A COMPREHENSIVE STUDY OF DCNN ALGORITHMS BASED TRANSFER LEARNING FOR HUMAN EYE CATARACT DETECTION

BY

SUSMOY PAUL

ID: 191-15-12910

AND

MD. OMAR JILANI JIDAN

ID: 191-15-12905

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

MS. SHARUN AKTER KHUSHBU
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

MS. SHARMIN AKTER
Senior Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 29th 2023

APPROVAL

This research project, titled "**A Comprehensive Study 0f DCNN Algorithms Based Transfer Learning for Human Eye Cataract Detection.**", submitted by **Susmoy Paul, ID: 191-15-12910** and **Omar Jilani Jidan, ID: 191-15-12905** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfilment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on January 29, 2023.

BOARD OF EXAMINERS

Chairman

Dr. Touhid Bhuiyan
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

29-1-23

Arif Mahmud
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

29.01.23

Saiful Islam
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

29/01/23

Dr. Shamim H Ripon
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that we completed this project under the supervision of **Ms. Sharun Akter Khushbu, Lecturer, and Department of the CSE**, Daffodil International University Department of CSE. We also certify that no component of this project, or any part of it, has been submitted to any other institution for the award of a degree or diploma.

Supervised by:

31.1.23

Ms. Sharun Akter Khushbu
Lecturer, Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Co-Supervised by:

Ms. Sharmin Akter
Senior Lecturer, Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Submitted by:

Susmoy.

Susmoy Paul
ID: -191-15-12910
Department of CSE
Daffodil International University

Md. Omar Jilani Jidan
ID: -191-15-12905
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First and foremost, we are grateful to God for granting us the ability to complete our final year project/internship successfully.

We extend our sincere thanks with deep gratitude to supervisor **Mst. Sharun Akter Khushbu,** Lecturer, Department of CSE, Daffodil International University, Dhaka and co-supervisor **Ms. Sharmin Akter**, Lecturer, Department of CSE, Daffodil International University, Dhaka for their guidance and expertise in the field of Deep Learning. Our supervisor, with his extensive knowledge and dedication to the subject, provided constant encouragement, supervision, and constructive criticism throughout the project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Touhid Bhuiyan**, Professor and Head, Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Cataract, a common eye disorder characterized by clouding of the lens, is a leading cause of vision loss. Each year numerous numbers of people are falling sufferer to visionary loss globally. The majority of the time, this problem arises as people age. This problem also can occur in young age people because of injury or certain clinical situations. It's far known as cataract while a dense and cloudy layer create on the eye lens and consequences the clear vision which can cause a problem like blurry eyesight, diminished vision and prescient. Additionally, they face difficulty seeing in robust light and gradually it could be the reason for full blindness. An excellent way to manipulate the hazard and avoid blindness is to stumble on cataracts well-timed and correctly before it become more complicated. In this study, we propose a cataract detection system using deep learning and image processing techniques. Our system aims to automatically analyze ocular images and predict the presence of cataracts with high accuracy. We are trying to pick out an efficient and accurate manner of detecting cataracts primarily based on a Deep Convolutional Neural network (DCNN) with the publicly accessed dataset. We used the transfer learning methods with DCNN models which are VGG19, NASnet, Resnet50 and MobileNetV2 achieving the highest accuracy across 2000 image sets. Also, MobilNnetV2 achieved accuracy rates of 97.75% on the test images. Compared to other models, the final result indicates that MobileNetV2 takes the least time to recognize images and classify them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

The human eye is the only visual organ that plays an imperative role in the proper functioning of the whole organism and helps to identify objects in all its actions. However, many types of eye diseases can cause blurred vision and complete blindness. A cataract is an eye disease in which a thick, opaque layer form on the lens of the eye. This disease mainly occurs in old age. As we age, the lens of the eye becomes less flexible, blurry, and thicker. Aging and other diseases cause the proteins and fibers in the lens to break down and stick together, making a cloudy layer over the lens. Cloudiness increases as the cataract progress. Cataracts scatter and block light as it passes through the lens, preventing sharp images from reaching the retina and blurring vision. Cataracts can also occur in some unusual cases, such as fetal trauma and incomplete development of the cornea. People with this disease have to deal with many vision problems, such as difficulty seeing, reading, diving, and recognizing others [1].

The World Health Organization (WHO) reports that by 2025, at least 2.2 billion people around the world will be blind or visually impaired, with approximately 1 billion of these cases being preventable. Cataracts are a major contributor to blindness, responsible for over 50% of cases, and it is estimated that 40 million people will lose their sight due to this condition [2]. In Bangladesh, more than 750,000 people over the age of 30 are facing eye issue, 80% of whom are due to cataracts. The official figures say, about 120,000 new cataract patients are added each year. Residents of rural and remote areas often suffer disproportionately from blindness due to a lack of access to medical services and ophthalmologists [5]. Early detection of cataracts can reduce cataract complaints and prevent complete blindness. It can save the eye lens from completely block. Only minor surgery can help the patients to gain his full vision. However, diagnosing cataracts with an ophthalmoscope is expensive and requires an ophthalmologist. Therefore, it is necessary to develop an intelligent cataract diagnosis system that can reduce the dependence on ophthalmologists or skilled one for eye care. In addition, by reducing the need for manual inspection by healthcare professionals, we hope to lower the cost of eye checkups and make them more accessible to a wider population and reducing the burden of eye diseases on society.

That is why our objective of making a cataract detection is to accurately classify fundus images as containing cataracts or not without an ophthalmologists or skilled person. By automating the process of detecting cataracts in fundus images, we hope to improve the efficiency and accuracy of cataract detection, which can ultimately lead to better patient care. The system should identify the cataract from fundus image and decide the judgment using the pretrained models. Furthermore, we hope to contribute to the advancement of medical science and the development of new technologies for improving healthcare. Our aim is to use this project as a starting point for further research and development in this area, with the ultimate goal of improving patient outcomes and reducing the burden of eye diseases on society.

To achieve this, we have compared the performance of four different deep learning models (VGG19, NASnet, Resnet50, and MobileNetV2) for cataract detection. We used a diverse dataset of fundus images and preprocessing techniques to train the models more effectively. We have also implemented various approaches and hyperparameters to increase the accuracy of classifying images as normal or cataract-affected. The results of all four models (VGG19, Resnet50, NASnet and MobileNetV2) were evaluated, and the best model was selected for further classification of normal and cataract images. Our experimentation and analysis on the Fundus dataset demonstrate the effectiveness of the proposed model and its potential to benefit underserved populations and advance the field of medical image analysis.

An introduction has been provided in Chapter 1. In this paper there are four more sections, organized as follows. In Chapter 2 Related work and literature review and the challenges of the work. Chapter 3 presents the materials and methodology of the research. Chapter 4 describes the Result and Discussion and the Conclusion is given in the Chapter 5. At the end of the paper references has been given.

CHAPTER 2

Background

2.1 Literature Review

The study and development of automated methods for cataract detection and grading have been going on for the past few years. There are various studies conducted to analyze the fundus image for the classification and detection of cataracts. Various authors have proposed a range of models and have achieved varying levels of accuracy. However, the limitations in feature extraction and preprocessing techniques have resulted in suboptimal performance for these approaches. In order to optimize the precision of our deep-learning model for cataract detection, we conducted a literature review of relevant publications and journals. We evaluated a variety of existing approaches, including different datasets, preprocessing techniques, feature extraction methods, feature selection criteria, classifiers, and models. Our analysis revealed that image processing methods show promise for detecting cataract using fundus images.

From the study, we discovered that a research team had employed two distinct techniques to analyze and classify fundus images on a dataset of 388 images. The first approach was the Novel Angular Binary Pattern (NABP), while the other method involved the use of Kernel-Based Convolutional Neural Networks. They were able to gain 97.39% accuracy in their proposed method [17]. Li *et al*. proposed system of cataract detection utilizes decision tree algorithm with a two-dimensional gaussian filter [6]. By training the model on 1355 fundus images, they were able to achieve an accuracy of 92.8%. Budai *et al.* have introduced another machine learning-based system with an SVM Classifier [7]. In this approach, image is divided into 17 segments which are subsequently fed into a support vector machine (SVM) system for training and able to acquire an accuracy of 87.82%.

Zhang *et al.* provide a multi-model ensemble method that uses ultrasound images for automatic cataract detection [9]. According to our review of relevant literature, this system achieved the highest accuracy among deep learning approaches, with a rate of 97.5%.

While these models show promise, they do have some limitations. In particular, they rely on the degree of blurriness in retinal images for cataract detection and classification. This blurriness could potentially be caused by other ocular conditions such as corneal edema or diabetes mellitus, which may affect the accuracy of the model. Ran *et al.* proposed a cataract detection method using a DCNN algorithms and a random forest classifier for cataract detection [18]. The model was trained using 3460 retinal fundus images. The model also tested using 1948 images and able to achieve an accuracy of 97%.

Though these methods are achieving high accuracy there are some limitations in these models. They require prior training and are not equipped to handle high-dimensional features. Additionally, they may struggle with issues such as the vanishing gradient problem, overfitting, and underfitting.

Zhoul *et al.* developed a model that employed a combination of discrete state transition and Resnet, which allowed them to address the vanishing gradient problem by implanting residual technique [8]. Their model also eliminated the need for preprocessing of the dataset and was able to process high-dimensional features. Its architecture was able to identify deep semantic features from fundus images, resulting in an accuracy of 94%. Khan et al. achieved a similar level of accuracy using a transfer learning approach based on a CNN with VGG19 [10]. Their model achieved 97.47% of accuracy, with 97.47% precision and 5.27% of loss on a dataset of fundus images available on Kaggle [3].

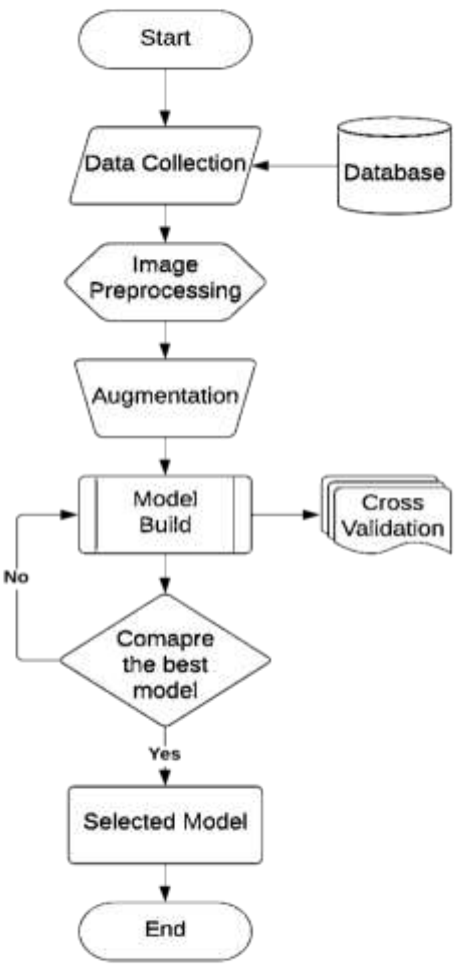
However, traditional methods may have limitations such as high-dimensional features and a tendency towards overfitting or underfitting. After observing this CNN-based recent work we find that there are still few works based on DCNN for cataract detection. Also, there are several challenges to deal with and the scope of improving the accuracy of the model. We can minimize the complexity by reducing the training paraments and overall model accuracy can be improved. Two state-of-the-art model architectures, MobileNetV2 and NASnet has been employed. These models are known for their high accuracy and efficiency, and were trained using a custom dataset of publicly available data. Our results show that while VGG19 and Resnet50 models have been widely used in previous studies, the combination of MobileNetV2 and NASnet can lead to even greater improvements in accuracy and efficiency. Our goal is to optimize the model by addressing the challenges in this field to improve overall performance.

CHAPTER 3

Research Methodology

This section outlines the tools used, the data collection process, data analysis, and the proposed model. Our proposed work utilizes a deep learning classification method, specifically a multi-layer Convolutional Neural Network (CNN) for classification. This model is able to distinguish Cataracts from fundus images. The overall architecture of the proposed method of this research is given in Figure 1. The cataract classification framework consists of several steps, including image acquisition, preprocessing, implementation of proposed models, and performance evaluation. These steps are described in more detail in the following sections.

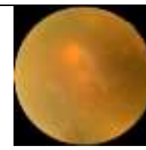
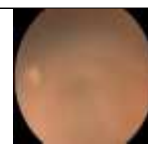
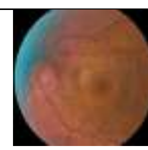
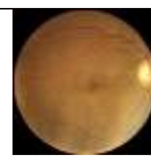
Figure 1: Workflow Diagram.



3.1. Data Collection

For this study, a collection of 6621 fundus images were collected from various sources, including Kaggle [3-4]. Besides that, we collected some high-resolution fundus images from HRF Image Database. The categories, sources of data, and samples are given in Table 1.

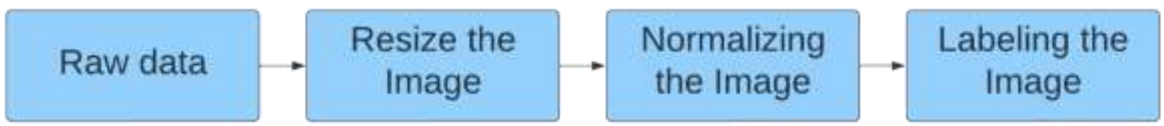
Table 1: Data Category, Sources, and Samples.

Category	Source	Sample
Cataract	Ocular Disease Detection, Kaggle	
Cataract	Ocular Disease Detection, Kaggle	
Cataract	Eye Disease_ Classification, Kaggle	
Cataract	Eye Disease_ Classification, Kaggle	
Normal	Eye Disease_ Classification, Kaggle	
Normal	Ocular Disease Detection, Kaggle	
Normal	Ocular Disease Detection, Kaggle	

3.2. Data Pre-processing

After the Collection process of the data from the datasets of the fundus image, we merge them and make them one for use. The merged dataset includes images of various diseases and abnormalities, such as diabetes retinopathy, cataract, hypertension, pathological myopia, glaucoma, age-related macular degeneration, and others, as well as normal photographs. As a result, we applied basic preprocessing steps to make the model ability better to evaluate the dataset. In the first step, we eliminate all fundus images class except cataract and normal. These fundus images had varying image sizes.

Figure 2: Image Preprocessing



For that reason, OpenCV is used to resize the images and used normalization by subtracting the mean of all pixels. Then we label each image with cataract and normal labels. After the labeling process, the dataset is converted into an array using the NumPy library for the further training process.

After preprocess and creating labeling, datasets divided into two class for test and train dataset. Train dataset has 80% and Test dataset has 20% data of the preprocessed dataset. Table 2 is describing the number of images in training dataset and testing dataset.

Table 2: Distribution of Dataset.

Dataset	Training Images	Testing Images	Total Images
Cataract	800	200	1000
Non-Cataract	800	200	1000

3.3 Data Augmentation

Augmentation procedures are used to create a range of modified images through techniques such as rotation, shifting, zooming, and cropping, with the goal of enhancing the model's ability to generalize to new images. “ImageDataGenerator” function of the Keras [23] framework is used for augmentation process. The arguments which are used with the “ImageDataGenerator” function are described in the Table 3. To increase the size of the training, testing, and validation sets, the augmentation process was utilized to generate a vast number of additional images. The techniques outlined in Table 3 were applied to the training dataset to improve the model's ability of learning with a larger amount of data.

Table 3: Functions Arguments Values.

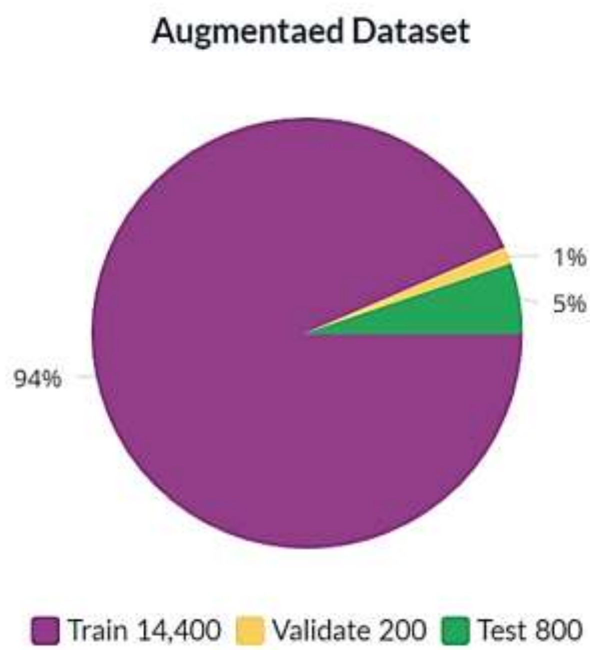
Augmentation Name	Values
Rescale	1/.255
Rotation Range	20
Width shift	0.2
Height shift	0.2
Zoom range	0.2
Horizontal Flip	True
Vertical Flip	True
Fill mode	Nearest

However, only validation and testing data will be employed with the first argumentation strategy from Table 3. Validation data are taken from the test images with the ratio of 0.5%. After the augmentation process data points are increased to 15400 which is used to train, test and validate the model when it was compiling. The total amount of data that will generate after augmentation is indicated in Table 4. Figure 3 describe the describe the segments of data points.

Table 4: Augmented Data

Dataset folder	Arguments	Generated data	Total Data
Train	8	12800	14400
Validation	1	200	200
Test	1	400	800

Figure 3: Total Data after Augmentation



3.4 Experimental DCNN Models

To identify the cataract form fundus eye image collected form ocular dataset we have applied different model architecture. We applied VGG19, NASnet and Resnet50 for attain better accuracy in the detection of cataract. A details analysis is discussed below for understanding the method architecture properly.

3.4.1 VGG19

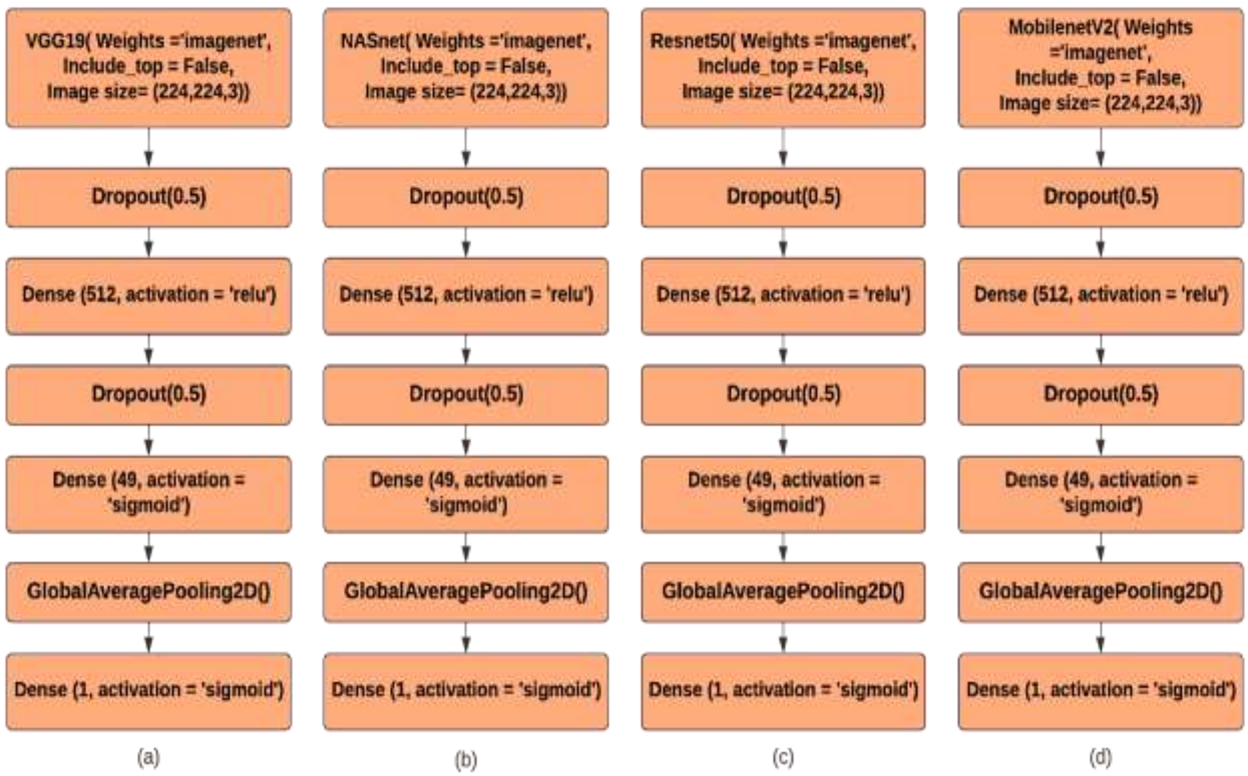
VGG19 is one of the advanced pre-trained convolutional neural network models that are used for image classifications and define an image in terms of shape, color, and structure. Visual Geometry Group developed the VGG19 at the oxford university [19]. One of the notable features of the VGG19 model is its depth. As the name suggests VGG19 has a total of 19 layers. In addition to its depth and use of small convolutional filters, the VGG19 model also includes max pooling layers, which down-sample the feature maps produced by the convolutional layers and help the model to focus on the most important features using a small convolutional filter with a size of 3x3 [19]. To further improve the model's generalization ability, we included a dropout layer with a rate of 0.5. This helps to prevent overfitting by randomly setting a portion of the inputs to zero during training. We also added a dense layer with 512 neurons and a ReLU activation function were implemented. Again, we the dropout with 49 neurons dense layer and sigmoid activation. A global average pooling layer reduces the spatial dimensions of the input by taking the average of the values in each channel. Finally, add a final dense layer to the model, with 1 unit and a sigmoid activation function (Fig: a). This layer is used to make the final prediction.

The final layer produces a probability score of 96.25% for the input image, indicating the likelihood that it belongs to a particular class.

3.4.2 NASnet

Neural Architecture Search or NAS is a type of convolutional neural network. It is the process of automating the network topology design of neural to achieve the best result. It was developed by the Google brain team to design the neural network architecture using fewer resources and minimal human interactions. Here we have used a pre-trained NASnet model with ImageNet weights and some additional layers. We randomly made half of the neurons through the activation using Dropout (0.5). After that, a fully connected layer architecture of 512 neurons using ReLU activation has introduced with the output and again drop half of the neurons. Then the output is passed through the fully connected layer of 49 neurons using sigmoid activation. Finlay a 2D global max pooling layer and downs ample the output according to the maximum height and width dimensions of the input layers. Then a dense layer extracts all combinational features from the previous layer and makes the prediction (Fig: b). The NASnet has given 96% of accuracy for detecting the cataract from the test image and able to classify the cataract and normal images.

Figure 4: Block diagram of VGG19, NASnet, Resnet50 and MobilenetV2



3.4.3 ResNet50

ResNet50 is a convolutional neural network (CNN) designed by Microsoft Research that has become a popular choice for image classification tasks. It is a deep neural network with 50 layers that utilizes a residual architecture. This model has achieved excellent results on various image classification benchmarks and is widely utilized in both industry and academia. It is called a "residual" network because of its residual functions regarding the input, rather than learning the desired functions directly. This makes it easier for the network to learn complex, highly non-linear functions and allows it to achieve much higher accuracy than would be possible with a traditional feed forward network. ResNet50 is trained on the ImageNet dataset, which consists of 1000 different classes containing 1.4 million labeled images. The model is capable of achieving very high accuracy on this dataset, making it a useful choice for many image classification tasks. We have used pre-trained Resnet50 with some more additional layers of dropout with 0.5, a dense layer of 512 neurons, Global average pooling layers, and a single node dense layer (Fig: c). The model has given us an accuracy of 77% for detecting the cataract from the test images.

3.4.4 MobileNetV2

MobileNetV2 is a convolutional neural network architecture designed for efficient on-device image classification. It was developed by Google and introduced in their 2018 publication "MobileNetV2: Inverted Residuals and Linear Bottlenecks". MobileNetV2 is an updated version of MobileNetV1, a popular lightweight model for mobile and embedded vision applications. MobileNetV2 improves upon the state-of-the-art performance of MobileNetV1 by using a combination of linear bottlenecks, inverted residual blocks, and short connection paths to reduce the number of computations required by the network. This makes MobileNetV2 faster and more efficient than its predecessor, while still maintaining a similar level of accuracy. One of the key features of MobileNetV2 is its ability to perform well on a variety of mobile and embedded devices, making it a useful choice for a wide range of applications. It is designed to be easily integrated into existing neural network architectures, allowing developers to build larger, more complex models using MobileNetV2 as a building block. We used the ImageNet weights for the model. Additionally, we have added some more layer with the pretrained MobilenetV2 making the top layer false (Fig: d). Using this proposed architecture for MobilenetV2 we have manage to achieve 97.75% accuracy on the test dataset for cataract detection.

3.5 Overview of the Models

For deep learning, most commonly CNN is used for analyzing or classifying images. In this paper we are working with the four deep CNN based algorithms and analyzing the best algorithm for the detection of cataracts, the comparison. Table 5 describe the performance score of the models.

Table 5: Performance score of VGG19, NASnet, Resnet50, MobileNetV2

Model	Test Accuracy	Test Loss	Specificity	Recall	F1 score	Precision
VGG19	0.9625	0.1169	0.96	0.96	0.96	0.96
NASnet	0.96	0.077	0.96	0.96	0.96	0.96
Resnet	0.755	0.560	0.76	0.76	0.75	0.76
MobileNetV2	0.9775	0.055	0.98	0.98	0.98	0.98

Table 6 provides the parameters that is compile with the models. The learning rate, optimizer, batch size, number of epochs, loss function, and environment data used are listed in Table 6. We used the Adam optimizer and binary cross entropy loss to train the model and calculate the loss between true and predicted labels. Each model has run through 15 epochs and 49 batch size. We trained and run the models on the google Colab with a GPU.

Table 6: Parameters in models compiling.

Parameters	Values
Epochs	15
Batch size	49
Learning rate	0.001
Decay	1*e*-3/epoch
Optimizer	Adam
Loss	Binary cross-entropy
Shuffling	Each epoch
Executable Environment	Colab

3.6 Activation Functions

Sigmoid function: The sigmoid function [11], also known as the logistic function, is a mathematical tool commonly used in machine learning for numerical measurement. It has the ability to convert any input value to a range between 0 and 1. The equation for the sigmoid function is represented as:

$$X = \frac{1}{1+e^{-y}} \qquad (1)$$

One of the main benefits of using this function is that its output is always between 0 and 1, making it useful for models that predict probability as an outcome.

ReLU: The rectified linear activation unit (ReLU) is a commonly used activation function in deep learning, particularly in the field of computer vision such as image classification, object detection, and more. Unlike the sigmoid function, which maps any input value to a range between 0 and 1, ReLU maps any input value greater than 0 to the same value, and any input value less than or equal to 0 to 0. This makes it computationally efficient and helps to alleviate the vanishing gradient problem, which can occur with sigmoid or other activation functions. The equation for ReLU is as follows:

$$f\,(input) = max\,(0, input) \qquad (2)$$

3.7 Evaluation Metrics

In this section, we will analyze the performance of our models using a performance matrix. This will give us a clear understanding of how well the models are able to predict the correct labels and identify areas for improvement. The evaluation metric will provide insights into the effectiveness of the model and allow us to adjust as needed.

Accuracy: Accuracy is one of the evaluations matrixes. It is a fraction of predictions, defined as the number of correctly predicts classifications divided by the total number of predictions made. It assessing the performance of a model using the equation

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (3)$$

Sensitivity: In order to evaluate the effectiveness of the models in detecting positive instances, sensitivity is commonly used as a metric. This metric, also referred to as the True Positive Rate (TPR) or Recall, quantifies the proportion of positive instances that are correctly identified by the model. [21]. Sensitivity or true positive rate can be calculated using following equation

$$Sensitivity/Recall = \frac{TP}{TP + FN} \tag{4}$$

Specificity: In assessing the performance of our model, we also consider the metric of specificity. This metric calculates the proportion of actual negatives that are correctly identified by the model. It can also be referred to as the True Negative Rate (TNR) and is a crucial measure for evaluating the model's ability to correctly identify non-cataract cases [21]. Specificity or true negative rate can be using following equation

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

Precision: Precision is one of the evaluations matrixes that calculate a machine learning model's performance by finding the ratio between the True Positives and all Positives. Precision is proportional to accurate positive predictions that are made, or the accuracy of minority class predictions.it is calculated as follows

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

F1 score: F1 score is an alternative evaluation metric that measures a model's accuracy. It is a combination of precision and recall, and is particularly useful when the classes of the dataset are balanced and have a similar number of data points. The F1 score is calculated by taking the harmonic mean of the precision and recall scores, as shown in the equation

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

CHAPTER 4

Result & Discussion

This section will briefly analyze and interpretation of the results that are obtained from performing four different models. We used “Google Colab” as our experimental environment. Below figures shows the results of the model's performance evaluation for each epoch using the Plot graph. These plots provide a visual representation of how well the model is performing during training and validation. Blue and Orange line respectively indicates the training loss and training accuracy. On the other hand, Green and Red lines are indicating the validation loss and accuracy respectively. The figure 5 -8 are the plot graph that describe the performance of the models for training and validation.

The Figures 5 and 6 that are provided below, helps to visualize the progress of the NASnet and VGG19 models throughout the training process. Specifically, Figure 5 displays the accuracy and loss during training and validation for NASnet, while Figure 6 presents the same information for the VGG19 model. These figures allow us to assess the performance of the models during training and evaluate how well they are able to generalize to new data.

Figure 5: Performance Evaluation of NASnet

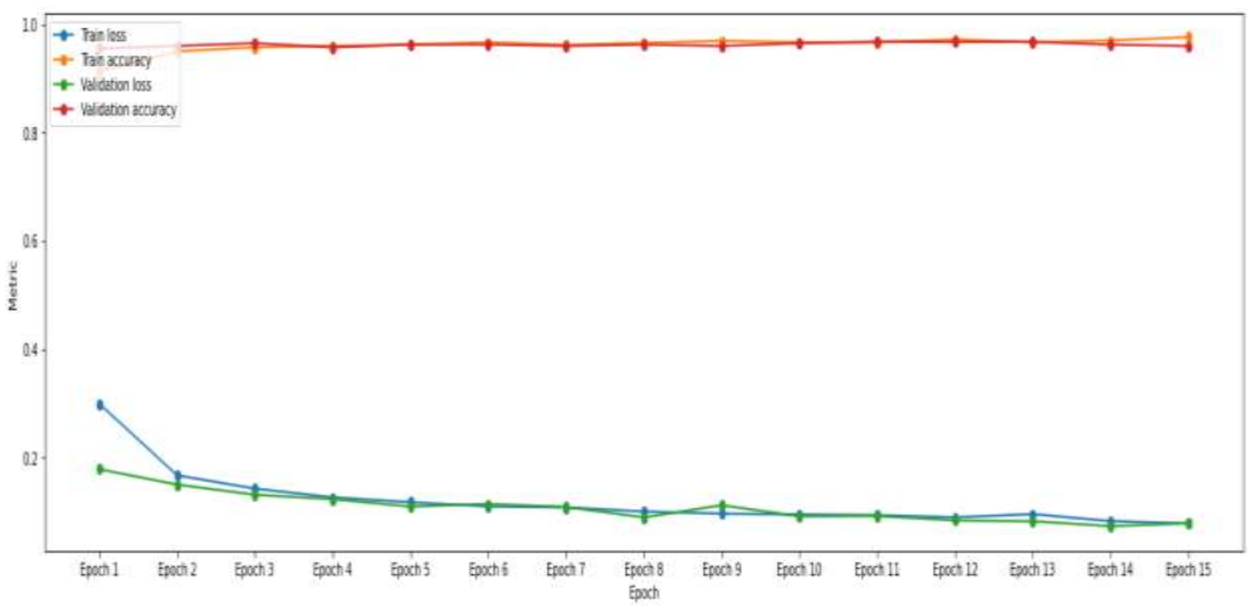
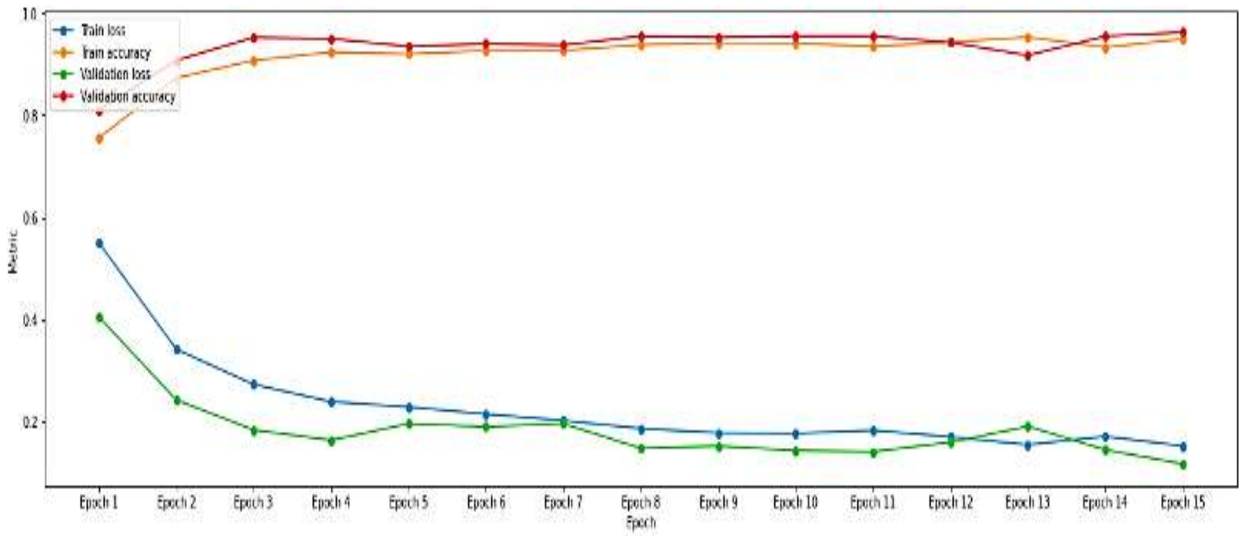
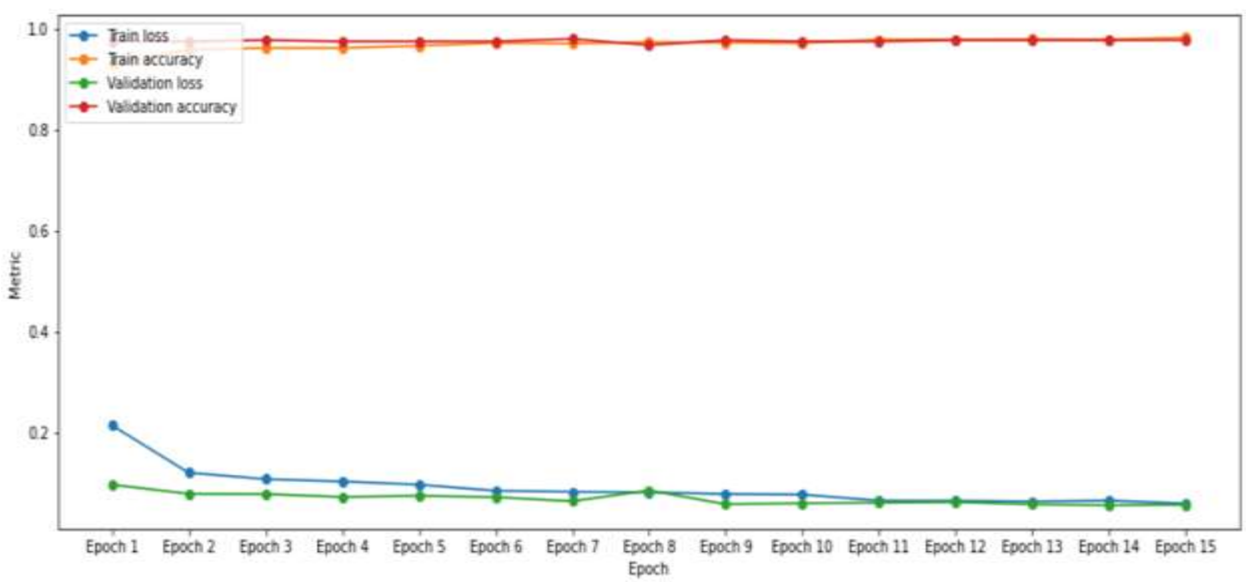


Figure 6: Performance Evaluation VGG19



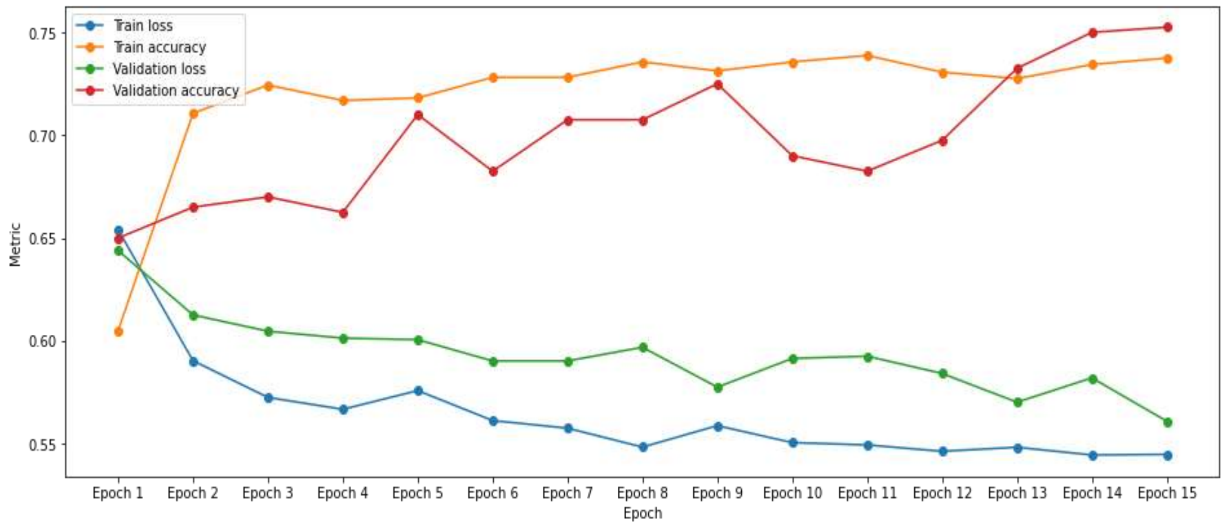
For the MobileNetV2 the maximum and minimum values of validation loss is respectively 0.05 and 0.09 which can be seen in the Figure 7. From the result we find that, the lowest and highest values of validation accuracy is 0.9675 and 0.98, respectively. And for the training accuracy, respectively 0.9294 and 0.9825. Also, overfitting or underfitting is not occurred in our model.

Figure 7: Performance Evaluation MobileNetV2



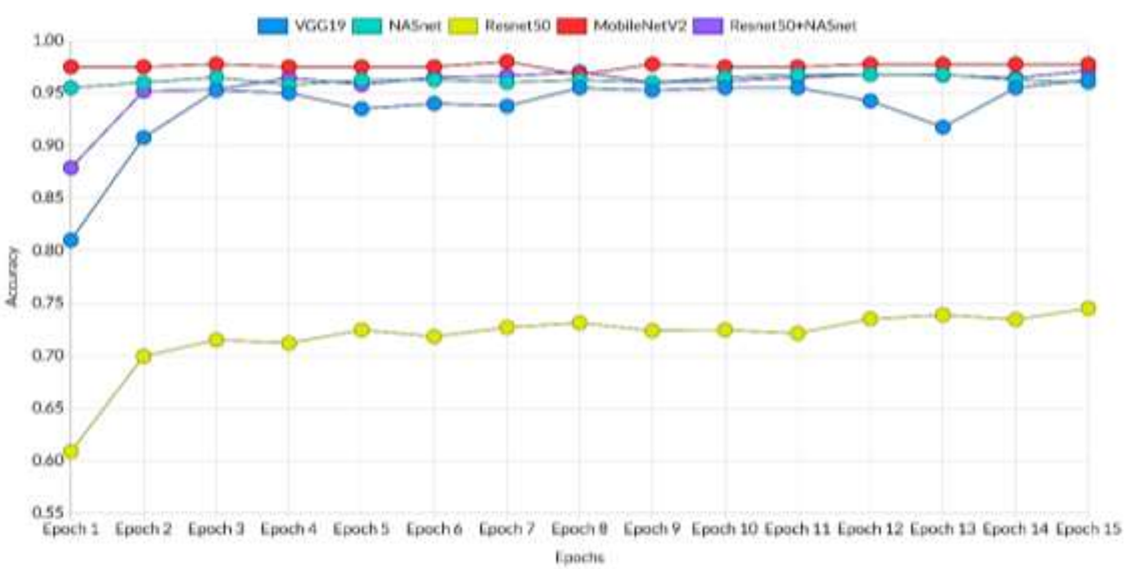
Resnet 50 has perform unparallel in the training and validation but its mange to achieve 74% accuracy in the training and 75% in validation accuracy. Result can be seen below in Figure 8.

Figure 8: Performance Evaluation of ResNet50



In Figure 5 we find that Resnet50 performs the lowest accuracy score than the other three models. The VGG19 also lag behind from the NASnet and MobileNetV2. In the beginning VGG19 was achieving accuracy around 81% but it rapidly grows and gave the highest accuracy of 96.25% for the test images. Though VGG19 and NASnet has achieve a good accuracy for test images, MobilenetV2 has outperform them all by achieving the accuracy of 97.75%.

Figure 9: Comparison of the model's Accuracy



In this section the below tables show the accuracy and loss of a models on a classification task, with each row representing the results for a different iteration or epoch of training. The accuracy is the percentage of correct predictions made by the model, while the loss is a measure of how well the model is able to learn the correct output for a given set of inputs.

When the accuracy increases as the loss decreases, indicating that the model is improving over time. This is a common trend in machine learning, as models generally become more accurate as they learn from the training data and reduce the error in their predictions.

Table 8 represent the graphs for the VGG19 model accuracy and loss.

Table 7: VGG19 model's Performance Graph.

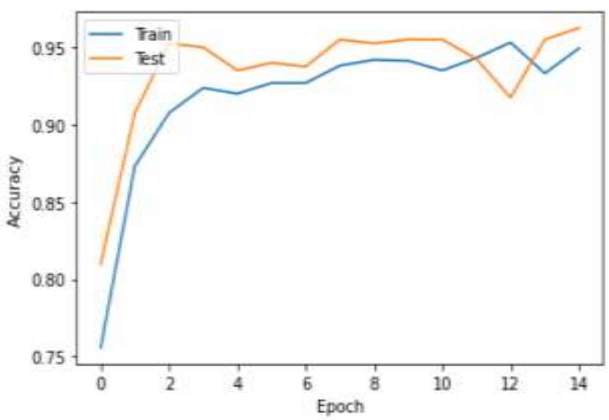
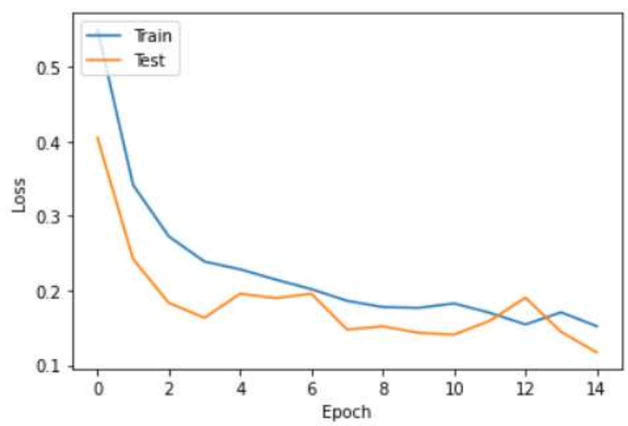
Accuracy	Loss
	

Table 8: NASnet model's Performance Graph.

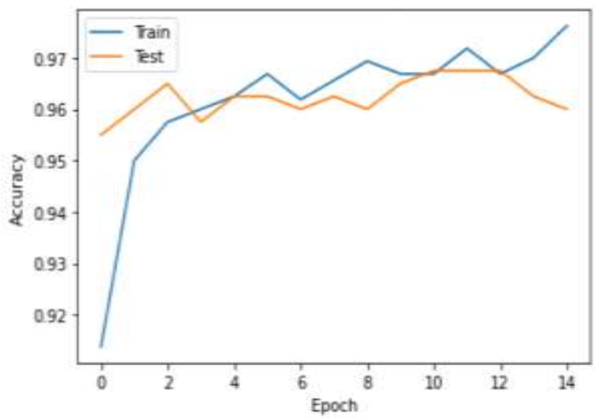
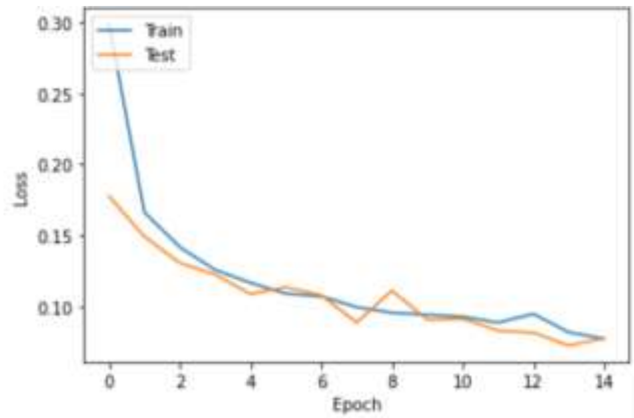
Accuracy	Loss
	

Table 9 and 10 are representing the accuracy and loss for both train and test data corresponding to each epoch. NASnet has able to achieve an accuracy of 96% where and loss is below the 0.1. Whereas Resnet50 is hardly able to get the accuracy of 77% and loss is high compared to the other models.

Table 9: ResNet50 model's Performance Graph.

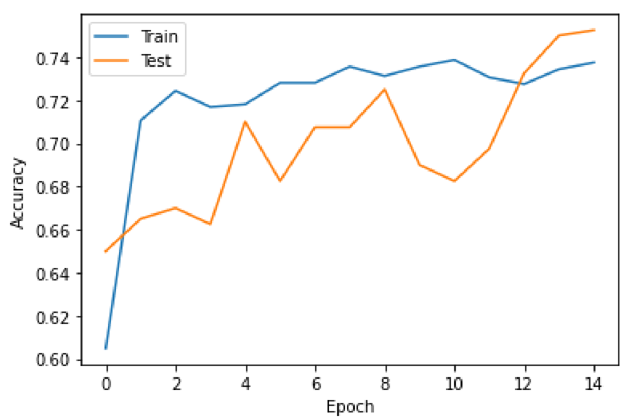
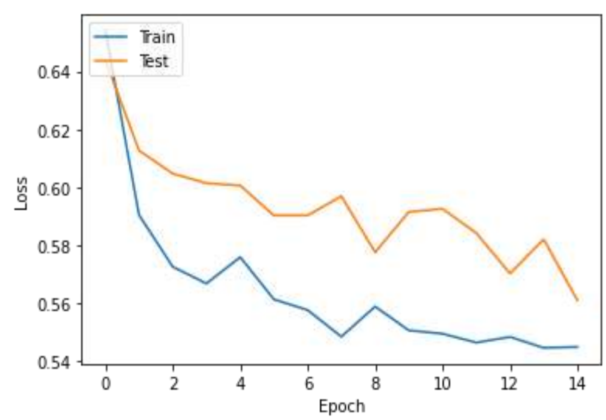
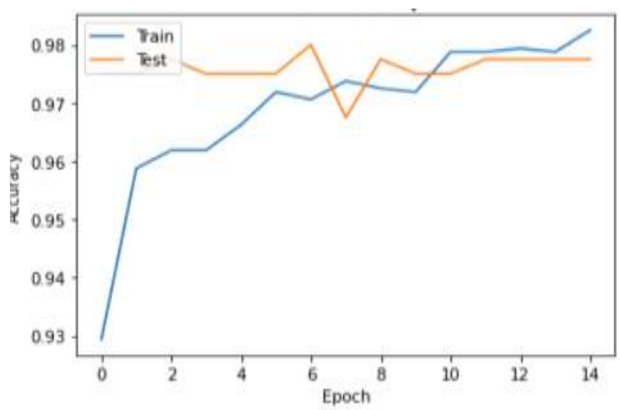
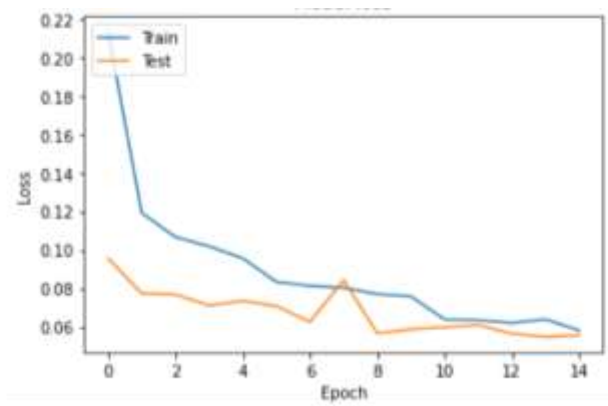
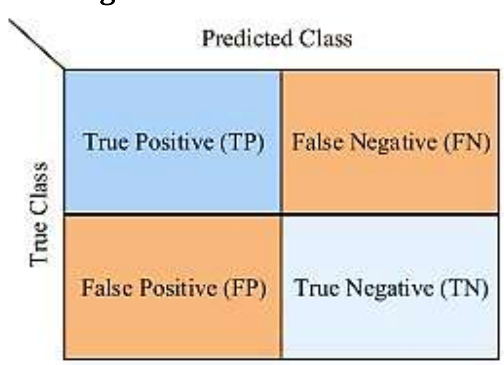
Accuracy	Loss
	

Table 10: MobileNetV2 model's Performance Graph.

Accuracy	Loss
	

Confusion matrix is a tool, widely used to assess the accuracy of a classification models in machine learning. The confusion matrix (figure 10) is a visualization that compares the predicted labeled data with the true labeled data. It provides a concise summary of the model's performance. The rows of the matrix correspond to the true labels of the data, and the predicted labels are represented by the columns. For example, the top left element in the matrix represents the instances number that were correctly classified as positive by the model. The bottom right element represents the number of instances which were correctly classified as negative. The other elements in the matrix represent incorrect predictions. By analyzing, it is possible to gain insights into the areas where the model is performing well, as well as areas for improvement. The cells of the table contain the count or frequency of each combination of true and predicted labels.

Figure 10: Confusion Matrix



In this context, "Positive" and "Negative" refer to the two classes that the model is trying to predict. A "True Positive" is an instance that was correctly classified as positive by the model, while a "False Negative" is an instance that was incorrectly classified as negative. Similarly, an incorrectly classified instance that was as positive is "False Positive". On the other hand, a correctly classified negative instance is "True Negative" [25].

The values in confusion matrix can be used for evaluate model's performance through various metrics such as precision, F1 score, and recall. The ratio of true positives among all positive instance made by the model measures by precision, while recall measures the ratio of true positives among all actual positive instances. The F1 score is the mean of recall and precision, with a higher number indicating better performance. These metrics can provide insight into the strengths and weaknesses of the classification model and help to identify areas for improvement. The below figures are presenting the confusion matrix score for the evaluate models.

Figure 11: Confusion matrix of VGG19

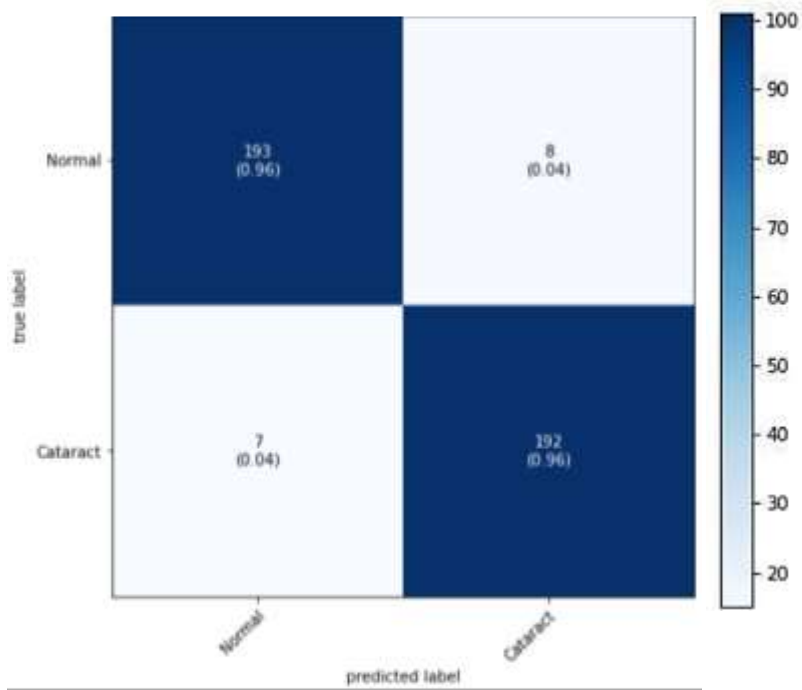


Figure 12: Confusion matrix of NASnet

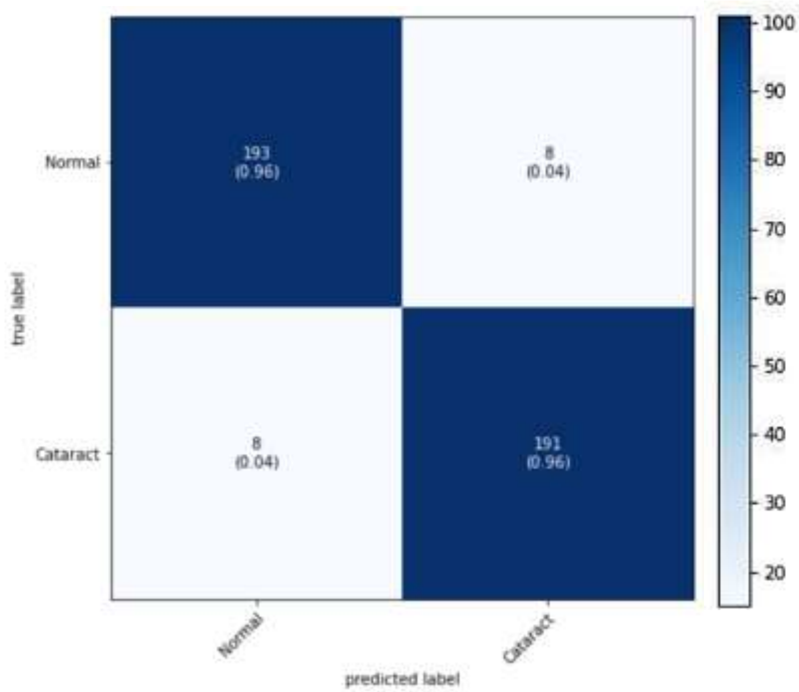


Figure 13: Confusion matrix of ResNet50

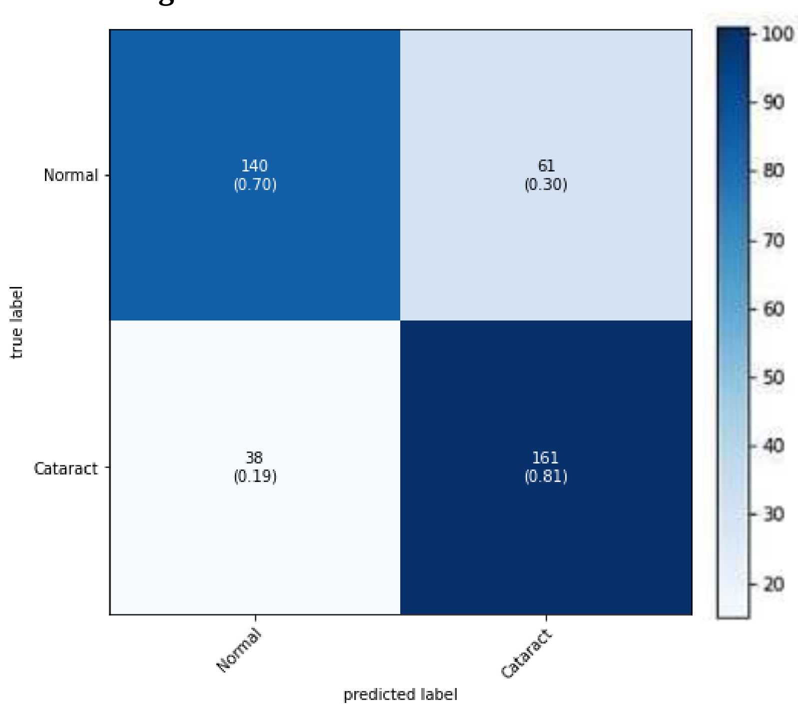
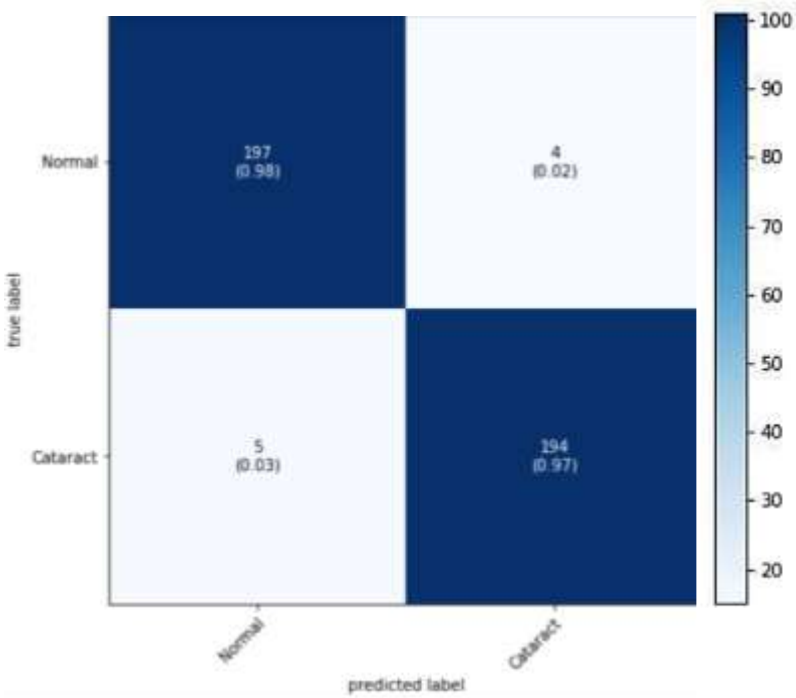


Figure 14: Confusion matrix of MobileNetV2



In this research, we compare several DCNN-based algorithms for the detection of cataracts and non-cataract eyes using retinal fundus images. We used publicly available fundus image data from Kaggle [3-4] to create a custom dataset of 6621 fundus images in three classes, which were preprocessed and enhanced. We trained the VGG19, Resnet50, and MobileNetv2 models using 1000 cataract and 1000 normal images (totaling 2000 images), reserving the remaining images for testing. Among the models VGG19, NASnet and MobileNetV2 performed really well, achieving an accuracy of more than 95%. Both VGG19 and MobileNetV2 had similar performance.

However, a detailed analysis of the model's statistical performance revealed that MobileNetv2 had performs slightly better with the highest accuracy, at 97.75%.

CHAPTER 5

Impact on Society, Environment and Sustainability

Cataract detection using deep learning can have a significant impact on society and the environment by improving access to early and accurate diagnosis and treatment for individuals suffering from cataracts. This can lead to improved quality of life for those affected by cataracts, as well as a reduction in the economic burden caused by the condition.

With the help of DCNN, cataract detection can be done more efficiently and accurately, which can lead to earlier diagnosis and treatment of the disease. This can help to prevent blindness and improve the quality of life for people affected by cataracts. Also, it helps to reduce healthcare costs by automating the process of cataract detection and enabling earlier treatment. This method can reduce the need for manual examination, which can be resource-intensive. With the help of this method, remote and low-resource areas, where access to healthcare is limited, can be covered easily. Increased efficiency in the detection process can also improve patient outcomes by providing more accurate and timely diagnoses.

In terms of sustainable development goals, Automatic cataract detection can contribute to several goals, such as,

- **Goal 3 - Good Health and Well-being:** By improving the speed and accuracy of cataract detection, machine learning can help to prevent blindness and improve the health and well-being of individuals affected by cataract.
- **Goal 9 - Industry, Innovation and Infrastructure:** By developing and implementing an automatic cataract detection system, it can help to increase access to healthcare, particularly in remote or under-served areas.
- **Goal 10 - Reduced Inequalities:** By reducing the cost and burden of cataract detection and treatment on healthcare systems, machine learning can help to reduce inequalities in access to healthcare.
- **Goal 11 - Sustainable Cities and Communities:** By increasing the speed and accuracy of cataract detection, machine learning can help to prevent blindness and improve the quality of life for individuals affected by cataract, particularly in urban areas.

CHAPTER 6

Conclusion and Future Work

The purpose of this paper is to present a survey of various methods for detecting cataracts. The approaches discussed in this paper are intended to aid in the diagnosis and treatment of cataracts. Deep convolutional neural network architecture-based models which are VGG19, NASnet, ResNet50 and MobilenetV2 with the transfer learning technique are trained using 2000 images to detect the cataract. As we can adjusted the overfitting, underfitting, and training period, deep learning is crucial. The highest accuracy was 97.75% given by MobileNetV2 and classify the images most accurately. In addition, MobileNetV2 should be used in conjunction. We have a proposed model that outperforms current transfer learning models. Transfer learning involves the fine-tuning of parameters and values [22]. Transfer learning is a training approach that aims to improve the overall performance of a model [24].

In future we will try more effective model implementation and preprocessing technique to achieve more improvements in accuracy and extend the model for cataract grading from the eye image. Aside from that, we'll add more data of funds imaging. We will also apply deep pre-processing approaches to increase the strength of the model.

APPENDIX

In this appendix, an overview of this work is given, which utilizes a deep convolutional neural network (CNN) architecture trained on a dataset of retinal fundus images. To enhance the model's performance and prevent overfitting, we incorporated a dense layer of 512 neurons with ReLU activation and a dropout layer with a rate of 0.5. Our system achieved an impressive accuracy of 96.5% on the test dataset and demonstrated potential for detecting cataract in its early stages. However, it's important to note that the dataset used in this study was limited to patients from limited hospital and may not be generalizable to the wider population. Additionally, the model's accuracy may be affected by variations in the quality and resolution of the images. Further research is needed to improve the robustness of the model and to evaluate its performance on more diverse and larger datasets.

REFERENCE

[1] Mayo Clinic Staff, Cataracts, www.mayoclinic.org/diseases-conditions/cataracts/symptoms-causes/syc20353790, Last Accessed: Jan. 20, 2023.

[2] World Health Organization, Blindness and vision impairment, www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment, Last Accessed: Jan. 20, 2023.

[3] Kaggle. (2023). Eye diseases classification. Last Accessed: Jan. 1, 2023. [Online]. Available: www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification, Last Accessed: Jan. 20, 2023.

[4] Kaggle. (2023). Ocular Disease Recognition. Accessed: Jan. 1, 2023. [Online]. Available: www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k, Last Accessed: Jan. 20, 2023.

[5] Hossain, Md Rajib, Sadia Afroze, Nazmul Siddique, and Mohammed Moshiul Hoque. "Automatic detection of eye cataract using deep convolution neural networks (DCNNs)." In 2020 IEEE region 10 symposium (TENSYMP), pp. 1333-1338. IEEE, 2020.

[6] L. Xiong, H. Li, and L. Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis," J. Healthc. Eng., vol. 2017, 2017.

[7] Z. Qiao, Q. Zhang, Y. Dong and J. Yang, "Application of SVM based on genetic algorithm in classification of cataract fundus images," 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, 2017, pp. 1-5

[8] Y. Zhou, G. Li and H. Li, "Automatic Cataract Classification Using Deep Neural Network With Discrete State Transition," in IEEE Transactions on Medical Imaging, vol. 39, no. 2, pp. 436-446, Feb. 2020.

[9] X. Zhang, J. Lv, H. Zheng, and Y. Sang, ''Attention-based multi-model ensemble for automatic cataract detection in B-scan eye ultrasound images,'' in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2020, pp. 1–10.

[10] M. S. M. Khan, M. Ahmed, R. Z. Rasel, and M. M. Khan, ‘‘Cataract detection using convolutional neural network with VGG-19 model’’ in Proc. IEEE World AI IoT Congr. (AIIoT), May 2021, pp. 0209–0212

[11] Hasan, Md Kamrul, Tanjum Tanha, Md Ruhul Amin, Omar Faruk, Mohammad Monirujjaman Khan, Sultan Aljahdali, and Mehedi Masud. "Cataract Disease Detection by Using Transfer Learning-Based Intelligent Methods." Computational and Mathematical Methods in Medicine 2021 (2021).

[12] Imran, Azhar, Jianqiang Li, Yan Pei, Faheem Akhtar, Ji-Jiang Yang, and Yanping Dang. "Automated identification of cataract severity using retinal fundus images." Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 8, no. 6 (2020): 691-698.

[13] Nur, Nahya, Sugiarto Cokrowibowo, and Rosalina Konde. "Cataract Detection in Retinal Fundus Image Using Gray Level Co-occurrence Matrix and K-Nearest Neighbor." In International Joint Conference on Science and Engineering 2021 (IJCSE 2021), pp. 268-271. Atlantis Press, 2021.

[14] Pratap, Turimerla, and Priyanka Kokil. "Deep neural network based robust computer-aided cataract diagnosis system using fundus retinal images." Biomedical Signal Processing and Control 70 (2021): 102985.

[15] Zhang, Linglin, Jianqiang Li, He Han, Bo Liu, Jijiang Yang, and Qing Wang. "Automatic cataract detection and grading using deep convolutional neural network." In 2017 IEEE 14th international conference on networking, sensing and control (ICNSC), pp. 60-65. IEEE, 2017.

[16] Weni, Indra, Pradita Eko Prasetyo Utomo, Benedika Ferdian Hutabarat, and Muksin Alfalah. "Detection of Cataract Based on Image Features Using Convolutional Neural Networks." Indonesian Journal of Computing and Cybernetics Systems) 15, no. 1 (2021): 75-86.

[17] Sirajudeen, A., and S. Karthikeyan. "Detection of cataract through feature extraction by the novel angular binary pattern (NABP) and classification by kernel based convolutional neural networks." (2021).

[18] J. Ran, K. Niu, Z. He, H. Zhang and H. Song, “Cataract Detection and Grading Based on Combination of Deep Convolutional Neural Network and Random Forests,” 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC), Guiyang, 2018, pp. 155-15

[19] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 2010

[20] Learning Transferable Architectures for Scalable Image Recognition https://arxiv.org/abs/1707.07012, Last Accessed: Jan. 20, 2023.

[21] Machine Learning, Sensitivity vs Specificity Difference, https://vitalflux.com/ml-metrics-sensitivity-vs-specificity-difference, Last Accessed: Jan. 20, 2023.

[22] Jabbar, H., and Rafiqul Zaman Khan. "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)." Computer Science, Communication and Instrumentation Devices (2015): 163-172.

[23] Manaswi, Navin Kumar. "Understanding and working with Keras." In Deep Learning with Applications Using Python, pp. 31-43. Apress, Berkeley, CA, 2018.

[24]] Torrey, Lisa, and Jude Shavlik. "Transfer learning." In Handbook of research on machine learning applications and trends: algorithms, methods, and techniques, pp. 242-264. IGI global, 2010

[25] Confusion Matrix - an overview, https://www.sciencedirect.com/topics/engineering/confusion-matrix#:~:text=Confusion%20matrix%20visualization.,negative%20and%20is%20actually%20negative, Last Accessed: Jan. 20, 2023.

ORIGINALITY REPORT

9%	3%	6%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	"Artificial Intelligence and Speech Technology", Springer Science and Business Media LLC, 2022 Publication	1%
3	"Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges", Springer Science and Business Media LLC, 2021 Publication	1%
4	Jonathan Janke, Mauro Castelli, Aleš Popovič. "Analysis of the proficiency of fully connected neural networks in the process of classifying digital images. Benchmark of different classification algorithms on high-level image features from convolutional layers", Expert Systems with Applications, 2019 Publication	1%
5	Submitted to Indian School of Business Student Paper	<1%