accuracy	Precision	Recall	F1 Score
Normal	7	8	85.7%	0.75	0.86	0.80
Moderate	20	20	95%	0.95	0.95	0.95
Sever	9	8	77.78%	0.88	0.78	0.82

The Unet++ model achieved overall classification accuracy of 0.916 and Unet model achieved an overall accuracy of 0.89. This test result illustrates that Unet ++ architecture performed better classification than Unet model. And also shows Unet++ have low validation loss than the Unet network. Both models show better classification performance metrics on moderate class. From 36 test image Unet++ model correctly classified 33 images and Unet also correctly classified 32 images out of 36 test images.

4.4 Age-related Macular Degeneration Image Classification

4.4.1 Preprocessing result for AMD

All images used are pre-processed using different techniques such as resizing, CLAHE, contrast equalization, and gamma correction before feeding into the model. The global and local contrast of the image is well adjusted and descriptive after pre-processing. The drusen deposits are clearly visible on maculas after pre-processing. Figure 4.11 shows the pre-processing result of the AMD images.

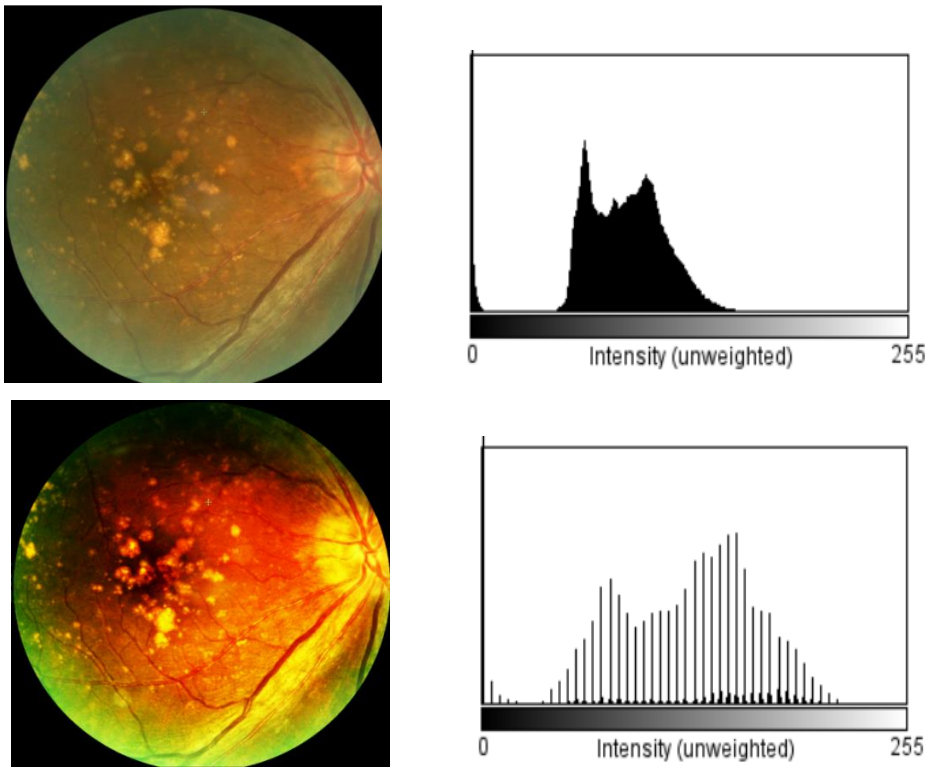

Figure 4. 11: The raw dataset images and pre-processed image with it's histogram plot.

Age-related macular degeneration disease classification indicates a good classification performance for a lower learning rate. Table 4.2 summarizes VGG16 architecture with different learning rates. For higher learning rate over-fitting has been a major problem. Modern convolutional networks use a repetitive convolution layer with pooling, flatten and dense layers.

VGG16 model uses few repeating blocks of elements when compared to other architecture. Architecture complexity is a major cause of over-fitting. It has been shown that choosing a higher epoch will also lead to over-fitting. The number of epoch and architecture complexity shows an inverse relation for optimizing results. Table 4.4 shows Training of the VGG16 model with different learning rates at 100 epoch and batch size of 10.

Table 4. 4: Training of VGG16 model with different learning rates.

Learning rate	0.001	0.0001	0.00001
Validation Accuracy	0.9	0.9	0.99
Validation Loss	0.63	0.024	0.016

The performance metrics of the AMD classification using VGG16 architecture is also discussed here using below confusion metrics. Based on the above learning performance model with 0.00001 learning rate is used for testing images. Because, it shows lower validation loss. The confusion matrix obtained for 0.00001 learning rate is shown in Table 4.5 below.

Table 4. 5: AMD classification confusion matrix.

	Normal	AMD
Normal	74	3
AMD	1	72

Table 4. 6: Performance metrics for VGG16 model classification.

Class	n(truth)	n(Classified)	Accuracy	Precision	Recall	F1 Score
Normal	75	77	97.33%	0.96	0.99	0.97
AMD	75	73	97.33%	0.99	0.96	0.97

The research result indicated that, Vgg16 model shows good performance for classifying the normal and AMD fundus image. The overall performance test obtained from the above metrics is 0.97 accuracy.

4.5 Discussion

This thesis presented an automated detection of both glaucoma and AMD from a fundus image. Unet ++ and Unet models were used for glaucoma detection and VGG16 model used for AMD detection from fundus image..

Effective pre-processing algorithms have been used for removing contrast problems and clarifying objects on the retina as shown in Figure 4.1 and Figure 4.11. It is has been inspected that retinal images are clear and well-articulated. So that, model can track the change and segment parts of the retinal like cup and disk.

Moreover, an application of state of the art Unet++ image segmentation techniques for cup and disk segmentation and the VGG16 model for Age-related macular degeneration detection is presented in these research. As inspected on the training performace section both the models learned well with training accuracy and validation accuracy greater than 0.96. But the validation loss is high for Unet model. During training the model using batch size of 10 shows a smooth curve than using batch sizes of 5. For reducing overfitting, the deeper architecture is not used and big dropout is not used during the training. Segmentation prediction shows clear and precise however sometimes small dot or object detected. It is clear that figure 4.10 shows the post-processing methods like erosion and dilution removed small hole and object artifacts from the segmented image. The VGG16 model has been trained with training and validation accuracy of greater than 0.9 with different learning rates. Lower learning rate performed better as shown in table 4.4 for classifying AMD and normal fundus image.

The previous research has focused on the traditional machine learning algorithm for the segmentation and detection of age-related disease [18, 19]. But these results demonstrated that a deep learning approach can accurately segment and classify retinal disease. In line with the problem statement, few researchers have applied deep learning methodology to develop an algorithm for automated detection of the various age-related eye diseases such as AMD and glaucoma [25, 26]. The Six-layer CNN model developed to detect glaucomatous fundus images

from normal fundus images. In their study, they have proposed network structure with two databases obtained AUC values of 0.831 and 0.887 respectively [28].

Instead of feeding the entire fundus image into the CNN model for classification. Zilly et al. segmented and measured the CDR from the fundus images and classified the images according to the CDR values. Similarly, this research also used the CDR feature. This research suggested convolutional neural network methods than other machine learning methods due to its high learning capability advantage. Similarly, Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images is presented as a multi-purpose retinal image analysis system and achieved AUC of 0.84 for AMD and 0.92 for all DR(Diabetic Retinopathy) cases [48].

This research integrated glaucoma and AMD detection from fundus images. While previous research has focused on only glaucoma or AMD. These experiments provide a new insight into the diagnosis of glaucoma diagnosis and AMD. Glaucoma classification has achieved 0.914 overall accuracy and AMD classification also achieved 0.97 overall accuracy. The research also demonstrated Unet++ model shows better performance than Unet model for the segmentation of the cup and disk from the fundus image. The performance metrics for normal, moderate and sever class of glaucoma also confirmed the performance of Unet++ over Unet model. Beyond these results, the researchers have to consider the complexity of the network created as an increase in the number of classes and data.

Overall, the research also points out new insight toward using multi-class age-related eye disease diagnosis from fundus image using the CNN approach is relevant for image segmentation and classification. Future research should be devoted to the development of one stand-alone retinal disease detection model that can improve the diagnostic performance of an eye in health care and make automated diagnostic systems more robust. This system model limited by need of large database and graphics processing unit to train the model. Future studies could fruitfully explore further based on these conclusions, practitioners should consider that age-related eye disease is like retinopathy and other retinal disease.

4.6 Graphical user interface(GUI) implementation of the research

Nowadays, digital image processing has been applied for many real problems solving like disease classification and segmentation of parts of the medical image processing area. Figure 4.12 shows GUI implementation of glaucoma and AMD diagnosis by loading Fundus image. Preprocessing steps is application of CLAHE, resizing, histogram equalization and Gamma correction. And the first part will classify loaded and preprocess image into normal and AMD. And the second part will classify glaucoma image after preprocessing and calculating the CDR from segmented image as discussed in the result section. Figure 4.12 shows an example of normal image diagnosis.

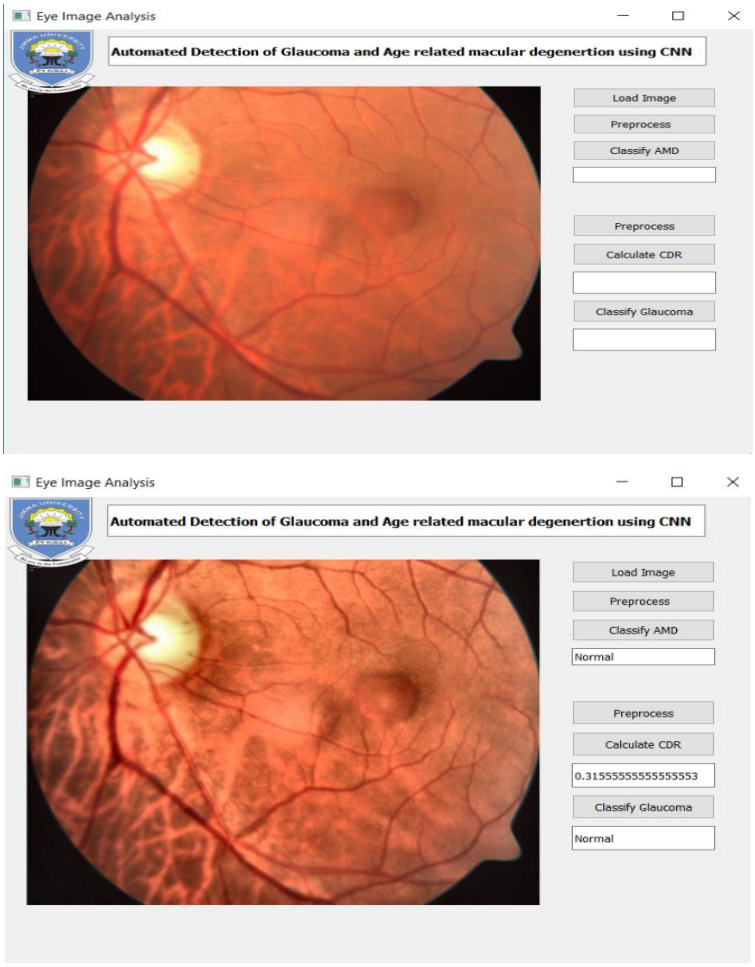

Figure 4. 12: GUI development for diagnosis of Glaucoma and AMD.

Chapter 5: Conclusion and Recommendation

5.1 Conclusion

The research has demonstrated CNN based age-related disease diagnosis from a fundus image shows remarkable results. Recent progress of CNN in advanced medical image diagnosis system by discovering hidden patterns from images by themselves. It's clear that the research is highly efficient and uses a multi-purpose diagnosis method for diagnosing glaucoma and AMD. All signs from these study points towards versatility and importance for retinal image analysis using CNN architecture.

This research aimed at the automatic detection and classification of glaucoma using the CDR feature and AMD disease classification from the retinal fundus image. This is very important for Age-related eye disease detection in its early stages save the patient from further vision loss and critical to use where lack of a skilled person is available in order to fill the knowledge gap.

All models are trained using a widely used databases like Drishti-GS1, ODIR-5K and also local image from Korean Hospital. Effective pre-processing algorithms have been used for removing contrast problems and clarifying objects on the retina. Unet++ model achieved overall performance accuracy of 0.914 and Unet model also achieved an overall accuracy of 0.89 by testing on 36 images. It is clearly illustrated Unet ++ architecture show better segmentation result than Unet model. Post-processing is necessary for removing small unnecessary objects and irregularity predicted by the segmentation models. AMD classification with VGG16 also achieved an overall accuracy of 0.97 by testing on 150 fundus images.

This work applies to widespread eye disease screening in hospitals and health centers. Also, allow the health care facilities to screen glaucoma and AMD patients with better accuracy and less time without the use of highly experienced ophthalmologists. It is possibly being progressively delicate to identify glaucoma in its beginning stages. In this manner, our discoveries should assist with making progressively precise glaucoma and AMD analysis from fundus image.

Despite of some limitations such as a data set and a dedicated graphical processing unit (GPU), it is possible to see the effects of this research. And also, another issue for the research was accessibility to local data.

5.2 Future works and Recommendation

Furthermore, this research is necessary to be a phase of building an algorithm to screen more diseases from fundus images. Nevertheless, looking at the association of retinal age-related disease it is important t incorporate others disease such as diabetic retinopathy. For the future, in addition to CDR, it is necessary to analyze other features and to test those features with different state-of-the-art classification techniques. The model efficiency can be further improved through model optimization and training with different hyper-parameters. Comparing with different classification architectures has been suggested for achieving better results. And also try AMD classification with other features. The model is capable of classifying glaucoma and AMD in the current state. As future work, it's strongly suggested the use of various state-of-the art deep learning architectures to further improve the classification accuracy.

References

[1] M. Zetterberg, "Maturitas Age-related eye disease and gender," *Maturitas*, vol. 83, pp. 19–26, 2016.

[2] A. A. Salam, T. Khalil, M. U. Akram, A. Jameel, and I. Basit, "Automated detection of glaucoma using structural and non structural features," *Springerplus*, vol. 5, p. 1519, 2016.

[3] L. Tamrat, G. W. Gessesse, and Y. Gelaw, "Original Article Adherence to Topical Glaucoma Medications in Ethiopian Patients," *Middle East Afr. J. Ophthalmol.*, vol. 22, no. 1, pp. 59–63, 2015.

[4] F. Jiang A.Grigorev, S.Rho, Z.Tian, Y.Fu, W.Jifara, K.Adil, and S.Liu, "Medical image semantic segmentation based on deep learning," *Neural Comput. Appl.*, no. July, pp. 1257–1265, 2017.

[5] D. A. Michael, M. K. Garvin, and S. Milan, "Retinal Imaging and Image Analysis Michael," *IEEE Rev Biomed Eng*, vol. 1, no. 3, pp. 169–208, 2011.

[6] R. R. A. Bourne *et al.*, "Causes of vision loss worldwide , 1990 – 2010 : a systematic analysis," *Lancet Glob Heal.*, vol. 1, no. 6, pp. 339–349, 2012.

[7] F. M. Cherinet, S. Y. Tekalign, D. H. Anbesse, and Z. Y. Bizuneh, "Prevalence and associated factors of low vision and blindness among patients attending St . Paul ' s Hospital Millennium Medical College , Addis Ababa , Ethiopia," *BMC Ophthalmol.*, no. 2018, pp. 10–12, 2020.

[8] C. K. N. S and K. Somashekar, "Image Processing Techniques for Automatic Detection of Glaucoma - A Study," *Int. J. Latest Technol. Eng. Manag. Appl. Sci.*, vol. VI, no. Vii, pp. 112–119, 2020.

[9] A. Anton, M. Fallon, and A. Morilla-grasa, "Cost and detection rate of glaucoma

screening with imaging devices in a primary care center," pp. 337–346, 2017.

[10] Y. Hagiwara, J. En, W. Koh, N. A. Polytechnic, J. H. Tan, and S. V Bhandary, "Computer-Aided Diagnosis of Glaucoma Using Fundus Images : A Review," *Comput. Methods Programs Biomed.*, vol. 168, no. July, pp. 1–12, 2018.

[11] U. R. Acharya *et al.*, "Biomedical Signal Processing and Control Decision support system for the glaucoma using Gabor transformation," *Biomed. Signal Process. Control*, vol. 15, pp. 18–26, 2015.

[12] "How Glaucoma is Treated." [Online]. Available: https://www.glaucomafoundation.org/treating_glaucoma.htm. [Accessed: 02-Jan-2020].

[13] B. N. Kumar, R. P. Chauhan, and N. Dahiya, "Detection of Glaucoma Using Image Processing Techniques: A Critique," *Semin. Ophthalmol.*, vol. 33, no. 2, pp. 275–283, 2018.

[14] M. Madhusudhan, N. Malay, S. R. Nirmala, and D. Samerendra, "Image processing techniques for glaucoma detection," *Commun. Comput. Inf. Sci.*, vol. 192 CCIS, no. PART 3, pp. 365–373, 2011.

[15] B. S. Kirar and D. K. Agrawal, "Empirical wavelet transform based pre-processing and entropy feature extraction from glaucomatous digital fundus images," *Int. Conf. Recent Innov. Signal Process. Embed. Syst. RISE 2017*, vol. 2018-Janua, pp. 315–319, 2018.

[16] S. Guo, K. Wang, H. Kang, Y. Zhang, Y. Gao, and T. Li, "BTS-DSN : Deeply Supervised Neural Network with Short," *Int. J. Med. Inform.*, vol. 126, pp. 105–113, 2019.

[17] T. Saba, M. W. Khan, M. Yasmin, and M. Sharif, "CDR based glaucoma detection using fundus images: a review," *Int. J. Appl. Pattern Recognit.*, vol. 4, no. 3, p. 261, 2017.

[18] M. Salman *et al.*, "A Novel Adaptive Deformable Model for Automated Optic Disc and Cup Segmentation to Aid Glaucoma Diagnosis," *J. Med. Syst. Vol.*, vol. 42, 2018.

[19] M. S. Haleem, L. Han, B. Li, and A. Nisbet, "Automatic Extraction of The Optic Disc Boundary For Detecting Retinal Disease," *Int. Conf. Comput. Graph. Imaging (CGIM)*, pp. 40–47, 2013.

[20] F. Yin, J. Liu, S. H. Ong, Y. Sun, D. W. K. Wong, and N. M. Tan, "Model-based Optic Nerve Head Segmentation on Retinal Fundus Images," *Annu. Int. Conf. IEEE*, no. 3, pp. 2626–2629, 2011.

[21] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep Learning Techniques for Medical Image Segmentation : Achievements and Challenges," *J. Digit. Imaging, Springer*, vol. 3, 2019.

[22] Z. Lai and H. Deng, "Medical Image Classification Based on Deep Features Extracted by Deep Model and Statistic Feature Fusion with Multilayer Perceptron," *Comput. Intell. Neurosci.*, p. 13, 2018.

[23] S. Liu *et al.*, "Deep Learning in Medical Ultrasound Analysis : A Review," *Engineering*, vol. 5, no. 2, pp. 261–275, 2019.

[24] M. Liang, L. Jiao, and Z. Meng, "A Superpixel-Based Relational Auto-Encoder for Feature Extraction of Hyperspectral Images," *Remote Sens*, pp. 1–18, 2019.

[25] N. O. Mahony *et al.*, "Deep Learning vs . Traditional Computer Vision," *Adv. Comput. Vision. CVC 2019. Adv. Intell. Syst. Comput.*, vol. 943, pp. 128–144, 2019.

[26] M. I. Razzak, S. Naz, and A. Zaib, "Deep Learning for Medical Image Processing : Overview , Challenges and Future," *Classif. BioApps. Lect. Notes Comput. Vis. Biomech.*, vol. 24, pp. 323–350, 2017.

[27] M. S. Haleem, L. Han, J. Van Hemert, and A. Fleming, "Glaucoma Classification using Regional Wavelet Features of the ONH and its Surroundings," *Eng. Med. Biol. Soc. (EMBC), Annu. Int. Conf. IEEE*, pp. 4318–4321, 2015.

[28] T. Y. Wong, X. Chen, J. Liu, D. W. Kee Wong, and Y. Xu, "Glaucoma detection based on deep convolutional neural network," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 715–718, 2015.

[29] Q. Abbas, "Glaucoma-Deep : Detection of Glaucoma Eye Disease on Retinal Fundus Images using Deep Learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. January, p. 6, 2017.

[30] A. Sevastopolsky, "Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network," *Pattern Recognition and Image Analysis*, vol. 27, no. 3, pp. 618–624,2017.

[31] P. M. Arabi, V. Deepa, T. S. Naveen, and D. Samanta, "Machine vision for screening of age-related macular degeneration using fundus images," *IEEE*, pp. 3–6, 2017.

[32] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, "Automated Grading of Age-Related Macular Degeneration From Color Fundus Images Using Deep Convolutional Neural Networks," *JAMA Ophthalmol.* vol. 135, no. 11, pp. 1170–1176, 2017.

[33] H. Fu *et al.*, "Joint Optic Disc and Cup Segmentation Based on Multi-label Deep Network and Polar Transformation," *IEEE*. vol. 0062, no. c, pp. 1–9, 2018.

[34] S. Liu *et al.*, "A Deep Learning-Based Algorithm Identi fi es Glaucomatous Discs Using Monoscopic Fundus Photographs," *Ophthalmol. Glaucoma*, pp. 1–8, 2018.

[35] J. Fan, C. Ma, and Y. Zhong, "A Selective Overview of Deep Learning," *arXiv Prepr. [stat.ML]*, no. April, pp. 1–37, 2019.

[36] "Convolutional neural network." [Online]. Available: https://en.wikipedia.org/wiki/Convolutional_neural_network. [Accessed: 16-Jan-2020]

[37] P. P. Chitte, "Analysis of Different Methodsfor Identification and Classification of

Cevical Spondylosis (CS): A Survey," *Int. J. Appl. Eng. Res. ISSN 0973-4562*, vol. 12, no. 21, pp. 11727–11737, 2017.

[38] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*, vol. 07 :SAGE Publishing; July, 2019.

[39] A. Chakravarty and J. Sivaswamy, "Joint optic disc and cup boundary extraction from monocular fundus images," *Comput. Methods Programs Biomed.*, vol. 147, pp. 51–61, 2017.

[40] "Odir2019 - Dataset-数据集." [Online]. Available: https://odir2019.grand-challenge.org/download. [Accessed: 10-July-2019].

[41] A. Allam, Ali; Youssif, Aliaa; Ghalwash, "Automatic Segmentation of Optic Disc in Eye Fundus Images : A Survey Automatic Segmentation of Optic Disc in Eye Fundus Images : A Survey," *ELCVIA Electron. Lett. Comput. Vis. image Anal.*, vol. 14, no. April, pp. 1–20, 2015.

[42] J. Staal *et al.*, "Ridge-Based Vessel Segmentation in Color Images of the Retina," *IEEE*, vol. 23, no. 4, pp. 501–509, 2004.

[43] D. Shen, G. Wu, and H. Suk, "Deep Learning in Medical Image Analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–48, 2017.

[44] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *Int. J. Multimed. Inf. Retr.*, vol. 7, no. 2, pp. 87–93, 2018.

[45] S. U. Akram and J. Kannala, "Mask-RCNN and U-net Ensembled for Nuclei Segmentation," *Dep. Comput. Sci. Aalto Univ.*, no. January, 2019.

[46] Z. Zhou and M. R. Siddiquee, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support*, vol. 11045, pp. 3–11, 2018.

[47] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Int. Conf. Med.*, pp. 1–8, 2015.

[48] C. Agurto *et al.*, "Automatic Detection of Diabetic Retinopathy and Age-Related Macular Degeneration in Digital Fundus Images," *Invest Ophthalmol Vis Sci.*, vol. 52, no. 8, pp. 5862–5871, 2011.

Appendix

Appendix A: Unet Custom Model Architecture Code

```
def conv2d_block(input_tensor, n_filters, kernel_size=3, batchnorm=True):
    # first layer
    x = Conv2D(filters=n_filters, kernel_size=(kernel_size, kernel_size), kernel_initializer="he_normal",
               padding="same")(input_tensor)
    if batchnorm:
        x = BatchNormalization()(x)
    x = Activation("relu")(x)
    # second layer
    x = Conv2D(filters=n_filters, kernel_size=(kernel_size, kernel_size), kernel_initializer="he_normal",
               padding="same")(x)
    if batchnorm:
        x = BatchNormalization()(x)
    x = Activation("relu")(x)
    return x
```

```
def get_unet(input_img, n_filters=16, dropout=0.5, batchnorm=True):
    # contracting path
    c1 = conv2d_block(input_img, n_filters=n_filters*1, kernel_size=3, batchnorm=batchnorm)
    p1 = MaxPooling2D((2, 2)) (c1)
    p1 = Dropout(dropout*0.5)(p1)

    c2 = conv2d_block(p1, n_filters=n_filters*2, kernel_size=3, batchnorm=batchnorm)
    p2 = MaxPooling2D((2, 2)) (c2)
    p2 = Dropout(dropout)(p2)

    c3 = conv2d_block(p2, n_filters=n_filters*4, kernel_size=3, batchnorm=batchnorm)
    p3 = MaxPooling2D((2, 2)) (c3)
    p3 = Dropout(dropout)(p3)

    c4 = conv2d_block(p3, n_filters=n_filters*8, kernel_size=3, batchnorm=batchnorm)
    p4 = MaxPooling2D(pool_size=(2, 2)) (c4)
    p4 = Dropout(dropout)(p4)

    cc1 = conv2d_block(p4, n_filters=n_filters*16, kernel_size=3, batchnorm=batchnorm)
    pp1 = MaxPooling2D(pool_size=(2, 2)) (cc1)
    p4 = Dropout(dropout)(pp1)

    c5 = conv2d_block(p4, n_filters=n_filters*32, kernel_size=3, batchnorm=batchnorm)
```

```
# expansive path
uu1 = Conv2DTranspose(n_filters*16, (3, 3), strides=(2, 2), padding='same') (c5)
uu1 = concatenate([uu1, cc1])
uu1 = Dropout(dropout)(uu1)
cc6 = conv2d_block(uu1, n_filters=n_filters*16, kernel_size=3, batchnorm=batchnorm)

u6 = Conv2DTranspose(n_filters*8, (3, 3), strides=(2, 2), padding='same') (cc6)
u6 = concatenate([u6, c4])
u6 = Dropout(dropout)(u6)
c6 = conv2d_block(u6, n_filters=n_filters*8, kernel_size=3, batchnorm=batchnorm)

u7 = Conv2DTranspose(n_filters*4, (3, 3), strides=(2, 2), padding='same') (c6)
u7 = concatenate([u7, c3])
u7 = Dropout(dropout)(u7)
c7 = conv2d_block(u7, n_filters=n_filters*4, kernel_size=3, batchnorm=batchnorm)

u8 = Conv2DTranspose(n_filters*2, (3, 3), strides=(2, 2), padding='same') (c7)
u8 = concatenate([u8, c2])
u8 = Dropout(dropout)(u8)
c8 = conv2d_block(u8, n_filters=n_filters*2, kernel_size=3, batchnorm=batchnorm)

u9 = Conv2DTranspose(n_filters*1, (3, 3), strides=(2, 2), padding='same') (c8)
u9 = concatenate([u9, c1], axis=3)
u9 = Dropout(dropout)(u9)
c9 = conv2d_block(u9, n_filters=n_filters*1, kernel_size=3, batchnorm=batchnorm)

outputs = Conv2D(1, (1, 1), activation='sigmoid') (c9)
model = Model(inputs=[input_img], outputs=[outputs])
return model
```

Appendix B: Unet ++ Custom Model Architecture Code

```
def standard_unit(input_tensor, stage, nb_filter, kernel_size=3):
    act = 'elu'
    x = Conv2D(nb_filter, (kernel_size, kernel_size), activation=act, name='conv'+stage+'_1',
               kernel_initializer = 'he_normal', padding='same', kernel_regularizer=l2(1e-4))(input_tensor)
    x = Dropout(dropout_rate, name='dp'+stage+'_1')(x)
    x = Conv2D(nb_filter, (kernel_size, kernel_size), activation=act, name='conv'+stage+'_2',
               kernel_initializer = 'he_normal', padding='same', kernel_regularizer=l2(1e-4))(x)
    x = Dropout(dropout_rate, name='dp'+stage+'_2')(x)

    return x
```

```
def Nest_Net(img_rows, img_cols, color_type=1, num_class=1, deep_supervision=False):
    nb_filter = [32,64,128,256,512]
    act = 'elu'
    # Handle Dimension Ordering for different backends
    global bn_axis
    if K.image_dim_ordering() == 'tf':
      bn_axis = 3
      img_input = Input(shape=(img_rows, img_cols, color_type), name='main_input')
    else:
      bn_axis = 1
      img_input = Input(shape=(color_type, img_rows, img_cols), name='main_input')
    conv1_1 = standard_unit(img_input, stage='11', nb_filter=nb_filter[0])
    pool1 = MaxPooling2D((2, 2), strides=(2, 2), name='pool1')(conv1_1)
    conv2_1 = standard_unit(pool1, stage='21', nb_filter=nb_filter[1])
    pool2 = MaxPooling2D((2, 2), strides=(2, 2), name='pool2')(conv2_1)
    up1_2 = Conv2DTranspose(nb_filter[0], (2, 2), strides=(2, 2), name='up12', padding='same')(conv2_1)
    conv1_2 = concatenate([up1_2, conv1_1], name='merge12', axis=bn_axis)
    conv1_2 = standard_unit(conv1_2, stage='12', nb_filter=nb_filter[0])
    conv3_1 = standard_unit(pool2, stage='31', nb_filter=nb_filter[2])
    up2_2 = Conv2DTranspose(nb_filter[1], (2, 2), strides=(2, 2), name='up22', padding='same')(conv3_1)
    conv2_2 = concatenate([up2_2, conv3_1], name='merge22', axis=bn_axis)
    conv2_2 = standard_unit(conv2_2, stage='22', nb_filter=nb_filter[1])
    up1_3 = Conv2DTranspose(nb_filter[0], (2, 2), strides=(2, 2), name='up13', padding='same')(conv2_2)
    conv1_3 = concatenate([up1_3, conv1_1, conv1_2], name='merge13', axis=bn_axis)
    conv1_3 = standard_unit(conv1_3, stage='13', nb_filter=nb_filter[0])
    nestnet_output_1 = Conv2D(num_class, (1, 1), activation='sigmoid', name='output_1', kernel_initializer = 'he_normal',
                              padding='same', kernel_regularizer=l2(1e-5))(conv1_2)
    nestnet_output_2 = Conv2D(num_class, (1, 1), activation='sigmoid', name='output_2', kernel_initializer = 'he_normal',
                              padding='same', kernel_regularizer=l2(1e-5))(conv1_3)
    if deep_supervision:
        model = Model(input=img_input, output=[nestnet_output_1,
                                               nestnet_output_2])
    else:
        model = Model(input=img_input, output=[nestnet_output_2])
    return model
```

Appendix C: VGG16 Model architecture

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
dense_1 (Dense)	(None, 2)	8194

Total params: 134,268,738
Trainable params: 134,268,738
Non-trainable params: 0