An Efficient Deep Learning Approach to Detect Retinal Disease Using Optical Coherence Tomographic Images

by

Farhan Sakib khan
18301176
Nowshin Ferdaus
18101113
Tamim Hossain
18301183
Quazi Sabrina Islam
19101673
Md. Iftakharul Islam
18301020

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
May 2022

© 2022. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Farhan Sakib khan

18301176

Nowshin Ferdaus

18101113

Tamim Hossain

18301183

Quazi Sabrina Islam

19101673

Md. Iftakharul Islam

18301020

Approval

The thesis titled "An Efficient Deep Neural Network Approach to Detect Retinal Disease" submitted by

1. Farhan Sakib khan (18301176)
2. Nowshin Ferdaus (18101113)
3. Tamim Hossain (18301183)
4. Quazi Sabrina Islam (19101673)
5. Md. Iftakharul Islam (18301020)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on May 24, 2022.

Examining Committee:

Supervisor:
(Member)

Dr. Md. Ashraful Alam

Assistant Professor
Department of Computer Science and Engineering
Brac University

Co-supervisor:
(Member)

Dewan Ziaul Karim

Lecturer
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Associate Professor
Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi

Chairperson and Associate Professor
Department of Computer Science and Engineering
BRAC University

Ethics Statement

It is imperative that the thesis be written strictly in compliance with the university's rules and regulations, as well as ethical principles for doing research. Original data has been incorporated into our thesis. We double-checked our citations and references. Each of the paper's five co-authors accepts responsibility for any violations of the thesis rule. There were several Questionnaire Free tools, articles, and YouTube videos that helped us address problems. Furthermore, we would want to take this opportunity to express our gratitude to everyone who has helped us along the way. Our thesis was completed without the use of unethical tactics. The Brac University's code of ethics guides our activities.

Abstract

Optical Coherence Tomography (OCT) is an effective approach for diagnosing retinal problems that can be used in conjunction with traditional diagnostic testing methods. We developed and implemented a deep Convolutional Neural Network (CNN) model, which has the capability to effectively identify and classify Optical Coherence Tomography (OCT) images into the following four categories: Normal, DMD, CNV, and DME. This 21 layered CNN model is built with three basic layers: a convolutional layer, a pooling layer, and a fully connected layer along with dropout and dense layers. Our proposed model is able to detect and differentiate between the OCT images with a high amount of accuracy. The 21 layer proposed CNN model was used for the classification and diagnosis of retinal sickness using OCT images. To justify the efficiency of our custom CNN model, seven pre-trained CNN models (VGG16, VGG19, MobNetV2, Resnet50, DenseNet121, InceptionV3, and InceptionResNetV2) were used and testified with the same amount of dataset. In terms of the accuracy, precision, recall, and f1 score, which are all tested in this paper, the suggested CNN model along with seven other pre-trained CNN architectures perform comparable on the available dataset. The proposed model has an accuracy rate of 98.37 percent, which is greater than any of the experimental results of the CNN models utilized in this research due to the fact that the recommended model was developed. When it comes to the diagnosis of retinal problems, the CNN model that was suggested performs far better than any other model that was previously used.

Keywords: Convolutional Neural Network, Optical Coherence Tomography, Deep Learning, Retinal Disease, VGG16, VGG19 MobNetV2, ResNet50, DenseNet121, InceptionV3, and InceptionResNetV2.

Dedication

This paper is dedicated to our families and fellow team members. The constant support of the family members and the determination of the team members played a great role in making this paper a success. We could not have finished our thesis without the help of our esteemed supervisor and co-supervisor, who have been a continual source of guidance and advice. We also dedicate this paper to them.

Acknowledgement

In the first place, we want to express our gratitude to Allah, the Almighty, for providing us with the resources, opportunities, and guidance necessary to complete this research on schedule. As a second tribute, we would like to thank our thesis supervisor, Dr. Ashraful Alam and co-supervisor, Mr. Dewan Ziaul Karim, for their tireless support and guidance as we tackled a difficult topic. To get through the challenges, we had their continuous support and comments. Despite the current pandemic, they made time for us and provided valuable insights to help us enhance our job. For that, we will be eternally thankful. Final words of thanks go out to the entire faculty and student body for establishing such a welcoming learning environment in which we were able to develop professionally while also doing this research to the best of our abilities.

Nomenclature

OCT Optical Coherence Tomography
CNN Convolutional Neural Network
DMD Drusen Macular Degeneration
DME Diabetic Macular Edema
CNV Choroidal Neovascularization
AMD Age-related Macular Degeneration
DR Diabetic Retinopathy
PRP Panretinal Photocoagulation
VA Visual Acuity
PPV Pars plan Vitrectomy
EMM Epiretinal Macular Membrane
BBG Brilliant Blue G
ERM Epiretinal Membrane
ILM Internal Limiting Membrane
RPD Reticular Pseudodrusen
RPE Retinal Pigment Epithelium

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background Information

Retinal diseases are a primary cause of visual impairment and blindness that is irreversible, particularly in less developed nations. The World Health Organization (WHO) estimates that there are now 253 million people living with some kind of visual impairment and that there are 36 million people living with a form of blindness that is permanent [1]. More than eighty percent of instances of visual impairment are preventable or curable if adequate retinal screening and treatment planning is carried out during the early stages of the condition [2]. However, in order to make an accurate early diagnosis, the retinal images that were produced using Optical Coherence Tomography (OCT) need to be examined by an experienced ophthalmologist [3]. Macular degeneration is a painless, progressive eye condition that may cause vision impairment in the center of one visual cortex. The World Health Organization (WHO) [4] cites it as the major cause of irreversible eyesight loss in affluent countries. The condition of macular degeneration does not result in complete blindness; the ability to see is retained. Patients suffering from severe instances of the illness may become completely blind.
Underneath the retina, the drusen are yellowish masses. The presence of fatty acids and proteins are the cause of beginning drusen. Though Drusen are not thought to be a cause of age-related macular degeneration (AMD), it raises a person's chance of getting AMD and could be a symptom of the disease. Various sorts of drusen exist. For a prolonged-term, if at all, little drusen may not cause visual issues. Drusen that are larger increases the chance of progressive AMD, which can lead to visual loss. Macular degeneration is characterized by drusen. An ophthalmologist can notice these tiny yellow or white patches on the retina during a dilated eye exam or using retinal imaging.
Along with that, Diabetic Retinopathy and Diabetic Macular Edema are caused by longterm deterioration to the retina's neurovascular systems. The etiology of retinal injury is unknown however, metabolic and neuroinflammatory disturbances are suspected. Thorough metabolic management of the clinical syndrome through the use of optical antiinflammatory medications, such as connective tissue growth factor inhibitors and corticosteroids, are used to treat these processes. Whereas improved knowledge of the optical and vascular processes that underpin diabetic retinopathy will result in better diagnostic and treatment options, as well as enhanced sight maintenance. Slightly out of focus or irregular images towards or in the core of

one's range of view is the most common sign of macular edema. Colors may also seem pale or blurred out. Macular edema causes symptoms that vary from slightly blurred vision to significant vision loss in the majority of persons. DR is a primary cause of disability and impaired vision and is one of the most serious consequences of diabetes. Retinopathy is more frequent in those who have type 1 diabetes, although people who have type 2 diabetes are more likely to develop the disorder.New blood vessels that originate in the choroid may enter the sub–retinal pigment epithelium (sub–RPE) or the subretinal gap when the Bruch membrane breaks. This causes the patient's vision to become distorted or hazy. The most common cause of vision impairment is choroidal neovascularization (CNV) [5]. A noninvasive imaging method that employs light to take pictures, optical coherence tomography (OCT) is an acronym [42]. When it comes to disease diagnosis, OCT makes use of light waves to produce cross-section images of the retina. This enables the ophthalmologist to observe more clearly each of the different layers that make up the retina. With the assistance of optical coherence tomography, the ophthalmologist will be able to map the retina and determine its depth. As a consequence, the measurements used to assist in the diagnostic procedure have grown more straightforward and accurate. Image recognition and classification are used by CNNs to identify objects, differentiate faces, and more. They are made up of neurons that can be programmed to take on different weights and biases. CNNs are often used to classify images, group them based on similarity, and then identify them objects within those images. This method will help us to detect the disease more accurately and will give the ophthalmologist a clear idea of the infected areas. Preventing these difficult and severe eye disorders requires early detection. Our technique may assist in better understanding and detecting retinal abnormalities, which can avoid blindness.

1.2 Problem Statement

It is a light-sensitive thin membrane that is located at the bottom of our eyes and that sends signals to the brain when we are exposed to intense light. These neuronal cells are located in the center of the retina, providing folks with the clear, focused sight that they need to browse, drive, and take notice of details. Any portion of the retina might be damaged by retinal disorders which can have a significant effect on one's life. Retinal disorders that go untreated can cause significantly blurred vision and possibly blindness. Certain retinal disorders are treated with rapid diagnosis, although it can be managed or delayed to maintain and perhaps even recover eyesight. Meanwhile, the challenge here is to detect the loss of eyesight in a stage from which we can prevent extensive vision loss. Worldwide, retinal disorders are a leading source of optical damage including vision loss. According to community research, the frequency of retinal diseases ranges from 5.35% to 21.02% in those aged 40 and more. In contrast, retinal illness is a major of severe sight in affluent countries. As per the BrightFocus Foundation[13], 11 million Americans have age-related macular degeneration in a certain way. Only 10% of AMD cases are caused by wet AMD, whereas 90% of AMD cases that result in substantial vision loss are caused by wet AMD. Every year, according to the Macular Degeneration Partnership, 200,000 additional patients are diagnosed with AMD. Similarly, Diabetic retinopathy affects 7.7 million Americans, according to the National Eye Institute, with the figure predicted to rise to 11.3 million by 2030. Diabetic retinopathy affects up to 45 percent

of all persons who have diabetes. According to the medical journal JAMA Ophthalmology, diabetic macular edema (DME) is the most prevalent type of diabetic retinopathy, affecting approximately 745,000 Americans [7].

Here it seems that the lack of detection has caused many people in different ways. Ignoring the symptoms or any other reasons have raised the number of blindness which can be prevented by simply detecting the problems at an early age. Along with that this will open the door for the ophthalmologists to protect people from retinal diseases in some cases redeeming the loss of eyesight.

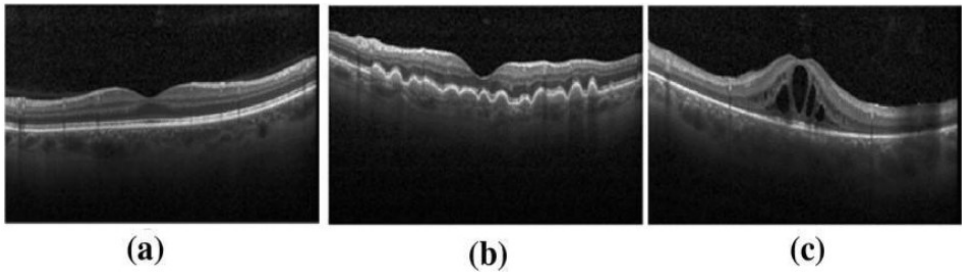

Figure 1.1: Retinal OCT images of a normal, b DMD and c DME conditions

1.3 Research Objective

The main focus of this study is to create a disease detection system for identifying prenatal retinal illnesses such as CNV, DME and DMD, which is currently under development. This system will take OCT images as input. We will use Deep learning to determine whether the images are diseased or not. After necessary pre-processing the images will go through the proposed CNN model and divide those in three categories. The objectives of this research are:

1. Understand image processing and how it works.
2. To understand data pre-processing techniques like Denoising and Reshaping.
3. To build a model to detect retinal disease based on OCT images.
4. To understand the impact of deep learning in our model.
5. To provide guidance and support for ophthalmologists to detect disease faster and accurately.

Chapter 2

Related Work

Drusen are little yellow deposits of lipids, which are fatty proteins. Drusen Macular Degeneration (DME) is characterized by the accumulation of drusen behind the retina. Slowly but surely the primary vision begins to fade as little stones of dirt build over time. Direct vision deformation (DMD) produces distortions of straight lines, difficulty adapting from bright to low lighting and blurry or fuzzy images in the center of vision. As intraretinal fluid accumulates in the inner and outer layers of the retina, it thickens the retina. This condition is known as Diabetic Macular Edema (DME). The retinal vasculature's hyperpermeability is assumed to be the reason. DME can be found in diabetic retinopathy at any stage. Blurred vision, washed-out colors, more floaters in vision, and double vision are some of the prominent symptoms. One cannot keep drusen from growing. It is regarded as typical to have some hard drusen. Routine medical exams can allow us to identify drusen early and determine if one has DMD. Not everyone who has drusen will get DMD later in life. If anyone has drusen then they do not need treatment unless they also have DMD. Early therapy for DMD can help to decrease the disease's development and reduce macular degeneration, stated in a journal [8]. Similarly, DME is also known as an incurable disease. Although the therapies for acute and distributed DME are different, each requires laser operations. The majority of clinicians treat focal DME with focused laser treatment and diffuse DME with grid laser treatment. The objective of these types of treatments is to stop macula leaking. A DME treatment usually takes 3-6 months to recuperate. The affected individual may feel light sensitivity, inflammation in the eye, and black patches in the middle of their vision while the eye recovers and the swelling in and around the macula decreases. All of those are common adverse effects that will likely fade away with time. However, surgical intervention does not often help people with DME see better. Although there is occasionally hardly anything that an individual can do to preclude diabetic retinopathy or DME through remaining consistent with their lifestyle, staying physically fit, ingesting plenty of nutritious food along with visiting the eye specialist once at a time to get updates regarding their eyes. Eventually, this will increase the chances of avoiding them. There was also an elderly guy who underwent eye surgery and panretinal photocoagulation (PRP) at a health center in [9] who had cataracts and progressive diabetic retinopathy in one eye. During the month of July 2017, the patient was sent to our clinic after having been diagnosed with severe DME in one of his eyes. He had previously been diagnosed with emmetropia in one eye and had a history of vision loss in another eye as a youngster, according

to his doctor. The VA (corrected visual acuity) of the patient's right and left eyes was not clearly discernible throughout the examination. One eye was found to be implanted with an intraocular lens. Aflibercept was injected into the eyes four times as part of the therapy. An MH was formed in December 2017 when hard exudates were collected on the macula's temporal side. On the other hand, a normal fluid cuff showed a lot of granular shadows in it, while the MH edge toward the vitreous cavity had grown a convex surface throughout the OCT test. Left eye visual acuity was 0.46. (logMAR). Patients diagnosed with MH were treated with PPV. After surgery to remove the vitreous core, an artificial separation between the macular area and the periphery of the vitreous was formed in the posterior portion of the eye. An epiretinal macular membrane (EMM) was seen surrounding the MH, although it was not stained with Brilliant Blue G (BBG), suggesting that this could be an unintended result of the anti-VEGF injections. An intraocular fluid-air exchange was performed while the ERM and ILM were separated. Immediately after surgery, the patient was instructed to remain in an erect position.Following surgery, the macular hole (MH) was sealed, but the foveal retina lost some of its strength, and VA remained at 0.46. (logMAR). In the year [10], our clinic was referred to a patient who was 61 years old and had hazy vision in both eyes. Both of the patient's eyes had a best-corrected vision of 1.2. Her medical history and test results were typical. Funduscopic examination, on the other hand, indicated macular atrophy in both eyes. RPD spread throughout the retina's posterior pole and midperiphery, enclosing the macular atrophy.It had been four years after her first visit when she was sent to our clinic because her visual acuity had deteriorated to the point where it was 0.1 in the right eye and 0.08 in the left eye. The pressure inside of the eye was not outside of the normal range at any point. Macular atrophy had advanced bilaterally, with the increasing visibility of the choroidal blood vessels, according to funduscopic examination. It was still possible to view the RPD in the locations that had been depleted. Imaging using fundus autofluorescence and nearinfrared reflectance indicated hypofluorescence in the posterior pole.Following the discovery of a transmission defect with granular hyperfluorescence by fluorescein angiography, an indocyanine green angiography was performed, which demonstrated hypofluorescence inside the lesion. In both eyes, there was no evidence of choroidal neovascularization. The Heidelberg Retinal Angiograph 2 was used to capture fluorescein angiograms, indocyanine green angiography, and near-infrared reflectance. It was plainly seen above the retinal pigment epithelium using optical coherence tomographyUsing optical coherence tomography, it was observable above the retinal pigment epithelium. Reduction in macular thickness and obscuration of the photoreceptor line in both eyes raised the choroidal signal. Multifocal electroretinograms showed a reduction in amplitude. Low vision therapy is still being provided to the patient. Here we discovered that DME is a disease that can develop at any phase of DR and is distinguished by edema and retinal thickening—the latter of which may also contain hard excess fluid, the most prevalent cause of loss of vision from DR. The eyes of 413 diabetes patients were analyzed in this research, which was conducted between January 2011 and July 2012. 15.3% of individuals had DME and the researchers discovered [11]. Approximately 746,000 adults over the age of 40 in the United States are at risk of developing DME, which has a 3.8 percent weighted likelihood. DME was more common in those with T1DM than those with T2DM (14 percent in global pooled data) (6 percent in worldwide pooled data). The Wisconsin

Epidemiologic Study of Diabetic Retinopathy, conducted over a period of 10 years, revealed that around 20% of patients with T1DM and 14% to 25% of patients with T2DM had DME. Patients with T1DM who participated in this study had a DME rate of 29% over the course of 25 years, with a clinically severe DME rate of 17% [12] Imaging of the fibrovascular complex and fluid accumulation may be used to assess the presence of CNV using OCT.Choroidal neovascularization, also known as CNV, refers to the development of new blood vessels that extend from the choroid into the subretinal or sub-RPE area. CNV causes blindness[5]. Despite this, it is unable to differentiate between vascular and fibrous tissues. The whereabouts and activities of CNV remain unclear as a consequence [44].
Age-related macular degeneration, more often referred to as AMD, is the most common cause of blindness that cannot be reversed. People who are 50 and older make up the vast majority of those who are diagnosed with AMD. You put yourself at a higher risk of getting age-related macular degeneration (AMD) as you become older. More than 14 percent of those aged 80 and older in the United States are affected by the disorder, making elderly white Americans the group most likely to suffer from it.Among the many age-related eye disorders that women are more likely to suffer from, Age-related Macular Degeneration(AMD) is one of the most prevalent. Women were reported to be 65 percent more likely than men to be diagnosed with AMD in 2010.. The majority of those diagnosed with AMD in the United States are white. In the United States in the year 2010, patients diagnosed with AMD were mostly Caucasian. On the other hand, AMD was found in 4 percent of people of African-American and Hispanic American descent in the United States [14]. AMD affects around 11 million individuals in the United States (US), whereas the total number of persons affected by the condition worldwide is 170 million. As a consequence of this, AMD is the primary reason why vision loss occurs in industrialized countries, and it is the third most common reason globally. It is anticipated that the number of people living with AMD in the United States will reach 22 million by the year 2050, while the number of people affected globally will reach 288 million by the year 2040 [15]. In addition, Krizhevsky et al. [16] make use of e ImageNet LSVRC-2010 to categorize around 1.2 million high quality photos. An architecture of a neural network typically consists of numerous layers. Layers like convolutional, pooling, and fully connected are used, among others. Pooling layers are used after the convolutional layer in a convolutional neural network.[17]. Every slice of the representation is independently treated throughout the pooling procedure.[18]. There is a residual learning strategy described in [43] that uses noisy images to learn noise and subtracts the residual noise from the noisy image to get a denoised image instead of using latent clean images to learn noise. To further enhance model learning, batch normalization and residual learning are used together. They get improved precision and a better outcome as a consequence of the whole procedure.

Chapter 3

Work Plan

3.1 Research Methodology

In this section, the technique that has been recommended for this thesis will be illustrated. In line with the approach that we have proposed, we want to make use of OCT photos in order to diagnose the sickness.OCT uses interferometry, which is a non-invasive imaging technique that does not need any contact with the patient [19]. The findings of the research indicate that OCT pictures may be used to acquire measurements between the layers of the retina. Because of this, it is possible that these images can be used to identify a broad variety of eye illnesses. The key advantages of optical coherence tomography (OCT) over earlier methods are its non-invasive nature, superior resolution, quick scan speed, and capability to deliver 3D information in a 2D environment. Residual learning, which is described in [20], uses noisy photographs to learn noise instead of latent clean photos; then, once a noisy image is denoised, the residual noise is subtracted from it. To enhance model learning even further, batch normalization and residual learning are often used in combination. As a result of the whole process, they are able to achieve higher levels of accuracy and satisfaction with the results.
The workflow consists of a conventional CNN model with a 21-layered architecture and a transfer learning approach that uses seven pre-trained CNN models (DenseNet121, ResNet50, InceptionResNetV2, InceptionV3, VGG16, VGG19, and MobileNetV2) to train the data set. The data set is then tested and validated to determine its level of accuracy. Then, their performance is compared based on the metrics of accuracy, precision, recall, F1 score in order to determine which model is the most effective for detecting and binary classifying intracranial bleeding on brain CT scan images. The stages that make up the process are broken down and presented below in a logical order.

Step 1: Collection of Data

Step 2: Data Augmentation

Step 3: Custom CNN model

Step 4: Existing CNN models

Step 5: Performance Evaluation

3.2 Dataset

Here, we will go through the methods we will be using to achieve our aim. In the beginning, we will utilize the Mendeley OCT dataset [41], which includes tagged photos. A large number of high-quality images are preferred in deep learning frameworks for model training and testing. A strong OCT picture site may be built using data reasoning. Through this, the process of classifying and differentiating between the images will be more accurate in rate.Using a technique known as "data augmentation," practitioners may greatly enhance the variety of data which is available for training models without having to acquire any additional data . By applying minor modification to the existing dataset helps to artificially generate images through controlling orientation, brightness, scale, location and so on. The way is simple to improve the model's forecast accuracy without requiring significant adjustments to the model itself. We are planning to use about 15,400 images in our experiment which will have different categories such as DMD, DME, CNV and Normal images.

3.3 Data Sample

There are OCT pictures of a damaged retina shown below (Figure 3.1), as well as an image of a normal retina. OCT images of the retina are shown in the images. If this is true, it means that retinal images with a wide range of changes in the area are affected, and the converse is true.

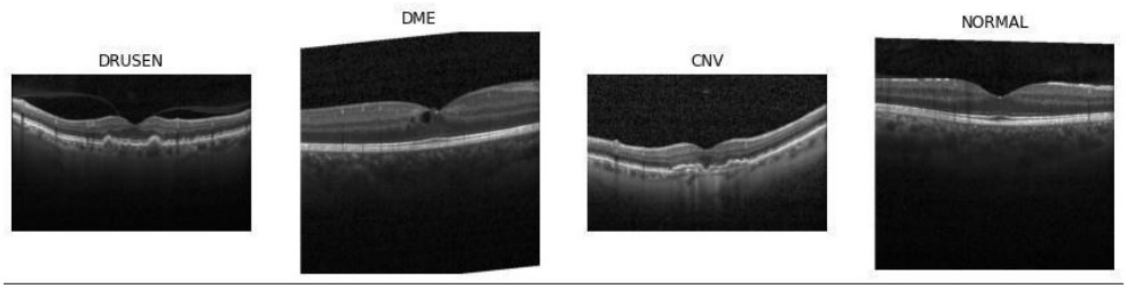

Figure 3.1: Sample Data

Training Set
For a successful input-to-output mapping, the model weights are updated using a training dataset. Neural networks are trained on real-world examples of data and solutions that help them generalize data into a consistent input-output relationship. Algorithms or output labels may be a source of input for a system. This training phase is handled by an optimization strategy that searches a space of potential values for the weights of the neural network model for a set of weights that perform well on the training dataset.

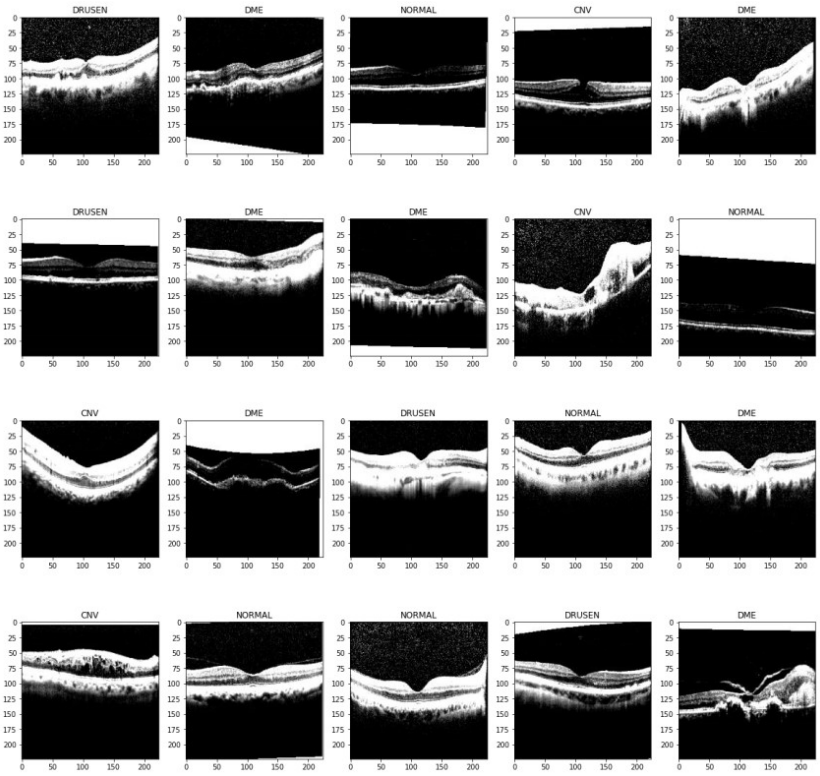

Figure 3.2: Data training

Testing Set
The algorithm that we deployed will learn from the training set in order to maximize the potential benefits that may be gained by using real-world data. The positive results for an unknown test collection will boost the algorithm's confidence in the actual world.

Validation Set
On our training dataset, an unbiased assessment of model fit is employed to tweak the model's hyperparameters. An unfair measurement occurs when the efficiency of the validation dataset on the selected model is mentioned in the design model.

3.4 Data Labels

Here, we have the option of dividing our dataset into two distinct labels. Human eyes with and without retinal disease are seen in these images. Therefore, a binary classifier may be used to characterize our dataset. To discriminate between two identical binary datasets, we use the "Class" attribute.
In Figure 3.3, we can see an example of label balance. There are two methods to depict data in a bar chart. In this scenario, 0 takes priority over 1. That is, our collection has more unaffected photos than affected ones. The class property indicates that the number of uninfected eyes is 0, whereas the number of diseased eyes is 1.

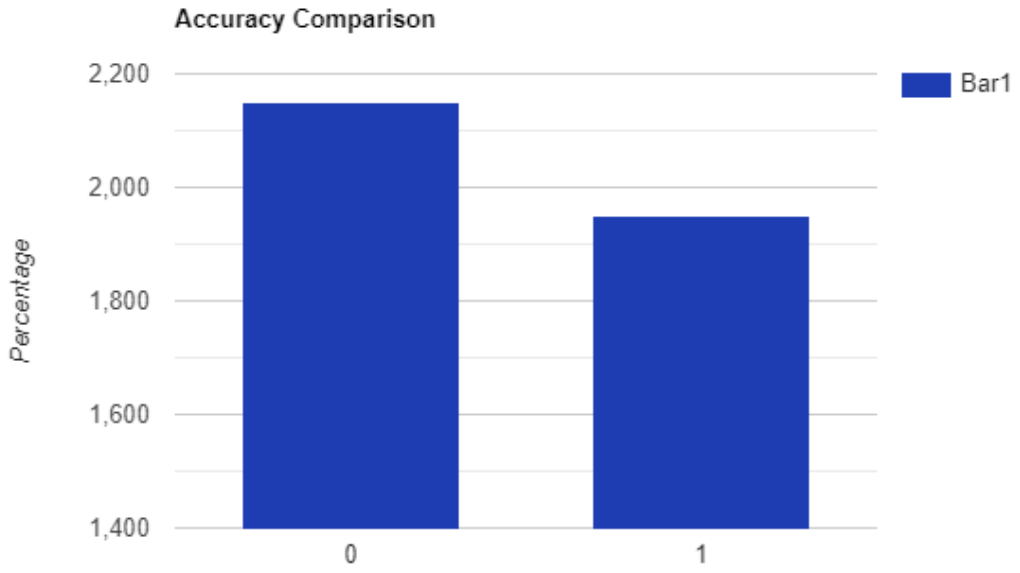

Figure 3.3: Bar-chart illustration of label balancing

3.5 Image Resizing

To function effectively, convolutional neural networks need equal picture sizes. All photos must be scaled to a constant size before to being fed into the CNN, since neural networks need inputs of the same size. The less the shrinkage needed, the greater the fixed size. Less shrinkage results in less distortion of the image's characteristics and patterns. Though CNNs need consistent image sizes, there are a few very simple solutions for datasets including images of varying sizes [21]. We may randomly crop and resize the photographs to the same dimensions. This technique delivers more robust data enhancement. In order to effectively execute the picture compression framework and prevent memory overflow, each original OCT image of the retina with a size of 150 x 150 pixels was scaled down along the top and bottom bounds during network training [22]. This was achieved by using the TensorFlow framework.

3.6 Data Augmentation

In order to train neural networks for application in real-world situations, it is necessary to supplement the data available. In order to increase our model's ability to generalize and make better, more accurate predictions on data that it has not been trained on, we may use data augmentation to supplement the data that it has been trained on. TensorFlow has two techniques for doing data augmentation, both of which are described here. Choosing the first option is much less complicated and time-consuming. It is somewhat more difficult to utilize the second method (often because you must read the TensorFlow documentation in order to discover the precise functions you want), but it allows you to take more responsibility for the

data augmentation process. By suddenly shifting the picture orientation, we were able to utilize data augmentation techniques to increase the output of our network in our system. Vertical and horizontal flipping and rotation were used in the first photographs by the augmentation operator at 90, 180, and 270 degrees vertical and horizontal rotation.

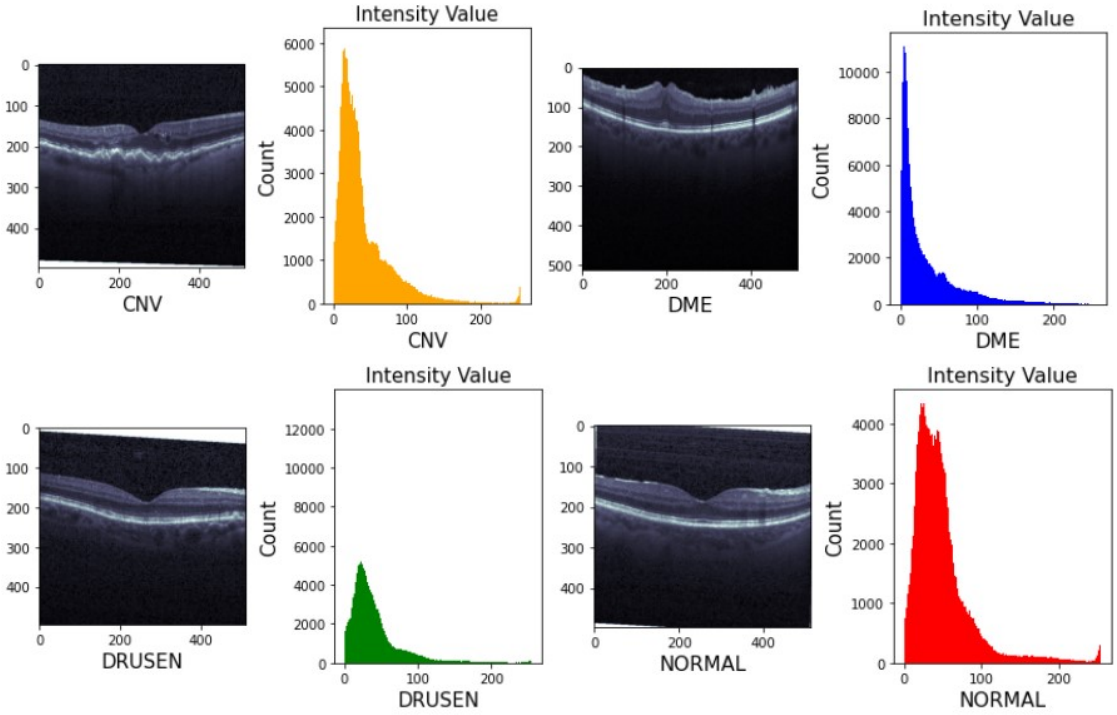

Figure 3.4: Pixel intensity values

3.7 Batch Normalization

Batch normalization is the process of normalizing system data that may be delivered to the activation functions of a preceding layer or the inputs themselves. Generalization error is reduced by accelerating training and giving some generalized linearity. Batch normalization solves the issue of internal covariate shift. Each layer of the network may adapt more independently and can be used to equalize the output of the prior levels using batch normalization.

Chapter 4

Implementation of CNN models

4.1 CNN

A convolutional neural network, sometimes known as a CNN, is a specific kind of neural network that is optimized for the processing of input that has an architecture similar to a grid, such as an image[18]. Neural networks are made up of many different parts, one of which is the convolutional neural network (CNN). In order to identify objects, recognize faces, and so on, CNNs employ visual recognition and classification. They are composed of neurons that may be trained to change their weights and biases. The most common usage of CNNs is to categorize pictures, group them into clusters based on similarity, and then identify specific objects. Faces, street signs, animals, and other recognizable objects may all be recognized by algorithms that use CNNs [6]. The convolutional, pooling, and fully linked layers of a CNN are the most common.The first layer of a CNN network, the Convolutional Layer, does the majority of the computing effort. Utilizing filters or kernels to generate convolutional data or images. By adjusting the slider, we may add filters to the data. If the RGB value of the image's depth is 4, a filter with the same depth would also be applied. For each sliding movement, a particular value is taken from each filter in the picture and added together. A 2D matrix is the result of applying a 3D color filter on a convolution with a 2D output. Down sampling features are the third step in the Pooling Layer. Every layer of the 3D volume is coated with it. Flattening is the last step in the process of creating a fully connected layer. The neural network is given a single column of the pooled feature map matrix, which is subsequently processed. We were able to develop a model by connecting all of the layers together. We can then use an activation function like SoftMax or Sigmoid to further categorize the data generated by the algorithm. Softplus units increase DNN performance and reduce convergence time compared to sigmoid and ReLU units.[25]

Optimizer Adam
An optimizer is a procedure or algorithm that alters the properties of a CNN architecture, such as the parameters and the learning rate. Examples of these properties include: As a result, it contributes to the total reduction of damage and actually increases efficiency. [23] Adam is an extension of stochastic gradient descent, which has gained popularity for deep learning in computer vision and NLP. These include approaches for image processing and voice recognition. [24]
Reiterating the optimizer is a deep-learning strategy. Equation(4.1) explains the

"Adam" optimizer function.

$$\omega t + 1 = \omega t - \alpha mt \tag{4.1}$$

mt = aggregate of gradients at time t, α = learning rate at time t, ωt = weights at time t, ωt+1 = weights at time t + 1.

Softmax A nonlinear softmax output layer is commonly used when neural networks are used for pattern classification tasks. This, as we all know, is standard procedure. Because of its non-linearity, the soft-max output layer of a neural network has the ability to make significant changes to the frequency at which the network generates outputs.

4.2 Proposed CNN Model

21-Layered CNN Model

Information processing that has a grid-like structure, such as an image, is the area of expertise of a class of neural networks called Convolutional Neural Network, which is sometimes abbreviated as CNN or ConvNet for short[18]. A binary representation of visual data is what we refer to as a digital picture. It comprises a sequence of pixels. In the solution that has been suggested, a multi-layered deep CNN model has been used in order to differentiate between real and fake images. Fully connected (FC), Convolutional layer, and Pooling layers are the foundation of a CNN model in addition to that, we will be using the Dropout layer. Below, we will go over the specifics of each layer.

Convolutional Layer: An essential component of a CNN is a convolutional layer. All the settings for these filters (or kernels) must be learned during the training process. It's common for filters to be smaller in size than the image they are intended to enhance. This layer uses kernel filters to extract essential information from the input images that are convolutionally processed. The kernel filters are similar to the input images, but they have lower constant parameters. Edge detection, blurring, and sharpening can be accomplished through the convolution of an image with several filters. In the convolutional layer, we used the Conv2D layer to construct this CNN model. The model was built with a total of ten Conv 2D layers. [45]

Pooling Layer: Following the convolutional layer in convolutional neural networks are layers known as pooling layers. In order to improve the efficiency of the computations being performed, pooling is used to reduce the amount of the features that are extracted, and therefore, the number of trainable parameters. The pooling filter defines the amount of the range that is summarized by the pooling procedure. If a filter's parameters are 2x2, then the summary section is also 2x2 in size. Here we can detect four layers in total with other layers. [46]

Fully Connected Layer(FC): This layer is comprised entirely of feed forward neural networks. Fully Connected Layers(FNN) are the layers that come after the final few in the network's architecture. After that, the output of the last pooling or convolutional layer is flattened before it is sent on to the fully connected layer as the input. This model carries the following layers :

Flatten Layer: Once the fourth MaxPooling layer has been used, a single flatten layer will be applied. In the end, this is beneficial for the network as a whole in general.

Dense Layer: In addition to the flattening layer, this model has two dense layers. The outputs of previous levels are sent to all neurons in this layer.

Dropout Layer: During the training process, this layer will periodically reset all of the inputs to zero as a means of reducing the possibility of the model being too accurate.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_10 (Conv2D)          (None, 124, 124, 60)      1560

 conv2d_11 (Conv2D)          (None, 120, 120, 60)      90060

 conv2d_12 (Conv2D)          (None, 117, 117, 30)      28830

 conv2d_13 (Conv2D)          (None, 114, 114, 30)      14430

 max_pooling2d_4 (MaxPooling  (None, 57, 57, 30)       0
 2D)

 dropout_4 (Dropout)         (None, 57, 57, 30)        0

 conv2d_14 (Conv2D)          (None, 53, 53, 60)        45060

 conv2d_15 (Conv2D)          (None, 50, 50, 30)        28830

 conv2d_16 (Conv2D)          (None, 47, 47, 30)        14430

 max_pooling2d_5 (MaxPooling  (None, 23, 23, 30)       0
 2D)

 dropout_5 (Dropout)         (None, 23, 23, 30)        0

 conv2d_17 (Conv2D)          (None, 20, 20, 60)        28860

 max_pooling2d_6 (MaxPooling  (None, 10, 10, 60)       0
 2D)

 conv2d_18 (Conv2D)          (None, 8, 8, 30)          16230

 conv2d_19 (Conv2D)          (None, 6, 6, 30)          8130

 max_pooling2d_7 (MaxPooling  (None, 3, 3, 30)         0
 2D)

 dropout_6 (Dropout)         (None, 3, 3, 30)          0

 flatten_1 (Flatten)         (None, 270)               0

 dense_2 (Dense)             (None, 5000)              1355000

 dropout_7 (Dropout)         (None, 5000)              0

 dense_3 (Dense)             (None, 4)                 20004

=================================================================
Total params: 1,651,424
Trainable params: 1,651,424
Non-trainable params: 0
_________________________________________________________________
```

Figure 4.1: Proposed Model Architecture

4.3 Pre-trained CNN Models

4.3.1 VGG16

It is generally agreed upon that VGG16 is among the very finest vision model architectures that have ever been developed. Karen Simonyan and Andrew Zisserman at Oxford introduced this CNN model. The approach was suggested in 2013 but presented during the 2014 ILSVRC ImageNet Challenge. They named it VGG after Oxford's Visual Geometry Group, where they worked. Authors suggested network designs depending on depth. All ImageNet Challenge settings use a stack of multiple (3 x 3, stride 1, padding 1) convolution layers followed by a 2 x 2 maxpooling layer. Multiple stack combinations were cycled to reach different depths. Each configuration's number indicates the number of weight-parameter layers. Throughout it uses convolution and max pool layers. Ending with 2 FC and a SoftMax. VGG16 has 16 weighted layers. This network contains 138 million parameters (estimated)[26].

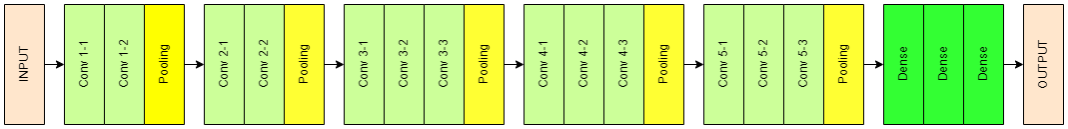

Figure 4.2: VGG16 Architecture

4.3.2 VGG19

This convolutional neural network was trained using ImageNet photos. VGG19 has pre-trained layers and a thorough grasp of form, color, and structure. To train VGG19, we used millions of pictures with tough categorization tasks. A keyboard, mouse, pencil, and a variety of animals [27] are among the more than 1000 items that may be classified by the network's 19 layers. Thus, the network has been able to learn a wide range of features for a wide range of images.
224x224 pixels is the VGG19's image input size. The convolutional layers of VGG make advantage of the smallest possible receptive field to capture up/down and left/right motion. Following this is a ReLU activation technique.The ReLU activation function is considered to be a piecewise linear activation function since it outputs the input when it is negative and returns zero when it is positive. The stride value for maintaining spatial resolution after convolution is 1 pixel[28].

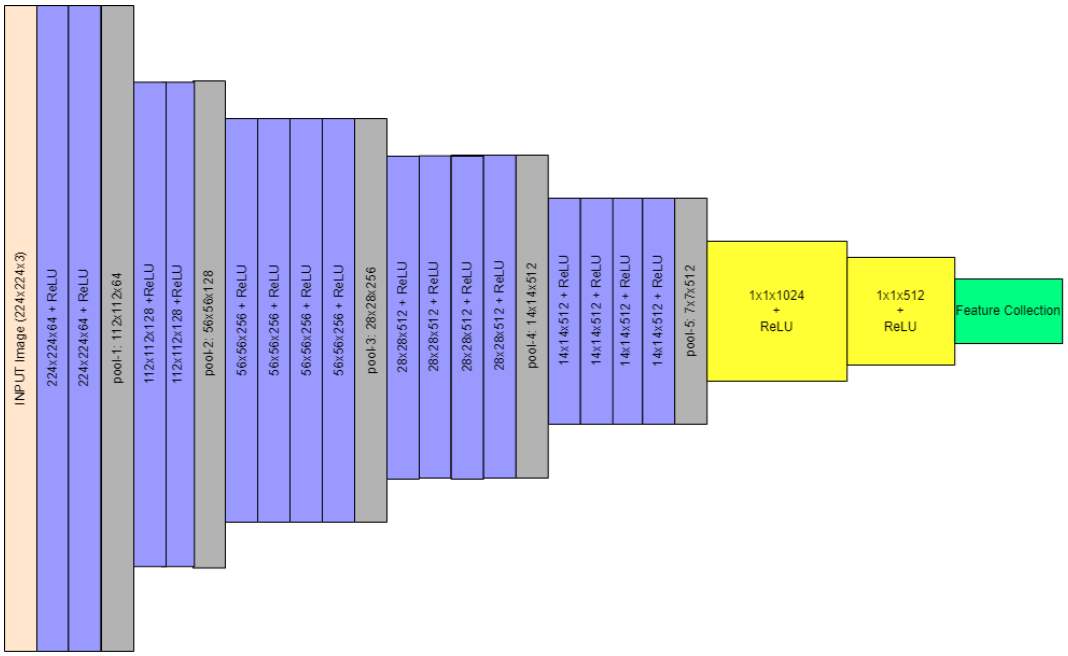

Figure 4.3: VGG19 Architecture

4.3.3 DensNet121

In a typical intake Convolutional Neural Network (CNN), each convolutional layer obtains overall information from the previous convolutional layer and produces the desired extracted features which is then passed on to the next convolutional layer, with the exception of the first (which takes in the input). 'L' layers have several direct 'L' connections between layers. The 'dwindling gradient' issue arises as CNN network depth grows. When the knowledge channel from source to destination layers lengthens, some data may 'disappear' or be lost, lowering the network's training capability[29]. DenseNets simplifies layer connections and changes CNN architecture to resolve this situation. Each layer in a DenseNet design is directly linked to every other layer, thus the name. L-levels have L(L+1)/2 direct connections[30].
Contrary to common assumption, DenseNets have fewer variables than conventional CNNs since they do not have to learn duplicate feature maps.
Additionally, several ResNet implementations have shown that many layers contribute relatively little and may be eliminated. ResNets contain several parameters since each layer must learn its own weights. DenseNets layers, on the other hand, are very limited and add a negligible number of new feature-maps[30].
Due to the previously outlined information flow and gradients, training certain very convolutional models proved to be challenging. Because each layer has direct connections to the gradient descent slopes and the input data picture, DenseNets addresses this issue.

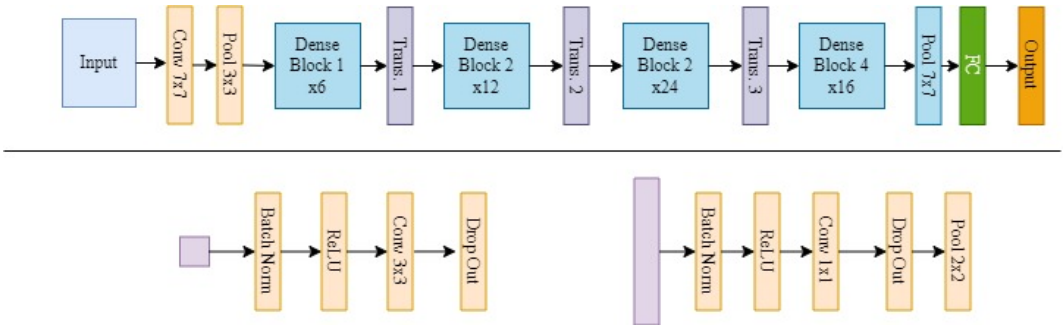

Figure 4.4: Densnet121 Architecture

4.3.4 MobNetV2

In order to reduce the size of the model and the complexity of the network, Mobile Nets are a depth-wise separable convolution design that reduces the number of connections. Embedded and mobile applications benefit from the technology. The author has included two global hyperparameters into this sort of network, which are as follows: A good balance between model latency and accuracy is achieved with this technique. In addition, the hyperparameters give the capability of selecting a suitably scaled model in accordance with the problem restrictions, if necessary.

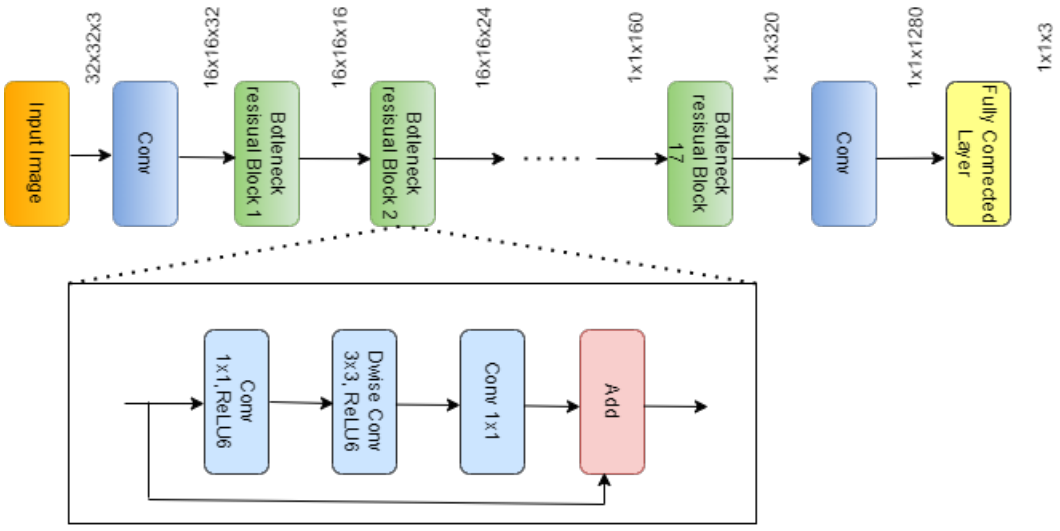

Figure 4.5: MobileNetV2 Architecture

4.3.5 ResNet50

As a Convolutional Neural Network (CNN) architecture known as ResNet, it is a strong backbone model for many computer vision applications. It's a 50-layer convolutional neural network. Skip connections are used by ResNet to transfer data from one layer to another. Because of this, convolutional networks with up to thousands of layers may be built, which outperform shallower networks. As a result, the issue of a vanishing gradient was lessened. 224 × 224 pixels is the constant size of the input pictures in this model. Several variations of ResNet are available, each with a different number of layers. It is possible to train networks with a high number of layers (up to tens of thousands) without increasing the proportion of training errors. The vanishing gradient issue may be addressed via ResNets and identity mapping. Using the ImageNet database, we can import a network that has already been trained on over a million photos.In addition to keyboards, pencils and mice the pretrained network can identify more than a thousand other things. Deep neural networks with more neural layers are more efficient and have a lower proportion of mistakes when using this approach. As a result, it is now feasible to train far deeper networks than was previously allowed, thanks to the use of skip connections.

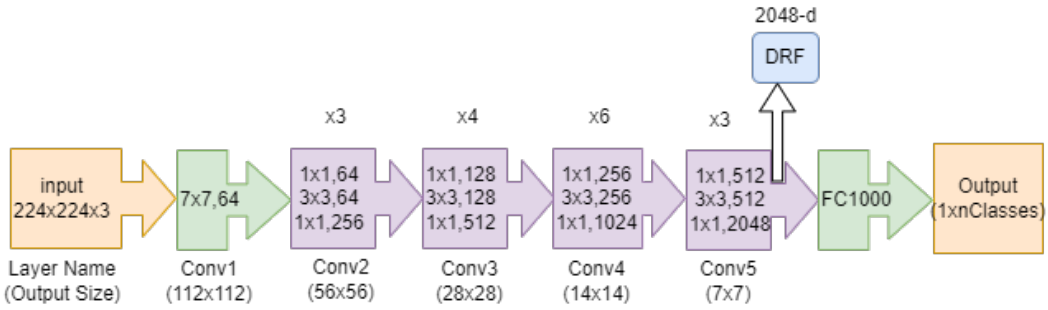

Figure 4.6: Resnet50 Architecture

4.3.6 InceptionV3

The Revised Inception for Computer Vision was published in 2015 and included Fig.1 as part of the proposal for a new version of Inception. Including both terms of total quantity of parameters produced by the system and in light of the financial cost incurred, Inception Networks (GoogLeNet/InceptionV1) outperform VGGNets (memory and other resources) [31]. Using this software, photos may be sorted into over a thousand different object groups. For transfer learning, the Inception-v3 model is amongst the most preferred options. With this, we can go back and retrain the final layers of current products, which saves a lot of time. Inception-v3 was trained on over a million pictures from the ImageNet database, demonstrating that the model may be utilized on a smaller dataset with high accuracy classifications without retraining. The Inception Layers are a set of layers (11 convolutional layer, 33 convolutional layer, 55 convolutional layer) that combine the result filters into a single output vector, which produces the parameters for the next stage. Changes to an Inception Network must be handled with care to avoid losing any operational advantages. Since the new network's performance is so uncertain, modifying an Inception network for various use cases becomes a challenge. Inception v3 has already provided a number of ways to improve the network in order to free up the restrictions for quicker model adoption. Batch normalizing, down sampling and parallel computation are only few of the techniques utilized in parametric convolution. Two sets of parameters: 5 million (V1) and 23 million (V2) (V3).[32]

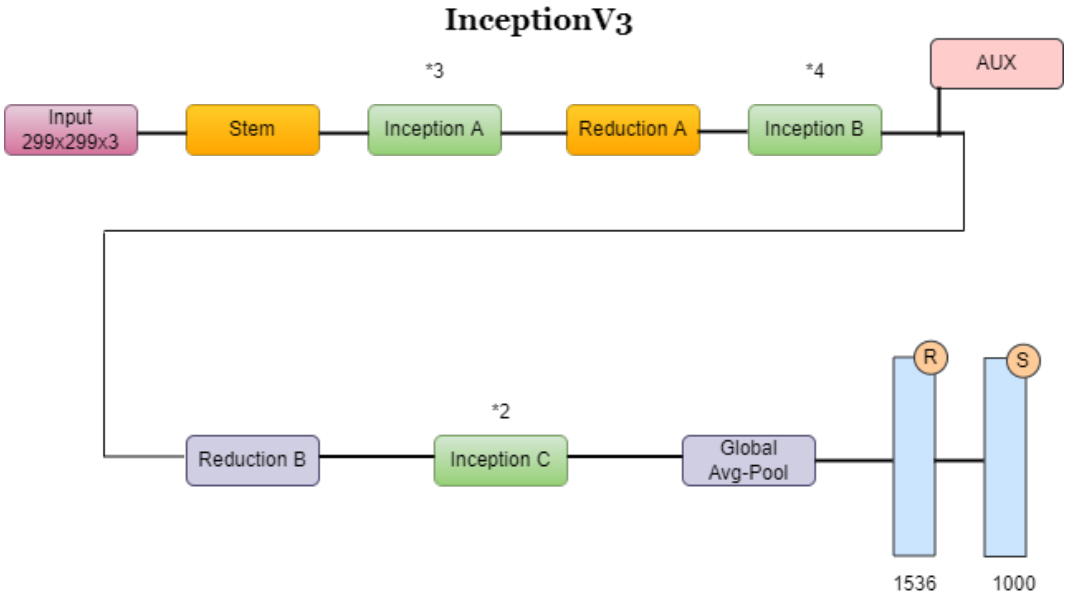

Figure 4.7: InceptionV3 Architecture.

4.3.7 InceptionResNetV2

As an extension of the Inception family of architectures, Inception-ResNet-v2 contains residual connections. Images from the ImageNet database are used to train it. One thousand item categories may be categorized by the network, which has 164 layers and is capable of classifying photos into 1000 different types of objects. Consequently, the network has learnt a broad variety of visual feature representations. Class probabilities are calculated using a collection of picture input dimensions of 299-by-299. It is based on the framework of Inception and the Residual link. Multiple convolutional filters and residual connections are used in the Inception-Resnet block. In addition to reducing training time, using residual connections prevents the degradation issue produced by deep structures. Resnet-fundamental v2's network architecture is shown in the diagram.

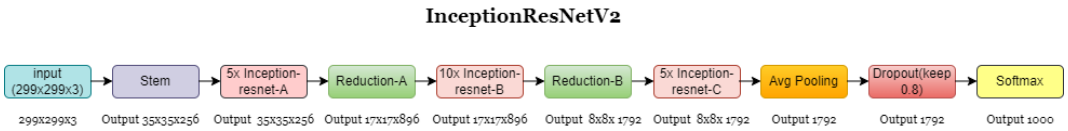

Figure 4.8: InceptionResnetV2 Architecture

Chapter 5

Performance Analysis

The objective of doing a performance analysis is to validate the innovative applications of technology that have been employed to increase performance. This gives us an overview of the entire work and highlights the achievements and flaws of a specific work which help improve existing approaches and learn new ones. It can also be used to evaluate strengths and flaws of others.Tactical and technical appraisal, movement analysis, and statistics compilation are the most important aspects of performance analysis.[34]

5.1 Performance Parameter

The F1 score, precision, accuracy, and recall have been computed for each of these models in order to evaluate and provide a comparative analysis of the findings. This was done in order to assess, analyze, and highlight a comparative analysis of the outcomes based on a comparison of the performances of both conventional and pretrained CNN models. This part will begin with a discussion of the equations for the performance measures employed throughout the investigation.

$$Accuracy = TP \ + \ \left(\frac{TN}{CP}\right) \ + \ CP \ + \ CN \tag{5.1}$$

$$Precision = TP \ * \ \left(\frac{TPR}{TP}\right) \ + \ FP \tag{5.2}$$

$$Recall = TP \ * \ \left(\frac{TPR}{TP}\right) \ + \ FN \tag{5.3}$$

$$f1 \ score = \frac{TP}{TP \ + \ \frac{1}{2} \ * \ (FP \ + \ FN)} \tag{5.4}$$

Here, TP=True Positive, TN= True Negative, CP= Condition Positive, CN= Condition Negative, TPR= True Positive Rate, FP= False Positive, FN= False Negative

The training dataset for this model is evaluated against the proposed network based on a number of parameters. These parameters include batch size, epoch, learning rate, optimizer, and callbacks. Following the completion of the pre-processing of the dataset, training can then begin. The Transfer Learning methodology allows for the modification of its properties prior to the start of training.Before beginning the machine learning process, the user chooses between custom CNN models and pre-trained models. Following this step, the directory containing both sets of newly formed categories is imported depending on the size of the first layer of input. There are two portions to the data: 80 percent is used for training, and 20 percent is used for testing. "Adam" [35] is the optimizer that has been employed, and it is a gradient-based method that focuses on novel forecasts of instances which are of relatively low. Randomized goal functions may be improved using this technique. The "Adam" optimizer is utilized because it is easy to design, efficient, takes minimal memory, and is resistant to gradient diagonal resizing. This is because the method is suitable for scenarios with enormous quantities of data and/or parameter values [36]. We have made use of 32 batches and 30 epochs altogether (each epoch, there are 450 iterations). We have utilized "Categorical CrossEntropy" in the part pertaining to the loss. Because "RGB" was chosen as the color mode in the section devoted to color modes, it was determined that the images would be split up into three distinct channels as a result. The execution environment has an effect on training completion times in addition to the various training options. In each and every experiment, the running environment consists of a graphics processing unit (GPU). For the sake of testing and gathering information, Google Colab Pro made use of random GPUs like Nvidia V100 or P100 [37]. On the command line of Google Colab, to investigate the GPU's settings, type "!nvidia-smi."

Table 5.1: Parameter Comparison

Parameter	21-layer proposed model	Pre-trained model
Training Data	80%	20%
Testing Data	80%	20%
Batch Size	32	32
Target Size	(224,224)	(224,224)
Epoch	30	30
Max queue size	100	100
Step	426	426
Execution Environment	GPU	GPU
Verbose	1	1
Optimizer	Adam	Adam
Loss Function	Categorical	Categorical
Initial Learning Rate	0.001	0.001
Color Mode	RGB	RGB
Class Mode	Binary	Binary
Matrics	Accuracy,Loss,Recall,Precision	Accuracy,Loss,Recall,Precision
Callback	ReduceLROnPlateau	ReduceLROnPlateau

5.2 Performance of proposed model

We decided to evaluate a total of 15,400 OCT scans, which were divided among four categories: CNV images, DMD images, DME images, and normal images. In conclusion, it was determined that the model that we had proposed was accurate 98.37 percent of the time. The table below presents the findings of accuracy and loss after training and testing.

Table 5.2: Accuracy and Loss of the Proposed CNN Model in Training and Testing

Testing Accuracy	Testing Loss	Training Accuracy	Training Loss
98.37%	10.94%	96.83%	14.18%

The table shows the testing and training accuracy of the proposed 21 layers model which is 98.37% and 95.83% respectively. The model has a loss of 10.94% in testing data and 14.18% while training data.

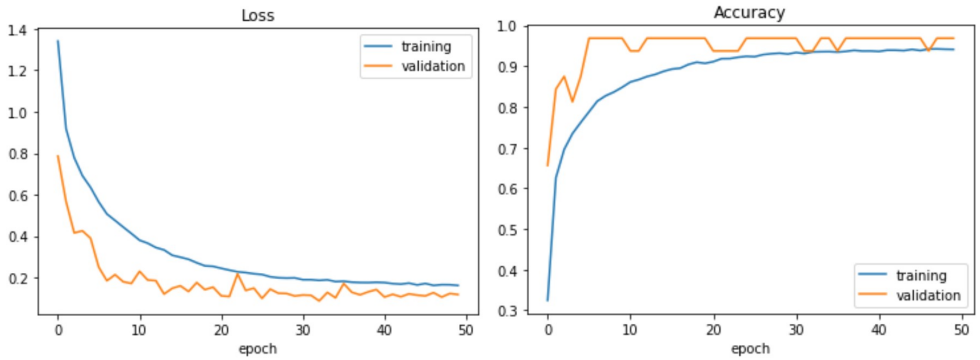

Figure 5.1: Training and validation graph of Proposed Model

From the above graphs we can estimate the accuracy and loss of the finding in terms of training and validation dataset. The accuracy graph shows the acceleration margin and scattered arbitrary lines on the training dataset. On the other hand the loss graph shows the same study in a decline manner.

5.3 Performance of pre-trained models

VGG16: Using the VGG16 model, we were able to obtain a testing accuracy of 94.21 percent. The accuracy and loss graphs for VGG16 are shown in Figure 5.2. We observed, in both the training and the validation loss graphs, that the loss lowers to a sufficient degree with time, which leads us to believe that regression is not occurring. Since we discovered that the training curves and the validation curves are similarly convergent, we are able to exclude the possibility that VGG16 overfit the data. Because of the quantity of data provided, there are occasional variations in the validation curve. Figure 5.2 shows Training and testing accuracy, loss of VGG16 model.

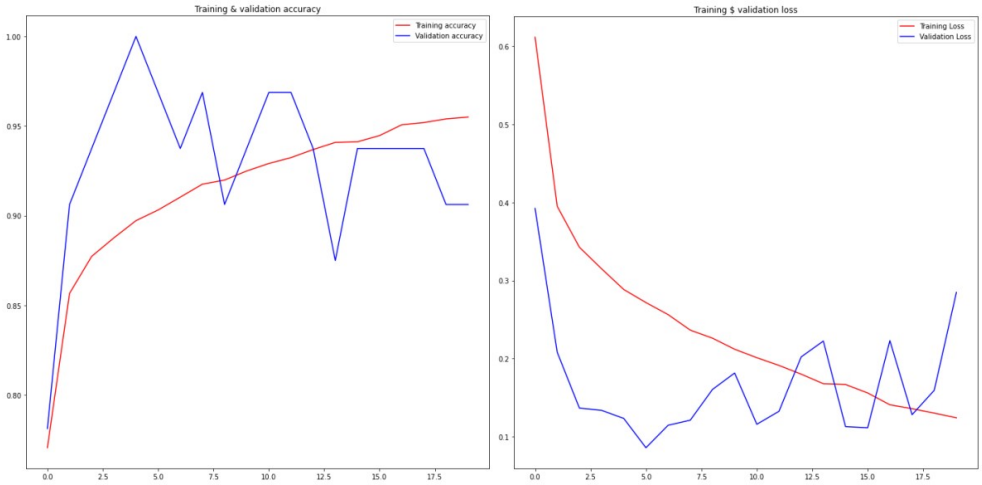

Figure 5.2: Training and validation graph of VGG16

VGG19: Using the VGG19 model, we were able to obtain a testing accuracy of 91.43 percent. The accuracy and loss graphs for VGG19 are shown in Figure 5.3. We observed that the loss lowers with time in the training and validation loss graphs, with a tiny peak point in the end.Therefore we estimate the acceleration and declination of VGG19 since the training and testing curves are both convergent. Because of the quantity of data provided, there are occasional variations in the validation curve.

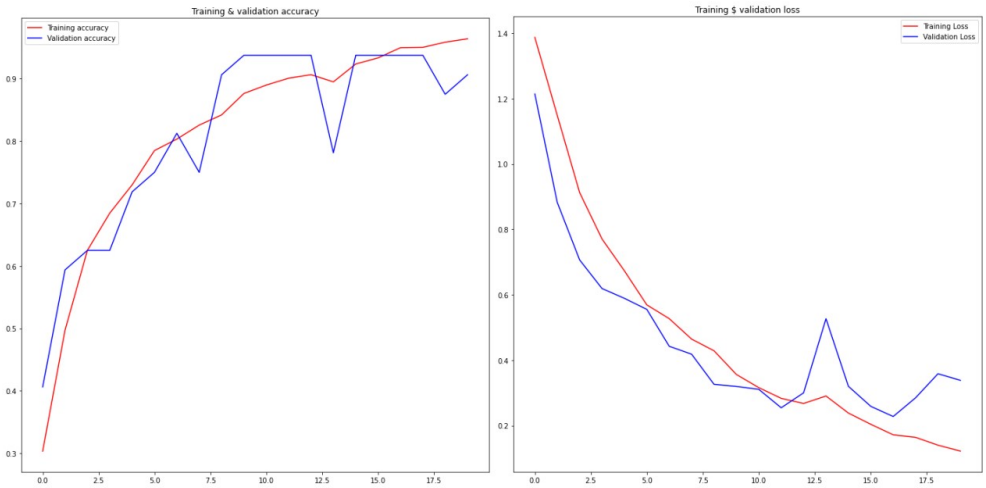

Figure 5.3: Training and validation graph of VGG19

MobNetV2: Here MobNetV2 model that we used was able to obtain a testing accuracy of 90.81 percent. The accuracy and loss graphs for MobNetV2 are shown in Figure 5.4. In the testing and training cost graphs, we can see that the damage decreases with time.Because we discovered that the training curve and the validation curve are similarly convergent, we are able to conclude that MobNetV2 did not exhibit overfitting. Because of the quantity of data provided, there are occasional variations in the validation curve.

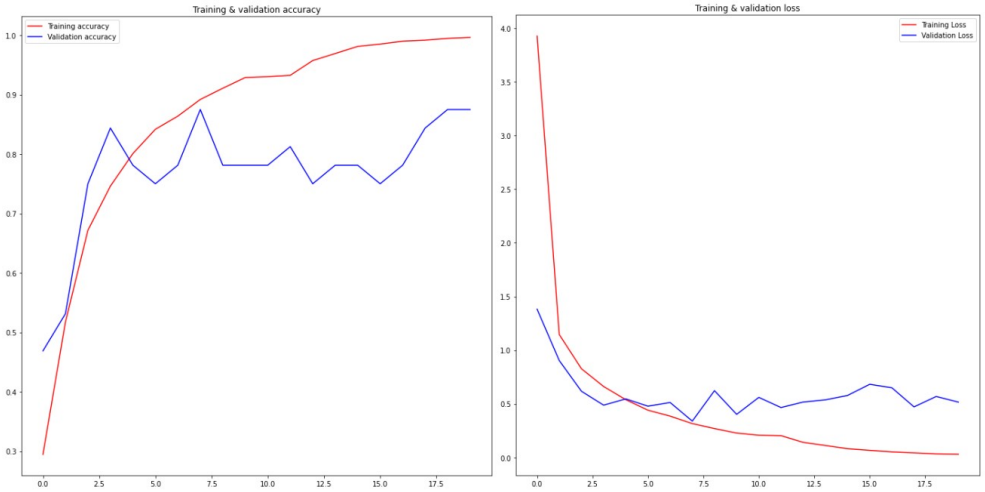

Figure 5.4: Training and validation graph of MobNetV2

Resnet50: Testing accuracy was 92.83 percent in the ResNet50 model. The accuracy and loss graphs for ResNet50 are shown in Figure 5.5. The loss value is heading towards zero. The curves also converge that shows training and testing accuracy, loss of the Resnet 50 model.

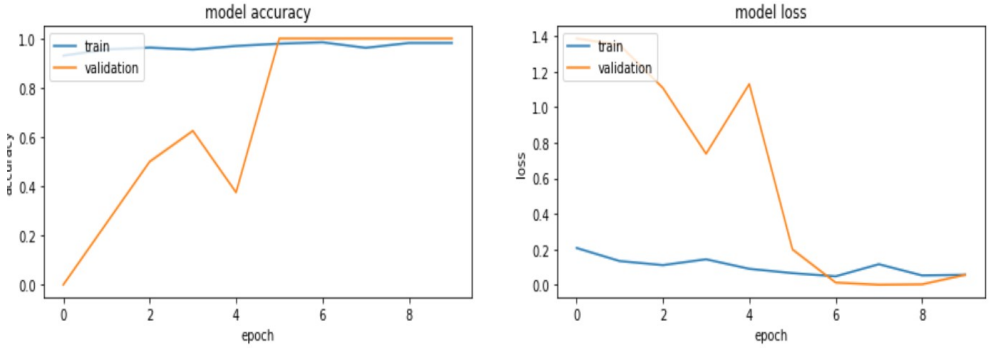

Figure 5.5: Training and validation graph of Resnet50

Densenet50: A trial efficiency score of 90.29 percent was obtained using the DenseNet121 model. The testing and training efficiency and cost of the DenseNet 121 model are shown in Figure 5.6. We saw that there is indeed a significant reduction in loss, so the curve is converging in the training and validation loss graphs. Figure 5.6 shows the training and testing accuracy and loss of the DenseNet 121 model.

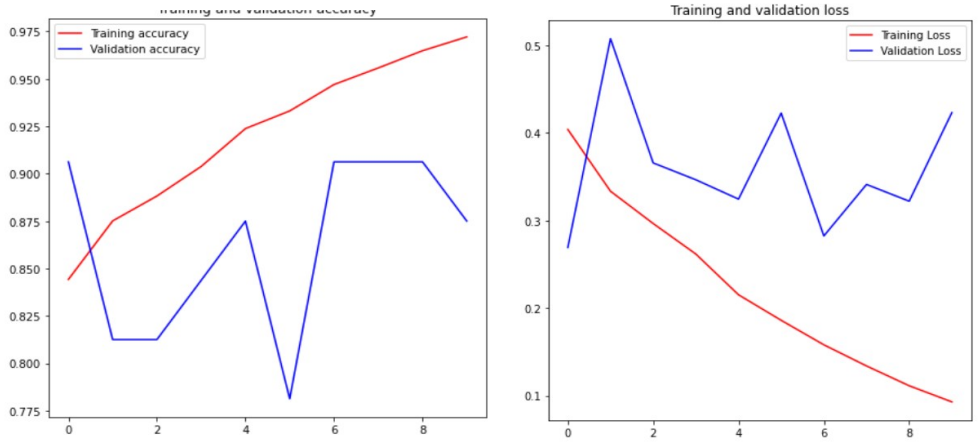

Figure 5.6: Training and validation graph of DenseNet121

InceptionV3: The testing accuracy was determined to be 89.36 percent using the InceptionV3 model. The accuracy and loss graphs for InceptionV3 are shown in Figure 5.7. Despite slight variations, the calibration and testing curves converge on each other. For the InceptionV3 model, the training and validation loss values both decrease in proportion. The curves also shows, resulting in a reduction in Inception V3 model training and testing accuracy.

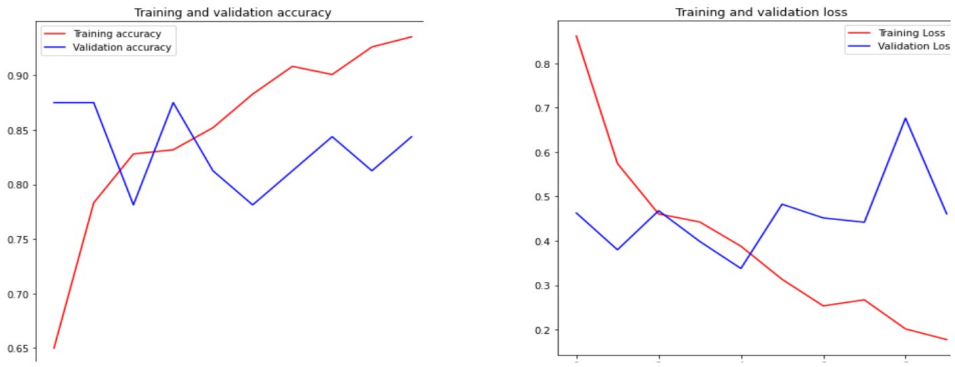

Figure 5.7: Training and validation graph of InceptionV3

InceptionResnetV2: The InceptionResNetV2 model has an accuracy of 84.46 percent in testing. Figure 5.8 depicts the accuracy and loss graphs for Inception-ResNetV2. Small variations are seen in this graph, indicating variances between the loss value training and validation curves. The curves also show the InceptionRes-netV2 model's training and testing accuracy as well as loss.

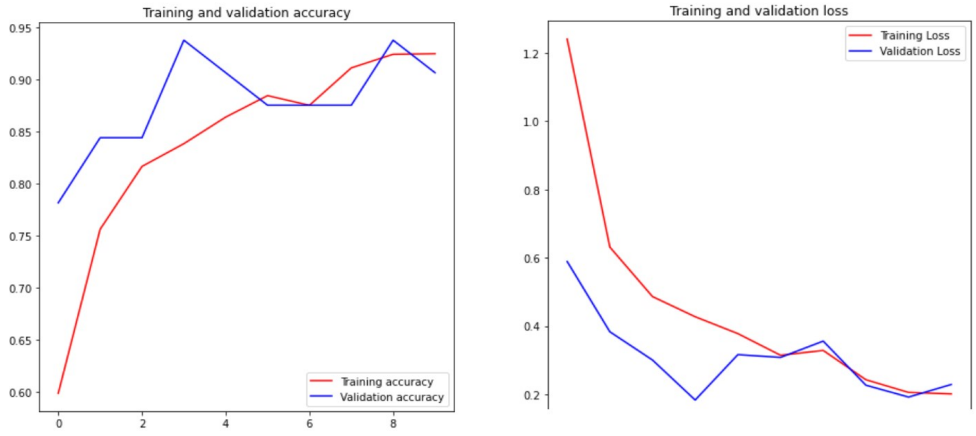

Figure 5.8: Training and validation graph of InceptionResnetV2

5.4 Compare and Analysis

There are seven pre-trained models included in this study: VGG16, VGG19, MobileNetV2, ResNet50, DenseNet121, InceptionV3, and InceptionResNetV2. A performance comparison is also given between the proposed 21-layer CNN model and seven pre-trained models. The statistics in the tables show that the suggested models perform significantly better than the existing models. The custom model and models that have already been trained Figure 5.9 shows a bar graph indicating the accuracy of the CNN models. The most accurate result possible was achieved by the proposed model which was 98.37 percent. VGG16 achieved an accuracy of 94.21 percent, while VGG19 achieved an accuracy of 91.43 percent. MobileNetV2, ResNet50, DenseNet121, InceptionV3, and InceptionResNetV2 each exhibit an accuracy that is proportionately 90.81 percent, 94.83 percent, 90.29 percent, 89.36 percent, and 84.46 percent, respectively. By Using the same dataset, we can see that the efficiency of detection the disease varies from model to model.
The below graph shows the difference between seven pretrained models along with the proposed CNN model.

Table 5.3: Comparison between CNN architectures

Architecture	Precision	Recall	F1 score	Accuracy
VGG16	93.38%	96.24%	98.12%	94.21%
VGG19	89.87%	92.68%	91.54%	91.43%
MobNetV2	90.80%	90.70%	90.80%	90.81%
ResNet50	93.98%	95.98%	93.92%	94.83%
DenseNet121	89.32%	92.77%	90.25%	90.29%
InceptionV3	88.12%	90.90%	89.41%	89.36%
InceptionResNetV2	82.72%	85.32%	84.43%	84.46%
Proposed Model	97.26%	98.98%	98.57%	98.37%

We may conclude that the customized CNN models outperform other pre-trained CNN models based on the experimental findings shown in the bar chart. The accuracy of CNN models are shown by a bar graph such in Figure 5.9 where the values are respectively VGG16 98.37 percent, VGG19 91.43 percent, MobNetV2 90.81 percent, ResNet50 94.83 percent, DenseNet121 90.29 percent, InceptionV3 89.36 percent, and InceptionResNetV2 84.46 percent are the accuracy levels achieved by these pre-trained CNN models. Therefore, the use of transfer learning models on the same dataset demonstrates how effective the proposed model can be because of the multiple layers in the architecture.InceptionV3 and InceptionResNetV2 CNN models have the lowest accuracy in the performance bar chart.

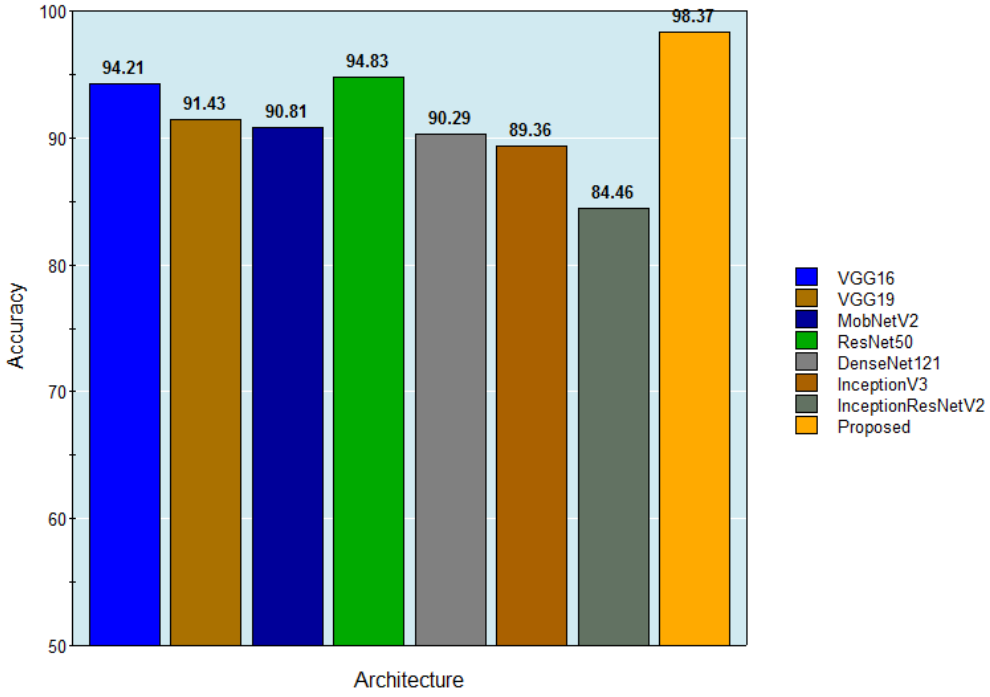

Figure 5.9: Accuracy Comparison

5.5 Accuracy Comparison on Related Work

A comparison is made in this part between the accuracy achieved by this thesis article and the accuracy of similar recent papers on datasets of OCT pictures. OCT-NET was used by Perdomo et al. in the study to diagnose retinal illness, and they had achieved an accuracy of 93.75 percent of the time. Both Lee et al. and Rajagopalan et al. showed CNN implementation for identifying retinal illness in their own study, where they reached an accuracy of 93.45 and 95.70 percent respectively. Awais et al., on the other hand, made use of a feedforward neural network, which allowed them to achieve an accuracy of 87.50 percent. An overview of the accuracy review that was accomplished by this study and other research efforts that were recently completed on the specific dataset that is being used in this paper can be found in the table that follows. When compared to the other approaches, it was discovered that the CNN model utilized in this study achieved the maximum level of accuracy possible, which was 98.37 percent. This accuracy review is represented by a bar graph which may be found in Figure 5.9.

Table 5.4: Accuracy Comparison

Approaches	Methods	Accuracy
This Paper	CNN	98.37%
Perdomo et al.[39]	OCT-NET	93.75%
Lee et al.[38]	CNN	93.45%
Awais et al.[33]	ConvNet	87.50%
Rajagopalan et al.[40]	CNN	95.70%

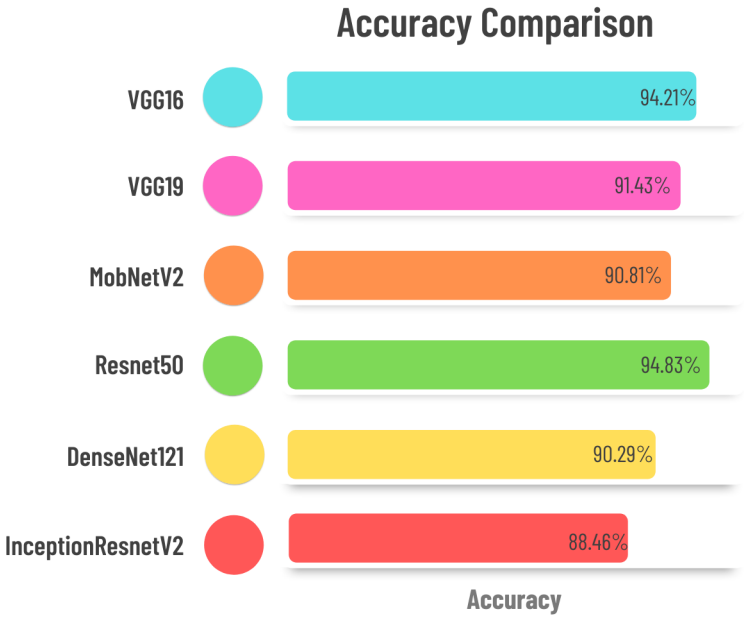

Figure 5.10: Accuracy Comparison on related works

Chapter 6

Conclusion

This study outlines a procedure that, when it comes to the treatment and prevention of eye diseases, can be of great use to ophthalmologists. According to the findings of the research article, this system has a detection rate of 98.37% and may detect these diseases in their earliest stages, hence assisting individuals in avoiding the potential consequences. We use our recommended CNN model to analyze the data, which consists of 15,400 OCT images that have been divided into three portions and pre-processed. This allows the model to distinguish between images that are clean and those that include faults. In conclusion, we are able to establish that the 21-layered CNN model that was suggested is an improved method for diagnosing eye disorders based on OCT data. The accuracy of the model that was proposed is noticeably greater than that of the other methodologies that were considered. It is possible that ophthalmologists who use this technology will be able to carry out effective retinal image analysis. This will allow them to provide improved treatment at an earlier stage in the progression of the disease, thereby preventing blindness. Because of this, we intend to increase the effectiveness of our model in the coming years.

6.1 Future Work

When it comes to diagnosing and treating retinal problems, OCT is becoming more and more popular. Future advances in technology and software will make it easier to use. To increase axial resolution, commercial scanners may use ultra-broad spectral bandwidth light sources as they become more affordable. OCT clinical use is still unclear. More research is required to establish the spectrum of macular illnesses for which OCT is effective and to compare its usefulness to other imaging modalities. Clinical practice needs more experience and research. However, further information is required to decide whether OCT testing alone will be enough. Moreover, the suggested approach might be applied to datasets from other domains. The suggested model may be extended to bigger datasets in the future to further improve its performance. In a few years, we will have answers to these remaining questions and other issues.

Chapter 7

Bibliography

1. R, R, A., Bourne, S., R, Flaxman., T, Braithwaite., "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," Lancet Glob. Heal., vol. 5, no. 9, pp. e888–e897, Sep. 2017

2. K, Attebo., P, Mitchell., R, Cumming., and W, Smith., "Knowledge and beliefs about common eye diseases," Aust. N. Z. J. Ophthalmol., vol. 25, no. 3, pp. 283–287, Aug. 1997

3. A, Al-Mujaini., U, K, Wali., and S, Azeem., "Optical coherence tomography: clinical applications in medical practice," Oman Med. J., vol. 28, no. 2, pp. 86–91, Mar. 2013.

4. Vision impairment and blindness. (2021, October 14). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visualimpairment

5. Wu, L., MD. (2021, December 15). Choroidal Neovascularization (CNV): Practice Essentials, Background, Pathophysiology. Medscape.

6. Bansari, S. (2019). Introduction to how CNNs Work. [online] Medium. Available at: https://medium.datadriveninvestor.com/introduction-to-how-cnns-work-77e0e4cde99b [Accessed 26 Mar. 2021].

7. The Eye Diseases Prevalence Research Group: The prevalence of diabetic retinopathy among adults in the United States. Arch Ophthalmol 2004;122:552-563.

8. Cheng, Y. S., Lin, S. H., Hsiao, C. Y., & Chang, C. J. (2019). Detection of Choroidal Neovascularization by Optical Coherence Tomography Angiography with Assistance from Use of the Image Segmentation Method. Applied Sciences, 10(1), 137. https://doi.org/10.3390/app10010137

9. Y, Yoshida., T, Sato., S, Oosuka., M, Mimura., M, Fukumoto., T, Kobayashi., T, Kida., & T. Ikeda, "Two cases of diabetic macular edema complicated by an atypical macular hole," BMC Ophthalmology, vol. 20, no. 1, 2020.

10. Makino, S. (2014, June 6). A Case of Dry Age-Related Macular Degeneration with Reticular Pseudodrusen.

11. Acan, D., Calan, M., Er, D., Arkan, T., Kocak, N., Bayraktar, F., & Kaynak, S.(2018). The prevalence and systemic risk factors of diabetic macular edema: a crosssectional study from Turkey. BMC Ophthalmology, 18(1). https://doi.org/10.1186/s12886-018-0753-y

12. Holekamp NM. Overview of diabetic macular edema. Am J Manag Care. 2016 Jul;22(10Suppl):s284-s291. PMID: 27668630.

13. Macular Degeneration Overview — BrightFocus Foundation. (2020b). BrightFocus Foundation. https://www.brightfocus.org/macular/overview

14. Age-Related Macular Degeneration (AMD) Data and Statistics — National Eye Institute. (2019). National Eye Institute. https://www.nei.nih.gov/learn-about-eyehealth/outreach-campaigns-andresources/eye-health-data-and-statistics/age-relatedmacular-degeneration-amd-data-andstatistics

15. Pennington, K. L., & DeAngelis, M. M. (2016). Epidemiology of age-related macular degeneration (AMD): associations with cardiovascular disease phenotypes and lipid factors. Eye and Vision, 3(1). https://doi.org/10.1186/s40662-016-0063-5

16. Krizhevsky, A. (2012). ImageNet Classification with Deep Convolutional Neural Networks. NeurIPS Proceedings.

17. F. Hasan, "What are some deep details about pooling layers in CNN?," Edpresso, 2020.

18. Mishra, M. (2020). Convolutional Neural Networks, Explained. [online] Medium. Available at: https://towardsdatascience.com/convolutional-neural-networks-explained- 9cc5188c4939.

19. Jifara, W., Jiang, F., Rho, S., Cheng, M., Liu, S. (2017). Medical image denoising using convolutional neural network: a residual learning approach. The Journal of Supercomputing, 75(2), 704–718. https://doi.org/10.1007/s11227-017-2080-0

20. Jaybhay, J., Shastri, R. (2015). A Study of Speckle Noise Reduction Filters. Signal Image Processing : An International Journal, 6(3), 71–80. https://doi.org/10.5121/sipij.2015.6306

21. Pengfei Guo, Dawei Li, and Xingde Li, "Deep OCT image compression with convolutional neural networks," Biomed. Opt. Express 11, 3543-3554 (2020).

22. Thakur, A. (2020, August 19). How to handle images of different sizes in a convolutional neural network. Wamp;B. Retrieved May 24, 2022.

23. A. Gupta, "A Comprehensive Guide on Deep Learning Optimizers," Analytics Vidhya, 7 October 2021.

24. J. Brownlee, "Gentle Introduction to the Adam Optimization Algorithm for Deep Learning," Machine Learning Mastery, 3 July 2017.

25. Hao Zheng, Zhanlei Yang, Wenju Liu, Jizhong Liang, & Yanpeng Li. (2015). Improving deep neural networks using softplus units. 2015 International Joint Conference on Neural Networks (IJCNN).

26. R. Thakur, "Step by step VGG16 implementation in Keras for beginners," Towards Data Science , 2019.

27. Sec, D.I. (2021). VGG-19 Convolutional Neural Network. [online] All about Machine Learning. Available at: https://blog.techcraft.org/vgg-19-convolutional-neural-network/

28. Quora. (n.d.). What is the VGG-19 neural network? [online] Available at: https://www.quora.com/What-is-the-VGG-19-neural-network [Accessed 21 May 2022].

29. P. Ruiz, "Understanding and visualizing DenseNets," Towards Data Science, 2018.

30. G. Huang, Z. Liu, van, and Weinberger, Kilian Q, "Densely Connected Convolutional Networks," arXiv.org, 2016. arxiv.org/abs/1608.06993.

31. Elsharif, Abeer Abed ElKareem Fawzi & Abu-Naser, Samy S. (2022). Retina Diseases Diagnosis Using Deep Learning. _International Journal of Academic Engineering Research (IJAER)_ 6 (2):11-37.

32. Obaid, T. F., & Eneizan, B. M. (2016). Transfer of training and post-training on job performance in Middle Eastern countries. Review of Public Administration and Management, 400(3786), 1-11.

33. Awais, M., Muller, H., Tang, T. B., Meriaudeau, F. (2017, September). Classification of SD-OCT images using a Deep learning approach. 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA).

34. May, C. (2020, October 13). Performance Analysis. Sport Australia. https://www.clearinghouseforsport.gov.au/kb/performance-analysis

35. Z. Zhang, "Improved adam optimizer for deep neural networks," in 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), IEEE, 2018, pp. 1–2.

36. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

37. Martineau, M., Atkinson, P., McIntosh-Smith, S. (2019). Benchmarking the NVIDIA V100 GPU and Tensor Cores. In: , et al. Euro-Par 2018: Parallel Processing Workshops. Euro-Par 2018. Lecture Notes in Computer Science(), vol 11339. Springer, Cham.

38. Lee, C. S., Baughman, D. M., Lee, A. Y. (2017). Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images. Ophthalmology Retina, 1(4), 322–327.

39. Perdomo, O., Otalora, S., Gonzalez, F. A., Meriaudeau, F., Muller, H. (2018). OCT-NET: A convolutional network for automatic classification of normal and diabetic macular edema using sd-oct volumes. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018).

40. Rajagopalan, N., Narasimhan, V., Kunnavakkam Vinjimoor, S., & Aiyer, J. (2020). Deep CNN framework for retinal disease diagnosis using optical coherence tomography images. Journal of Ambient Intelligence and Humanized Computing, 12(7), 7569–7580. https://doi.org/10.1007/s12652-020-02460-7

41. Retinal OCT Images (optical coherence tomography). (2018, March 25). Kaggle. https://www.kaggle.com/paultimothymooney/kermany2018

42. Fujimoto, J. G., Brezinski, M. E., Tearney, G. J., Boppart, S. A., Bouma, B., Hee, M. R., Southern, J. F., Swanson, E. A. (1995). Optical biopsy and imaging using optical coherence tomography. Nature Medicine, 1(9), 970–972. https://doi.org/10.1038/nm0995-970

43. Jifara, W., Jiang, F., Rho, S., Cheng, M., Liu, S. (2017). Medical image denoising using convolutional neural network: a residual learning approach. The Journal of Supercomputing, 75(2), 704–718. https://doi.org/10.1007/s11227-017-2080-0

44. Cheng, Y. S., Lin, S. H., Hsiao, C. Y., & Chang, C. J. (2019). Detection of Choroidal Neovascularization by Optical Coherence Tomography Angiography with Assistance from Use of the Image Segmentation Method. Applied Sciences, 10(1), 137. https://doi.org/10.3390/app10010137

45. Ñanculef, R., Radeva, P., Balocco, S. (2020). Training Convolutional Nets to Detect Calcified Plaque in IVUS Sequences. Intravascular Ultrasound, 141–158. https://doi.org/10.1016/b978-0-12-818833-0.00009-6

46. Hasan, F. (2022, May 24). What are some deep details about pooling layers in CNN? Educative: Interactive Courses for Software Developers.