ENHANCING RETINAL DISEASE DIAGNOSIS THROUGH ABLATION STUDIES: A ROBUST DEEP-LEARNING CNN MODEL WITH IMPROVED IMAGE QUALITY AND GENERALIZABILITY

BY

MD. ABDULLAH AL MOSADDIK

ID: 192-15-2889

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

Computer Science and Engineering

Supervised by

Mr. Mohammad Jahangir Alam
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by

Mr. Mushfiqur Rahman
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL

UNIVERSITY

Dhaka, Bangladesh

May 14, 2025

APPROVAL

The work that Md. Abdullah Al Mosaddik submitted to the Department of Computer Science and Engineering at Daffodil International University, titled **"ENHANCING RETINAL DISEASE DIAGNOSIS THROUGH ABLATION STUDIES: A ROBUST DEEP-LEARNING CNN MODEL WITH IMPROVED IMAGE QUALITY AND GENERALIZABILITY,"** has been approved in terms of both style and content and has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering. In 2025, the presentation took place on May 14.

BOARD OF EXAMINERS

Dr. Fizar Ahmed (FZA) — **Chairman**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Amatul Bushra Akhi (ABA) — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Lamia Rukhsara (LR) — **Internal Examiner**
Sr. Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammed Nasir Uddin (MNU) — **External**
Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

I hereby declare that this project has been done by us under the supervision of **Mr. Mohammad Jahangir Alam,** Supervisor's Designation, Department of Computer Science and Engineering, Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

14.05.25

....

Mr. Mohammad Jahangir Alam

Assistant Professor

Department of Computer Science and Engineering

Daffodil International University

Co-supervised by:

18.05.25

....

Mr. Mushfiqur Rahman

Assistant Professor

Department of Computer Science and Engineering

Daffodil International University

Submitted by:

....

Md. Abdullah Al Mosaddik

ID: 192-15-2889

Department of Computer Science and Engineering

Daffodil International University

ACKNOWLEDGEMENTS

Firstly, I would like to thank God Almighty for His divine blessing, which enabled me to successfully finish my final year project/internship.

I express my gratitude and deepest debt to Professor **Mr. Mohammad Jahangir Alam**, who is an **assistant professor in the CSE department at Daffodil International University** in Dhaka. To carry out this project, our supervisor has extensive knowledge of and a strong interest in the field of "computer vision." Completing this project has been made possible by his unending patience, academic guidance, constant encouragement, energetic and constant supervision, constructive criticism, valuable advice, and reading numerous subpar drafts and correcting them at every stage.

In addition to other faculty members and staff at Daffodil International University, I would like to extend our sincere gratitude to the Department of CSE for their helpful assistance in completing our project.

I would like to express my gratitude to all of our classmates at Daffodil International University who participated in this discussion during the coursework.

Lastly, I must respectfully thank our parents for their unwavering patience and support.

ABSTRACT

Retinal disease is a global health crisis that impacts millions of individuals annually. Early illness diagnosis is an important field of medical research since eye problems can greatly affect patients' eyes and overall well-being. A plethora of data for model training and improved model designs have led to the rapid development of deep learning techniques for computer vision applications. This research will describe a powerful convolutional neural network (CNN) model that uses ablation studies to categorize eye disease illnesses. The disease is colored, and the overall quality is increased by reducing noise, improving the image, and removing artifacts. The ODIR 5k dataset of retinal images is freely available to the public. Due to improvement techniques, the quantity of images depicting eye diseases has increased. At the outset, the enhanced dataset suggests a basic CNN model. The robust CNN model, which I propose, was derived from an ablation investigation. To train the model, a library of publicly available images depicting eye illnesses is utilized. In terms of accuracy (92%), area under the curve (AUC) (99%), and KAPPA (91%), the recommended robust model outperformed the competition. The model is robust and performs admirably when applied to fresh data. The model also achieves good recall and precision when used to identify binary and multi-class eye disorders.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

When the eye's ability to function normally is impaired or visual acuity is negatively impacted, I say that there is an ocular ailment. [1]: Yes. Problems with one's eyesight affect nearly everyone. Children who aren't covered by insurance or who can receive care at home are among those who need a specialist's attention. Eye problems that don't go away might damage the retina, which can cause irreversible vision loss and possibly blindness if left untreated. Many factors, including demographics (gender, age, occupation, lifestyle, socioeconomic level, cultural norms, cleanliness habits, etc.), contribute to the wide variation in the incidence of eye illnesses across nations. Irresistible eye infections are more likely in tropical communities because of factors like dust and daylight, according to a study that studied populations in temperate and tropical countries [2]. Furthermore, the prevalence of eye diseases in developing and industrialized nations is distinct. High rates of underdiagnosed and untreated ocular morbidity are prevalent in many underdeveloped nations, particularly in Asia. Of the 285 million persons on Earth who are visually impaired, 246 million have some degree of visual impairment and 39 million are completely blind. About 2.2% of the global population has nearsightedness or farsightedness, says the World Health Organization (WHO) [3]. You could have avoided or dealt with half of these situations differently, according to estimates. Untreated cataract (94 million), glaucoma (7.7 million), diabetic retinopathy (3.9 million), corneal opacities (4.2 million), trachoma (4.2 million), and 826 million cases of uncorrected presbyopia (2 million) combine to cause moderate to- severe distance vision impairment or blindness in 1 billion people. Many diseases and conditions can limit one's ability to see clearly. These include uncorrected refractive errors, cataracts, glaucoma, diabetic retinopathy, corneal opacity, trachoma, hypertension, and many more. There has been scant research on the incidence of blindness and visual impairment in Bangladesh. Rural areas are home to the vast bulk of the country's population. Even though a large majority of city dwellers require medical treatment, ophthalmology services are in short supply. Blindness is relatively uncommon, even though more companies are offering assistance to those who are visually impaired [4]. According to the Bangladesh National Blindness and Poor Vision Survey, about one-fifth of the population has low vision, which is characterized as a single-eye visual acuity of less than 6 inches. The increasing incidence of noncommunicable disorders, such as smoking and diabetes, may put Bangladeshis at a greater risk of eyesight loss. A substantial portion of the urban poor population resides in poor mental and physical health [5]. Providing these individuals with affordable or free comprehensive eye care treatments is of the utmost importance. By 2030, more than 400 million people will be impacted by DR, according to polls that have been conducted on the topic. Furthermore, by 2020, the number of people affected by glaucoma is projected to exceed 80 million worldwide. In recent years, eye illnesses have gained prominence as a major issue in public health around the world. Ophthalmic disease is important since it can cause persistent vision loss and has irreversible symptoms. Ocular illnesses such as age-related macular degeneration, diabetic retinopathy (DR), and cataract are major contributors to vision loss [6].

As CAD based early detection is important, deep-learning-based algorithms are increasingly used in medical image analysis recently. It has been demonstrated that deep-learning-based models excel at a number of tasks, including disease detection, sentiment analysis, and object detection [7]. CNNs have demonstrated promising performance in a number of areas, from disease classification to object detection, regarding diagnosing ocular diseases. However, high-performing classification models may be constructed without convolutions [8,9], a solution that solves the problem of computational complexity. In this context, the attention mechanism is fast rising in importance as a topic of study in the field of current ML. Vision Transformer (ViT) is a standout piece of research due to its application of a pure self-attention-based model to sequences of picture patches and its subsequent competitive performance on ImageNet classification [10] when compared to CNNs. So, I tried to use the attention mechanism with MobileNet.V2 architecture.

One of the most important steps in reducing an ophthalmologist's workload is the automatic diagnosis of illnesses. Without the need for human intervention, diseases can be detected using deep learning and computer vision technologies. Only a small number of these studies have been able to fully diagnose [11] more than one eye illness, even though many of them have produced encouraging results. To accurately diagnose various eye conditions, more research is required to analyze the various aspects of fundus imaging. This study suggests a system that uses deep learning to recognize different eye diseases. Multi class classification has been used as a different strategy [12].

The taken ocular disease's dataset is very unbalanced. This imbalance makes it difficult to accurately detect or classify disease or even a normal image. This method is not recommended for generalized classification tasks due to its low accuracy [13]. The classification of ocular diseases was the goal of this work. The study's dataset was incredibly unbalanced.

1.2 Motivation

Early ocular disease detection is a crucial step in preventing irreversible vision loss or even blindness. The identification of eye illnesses has been greatly aided by the development of several imaging techniques. Common diagnostic tools include color fundus photography (CFP) in addition optical coherence tomography (OCT) [14]. OCT produces retinal cross-section images, and retinal thickness can be measured to assess the health of the eyes. CFP keeps track of the surfaces inside the eyes to look for any potential issues. It has been demonstrated that both tools are useful for diagnosing ocular diseases in their initial stages. However, CFP is a more cost-effective and effective method, and it is advised for asymptomatic adults, particularly for the elderly populations, to undergo periodic fundus examinations with CFP [15]. Unfortunately, common ocular conditions like diabetic retinopathy, cataracts, macular degeneration caused by age, etc. progress with few early telltale symptoms, making it challenging to make a precise diagnosis in the initial stages. Additionally, manual inspection of the massive amounts of CPFs generated takes a lot of time and effort. The ratio of ophthalmologists to patients is imbalanced, with a shortage of ophthalmologists relative to the number of patients requiring care. In addition, the process of manually conducting fundus screening is a time-intensive endeavor and heavily dependent on the expertise of ophthalmologists [16]. These factors express challenges in conducting retinal screening on a large scale. Hence, the development of an automated computer-aided diagnostic (CAD) algorithm for the screening of ophthalmic diseases holds significant importance.

1.3 Rationale of the Study

The main objectives of the proposed work are given below:

- **i)** The objective of the system is to detect eye disease by classifying it into eight different classes.
- **ii)** Develop a reliable algorithm that can classify eye disease across the dataset with various properties.

iii) As the dataset is very imbalanced, I performed data augmentation to solve the data imbalance problem.

iv) To make the image more readable, I applied various image preprocessing algorithms to perform denoising, image enhancing and artifact removal.

v) Implemented several pre-trained models to find out the best transfer learning model for eye disease prediction.

vi) I further proposed a more robust model to perform the eye disease prediction more effectively by performing an ablation study.

1.4 Research Questions

- ❖ To better forecast the occurrence of eye diseases, how can I identify the areas where current machine vision-based systems are lacking in their ability to accurately categorize these conditions?
- ❖ When it comes to predicting various eye disorders, which augmentation and preprocessing will work best?
- ❖ How can I build a reliable model to improve the precision of disease classification for the eyes?

1.5 Expected Output

The focused challenges for this research are:

- ❖ **Dataset**: The dataset is imbalanced. For training and testing the models, the dataset needs to be well-balanced.
- ❖ **Data Augmentation:** To expand the size of the datasets, the selection of the proper data augmentation technique is challenging. Using a variety of augmentation methods to know about the optimal data augmentation.
- ❖ **Image Processing:** Because images from different sources can occasionally be noisy or have low contrast, image preprocessing is required. Producing classification images with no noise and improved contrast is challenging.
- ❖ **Base Model Selection:** To overcome lengthy training periods and a lack of data, an ideal base model is required for ablation investigations.
- ❖ **Improve Accuracy and Provide a Robust Model:** I employ a modified CNN model to increase the accuracy of my classification of eye diseases. The primary contribution of this study is a technique and a strong model proposed for more accurate prediction of eye diseases. To improve the classification, I have

performed some image preprocessing and to balance the dataset, I performed data augmentation. Finally, I have developed a modified CNN model to classify eye disease.

1.6 Project Management and Finance

No organization or business provided financial support for this research-based project.

1.7 Report Layout

The study's introduction, its aims, and its core research issues are all laid out in Chapter 1. Chapter 2 provides brief summaries of the reviewed material. In Chapter 3, the author lays out the specific methodology and models. Procedures and Resource Guide The experimental data and discussion of the study are presented in Chapter 4. Chapter 5 discusses the impact, effect and sustainability of the study. The current study is wrapped up in Chapter 6, and a roadmap for future research is provided.

CHAPTER 2
BACKGROUND

2.1 Preliminaries/Terminologies

Researchers have proposed various deep learning and machine learning methods to classify eyes. I present a few studies from this field's literature that classified diseases using datasets for common eye conditions.

2.2 Related Works

For use with individual patients, Junjun et al. [17] suggested using a convolutional neural network to create a multi-label eye illness classification model. To address this issue, I employ a dense correlation network (DCNet). DCNet is built upon a CNN and includes a classifier and a spatial correlation module. Features are extracted from left and right color fundus pictures using CNN's core program. Pixel-by-pixel associations between the two feature sets will be recorded by the spatial correlation module. A patient's representation is built by splicing together processed characteristics. Patients' voices are heard when diseases are categorized. Their proposed model has an AUC of 0.93. The suggested model outperforms a large number of baseline classification methods on a publicly accessible dataset using a multi-label soft margin loss.

In a separate study, Ning et al. [18] conducted benchmark experiments with innovative deep neural networks on a shared dataset of eye diseases. In their publication, they showed off a collection of fundus pictures labeled with a variety of diseases. A total of 10,000 pictures, representing a wide range of variances, were captured from the left eyes and right eyes of 5,000 clinical patients. I benchmark the previous in deep learning by comparing the results of many current models on our dataset. According to experimental results, increasing the neural network's depth alone does not result in a performance boost, but expanding the network can. Inception v4, with an accuracy of 72.42 percent, outperformed other deep learning models.

According to Nakhim et al. [19], a multi-category DNN technique can successfully classify the three most prevalent eye diseases. Data augmentation, region of interest reduction, and iso-luminance histogram equalization were only some of the picture preprocessing approaches and top-performing residual deep neural networks employed. When applied to our categorization of three retinal disorders from public databases, they attained peak and average accuracy of 91.16 and 85.79 percent, respectively. Specificities were 90.06 percent, 99.63 percent, 99.82 percent, and 91.90 percent, respectively, for images of healthy, AMD, DR, and GLC patients' eyes.

The study conducted by Smitha et al. [20] aims to evaluate the efficacy of classification of retinal fundus pictures using semi-supervised Generative Adversarial Networks (GANs). Additionally, the nonlocal retina framework has been utilized to enhance fundus pictures without resorting to gauzy smoothing at the expense of detail. The proposed research makes use of a large body of publicly available raw fundus data collected from a variety of eye institutions. When the acquired results were compared with some approaches of transfer learning, an 87% average accuracy was found. The findings indicate that semi-supervised GANs have the potential to be utilized for the classification of diverse retinal disorders.

It was proposed by Akhilesh et al. [21] Inception-ResNet-v2 is pre-trained using transfer learning to create the hybrid model, and then a custom block of CNN layers is added on top of it. The proposed model was evaluated using the Messidor-1 dataset for diabetic retinopathy and the APTOS-2019 blindness detection dataset (both from Kaggle). They compared the results of their model to those of other researchers. Test accuracy on Messidor-2 and APTOS was 72.33 and 82.13 percent, respectively.

The deep learning model presented by Yang et al. [22] consists of a data preprocessor, the DSRA-CNN feature extraction network, and a classifier for the categorization of eight distinct fundus disorders, making it possible for automated diagnosis of numerous fundus diseases. Features are extracted from input data using the DS block, DSR block, and SE block in the Exception-based DSRA-CNN network. The proposed model outperformed existing advanced convolutional neural networks in terms of accuracy (87.90%), precision (88.50%), F1 value (88.16%), and kappa score (86.17%). The proposed model can contribute to the development of automated diagnosis systems, which is crucial for preventing severe visual impairment.

Eight distinct categories of eye illnesses were proposed and analyzed by Islam et al. [23] using CNN. The model achieved an F-score of about 85%, a Kappa of about 31%, and an AUC of around 80%. The proposed approach is the first to forecast several illnesses in an eye using "real-world" datasets. The conclusion of the article is that artificial detection allows physicians to predict diseases early, treat patients without needless delay, and prevent total blindness.

In a different study, automated diagnosis and categorization of ophthalmological illnesses from fundus pictures using deep learning approaches is discussed by Faith et al. [24]. This includes the use of CNN (convolutional neural networks) and LSTM (long short-term memory) models. Classification accuracy is enhanced by the suggested technique, which combines R-CNN+LSTM architecture with a multilevel feature selection algorithm termed NCAR. Using the SVM classifier, the method achieved an accuracy of 89.54 percent, outperforming existing models such as ResNet, InceptionV3, MobileNet, and EfficientB3. The article emphasizes the significance of early detection and diagnosis of ophthalmological diseases in preventing vision impairment and blindness.

Neha et al. [26] proposed a method for automatically classifying fundus images into multiple classes and labels. Utilizing four convolutional neural network architectures based on transfer learning is the strategy. In this experiment, eight classes of left and right fundus images from the ODIR dataset were utilized. Using the VGG16 pre trained architecture in conjunction with the SGD optimizer produces superior results, as demonstrated by an AUC of 0.91 in the classification of multi-class fundus images on the ODIR database.

To analyze fundus images in color Vinay et al. [27] proposed an automated CNN-based rapid system with multiple labels. This experiment utilized the ODIR dataset, which contains seven classes of left-to-right fundus images. In addition to CLAHE (Contrast Limited Adaptive Histogram Equalization) preprocessing, augmentation and concatenation are utilized in preprocessing. Several transfer learning models, including VGG-16, InceptionV3, and ResNet50, had their performance evaluated. In the Hamming loss (HL) metrics, ResNet50 received 0.1793, InceptionV3 received 0.1462 and VGG16 received 0.236.

The suggested research by Smitha et al. [20] seeks to analyze how well semi-supervised GANs classify pictures of the retinal fundus. To further improve the fundus picture without softening the margins, the nonlocal retinex system is employed. The approach suggested here makes use of a large fundus data collection that has been released after being collected from a variety of eye facilities. The results are determined to be 87% accurate when compared to those obtained using the transfer learning method. This finding suggests that semi-supervised GANs might be used for the categorization of a wide variety of retinal illnesses. In this study, I investigate how well semi-supervised GAN can categorize pictures of the retina and optic nerve. Furthermore, the nonlocal retinex structure is used to improve fundus images without too softening the margins. The approach suggested here makes use of a large fundus data collection that has been released after being collected from a variety of eye facilities. Accuracy of 87% is found when results are compared to those obtained via transfer learning. It implies that semi supervised GANs might be used to classify a wide variety of retinal illnesses.

To analyze diabetic retinopathy (DR), Gayathri et al. [28] used three separate datasets: Messidor, Kaggle and IDRid. These datasets were used for multi-class and binary classification. To better extract characteristics from retinal pictures, a new convolutional neural network (CNN) model was presented. Six different machine learning classifiers—including SVM, AdaBoost, NB, RF, and J48—are fed the CNN model's output characteristics. The J48 classifier showed excellent accuracy in binary classification, reaching 99.89% of the time. The accuracy percentage was 99.59 percent in uniclass classification and 99.59 percent in multiclass classification. By excluding include ROIs that could have been affected by diabetic retinopathy (DR), the CNN feature extractor used by the researchers was able to reduce computation time and complexity while still using the full picture. To achieve a high degree of accuracy in multiclass and binary classification, it was determined that machine learning classifiers were required. A CNN (Convolutional Neural Network) was used to extract features; however, the suggested CNN's classification performance was not evaluated.

In [29], Yaqoob et al. presented a deep learning-based technique to DR picture classification and evaluation. In this technique, the RF is used to make classifications based on data derived from the ResNet-50 model. The suggested technique was compared to various transfer learning using two categories from the Messidor2 dataset and five classes from the ResNet50 dataset. EyePACS. For the Messidor2 and EyePACS datasets, the suggested technique attained an accuracy of 96% and 75.09. The suggested architecture achieves 96% accuracy on the two-category Messidor-2 dataset by integrating the deep features of ResNet-50 combined with a classifier from Random Forest. For the five-category EyePACS dataset, this drops to 75.09 percent due to the dataset's severely unbalanced data and the lack of suitable preprocessing procedures.

Five deep CNN models were presented by Sehrish et al. in [30]. Using the publicly available Kaggle dataset of retinal images, some transfer learning models enhance classification for different stages of DR and encode the rich characteristics. Compared to previous transfer learning approaches on the same Kaggle dataset, the proposed model recognizes all stages of DR, as shown by experimental findings.

Using CNN (Convolutional Neural Network) methods, which involve deep learning as a main component and are accelerated by GPUs, Lifeng et al. [31] determine if a microaneurysm is present in a fundus picture. Using low latency and fast inference, this system will recognize and segment medical pictures. Semantic segmentation is applied to analyze the fundus picture for signs of contamination. It has been proposed that the efficiency with which a deep CNN can be trained for segmentation of fundus pictures be leveraged by the Prognosis of Microaneurysm and Early Detection System for Nonproliferative Diabetic Retinopathy to increase the reliability and precision of NPDR prediction.

The necessary treatment procedures have been annotated on a dataset of DR fundus pictures produced by Zenato et al. [32]. They trained deep convolutional neural network models on this data set to categorize the severity of DR fundus pictures. The results of a four-degree categorization test showed an experimental accuracy of 88.72 percent. Clinical testing showed that 91.8% of ophthalmologists used the same criteria, validating the system's accuracy. They deployed their models to the cloud and offered preliminary DR diagnostic services to many medical facilities.

Timchenko et al. [33] presented the problem using a deep learning algorithm that automatically recognizes the pattern and classifies retinal images into one of five classes based on the pattern. On the Messior-2 dataset, they achieved the highest validation accuracy of 74.4%.

The potential of employing CFPs from a single patient visit to foresee the development of DR was established by Filippo et al. [12]. If further refined on bigger and more diversified datasets, including the possibility of early intervention, such an algorithm might ease early detection and referral to a retina expert for more frequent monitoring and treatment. The number of people willing to participate in clinical trials related to DR might also increase. For sensitivity of 66% and specificity of 77%, an AUC (Area Under Curve) of 0.68 0.13, 0.79 0.05 for 91% sensitivity and 66% specificity, and 0.77 0.04 for 79% sensitivity and 72% specificity, respectively.

The work of Yashal et al. [34] centers on the idea of using a deep learning model to the task of categorizing images of DR fundi according to their degree of severity. In this study, they present a deep learning based automatic classification and identification technique for DR (diabetic retinopathy) fundus pictures. The suggested technique includes the phases of preprocessing, segmentation, and classification. First, in the preprocessing stage, I get rid of the background noise around the image's edges. Histogram-based segmentation is then used to cut out the relevant parts of a picture. The next stage is to grade DR fundus images by degree using a synergic deep learning (SDL) model. Using the Messidor-2 DR data set, the supplied SDL model is defended. Additionally, it is worth noting that the AlexNet model has the lowest classification accuracy, with a classifier accuracy of 89.75%. Table data shows that the offered SDL model achieves an accuracy of 99.28%, a sensitivity of 98.54%, and a specificity of 99.38% in its classifications, making it obviously superior to other models.

Researchers Saja and colleagues [35] used photos from the ODIR dataset to offer an automated solution based on machine learning (CNN) technology for identifying individuals with macular degeneration (AMD). Diagnosing disorders like AMD requires pinpoint accuracy when identifying the disease's location in retinal scans. The collection of fundus pictures was processed to extract high-level features by a CNN (Convolutional Neural Network) for AMD and normal classification. Specificity, sensitivity, accuracy, F-score, and precision were among the measures used to evaluate the classification efficacy. The maximum achievable rate of classification accuracy was 99.1%.

Using deep learning algorithms on fundus pictures, Delavari et al. [36] want to create and validate a method for offering reliable Alzheimer's disease (AD) diagnosis. Since there aren't many cases with positive AD labeling in the available fundus datasets, I used the study of patient gender as a stand-in for the case study. This is because CNNs (convolutional neural networks) can reliably anticipate this property, even though humans can't see it. In this article, I provide a novel three-stage methodology for analyzing gender categorization in retinal fundus pictures using CNNs (Convolutional Neural Networks). In the first step, I train and test a CNN (Convolutional Neural Network) model for sex categorization. This was achieved by analyzing 3,146 photos from the Ocular Disease Recognition (ODIR) and 1,600 images from the Diabetic Retinopathy Online Challenge (DOVS) databases, respectively. The next phase will involve the use of deep learning strategies.

The purpose of this investigation [37] was to assess the performance of the classification of current convolutional neural network (CNN) models or structures using fundus dataset images annotated with eight different illness diagnoses. A total of eight unique eye disorders were correctly identified using an openly available database for intelligent recognition. Approximately 10,000 fundus photos from both eyes of 5,000 individuals were used to create an intelligent recognition database for ocular illnesses, as described in the current work. There are eight distinct eye conditions represented here. The goal of this research was to compare the effectiveness of three pre-trained convolutional neural network architectures in classifying ocular disorders. The models were run using Google Colab, simplifying the process by eliminating the requirement to set up the necessary environment and libraries from scratch. Partitioning the dataset into 70 percent for training, 10 percent for validation, and 20 percent for testing allowed for a more accurate evaluation of the models' performance. Ten thousand fundus pictures were created for each categorization by augmenting the training photos. The ResNet50 model showed remarkable performance achieving an accuracy of 97.1%, sensitivity of 78.5%, specificity of 98.5%, and precision of 79.7% in cataract classification. It also had the highest AUC of 0.964 and final score of 0.903 of the bunch. Compared to these, the VGG16 model achieved 96.2 percent accuracy, 56.0 percent sensitivity, 99.2 percent specificity, 84.1 percent precision, 0.949 percent area under the curve, and 0.857 percent overall score.

DeepRetino [38], an automated multi-classification technique for six ocular disorders (CNNs), was proposed by Fatima et al. using recent advances in deep learning, specifically convolutional neural networks. The CLAHE method was used to boost contrast in the fundus pictures during preprocessing. Xavier orthogonal and Adam optimizer are used to initialize and adjust the network weights throughout the training phase. Finally, the DeepRetino model is evaluated on the ODIR dataset to see if it is production-ready.

2.3 Summary and Comparative Analysis

The presented studies collectively showcase diverse methodologies and advancements in the field of automated diagnosis of eye diseases through deep learning models. Junjun et al. proposed a dense correlation network (DCNet) built upon a CNN, achieving superior performance in multi-label eye illness classification. Ning et al. conducted benchmark experiments, emphasizing the importance of network expansion over increased depth. Nakhim et al. successfully classified prevalent eye diseases using a multi-category deep neural network. Smitha et al. explored semi-supervised Generative Adversarial Networks (GANs) for retinal fundus picture classification, surpassing some transfer learning approaches. Akhilesh et al. introduced a hybrid model with Inception-ResNet-v2, exhibiting promising results on Messidor-2 and APTOS-2019 datasets. Yang et al. proposed a DSRA-CNN model, outperforming existing convolutional neural networks. Islam et al. utilized CNN for the categorization of eight distinct eye illnesses. Fatih et al. employed CNN and LSTM models, surpassing traditional deep learning approaches. Neha et al. focused on multi-class fundus image classification with transfer learning. Vinay et al. presented an automated CNN-based system for multiple labels, evaluating various models. These studies collectively highlight significant strides in leveraging advanced neural network architectures and preprocessing techniques for accurate and efficient eye disease diagnosis.

2.4 The scope of the issue

According to the research that has come before, there are a number of methods used to categorize eye illnesses that rely on machine learning and deep learning. But there are downsides, like inefficient use of time and data, poor accuracy, and no use of data augmentation or pre-processing techniques. Our research aimed to solve these problems by creating a dependable model to categorize several types of eye diseases.

2.5 Challenges

The research's specific challenges are:

- ❖ Dataset: The dataset is imbalanced. For training and testing the models, the dataset needs to be well-balanced.
- ❖ Data Augmentation: To expand the size of the datasets, the selection of the proper data augmentation technique is challenging. Using a variety of augmentation methods to know about the optimal data augmentation.

- Image Processing: Because images from different sources can occasionally be noisy or have low contrast, image preprocessing is required. Creating noise-free, contrast-enhanced classification images is difficult.
- Selecting the Base Model: To deal with lengthy training periods and a lack of data, ablation studies require the ideal base model.
- Improve Accuracy and Provide a Robust Model: I employ a modified CNN model to increase the accuracy of my classification of eye diseases.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

The research subject of this study focuses on the development and evaluation of a deep learning-based multi-classification and categorization system for ocular diseases. The primary goal is to employ convolutional neural network (CNN) models to effectively classify various eye ailments, providing accurate and timely diagnoses. The instrumentation utilized in this research involves the Ocular Disease Intelligent Recognition (ODIR) dataset, sourced from Kaggle, which comprises eight distinct classes and 5,000 patients. The dataset includes color fundus photographs of both left and right eyes, along with diagnostic keywords provided by medical professionals. The study leverages innovative models, including CNN, MobileNet V2, EfficientNet V2b3, and MobileNet V2+ soft attention CNN+LSTM, to evaluate their performance in ocular disease classification. The research subject underscores the significance of advanced deep learning techniques in enhancing the diagnosis and categorization of ocular diseases for improved healthcare outcomes.

3.2 Data Collection Procedure/Dataset Utilized

The Ocular Disease Intelligent Recognition (ODIR) dataset was employed from the widely used data repository, Kaggle. The dataset comprises eight distinct categories. ODIR is a systematically organized database of ophthalmic data that encompasses a comprehensive record of 5,000 patients and annotations of 3500 patient’s images are made publicly available. The database includes pertinent information such as the patients' age, color fundus photographs of both the left eyes and right eyes, and diagnostic keywords provided by medical professionals. The dataset under consideration is designed to serve as a realistic compilation of patient information gathered by Shanggong Medical Technology Co., Ltd. from various medical facilities and hospitals located in China. Fundus images are taken in diverse organizations using many cameras including Zeiss, Canon, and Kowa, leading to variations in image resolutions.

3.3 Statistical Analysis

The process of tagging annotations was carried out by human readers who had received training in quality control management.

The patients are segregated into eight distinct categories, comprising:

1. Diabetes (D),
2. Normal (N),
3. Cataract (C),
4. Glaucoma (G),
5. Hypertension (H),
6. Age related Macular Degeneration (A),
7. Other disorders/diseases (O)
8. Pathological Myopia (M),

Each class contains various number of images, like Normal (N): 2873 images, Diabetes (D): 1609 images, Glaucoma (G): 284 images, Cataract (C): 293 images, Age-Related Macular Degeneration (A): 266 images, Hypertension (H): 128 images, Pathological Myopia (M): 244 images, and other disorders/diseases (O): 982 images.

Table 3.1: Number of images in each class

Dataset Classes	**Images**
Diabetes (D),	1609
Glaucoma (G)	284
Cataract (C)	293
Age-Related Macular Degeneration (A)	266
Hypertension (H)	128
Normal(N)	2873
Pathological Myopia (M)	244
Other disorders/diseases (O)	982
Total Images	5699

Nevertheless, the sample distribution is uneven, and the pictures contain artifacts. In this work, I performed multiple image preprocessing operations to increase the quality of the image.

3.4 Proposed Methodology/Applied Mechanism

The whole working process is shown in figure 3.1 below.

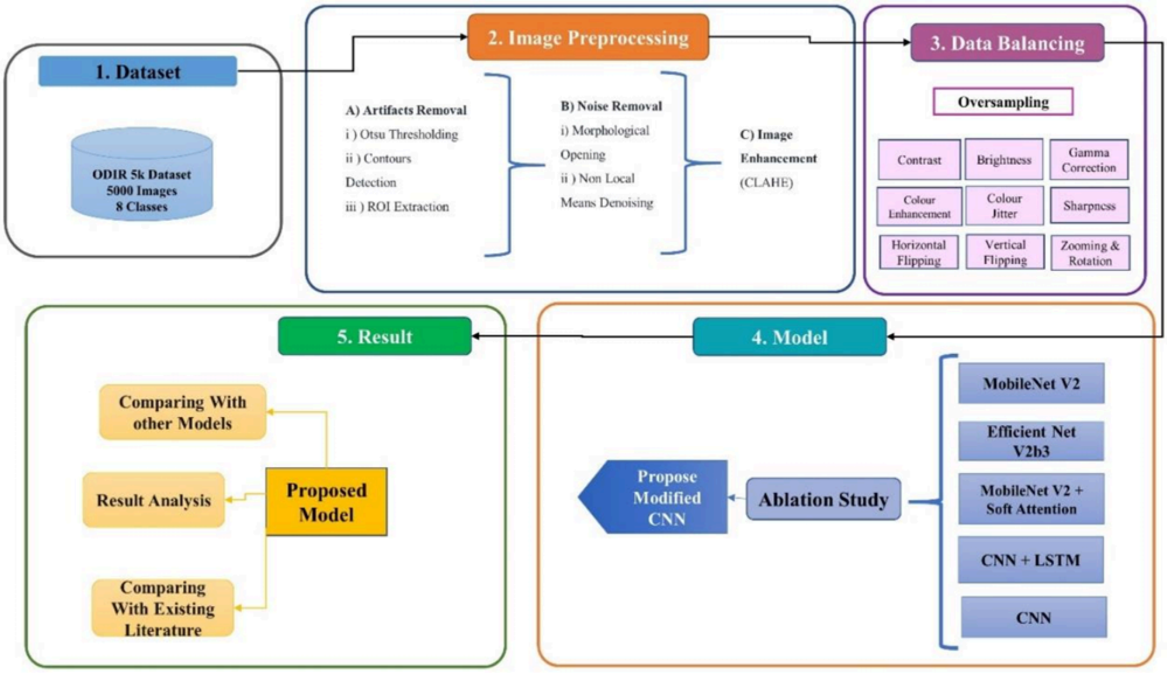


Figure 3.1: The entire process for predicting eye diseases.

A publicly accessible ODIR-5K dataset made up of eight classes and 5000 patients was obtained from Kaggle in order to assess the complete methodology. Then, several image preprocessing methods, such as artifact removal, noise reduction, and image enhancement, are applied to boost the image quality. The artifacts were removed using Otsu removal, contours detection, and ROI extraction. CLAHE has been used to improve image quality, while morphological opening and non-local means denoising have been used to reduce picture noise. Oversampling is done after picture preprocessing to balance the dataset. Many augmentation techniques, including contrast, brightness, gamma correction, color enhancement, color jitter, sharpness, horizontal and vertical flipping, zooming, and rotation are used in oversampling. The enhanced data set is sent to the models when data augmentation has been completed. Several innovative models are used in this work, including CNN, MobileNet V2, EfficientNet V2b3, and MobileNet V2+ soft attention CNN+LSTM. An ablation study is then conducted on the basic model to create a more robust model based on the best performance. The updated CNN model I suggest is the robust model. To demonstrate how well the suggested model works, several outcome analyses have been conducted. The parts that follow will go into depth about each step of the technique.

Image Preprocessing One of the key steps in getting good accuracy is image preprocessing, which is done before using images as neural network input. It entails a number of steps, including the removal of artifacts, the elimination of unwanted noise, and the enhancement of unclear but significant objects. precise and High-resolution color retinal images are necessary for the quick classification of eye images. Publicly available retinal fundus image datasets may contain background noise due to variations in resolution and compression methods employed during their creation. The classification of eye images poses a challenge due to the neural network model's reliance on enhanced, clean, and moderately symmetrical data. As such, preprocessing techniques are often necessary to facilitate accurate classification. The initial phase of image preprocessing in this study involved a thorough examination of the eye fundus images. Subsequently, to eliminate artifacts, first Otsu thresholding, then contour detection, and then ROI (region of interest) extraction was employed. The result of this step is utilized as the input for the subsequent step that is responsible for eliminating noise. As a part of the noise elimination procedure, the ROI (Region of Interest) picture undergoes morphological opening after its conversion into a binary image. The utilization of the bitwise AND function in conjunction with a binary mask facilitates the conversion of a binary image to an RGB image. Subsequently, the image is subjected to denoising through the utilization of the fastNlMeansDenoisingColored function. The image devoid of noise is subjected to a YUV transformation, wherein solely the Y channel undergoes CLAHE application. The image that has been improved is subsequently resized to a 512 by 512-pixel image. Figure 9 illustrates the preprocessing techniques utilized in the present investigation.

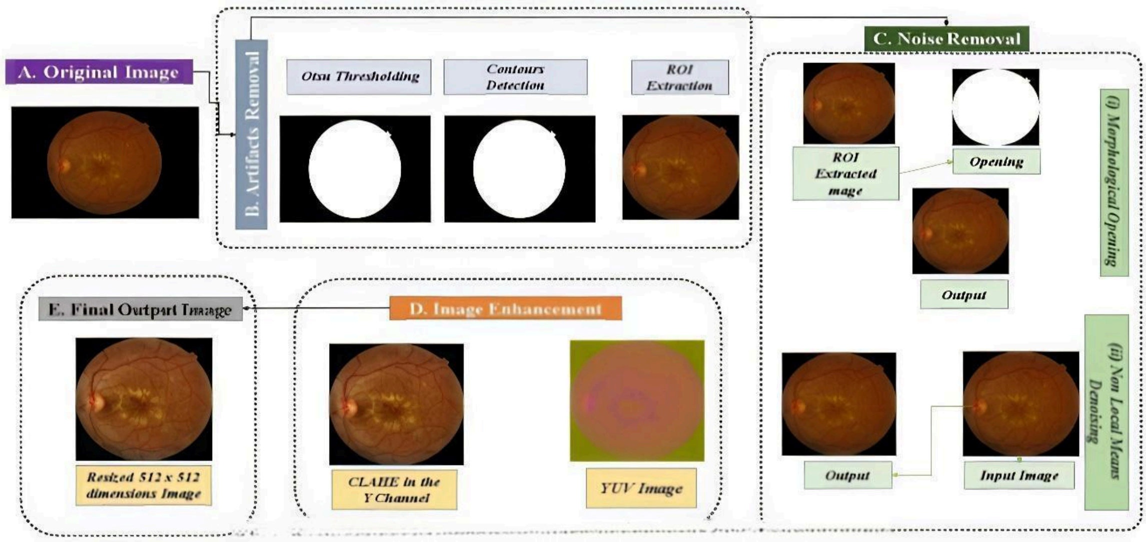


Fig 3.2: Whole preprocessing methods of fundus images.

Removal of artifacts

Unwanted areas or objects may unintentionally appear in the images. For the classification of ocular disease, dominant artifacts must be eliminated [25]. Artifacts like a dark background in our collection of retinal fundus pictures are not necessary for the classification job. Methods such as region Otsu thresholding technique, contour detection technique and sorting, locating boundary boxes, and Otsu thresholding are utilized.

i) Otsu thresholding

The Otsu method uses image thresholding as a technique to separate related data. The ideal global threshold value is determined by the Otsu method using an image histogram. The present study involves the utilization of the Otsu method to threshold retinal fundus images, with the aim of effectively separating the background from the region of interest (ROI). The process of converting a grayscale image into a binary image is achieved through a nonlinear method. Typically, the algorithm takes a grayscale image as input and uses the pixel intensity of the input image to create a binary image. The associated output pixel will be white if the intensity of a pixel exceeds the threshold. The output pixel will have a value of zero or black if the intensity of an input pixel is equal to or lower than the threshold value. The determination of the threshold value is possible. The equation utilizes the symbol "T" to represent the threshold value, while the numerical values "1" and "2" correspond to the mean intensity. The method cv2.threshold() is utilized to implement the binarization with the Otsu procedure in our research. As source parameter grayscale retinal fundus image is acknowledged, and the implementation of cv2.THRESH_OTSU is recognized as an auxiliary flag denoting the commencement of Otsu's method. The Otsu algorithm is employed by cv2.THRESH_OTSU parameter to select the optimal threshold level, while the thresholding technique cv2.THRESH_BINARY ascertains the pixel intensity. The threshold mask generated by the Otsu method is depicted in the uppermost figure.

ii) Detection and Sorting of Contours

Finding and extracting curves that correspond to the shapes of objects visible in pictures is known as contour detection. A contour refers to a delineation that represents the configuration or structure of an entity. The binary image obtained through Otsu thresholding technique serves as the input image for the cv2.findContours() function, which enables the identification of contours in the retinal fundus image. After the identification of the contours, a mathematical function is employed to arrange them in descending order based on their respective areas. Two arguments, namely the contour list and the region determined by cv2.contourArea, are passed to the function.

iii) **Extraction of Regions of Interest (ROI)**

The retinal fundus image's target area, known as the region of interest area, can be used to categorize diabetic retinopathy. I employ the cv2.boundingRect() method to divide this area. The input is a list of sorted contours. Numbers matching x, y, w, and h are returned by the cv2.boundingRect() function, respectively. The width, height, x, and y coordinates are all represented by these values. These pixel coordinates can be used to crop the area of interest. Figure xB(iii) shows how to extract the ROI portion and remove the extra black background padding.

Noise Removal

Fundus images typically have noise issues, and low contrast can also cause problems. Due to these issues, it is difficult to recognize and understand diseases from retinal fundus images. I use non-local means denoising after morphological opening to remove the noise from the images in the dataset. The steps for noise reduction are shown in Fig. 3.2 C.

i) Opening Morphologically

Morphological opening is utilized to smooth bright lesions and optic discs. Furthermore, it has the potential to aid in the identification of exudates and microaneurysms. Morphological opening is used to the extracted region of interest (ROI) as depicted in the acquired image. Prior to the application of morphological opening (as shown in Fig. 3.2 C(i)), the cv2.threshold function is utilized to binarize the image. The binary image undergoes morphological opening through the application of a kernel. Among the various kernel sizes that were tested, it was observed that the kernel size of (10, 10) yielded the most optimal outcomes. The process of acquiring a mask through the morphological opening technique involves the provision of cv2.MORPH_OPEN and kernel size as arguments to the cv2.morphologyEx method.

The ROI (Region of Interest) image is integrated with the mask through the utilization of the bitwise AND function, resulting in the production of a retinal fundus image that shows less noise. Figure 3.2 C(i) provides a concise overview of the methodology employed in the morphological opening technique. Contour detection is a process that involves identifying and extracting the curves that match the shapes of objects present in images. A contour refers to a delineation that represents the configuration or structure of an entity. The binary image obtained through Otsu's thresholding technique serves as the input image for the cv2.findContours() function, which enables the detection of contours in the retinal fundus image. After the identification of the contours, a mathematical function is employed to arrange the contours in descending order based on their respective areas. The function receives two arguments: the contour list and the area discovered by cv2.contourArea.

ii) Non-Local Means Denoise (NLMD)

The NLMD algorithm's basic concept is to substitute the average color of a pixel's neighbors for the original pixel's color. This leads to significantly less loss of image detail and improved post filtering clarity when compared to local mean approaches. NLMD reduces the noise in the pictures. To denoise an image with dimensions z = (z1; z2; z3) in channel I to pixel j, follow these steps.

$$\widehat{z_i}(x) = \frac{1}{C(x)} \sum_{k \in B(x,r)} \quad z_i(x)\omega(x,k), \qquad (1)$$

$$C(x) = \sum_{k \in B(x,r)} \quad \omega(x,k) \qquad (2)$$

The area of pixels x inside a radius of r is indicated here by $B\ (x,r)$. The squared Frobenius average distance between the centers of color patches at x and k that deteriorate under a Gaussian kernel is used to calculate the weight $\omega(x,k)$.

Image Enhancement

Image enhancement is the process of modifying digital images to produce output that is more suited for display or further analysis. One approach to enhancing the visibility of noteworthy features within an image involves the implementation of noise reduction techniques, as well as adjustments to the image's sharpness and brightness levels. The Contrast Limited Adaptive Histogram Equalization (CLAHE) technique was applied to enhance the visual quality of the images.

I. Adaptive histogram equalization with contrast limitation (CLAHE)

The utilization of CLAHE is employed to rectify the over-amplification of contrast and reinstate an equilibrium of general contrast. A modified form of adaptive histogram equalization is called contrast limited adaptive histogram equalization (CLAHE). that effectively mitigates the issue of noise amplification by constraining the amplification of contrast. The contrast enhancement in the vicinity of a particular pixel value is determined by the slope of the transformation function in CLAHE.

Data Augmentation

To ensure a more balanced dataset for our inquiry, data augmentation was employed. There are several ways to prevent overfitting issues, however data augmentation is considered an important responsibility. It is usual practice to increase variation and decrease the likelihood of overfitting by using techniques for augmentation-based oversampling. The class frequency of the datasets should be balanced, and samples should be produced that are representative of all possible photos, so it is necessary to successfully increase the amount of data. The preservation of visual details is another obstacle, especially in the medical area where such details may be vital. Consequently, the plan for augmentation must generate the images while preserving their quality.

I employ oversampling approaches in our study. Oversampling is a method used in image data balancing to manage the issue of underrepresented image classes in imbalanced datasets. In oversampling, the number of images in the underrepresented class is increased through the duplication of existing images or the creation of new images through the application of various image transformation techniques. This ensures that the model is trained on an equal number of images from each class, thereby enhancing the model's overall performance and precision. However, oversampling can occasionally result in overfitting, in which the model becomes overly specific to the training data and lacks the ability to generalize well to new data. To generate additional training samples by modifying existing data to increase the size of a dataset artificially. Here are brief explanations of several prevalent data augmentation techniques I used:

- ❖ Contrast: Adjusting the contrast of an image to make its features more prominent.
- ❖ Brightness: Brightness is the process of adjusting the overall lighting of an image by increasing or decreasing its brightness.
- ❖ Gamma Correction: Gamma Correction is the process of adjusting the gamma value of an image to enhance its overall quality.
- ❖ Color Enhancement: The process of increasing or decreasing the saturation of colors to make them more prominent or subdued.
- ❖ Color Jitter: Color Jitter is the process of altering the hue or value of an image to generate variations in color.
- ❖ Sharpness: The process of making an image appear sharp.

- ❖ Horizontal flipping: Horizontal flipping is the horizontal mirroring of an image to create a new sample with the same characteristics.
- ❖ Vertical Flipping: The process of mirroring an image vertically to create a new image with identical characteristics.
- ❖ Cropping and resizing an image to simulate a zoom-in or zoom-out effect.
- ❖ Rotation is the process of rotating an image to generate different perspectives and orientations of the same object.

Table 3.2: Class frequency details of original and augmented datasets

Dataset Classes	**Images without augmentation**	**Images with augmentation**
Diabetes (D),	1609	2873
Glaucoma (G)	284	2873
Cataract (C)	293	2873
Age-Related Macular Degeneration (A)	266	2873
Hypertension (H)	128	2873
Normal(N)	2873	2873
Pathological Myopia (M)	244	2873
Other disorders/diseases (O)	982	2873
Total Images	5699	22984

The imbalance in the dataset was already pointed out in the section where I described the dataset. With 128 images, Hypertension (H) has the fewest, while Normal (N) has the most, with 2873 images. To perform the oversampling, I increased the number of photos from each image class to 2873, which is equal to the maximum number of photographs for each image class.

Method and Proposed Model

In this study, five alternative models were tried to discover the most effective one based on accuracy before arriving on the best transfer learning model for the classification task.

i) EfficientNet v2b3

EfficientNetv2b3 is a family of convolutional neural networks introduced in 2019 by a team of Google researchers. EfficientNetv2b3 models are intended to achieve innovative accuracy on image classification tasks while remaining computationally efficient. The EfficientNetv2b3 architecture is founded on a method of compound scaling that scales all network dimensions uniformly, including depth, width, and resolution.

With only 4,000,000 parameters, EfficientNetv2b3 is the most compact member of the EfficientNet family. EfficientNetv2b3 achieves state of the art accuracy on the ImageNet dataset, a benchmark dataset for image classification tasks, despite its small size. EfficientNetv2b3 is optimal for use cases with limited computational resources, such as mobile devices and embedded systems.

EfficientNetv2b3 can achieve high accuracy with fewer parameters than other models, which is one of its primary advantages. This makes it a desirable option for use cases with limited memory and processing power. In addition, EfficientNetv2b3 can be fine-tuned on other datasets with relatively few examples, making it a flexible architecture for a variety of computer vision applications.

ii) MobileNetV2

Google introduced the MobileNetV2 convolutional neural network architecture in 2018. It is designed to be lightweight and effective, making it suitable for use in resource-constrained circumstances, such as on portable devices. MobileNetV2 achieves state of the art accuracy on tasks that classify images despite being significantly smaller and quicker than competing models.

MobileNetV2's architecture is a combination of depth-separable convolutions and linear bottleneck layers. Depth wise separable convolutions are a type of convolutional layer that divides the conventional convolution into two sperate layers: a pointwise convolution and a depth wise convolution. This reduces the number of required parameters and computations without sacrificing precision. Layers of linear bottlenecks are utilized to further reduce the number of model parameters. MobileNet V2 is highly optimized for mobile devices, utilizing low-precision arithmetic and other techniques to support hardware acceleration. This enables it to operate efficiently on a vast array of devices, from low-end smartphones to high-end embedded systems. MobileNet V2 is an excellent option for applications with limited computational resources that require high accuracy on image classification tasks. It is utilized extensively in mobile applications, autonomous vehicles, and other edge computing applications where performance and efficiency are crucial.

iii) MobileNetV2 + Soft Attention

MobileNetV2+SoftAttention is an extension of the well-known MobileNetV2 convolutional neural network architecture that adds a soft attention mechanism. Soft attention is a technique that enables the network to selectively focus on regions of an image, thereby enhancing its capacity to recognize objects and features.

MobileNetV2+SoftAttention's architecture is comparable to that of MobileNetV2, with the addition of a soft attention module. The soft attention module consists of a collection of convolutional layers used to compute attention maps for each feature map in the network. These attention maps are then used to weight the feature maps, enabling the network to selectively focus on image regions that are most pertinent to the current task.

Empirical evidence has shown that the utilization of MobileNetV2 in conjunction with Soft Attention has resulted in achieving the highest level of ability to perform on a variety of computer vision tasks, such as object detection, image classification, and semantic segmentation. The technique is notably efficacious in situations that entail images with intricate backgrounds or multiple objects, as it enables the network to concentrate on the image's most prominent areas.

In general, MobileNetV2+SoftAttention is an excellent option for image recognition applications that require high precision and efficiency. It is utilized extensively in mobile applications, autonomous vehicles, and other edge computing applications where performance and efficiency are crucial.

iv) CNN+LSTM

A complex deep learning architecture known as CNN+LSTM blends networks with Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs). This combined model is commonly used for sequential data analysis and has demonstrated exceptional performance in a variety of applications such as time series forecasting, natural language processing (NLP), and computer vision. The CNN component of the CNN+LSTM architecture oversees extracting spatial information from the input data. Through convolutional filters and pooling layers, CNNs excel at capturing local patterns and structures. This allows them to recognize notable features in photos, movies, and other multidimensional data.

In contrast, the LSTM component is intended to model sequential dependencies and capture temporal trends. LSTMs are a sort of recurrent neural network (RNN) that can process and recall data across lengthy sequences. They are especially effective for activities that need a comprehension of the context and temporal dynamics.

The CNN+LSTM design takes advantage of the strengths of both models by integrating them. The CNN component pulls spatial information and pertinent features from the input data, while the LSTM component processes sequential information and learns temporal connections. As a result, the model can study and comprehend complicated patterns in sequential data, making it suitable for applications such as video categorization, sentiment analysis, and speech recognition.

iv) **Base Convolutional Neural Network (CNN)**

The convolutional neural network (CNN) is a powerful deep learning algorithm that is frequently employed for image classification assignments, including the classification of eye images. The system is tailored to manage and evaluate visual information, rendering it exceptionally proficient for assignments related to computer vision.

CNN's design is informed by the hierarchical approach to information processing observed in the human visual system. The architecture consists of fully connected, pooling, and convolutional layers. The application of filters to the input image by the convolutional layers enables the capture of spatial relationships and the extraction of features. Pooling layers are utilized to decrease the dimensionality of the data and extract the most significant information by performing downsampling on the outcome of the convolutional layers. Finally, the FC (fully connected) layers employ the extracted features to perform classification.

The significance of CNN in the classification of eye images is attributed to its capacity to autonomously acquire and extract crucial features from the said images. The hierarchical architecture and robust feature extraction capabilities of CNN facilitate the recognition of intricate patterns and structures within eye images, leading to precise and dependable classification outcomes. In the realm of ophthalmology and medical diagnostics, timely identification of ocular ailments and disorders is imperative for optimal therapeutic intervention and patient management, rendering this aspect especially significant.

Through the utilization of a Convolutional Neural Network (CNN) model that has been trained on a substantial dataset of labeled ocular images, it is possible to acquire the ability to recognize a diverse range of ocular conditions like cataract, hypertension, macular degeneration and glaucoma. The utilization of a trained Convolutional Neural Network (CNN) model enables the classification of novel and unobserved eye images, which in turn speeds up diagnosing and monitoring ocular ailments. The utilization of deep learning techniques by CNN, coupled with its ability to manage intricate visual data, positions it as an essential tool for the classification of eye images and the progression of ophthalmic healthcare.

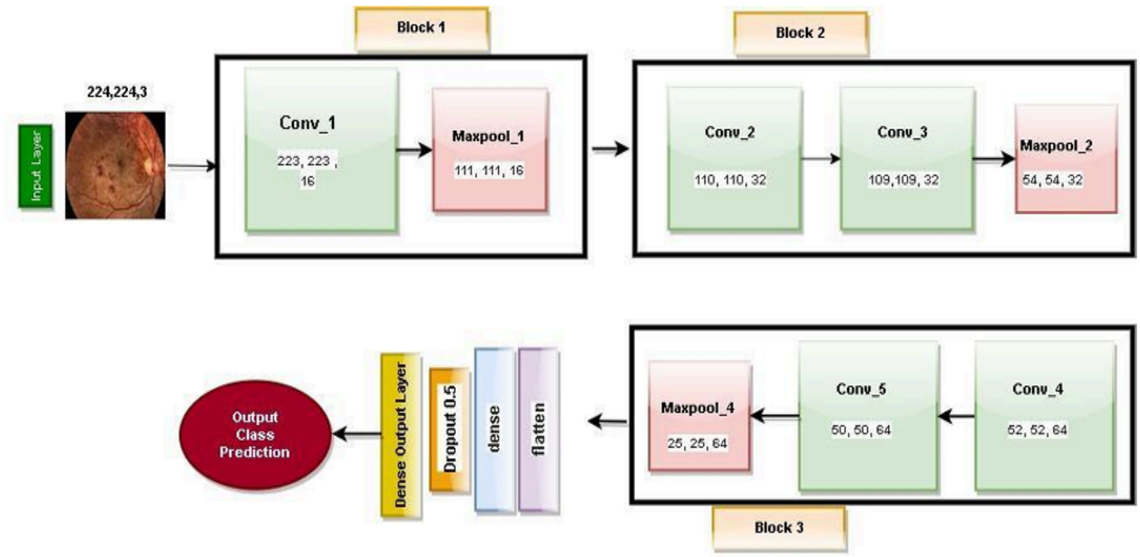


Fig 3.3: Base CNN model architecture

vi) Proposed Convolutional Neural Network

As the base CNN model performed well for its classification so I have developed a modified CNN classification by performing ablation studies and changing some layers. The architecture of base CNN. In the following figure I can see the architecture of modified CNN. at past in the input layer, I have taken the input image and there are four blocks in the modified CNN and in each block, I have one comb layer and one maxpool layer at the end I have flatten layer dense layer drop out layer and output class prediction so inthis way I have modified hour proposed the and model from the base model.

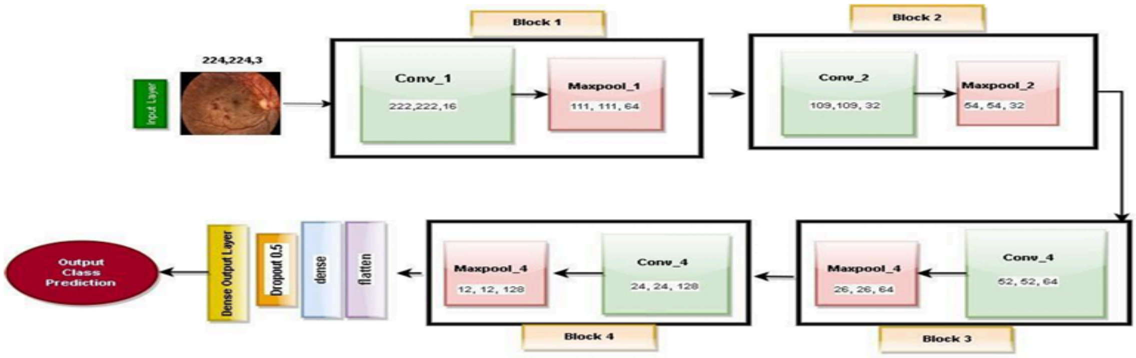


Fig 3.4: Proposed CNN model architecture

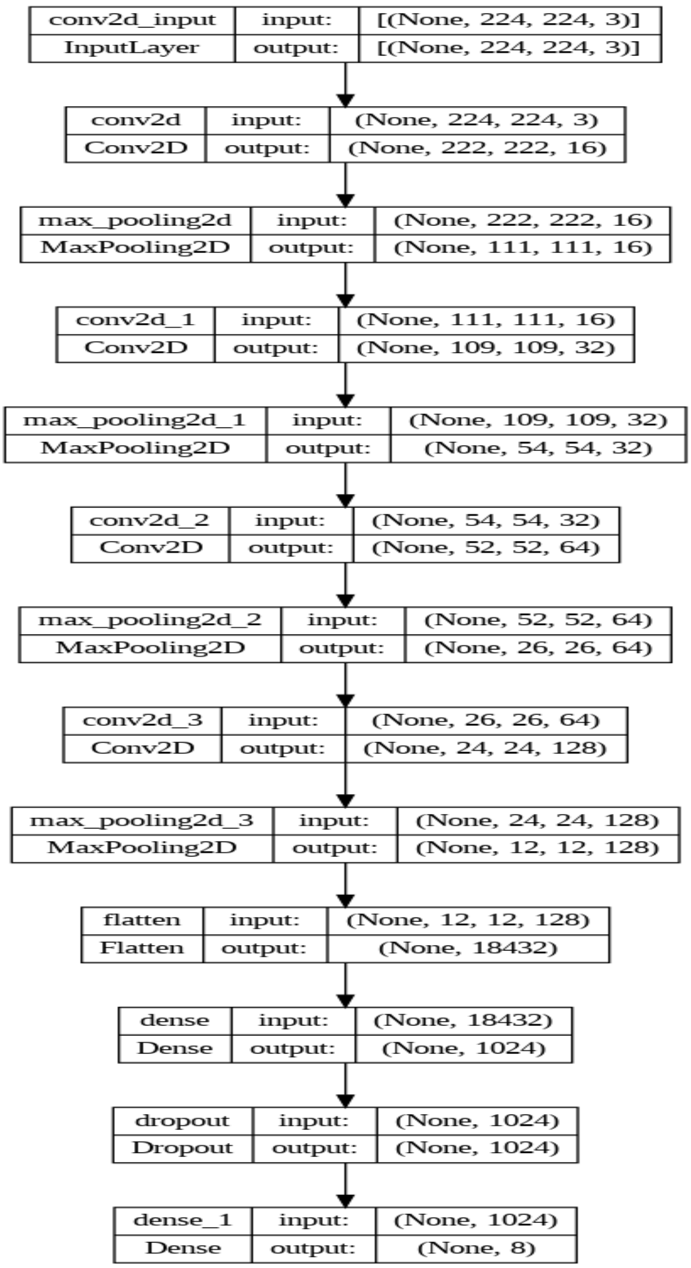


Fig 3.5: The Final Architecture of proposed model

Ablation Study

Adding a large number of concepts to a custom deep learning model usually increases the final model's efficiency. However, it is beneficial to understand the impacts of these enhancements independently in research. Researchers often evaluate their models without any features enabled so they can measure the impact of individual components and the overall performance loss. You can read all about the ablation study that was conducted to improve the model's ability to forecast eye diseases in the results section. I have made the following changes: updated the polling layer, batch size, loss function, optimizer, learning rate, and epochs.

Changing the Epochs

An epoch is a complete traversal of the entire training dataset. Changing the value of epochs affects the number of times the model views the training data. Increasing the number of epochs can improve convergence and accuracy, but it may also increase the danger of overfitting if the model becomes overly specific to the training data. Reducing the number of epochs may lead to underfitting and diminished performance.

Changing the Batch size

The batch size is the amount of training samples which are processed in a single iteration. Changing the batch size influences the training process's effectiveness and convergence. Larger batch sizes can speed up the training process since they leverage parallel computation, but they may demand more memory. Smaller batch sizes may lead to greater noise during training, but they can facilitate better generalization.

Changing the pooling layer

In a convolutional neural network (CNN), the pooling layer performs downsampling and decreases the dimension of the input data. By modifying the pooling layer, you can adjust the network's level of abstraction and the quantity of retained spatial information. Various pooling approaches, including maximum pooling and average pooling, can be utilized, and each has a unique effect on the network's ability to capture significant information.

Changing the loss function

The loss function quantifies the deviation between expected and actual values. Different loss functions are appropriate for various jobs. For instance, cross-entropy loss is widely used in classification problems, whereas mean squared error is utilized in regression jobs. Altering the loss function can influence the training procedure and the model's capacity to optimize and generalize.

Changing the Learning rate

The rate of learning influences the size of each gradient descent optimization step. It determines how much the parameters of the model are modified after each iteration. A high learning rate can expedite convergence, but it can also create instability and overshoot. A poor learning rate, on the other hand, can result in sluggish convergence or getting stuck in local optima. Adjusting the learning rate is essential for achieving a balance between convergence speed and optimization stability.

Changing the optimizer

During training, the optimizer determines how the CNN modifies its parameters depending on the computed gradients. Stochastic gradient descent (SGD), Adam, and RMSprop are common optimizers. Each optimizer has its own update algorithms and hyperparameters, which can influence the convergence rate, model stability, and performance.

Conclusion

The batch size, pooling layer, loss function, learning rate, optimizer, and epochs all play significant roles in the training and optimization of a CNN. Adjusting these factors can have a substantial effect on the model's performance, convergence rate, and generalizability. Even though the learners' ability to generalize decreases as the number of learners in a batch increase, in this model, the number of learners in a batch goes down. In this neural network the more the weights are modified, the more the curve shifts from underfitting to optimal to overfitting as the number of epochs increases. As far as I am concerned, the more a model is over-fitted, the higher its F1 score. Our main goal was to get the F1 score to go up. I used hyper-tuning in our model because of this.

3.5 Implementation Requirements

A number of particular conditions must be met for the suggested methodology for automated eye disease diagnosis to be implemented successfully. Firstly, a robust computational infrastructure with sufficient processing power and memory is essential to accommodate the training and evaluation of deep learning models. A programming environment, preferably using frameworks like TensorFlow is required for seamless model development and training. Access to the ODIR-5K dataset, obtained from Kaggle, is crucial for obtaining diverse and representative eye images for training and testing purposes. Additionally, the implementation requires image processing libraries such as OpenCV for artifact removal, noise reduction, and enhancement. The oversampling and augmentation procedures involve careful handling of the dataset and the incorporation of techniques like contrast adjustment, brightness variation, and geometric transformations. Evaluation metrics, including accuracy, precision, and kappa score, are necessary for assessing the performance of the models. The implementation process should be well-documented, and version control tools can be employed to track changes and facilitate collaboration among researchers and developers. Overall, a well-equipped computational environment, access to relevant datasets, and adherence to best practices in deep learning development are crucial for the successful implementation of the proposed automated eye disease diagnosis methodology.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup

Technical accuracy, precision, and F1 scores are some of the confusing metrics used in the examination. TP values are correct. The mislabeling of incorrect results causes false positives (FP). False negatives (FNs) are the third type, and they happen when a positive number is mistakenly thought to be negative. The fourth and fifth options are TN and FN, respectively. A positive value that is mistakenly thought to be negative is called a true negative (TN). The fourth one is TN, or true negative.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$\text{F1 Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{6}$$

Some more evaluation metrics are false positive rate (FPR), false negative rate (FNR), negative predicted values (NPV), mean absolute error (MAE), root mean square error (RMSE).

4.2 Experimental Results & Analysis

The findings of this paper will be reviewed in this section. The preprocessed and balanced ODIR 5k dataset is evaluated with the following models. Table 4.1 shows the Recall, Precision, FPR, FNR, NPV, FDR, Specificity, F1, and Test Accuracy of the applied and proposed model.

Table 4.1: Performance metrics of various models.

Model	Rec (%)	Spe (%)	Pre (%)	FPR (%)	FNR (%)	NPV (%)	FDR (%)	F1 (%)	Test Accuracy (%)
EfficientNet v2b3	76.90	96.70	77.18	3.29	23.09	96.70	22.81	77.04	77.10
MobileNet v2	76.60	97.08	78.9	2.91	20.4	97.15	21.08	79.25	80.10
MobileNet v2 + soft attention	87.98	98.28	88.02	1.71	12.02	98.28	11.98	88.00	88.01
CNN+LSTM	90.36	98.62	89.98	1.37	9.63	98.64	10.02	90.17	90.12
Modified CNN	**92.10**	**98.87**	**91.89**	**1.12**	**7.89**	**98.87**	**8.104**	**91.99**	**92.42**

From the table, I can see that the accuracy of the models is EfficientNet v2b3 achieved 77.10%, MobileNet v2 achieved 80.10%, MobileNet v2 + soft attention achieved 88.01% and CNN+LSTM 90.12%. The highest recall, specificity, precision, FPR, FNR, NPV, FDR, F1, and Test Accuracy are achieved of 92.10%, 98.87%, 91.89%, 1.12%, 7.89%, 98.87%, 8.104%, 91.99% and 92.42% respectively by the proposed modified CNN model.

Result of the Ablation Study

Several design components can be altered to improve classification accuracy. The base CNN architecture is modified in some experiments that are conducted as ablation research. The changes I have been making are Changing polling Layer, Changing the batch size, Changing the Loss Function, Changing the Optimizer, Changing the Learning Rate, Changing Epochs. Following an ablation study, table 4.2 presents the suggested model's final configuration.

Table 4.2: Configuration of the proposed model

Configuration	**Value**
Image sizes	224 x 224
Epochs	100
Optimization Functions	Adam
Learning rates	0.001
Batch sizes	32
Activation functions	Softmax
Dropouts	0.5
Momentum	0.5
Accuracy	92.42%

Table 4.2 displays the suggested model's final configuration. The image is 224 x 224, the epoch is 200, the optimization functions are Adam, the learning rate is 0.001, the batch sizes are 32, the activation function is Softmax, the dropout is 0.5, and the momentum is 0.9. Lastly, 93.28% accuracy was attained with this setup.

Confusion Metrics and Accuracy curve and loss curve

Confusion metrics are tables that show how well a classification model performed on test data where the true values were already known. It is possible to display the results of algorithms graphically. With N being the total number of target classes, the confusion metrics are N x N. Every metric item represents the number of observations that are known to belong to one class but are anticipated to belong to another. Confusion Metrics summarizes how well a categorization method performed. Within the cells of Metrics, you can see the success rate of the algorithm.

In the rows you'll find actual classes, and in the columns, you'll find predicted ones. True positives, or observations that were correctly identified as members of the target class, are in the top left cell of the Metrics table. In the top right cell, you can see the number of false positives, which are observations that were incorrectly allocated to the target class. In the bottom left cell, you can see the false negatives, which are observations that were wrongly labeled as not belonging to the target class. In the bottom-right cell of the Metrics, you may see positive numbers and negative numbers. In the accompanying figures, you can see the models' confusion metrics, accuracy and loss curves.

a) Mobilenetv2+soft attention

For the Mobilenetv2+soft attention, the acquired validation accuracy is 0.87 and test accuracy is 0.88.

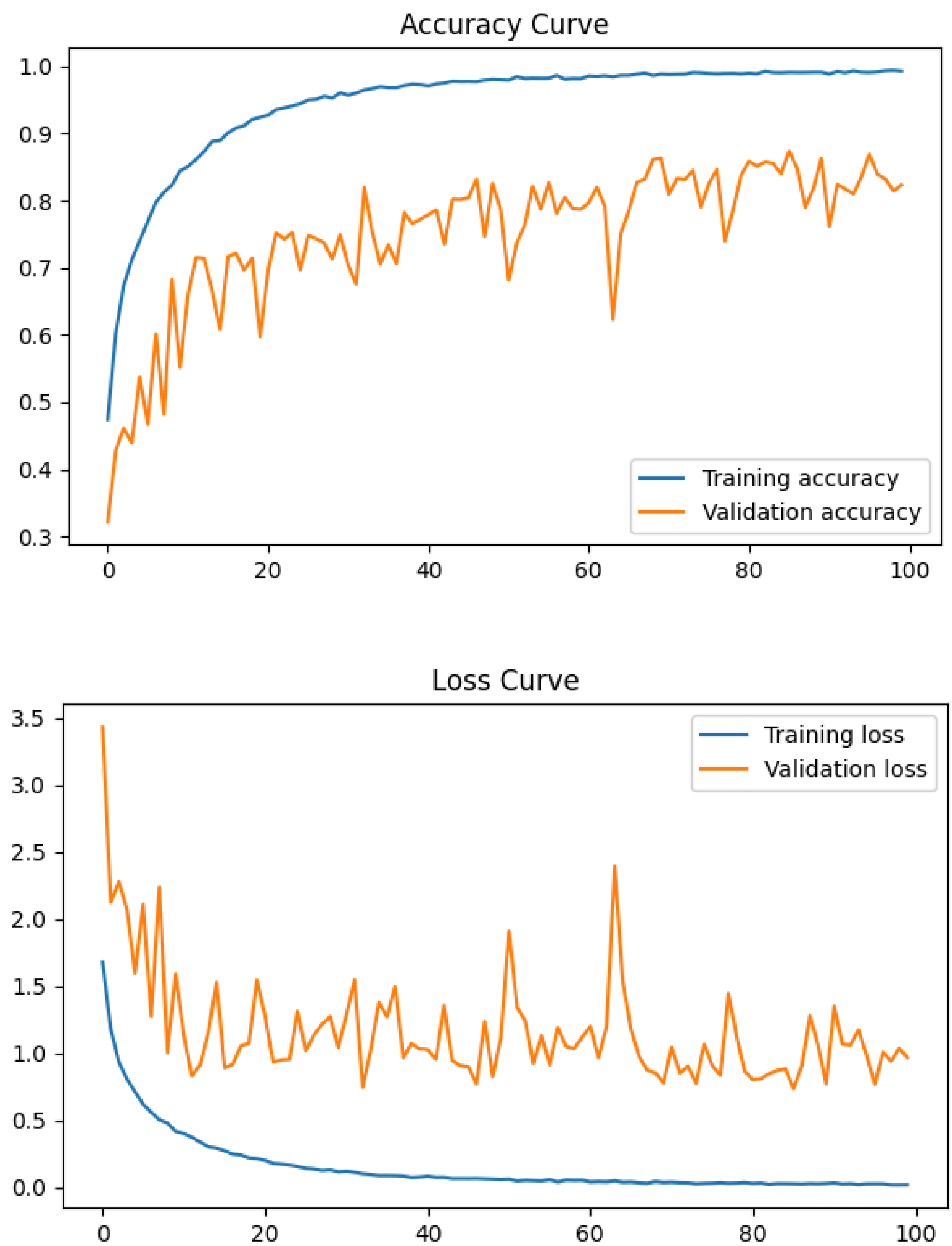


Fig 4.1. Loss and accuracy curve of the Mobilenetv2+soft attention model.

Figure 4.1. Showing the Loss and accuracy curve of the Mobilenetv2+soft attention model. The blue line in the graph indicates the training accuracy while training per epoch and the orange curve show the validation accuracy and validation loss per epoch.

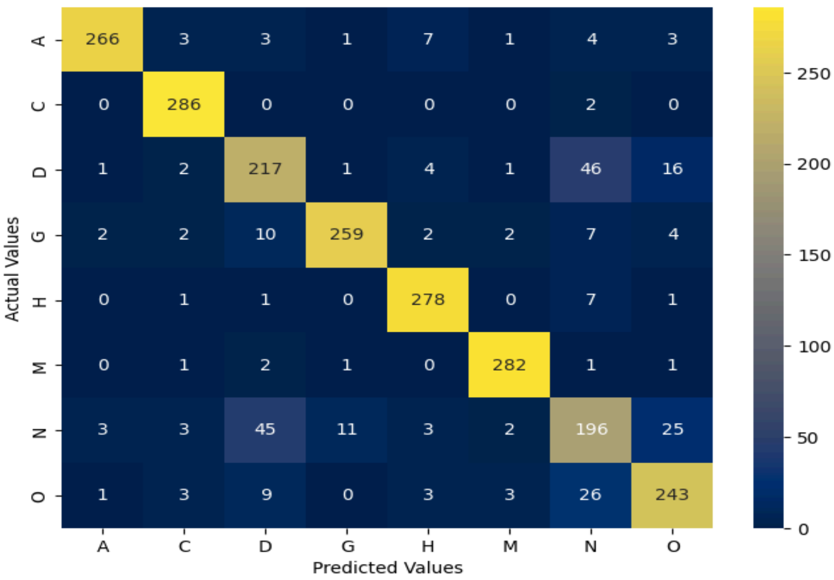


Fig 4.2. Confusion Metrics of the Mobilenetv2+soft attention model

	precision	recall	f1-score	support
A	0.97	0.92	0.95	288
C	0.95	0.99	0.97	288
D	0.76	0.75	0.75	288
G	0.95	0.90	0.92	288
H	0.94	0.97	0.95	288
M	0.97	0.98	0.97	288
N	0.68	0.68	0.68	288
O	0.83	0.84	0.84	288
accuracy			0.88	2304
macro avg	0.88	0.88	0.88	2304
weighted avg	0.88	0.88	0.88	2304

```
Rec=    87.97743055555556
Spe=    98.28249007936509
Pre=    88.02489907328041
FPR=    1.7175099206349207
FNR=    12.022569444444445
NPV=    98.28365299403004
FDR=    11.975100926719588
F1=     88.0011584131954
MAE   = 37.76041666666667
RMSE  = 121.90615698606496
```

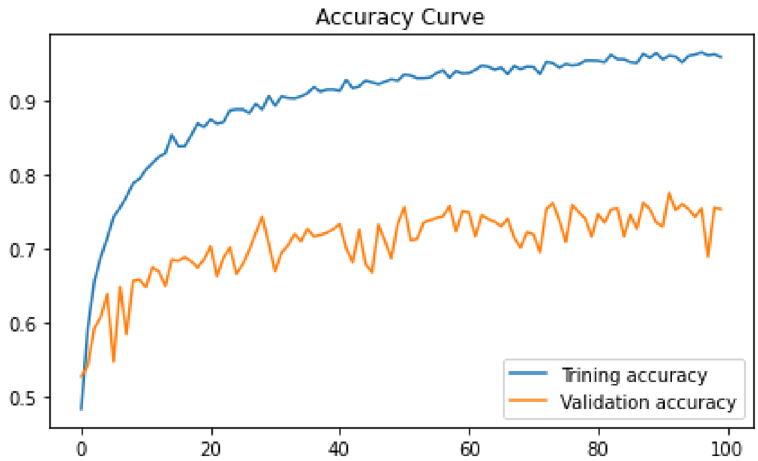
Fig 4.3. Statistical results of the Mobilenetv2+soft attention model

Figure 4.2. Showing the Confusion Metrics and figure 4.3 showing the achieved statistical results of Mobilenetv2+soft attention model. In the confusion matrix the diagonal yellow line shows that the true positive and negative are much higher than false positive and negative, which indicates that our model performs well in terms of class prediction. From the

classification report I got an overall overview of our developed Mobilenetv2+soft attention model by getting the value of F1-score, precision, and recall.

b) Mobilenetv2

For the Mobilenet v2, the acquired validation accuracy is 0.79142 and test accuracy is 0.80.



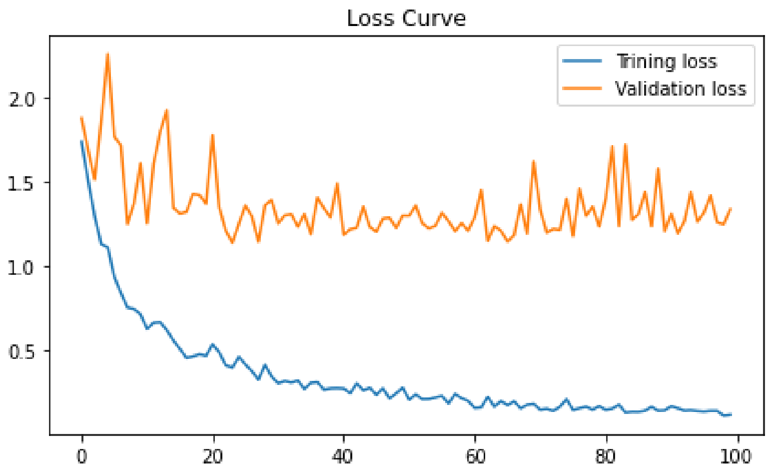


Fig 4.4. Loss and accuracy curve of the Mobilenetv2 model.

Figure 4.4. Showing the loss curve and accuracy curve of the Mobilenetv2 model. The blue line in the above graph indicates the training accuracy while training per epoch and the orange curve shows the validation accuracy and validation loss per epoch.

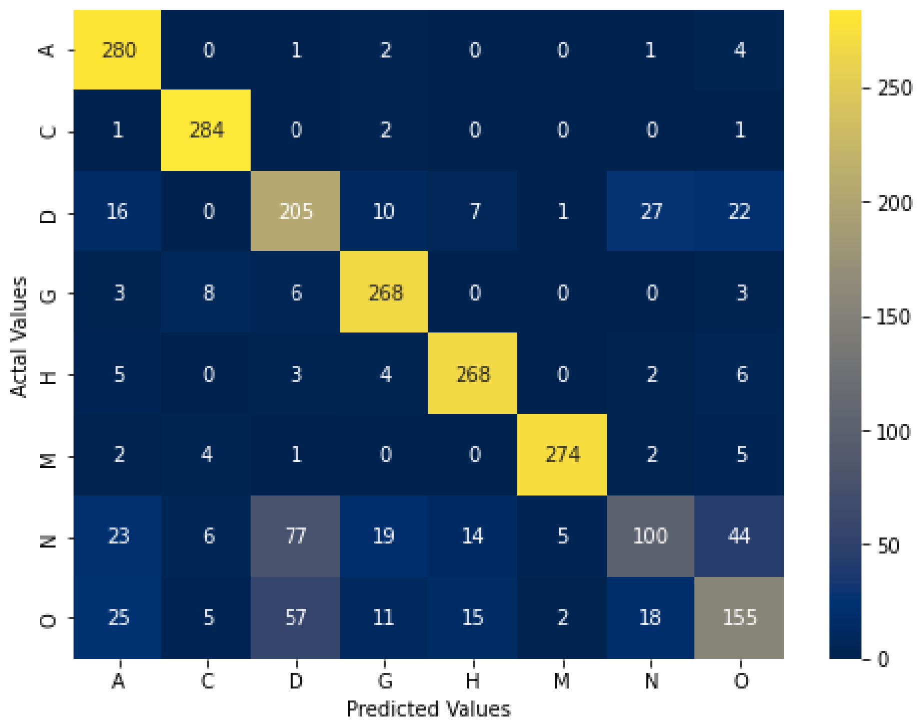


Fig 4.5. Confusion Metrics of the Mobilenet v2 model

	precision	recall	f1-score	support
A	0.79	0.97	0.87	288
C	0.93	0.99	0.95	288
D	0.59	0.71	0.64	288
G	0.85	0.93	0.89	288
H	0.88	0.93	0.91	288
M	0.97	0.95	0.96	288
N	0.67	0.35	0.46	288
O	0.65	0.54	0.59	288
accuracy			0.80	2304
macro avg	0.79	0.80	0.78	2304
weighted avg	0.79	0.80	0.78	2304

```
Rec=     79.60069444444444
Spe=     97.0858134920635
Pre=     78.91674415212754
FPR=     2.914186507936507 5
FNR=     20.399305555555557
NPV=     97.15892114530263
FDR=     21.083255847872458
F1=      79.2572437886938
MAE    = 72.78645833333334
RMSE   = 179.9522506110996
```

Fig 4.6. Statistical results of the Mobilenet v2 model

Figure 4.5 is showing the Confusion Metrics and figure 4.6 showing the achieved statistical results of Mobilenetv2. From the confusion matrix's yellow diagonal line, I can see that our developed mobilenetv2 model was able to classify correct class most of the time. And the classification report gives an overall overview by showing the recall, precision, and f1-score.

c) EfficientNet v2b3

For the EfficientNet v2b3, the acquired validation accuracy is 0.77439 and test accuracy is 0.77.

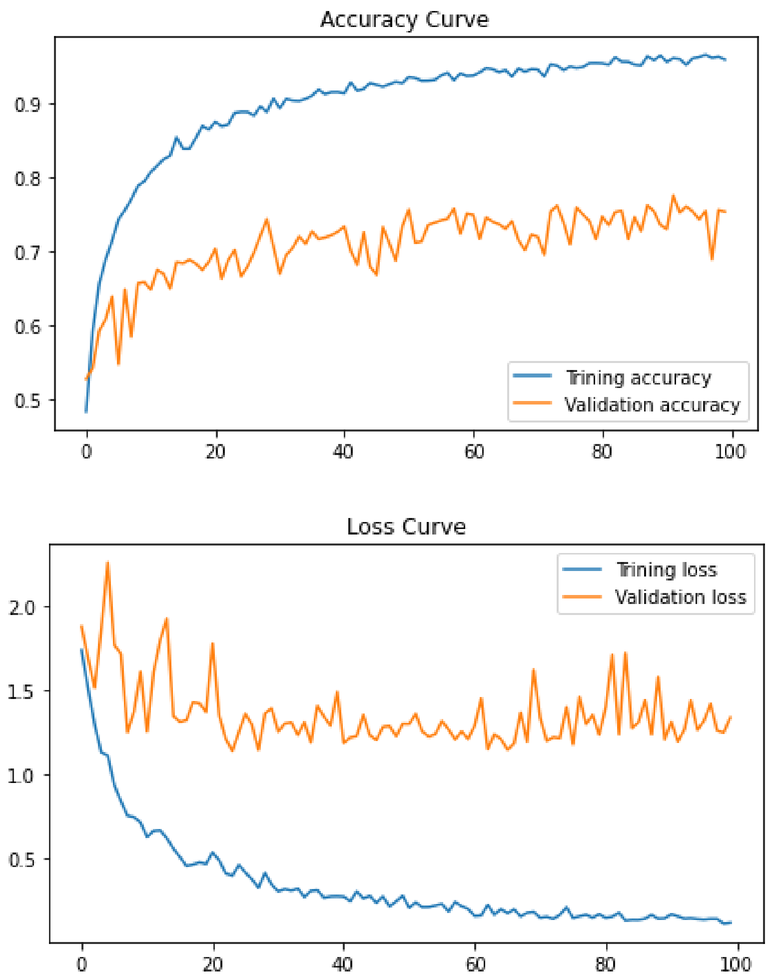


Fig 4.7. Loss and accuracy curve of the EfficientNet v2b3 model.

Figure 4.7. Showing the Loss and accuracy curve of the EfficientNet v2b3 model. In the accuracy curve and loss curve the blue line represents the training accuracy and training loss per epoch while training and the orange line shows the validation accuracy and validation loss per epoch which I can see fluctuates while training each epoch.

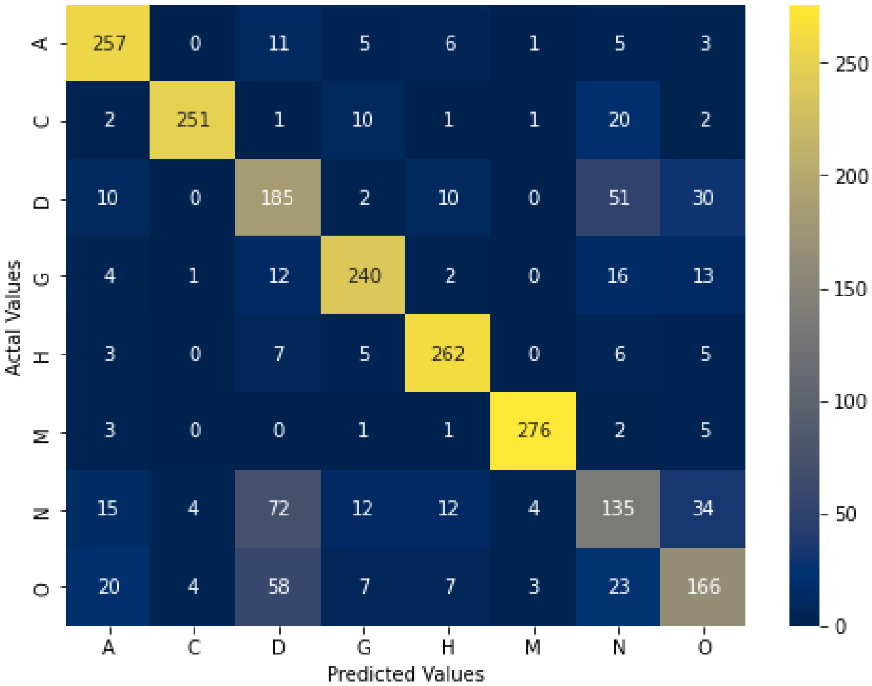


Fig 4.8. Confusion Metrics of the EfficientNet v2b3model

A	0.82	0.89	0.85	288
C	0.97	0.87	0.92	288
D	0.53	0.64	0.58	288
G	0.85	0.83	0.84	288
H	0.87	0.91	0.89	288
M	0.97	0.96	0.96	288
N	0.52	0.47	0.49	288
O	0.64	0.58	0.61	288
accuracy			0.77	2304
macro avg	0.77	0.77	0.77	2304
weighted avg	0.77	0.77	0.77	2304

```
Rec=     76.90972222222223
Spe=     96.70138888888889
Pre=     77.18901845827844
FPR=     3.2986111111111107
FNR=     23.09027777777778
NPV=     96.70805138079692
FDR=     22.81098154172155
F1=      77.04911723504281
MAE    = 82.421875
RMSE   = 189.262549121584
```

Fig 4.9. Statistical results of the EfficientNet v2b3 model

Figure 4.8 is showing the Confusion Metrics and figure 4.9 showing the achieved statistical results EfficientNet v2b3 model. The diagonal orange line in the confusion matrix shows that our model EfficientNet v2b3 performs well in terms of class prediction and classification matrix shows an overall overview of the model by getting the value of precision, recall and f1-score.

d) CNN+LSTM

For the CNN+LSTM, the acquired validation accuracy is 0.90 and test accuracy is 0.90.

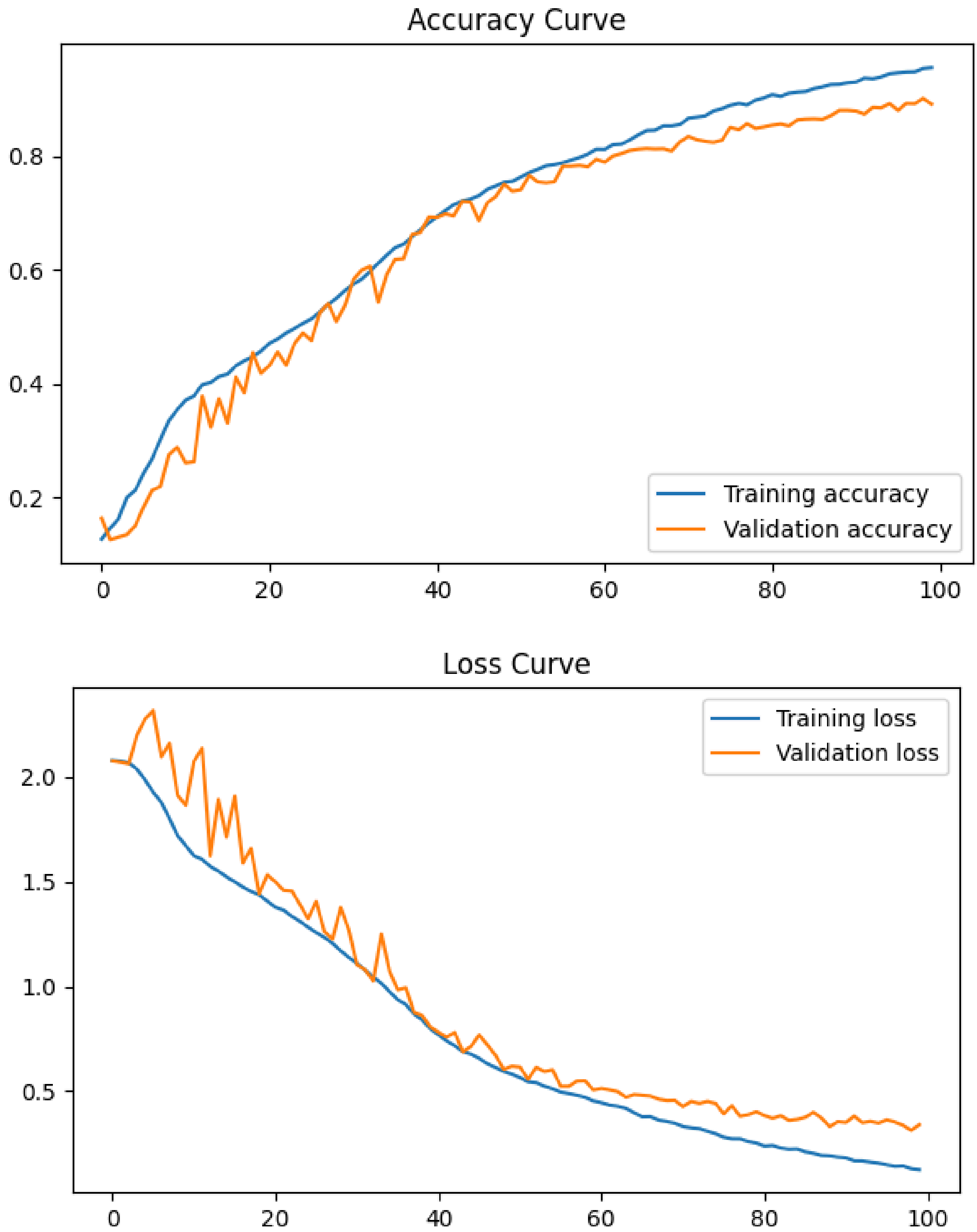


Fig 4.10. Loss and accuracy curve of the CNN+LSTM model.

Figure 4.10. Showing the Loss and accuracy curve of the CNN+LSTM model.

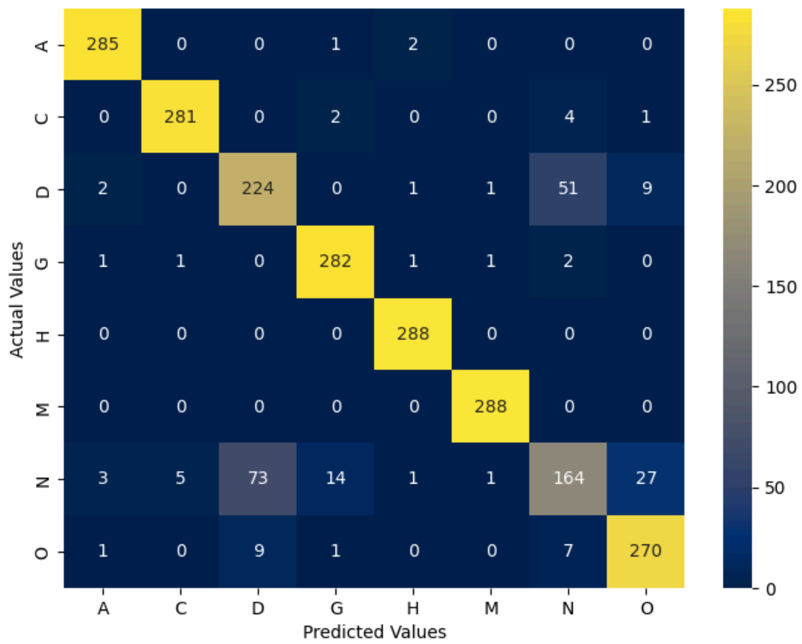


Fig 4.11. Confusion Metrics of the CNN+LSTM model

	precision	recall	f1-score	support
A	0.98	0.99	0.98	288
C	0.98	0.98	0.98	288
D	0.73	0.78	0.75	288
G	0.94	0.98	0.96	288
H	0.98	1.00	0.99	288
M	0.99	1.00	0.99	288
N	0.72	0.57	0.64	288
O	0.88	0.94	0.91	288
accuracy			0.90	2304
macro avg	0.90	0.90	0.90	2304
weighted avg	0.90	0.90	0.90	2304

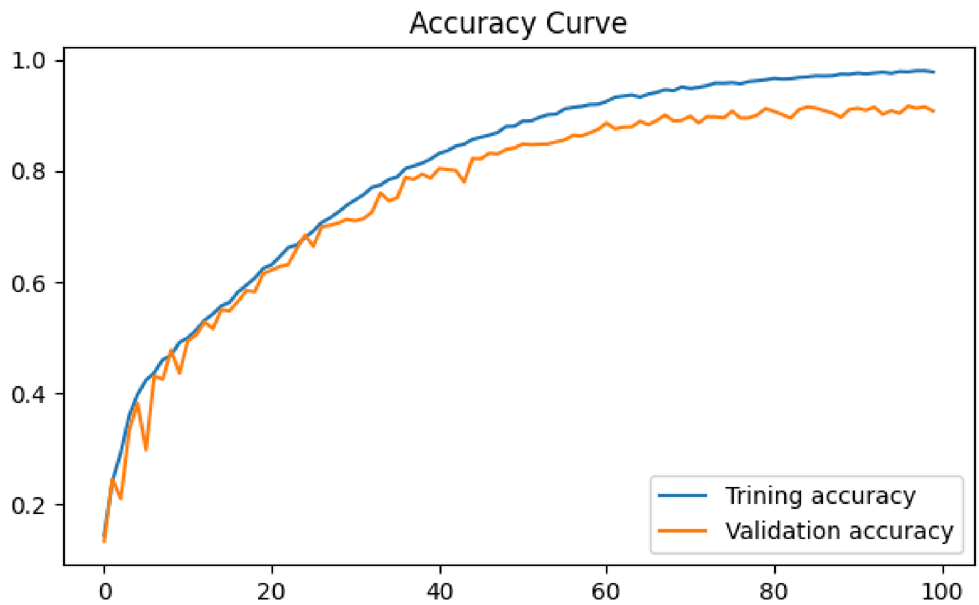
```
Rec=     90.36458333333334
Spe=     98.62351190476191
Pre=     89.98188207393596
FPR=     1.3764880952380951
FNR=     9.635416666666668
NPV=     98.6408578851024
FDR=     10.01811792606404
F1=      90.17282665116888
MAE    = 33.984375
RMSE   = 116.72245689859533
```

Fig 4.12. Statistical results of CNN+LSTM model Figure

4.11. showing the Confusion Metrics and figure 4.12 showing the achieved statistical results of CNN+LSTM model. From the confusion matrix I can see that our CNN+LSTM model was able to perform well in terms of class prediction and the classification reports result I got an overall overview of our CNN+LSTM model.

e) Modified CNN

For the Modified CNN, the acquired validation accuracy is 0.916 and test accuracy is 0.92.



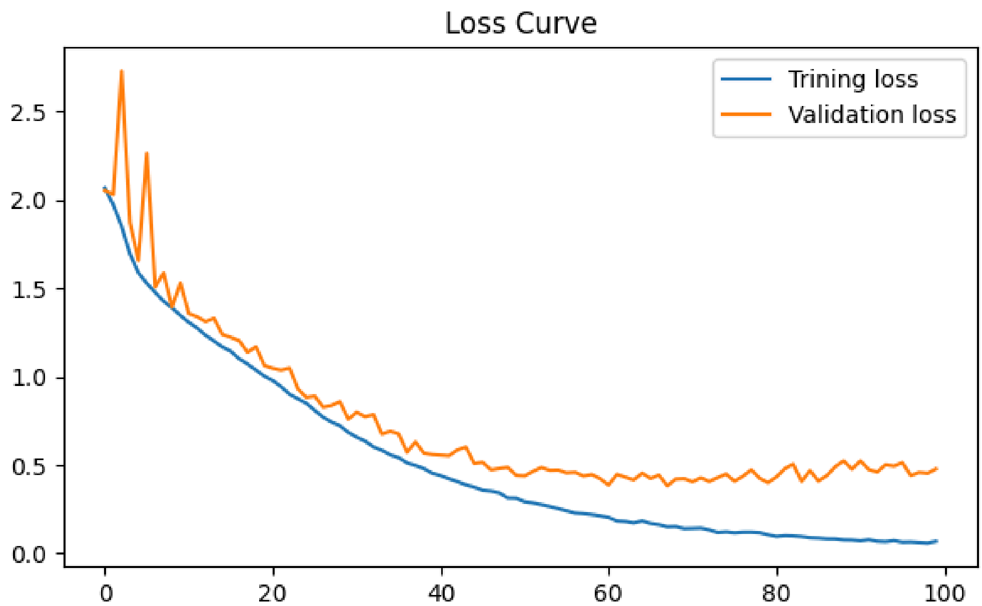


Fig 4.13. Loss and accuracy curve of the Modified CNN model.

Figure 4.13. Showing the Loss and accuracy curve of the modified CNN model. I can see that training accuracy and validation accuracy is getting higher while training per epoch and

becomes stable after 85 epochs and the training and validation loss becomes stable after a specific period of epoch training.

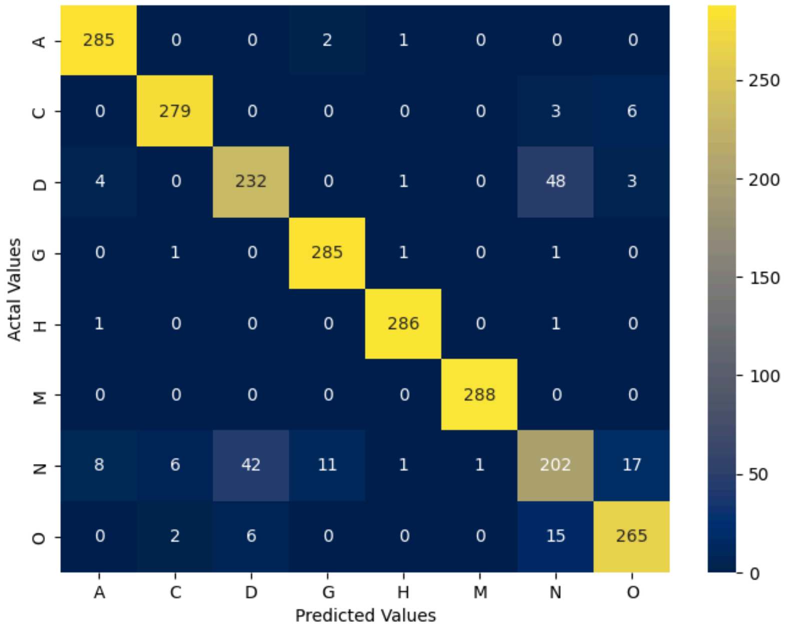


Fig 4.14. Confusion Metrics of the Modified CNN model

	precision	recall	f1-score	support
A	0.96	0.99	0.97	288
C	0.97	0.97	0.97	288
D	0.83	0.81	0.82	288
G	0.96	0.99	0.97	288
H	0.99	0.99	0.99	288
M	1.00	1.00	1.00	288
N	0.75	0.70	0.72	288
O	0.91	0.92	0.92	288
accuracy			0.92	2304
macro avg	0.92	0.92	0.92	2304
weighted avg	0.92	0.92	0.92	2304

```
Rec=     92.10069444444444
Spe=     98.87152777777779
Pre=     91.89526080591742
FPR=     1.128472222222222
FNR=     7.899305555555555
NPV=     98.87720878134733
FDR=     8.104739194082594
F1=      91.99786294064947
MAE    = 28.038194444444443
RMSE   = 108.05252066574951
```

Fig 4.15. Statistical results of Modified CNN model

Figure 4.14. showing the Confusion Metrics and figure 4.15 showing the achieved statistical results of the modified CNN model. From the diagonal blue line of the confusion matrix, I can see that our modified CNN model performs well in terms of classification report. And the classification report gives an overall overview of modified CNN model.

Table 4.3: Showing the F1, Kappa and AUC score of modified CNN.

Model	F1	KAPPA	AUC
CNN	92	91	99

4.3 Discussion

The table shows a Comparative analysis with previous work on ODIR dataset.

Table 4.4: Comparative analysis with Previous work on ODIR:

Model	AUC (%)	KAPPA (%)	Precision (%)	F1 (%)	Accuracy (%)
Resnet-101 [17]	93	63	-	91	-
Semi-supervised GAN [20]	84	81	-	88.16	87
DSRA-CNN [22]	-	86.17	88.50	88.16	87
(R-CNN+LSTM) +NCAR+SVM [24]	97	-	-	89.97	89.54
Our Proposed modified CNN	98	91	91.89	91.99	92

From table 4.4, comparing the other model mentioned in the table, I can clarify that our proposed modified CNN model gives better AUC, KAPPA, precision, f1-score and accuracy than most other models.

Table 4.5: Accuracy comparison between our developed models.

Model	**Recall**	**Precision**	**F1**	**Test Accuracy**
EfficientNet v2b3	76.90	77.18	77.04	77

MobileNet v2	79.60	78.9	79.25	80
MobileNet v2 + soft attention	87.98	88.02	88.00	88
CNN+LSTM	90.36	89.98	90.17	90
Modified CNN	**92.10**	**98.87**	**91.99**	**92**

From the above table, I can see that our developed modified CNN model performs better than the other model I developed. The modified CNN model provides better recall, precision, f1-score and test accuracy compared to the other model mentioned in the table.

CHAPTER 5

IMPACT ON SUSTAINABILITY, ENVIRONMENT, AND SOCIETY

5.1 Impact on Society

The repercussions of diabetes Loss of vision or blindness may occur because of diabetic retinopathy. Diabetics can significantly lower their risk of vision loss and enhance their quality of life by detecting and treating diabetic retinopathy early.

Early identification of diabetic retinopathy has benefits that go beyond the patient and can impact society positively. Preventing the social and economic toll that vision loss can take requires prompt diagnosis and treatment of diabetic retinopathy.

Loss of productivity and increased healthcare costs could occur, for example, when people with diabetic retinopathy need help with everyday tasks or are unable to work due to their condition. Because early therapy is often more effective and cheaper than treatment for advanced disease stages, early detection and treatment can also reduce the overall healthcare expenditures linked with diabetic retinopathy.

Ultimately, detecting diabetic retinopathy can greatly benefit both those living with diabetes and society at large.

5.2 Impact on Environment

Several beneficial environmental consequences may result from using computer vision to identify diabetic retinopathy at an early stage.

Patients, particularly those in underserved areas or those in more distant places, may find the diagnosis process easier and more pleasant if computer vision is used for early identification. By reducing transportation-related emissions, this can help patients avoid long diagnostic visit travel times, which is good for the environment.

In general, by making the diagnostic procedure more accessible and efficient and by doing away with the need for specific diagnostic gear and transportation, computer vision can have a positive impact on the environment when used for the early detection of diabetic retinopathy.

5.3 Ethical Aspects

The integration of computer vision in diabetic retinopathy screening introduces a spectrum of ethical considerations that demand careful attention to ensure responsible and equitable deployment of this technology. Safeguarding patient privacy is paramount, necessitating the implementation of robust encryption, anonymization techniques, and secure storage protocols to protect sensitive medical information. Informed consent becomes a crucial component, requiring transparent communication with patients about the purpose, potential benefits, and associated risks of using computer vision in screening. Mitigating biases in algorithms to ensure fairness across demographic groups is imperative, as is the need for transparent decision-making processes to build trust between healthcare providers and patients. Equitable access to screening must be prioritized, addressing disparities in technology access to prevent further healthcare inequalities. Continuous monitoring, professional oversight, and anticipation of unintended consequences are integral to a comprehensive ethical framework, allowing the responsible integration of computer vision in diabetic retinopathy screening and contributing to ethical healthcare practices.

5.4 Sustainability Plan:

The following are some ideas for a long-term strategy to use computer vision for diabetic retinopathy screening:

a) **Energy efficiency:** The diagnostic process can have less of an impact on the environment if the computers and other equipment utilized are more energy efficient.

b) **Use of renewable energy:** The environmental impact of diagnostics can be further mitigated by using renewable energy sources like solar or wind power to power the computers and equipment utilized in the procedure.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.2 Summary of the Study:

This study introduces a robust strategy for the effective deep learning-based multi classification and categorization of ocular diseases using a convolutional neural network (CNN) approach. The proposed model exhibits a high accuracy rate of 92%, as evidenced by an impressive Area Under the Curve (AUC) value of 98%. The innovation lies in a technique that aids both the public and medical professionals in identifying various eye diseases, irrespective of their presence. The model, after rigorous testing and evaluation, shows promise as a valuable reference tool for medical professionals in diagnosing eye problems. However, the study acknowledges limitations, such as an insufficient dataset and the need for more extensive testing with unstudied medical images and real-time data. The proposed fine-tuned modified CNN model demonstrates proficiency in classifying four distinct types of eye diseases, indicating potential for further refinement. Future work includes exploring therapeutic applications of deep learning for diagnosing other disorders and reducing preprocessing steps as models improve. The study envisions improved diagnostic accuracy and efficiency, paving the way for advancements in ocular disease diagnosis and potentially extending to other medical domains.

6.2 Conclusion:

This study presents a strategy for the best deep learning-based multi-classification and categorization of ocular illnesses, based on a convolutional neural network-based approach. Regardless of whether the eye ailment is present or not, a new technique has emerged that could help both the public and medical professionals identify distinct kinds of eye diseases. The proposed model has a 92% accuracy rate, as shown by an AUC value of 98%. I have concluded that the model shows promise as a reference tool for medical professionals to utilize in the diagnosis of eye problems after passing through the stages of testing and evaluation. Even though it only takes a few random photos to establish if a case has been correctly diagnosed, the old method takes too long to find out if a case has been correctly detected.

6.2 Limitation and Future work

Compared to traditional classifiers, CNN models trained using deep learning performed better in this study's multiclass classification task. Despite a significant loss of accurate medical data, the dataset intended for the proposed model is inadequate; this is a critical error in the study. Using bigger amounts of unstudied medical pictures and real-time medical data, it is feasible to assess the effectiveness of the suggested methodology soon. The research model, however, correctly classifies the four distinct kinds of eye diseases, according to most testing. While there might be some bumps along the road, reaching this objective is certainly within reach. There are a few little issues, but it's easy to ensure that the proposed fine-tuned modified CNN model is better at diagnosis overall and more accurate. It might not be too tough to do this. Achieving this objective is within reach. Researchers may investigate the therapeutic uses of deep learning for diagnosing other disorders in future studies. It is possible that rare diseases could be better diagnosed with the help of deep neural network learning. Also, the number of preprocessing steps needed might go down as the models get better. Improving the performance of deep learning models could also be achieved by gaining a better understanding of the reconstruction kernel or the picture thickness.

References

[1] Li T, Guo S, Wang K, Liu H, Kang H, Gao Y, and Li T. diagnostic evaluation of deep learning algorithms to screen for diabetic retinopathy. N Y. Inf Sci. 2019;501: 511. 10.1016/j.ins.2019.06.011 (522). doi

[2] Gangwar AK, Ravi V. Diabetic Retinopathy Detection with Transfer Learning as well as deep learning, advancements in intelligent systems and computing. Science and Business Media Deutschland GmbH, Springer Publishers, 2021, pp. 679–689. 10.1007/978-981-15-5788-0_64 is the doi.

[3] Folk JC, Clarida W, Amelon R, Erginay A, Lou Y, Abràmoff MD, et al. enhanced automated identification of diabetic retinopathy using deep learning on a publicly accessible dataset. doi:10.1167/iovs.16-19964; Invest Ophthalmol Vis Sci. 2016;57: 52005206

[4] Articles Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study, Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, et al. in 2019. www.thelancet.com/ is accessible.

[5] Krause J, Rahimy E, Karth P, Widner K, Corrado GS, Rahimy Gulshan V, et al. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. The doi is 10.1016/j.ophtha.2018.01.034 for 2017.

[6] Wilson GA, Sime MJ, Hong SC, and Ramachandran N. Deep neural networks are used to screen for diabetic retinopathy. Ophthalmology Clin Exp. 2018;46: 412–416. Reference: 10.1111/ceo.13056

[7] Rakibul Haque, Azam S, and Montaha S Hasan MZ, Rafid AKM, Ghosh P, Jonkman M, et al. BreastNet18: An Ablation Study-Evaluated High Accuracy Fine-Tuned VGG16 Model for Breast Cancer Diagnosis from Enhanced Mammography Images. Page 1347, Biology 2021, Vol. 10, 2021: 1347. doi:10.3390/BIOLOGY10121347

[8] Vocaturo E, Zumpano E, Caroprese L. Convolutional Neural Network Techniques on X-ray Images for Covid-19 Classification. Proceedings - 2021 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2021. 2021; 3113–3115. doi:10.1109/BIBM52615.2021.9669784

[9] Hussain E, Hasan M, Rahman MA, Lee I, Tamanna T, Parvez MZ. CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. Chaos Solitons Fractals. 2021;142. doi: 10.1016/j.chaos.2020.110495

[10] Paul S, Chen P-Y. Vision Transformers are Robust Learners. 2021.

[11] Jin L, Tan F, Jiang S. Generative adversarial network technologies and applications in computer vision. Comput Intell Neurosci. 2020;2020: 1459107.

[12] Arcadu F, Haskova Z, Prunotto M, Willis J, Maunz A, and Benmansour F. Individual patients' diabetic retinopathy progression is predicted by a deep learning algorithm. 2019;2. doi:10.1038/s41746-019-0172-3 NPJ Digit Med

[13] Ting DSW, Cheung CYL, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA - Journal of the American Medical Association. 2017;318:2211–2223. doi:10.1001/jama.2017.18152

[14] Dhar P, Department of Computer Science, Engineering Cox's Bazar City College B. A method to detect breast cancer based on morphological operation. Int j educ manag eng. 2021;11: 25–31.

[15] Determining the Severity of Diabetic Retinopathy through Analyzing Retinal Images Using Machine Learning algorithm. doi:10.13140/RG.2.2.23178.39364

[16] Institute of Electrical and Electronics Engineers. Bangladesh Section Institute of Electrical and Electronics Engineers. 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON): 28th30th November, FARS Hotel and Resorts, Dhaka, Bangladesh.

[17] He J, Li C, Ye J, Qiao Y, Gu L. Multi-label ocular disease classification with a dense correlation deep neural network. Biomed Signal Process Control. 2021;63. doi: 10.1016/j.bspc.2020.102167

[18] Li N, Li T, Hu C, Wang K, Kang H. A Benchmark of Ocular Disease Intelligent Recognition: One Shot for Multi-disease Detection. 2021. Available: http://arxiv.org/abs/2102.07978

[19] Chea N, Nam Y. Classification of Fundus Images Based on Deep Learning for Detecting Eye Diseases. Computers, Materials & Continua. 2021;67: 411–426. doi:10.32604/cmc.2021.013390

[20] Smitha A, Jidesh P. Classification of Multiple Retinal Disorders from Enhanced Fundus Images Using Semi-supervised GAN. SN Computer Science 2021 3:1. 2021;3: 1–11. doi:10.1007/S42979-021-00945-6

[21] Gangwar AK, Ravi V. Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning. Advances in Intelligent Systems and Computing. Springer Science and Business Media Deutschland GmbH; 2021. pp. 679–689.
doi:10.1007/978-981-15-5788-0_64 Biomed
[22] Yang X lian, Yi S li. multi-classification of fundus diseases based on DSRA CNN. Signal 10.1016/J.BSPC.2022.103763 Process Control. 2022;77: 103763. doi:
[23] Islam MT, Imran SA, Arefeen A, Hasan M, Shahnaz C. Source and Camera Independent Ophthalmic Disease Recognition from Fundus Image Using Neural Network. 2019 IEEE International Conference on Signal Processing, Information Communication and Systems, SPICSCO doi:10.1109/SPICSCON48833.2019.9065162 2019. 2019; 59–63.
[24] Demir F, Taşcı B. An Effective and Robust Approach Based on R-CNN+LSTM Model and NCAR Feature Selection for Ophthalmological Disease Detection from Fundus Images. Journal of Personalized Medicine 2021, Vol 11, Page 1276. 2021;11: 1276. doi:10.3390/JPM11121276
[25] Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Abul Kashem S Bin, et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Comput Biol Med. 2021;132: 104319.
[26] Gour N, Khanna P. Multi-class multi-label ophthalmological disease detection using transfer learning- based convolutional neural network. Biomed Signal Process Control. 2021;66. doi: 10.1016/j.bspc.2020.102329
[27] Nair V, Suranglikar S, Deshmukh S, Gavhane Y. Multi-labeled ocular disease diagnosis enforcing transfer learning. 2021 55th Annual Conference on Information Sciences and Systems, CISS 2021. Institute of Electrical and Electronics Engineers Inc.; 2021. doi:10.1109/CISS50987.2021.9400227
[28] Gian MKS, Raman V, Then PHH. Deep Learning for The Automatic Diagnosis and Detection of Multiple Retinal Abnormalities. Journal of Integrated Design and Process Science. 2019;23: 5–41. doi:10.3233/JID190017
[29] Yaqoob MK, Ali SF, Bilal M, Hanif MS, Al-Saggaf UM. ResNet Based Deep Features and Random Forest Classifier for Diabetic Retinopathy Detection Sensors 2021, Vol 21, Page 3883. 2021;21: 3883. doi:10.3390/S21113883
[30] Shamshirband S, Rehman ZU, Khan FG, Shah S, Khan A, Qummar S, et al. An Ensemble Method for Diabetic Retinopathy Detection Using Deep Learning. IEEE Access. 7: 15530–150539, 2019. The title is 10.1109/ACCESS.2019.2947484.

[31] Qiao L, Zhu Y, Zhou H. Diabetic Retinopathy Detection Using Multianeurysm Prognosis and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Using Deep Learning Algorithms. Access, IEEE. 2020;8: 104292104302; doi:10.1109/ACCESS.2020.2993937

Using Deep Neural Networks to Diagnose Diabetic Retinopathy [32] Gao Z, Li J, Guo J, Chen Y, Yi Z, Zhong J. In 2019, IEEE Access, 7: 3360–3370. 10.109/ACCESS.2018.2888639 is the doi.

[33] Tymchenko B, Marchenko P, Spodarets D. Diabetic Retinopathy Detection Using Deep Learning. 2020. You can access it at http://arxiv.org/abs/2003.02261.

[34] Engineering College Sri Venkateshwara. Institute of Electrical and Electronics Engineers, Department of Electronics and Communication Engineering. IEEE Computer Society, Institute of Electrical and Electronics Engineers, Bangalore Section. Proceedings of the 2nd IEEE International Conference on Recent Trends in Electronics, Information, and Communication Technology (RTEICT-2017) were held May 19–20, 2017.

[35] Mahdi S, Enas H; Al-Saadi H, Ali; Al-Sultan Y. Automatic classification of AMD in retinal images. AIP Conf Proc. 2022;2394: 050001. doi:10.1063/5.0121323

[36] Delavari P. Explainable diagnosis using deep learning: classification of fundus images as a promising tool for early detection of Alzheimer's disease. 2022 [cited 16 May 2023]. doi:10.14288/1.0419380

[37] Emir B, Colak E. Performance analysis of pretrained convolutional neural network models for ophthalmological disease classification. Arq Bras Oftalmol. 2023 [cited 16 May 2023]. doi:10.5935/0004-2749.2022-0124

[38] Zahra Belharar F, Zrira N. DeepRetino: Ophthalmic Disease Classification from Retinal Images using Deep Learning. 2022 IEEE 9th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, SETIT 2022. 2022; 392–399. doi:10.1109/SETIT54465.2022.9875570

192-15-2889

ORIGINALITY REPORT

25%	20%	17%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	7%
2	www.mdpi.com Internet Source	2%
3	Submitted to Daffodil International University Student Paper	1%
4	www.wjgnet.com Internet Source	1%
5	Submitted to Jacksonville University Student Paper	1%
6	Mohaimenul Azam Khan Raiaan, Kaniz Fatema, Inam Ullah Khan, Sami Azam et al. "A Lightweight Robust Deep Learning Model Gained High Accuracy in Classifying a Wide Range of Diabetic Retinopathy Images", IEEE Access, 2023 Publication	1%
7	link.springer.com Internet Source	<1%
8	"Innovations in Computational Intelligence and Computer Vision", Springer Science and Business Media LLC, 2023 Publication	<1%
9	Submitted to CSU Northridge Student Paper	<1%