EDNET-20: AN EYE DISEASE CLASSIFICATION APPROACH COMBINING CNN AND HYBRID ATTENTION

BY

MOHAMMAD RIADUR RASHID
ID: 201-15-14266

FINAL YEAR DESIGN PROJECT REPORT

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

Supervised By

Shayla Sharmin
Senior Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Md. Sadekur Rahman
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "EDNet-20: An Eye Disease Classification Approach combining CNN and Hybrid Attention", submitted by Mohammad Riadur Rashid to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 13-07-2024.

BOARD OF EXAMINERS

Dr. Md. Ismail Jabiullah (MIJ) — **Board Chairman**
Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Abdus Sattar (AS) — **Internal Examiner 1**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Ms. Zahura Zaman (ZZ) — **Internal Examiner 2**
Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Ahmed Wasif Reza (DWR) — **External Examiner**
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

I hereby declare that this project has been done by me under the supervision of **Shayla Sharmin, Senior Lecturer, 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:

Shayla Sharmin 13.7.24

Shayla Sharmin
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

M.Riadur Rashid

Mohammad Riadur Rashid
ID: 201-15-14266
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty for His divine blessing making it possible for me to complete the final year project successfully.

I am grateful and wish my profound indebtedness to **Shayla Sharmin, Senior Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of deep learning to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to the **Head of the Department of CSE**, for his kind help in finishing my project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

In order to optimize clinical management and treatment, it is crucial to accurately diagnose and classify eye diseases. The significant increase in the occurrence of eye diseases and the frequency of visual impairment recently has highlighted the urgent need for enhanced diagnostic techniques and therapies. Deep learning algorithms have shown significant promise in addressing the issue of disease classification. Current diagnostic methods for eye diseases increasingly include advanced medical imaging techniques, including optical coherence tomography angiography (OCTA), optical coherence tomography (OCT), color fundus photography (CFP), and fundus fluorescein angiography (FFA). This study aims to construct a robust framework for accurately classifying eye diseases by using two distinct color fundus photography (CFP) datasets. To classify the diseases this study proposes a novel classifier model called EDNet-20. The EDNet-20 design represents a novel convolutional neural network. The EDNet-20 architecture saw significant enhancements via a rigorous model tweaking of the Convolutional Neural Network (CNN) model using the Mendeley dataset. The study meticulously used data augmentation and pre-processing techniques to enhance the quality and quantity of images in both datasets. The recommended model attained an accuracy of 94.26% on the Mendeley dataset and 91.80% on the private dataset. Compared to eight high-quality transfer learning models, three models that use attention mechanisms, and the basic convolutional neural network (CNN), the recommended model outperformed in terms of accuracy, precision, specificity, recall, and f1 score. These results provide evidence that deep learning technologies may successfully reduce the incidence of vision loss caused by eye diseases. EDNet-20 will be valuable for medical practitioners in enhancing patient outcomes because of its streamlined design and exceptional ability to accurately classify eye diseases.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

CSCR	Central Serous Chorioretinopathy
DR	Diabetic Retinopathy
DE	Disc Edema
GC	Glaucoma
MS	Macular Scar
MP	Myopia
NM	Normal
RD	Retinal Detachment
RP	Retinitis Pigmentosa
DS	Depthwise Separable
FC	Fully Connected
CONV	Convolutional
BN	Batch Normalization
CNN	Convolutional Neural Network
CAD	Computer Aided Diagnosis

CHAPTER 1
Introduction

1.1 Overview

The eye, a marvel of biology, captures the world's beauty in its intricate dance of light and perception. The eye holds paramount importance as the primary sensory organ in the human body, as approximately 80% of sensory input is received through it [1]. Eye diseases can pose obstacles to both clear vision and sensory perception, frequently disrupting people's visual capabilities [2]. Eye diseases have been challenging the public health system worldwide for a long time. According to the WHO, blindness and vision impairment pose a significant public health challenge, impacting a minimum of 2.2 billion individuals globally, and the numbers are increasing rapidly [3]. The most common eye conditions comprise diabetic retinopathy (DR) [4, 5, 6], glaucoma [4, 5, 6], age-related macular degeneration (AMD) [4, 5, 6], myopia [6], retinitis pigmentosa (RP) [7], retinal detachment (RD) [8], disc edema [9], and central serous chorioretinopathy (CSCR) [10], which cause visual impairment. In 2010, statistical analysis indicated that 1 in every 39 individuals who were blind experienced blindness because of DR, while 1 in every 52 individuals faced visual impairment due to DR [11]. By 2030, the number of individuals with diabetes retinopathy is projected to exceed 400 million, while the number of those suffering from glaucoma is expected to reach 80 [12]. Around 67 million individuals in the EU are presently impacted by some form of AMD, and projections suggest this figure will rise by 15% by 2050 due to the aging population [13]. The prevalence of myopia and high myopia is expected to notably rise on a global scale, impacting close to 5 billion individuals with myopia and 1 billion with high myopia by 2050 [14]. RP affects more than 1.5 million patients worldwide and is the most common type of inherited retinal dystrophy, occurring in approximately 1 in 4000 people globally [15]. Rhegmatogenous retinal detachment stands out as the predominant urgent condition affecting the retina, occurring at a rate of 6.3 to 17.9 cases per 100,000 people annually [16]. Papilledema, characterized by optic disk edema stemming from idiopathic intracranial hypertension (IIH), represents the most common cause of this condition. Its occurrence is notably high, reaching up to 3.5 cases per 100,000 females aged 15 to 44 years old [17]. A study conducted in South Korea comparing corticosteroid users and non-users found that the overall occurrence of central serous chorioretinopathy was 5.4 cases per 10,000 person-years among men and 1.6 cases per 10,000 person-years among women [18]. Nearly 1.1 billion vision impairment effects can be prevented worldwide

[19]. If eye diseases aren't caught early, they can harm the retina and lead to permanent vision loss [20]. Detecting eye diseases early and providing appropriate treatment is essential to preventing vision loss [21]. Deep learning methods have attained cutting-edge accuracy in distinguishing between healthy and diseased retinal fundus images but effectively classifying multiple retinal eye diseases remains a significant hurdle [22]. 75% of patients with DR reside in regions lacking adequate development, where there is a shortage of specialists and the necessary infrastructure for detection. As a result, millions of people around the globe still suffer from vision problems due to inadequate predictive diagnosis and a lack of proper eye care [23]. The increasing prevalence of eye diseases is paralleled by a rise in the acquisition of retinal fundus images through screening. This leads to a significant increase in the workload and time required for medical professionals due to the complexity involved [24].

1.2 Background and Present State

Presently, medical imaging techniques stand out as one of the most valuable tools for diagnosing various eye diseases reviewed [25]. Analyzing these medical images necessitates the use of cutting-edge technological advancements. Ophthalmologists utilize a range of imaging techniques, such as Optical Coherence Tomography (OCT) [26], Fundus Fluorescein Angiography (FFA) [27], Color Fundus Photography (CFP) [28], and several other scans, to assess eye conditions. These advanced methods enable precise and swift categorization, significantly enhancing patient prognosis in eye diseases.

Through advancements in computational capabilities and the refinement of neural networks, deep learning has established the pinnacle of autonomous learning [29]. Deep learning finds application in ophthalmology for tasks like analyzing data, segmenting images, diagnosing conditions automatically, and potentially predicting outcomes [30]. CNNs, a form of deep learning architecture comprised of numerous layers arranged hierarchically, are employed to convert pixel data from images into significant features. These extracted features are then instrumental in the process of identifying and categorizing various diseases. Leveraging deep learning advancements holds the potential to expedite the analysis of color fundus images, thereby enhancing the classification of eye [31, 32, 33]. It is possible to enhance the performance of CNN models through the attention mechanism. The attention mechanism holds the potential for enhancing CNNs' segmentation performance by emulating human behavior, concentrating on the most pertinent information in the feature maps, and downplaying irrelevant parts [34]. The core idea of the attention mechanism is that the model learns by concentrating on specific areas of interest. Recently, the attention mechanism has proven to be effective in image captioning [35].

This study aims to classify different types of eye diseases using color fundus images to help medical professionals and radiologists develop better treatments and reduce blindness and vision impairments. The study employs two datasets containing variations in the image size and unwanted noise. Removing these undesirable elements from the images is essential for optimizing the performance of a CNN model. Image augmentation, which involves modifying images to enhance training data, utilizes various techniques such as photometric, geometric, and elastic deformation. Of these, the most effective photometric augmentation techniques are employed. The study proposes an automated and reliable deep learning model called EDNet-20, based on CNN architecture. The model is fine-tuned through model optimization and the overall process of color fundus image classification. The study deploys eight transfer learning models, three models that use attention mechanism, and the basic convolutional neural network (CNN) to compare the robustness of the proposed model. Various performance metrics are employed to assess how well the model works while investigating different image pre-processing methods to achieve the best results. The research conducts a thorough experiment on the dataset to determine the most effective pre-processing algorithms and their settings, known as model optimization. Once optimal performance is achieved, the proposed model undergoes k-fold cross-validation to prevent overfitting. Consistent findings from the study highlight that integrating image processing techniques with model optimization and transfer learning can accelerate the diagnosis of eye diseases, potentially benefiting clinicians in their treatment approaches.

1.3 Problem Statement

The increasing accessibility of digital imaging technologies and the prevalence of eye diseases worldwide have generated an urgent need for precise and rapid methods of diagnosis and classification. Fundus photography may enhance our understanding of different eye illnesses by capturing detailed images of the posterior segment of the eye. The human study of these photographs is arduous, subjective, and susceptible to errors when it comes to recognizing little abnormalities that could otherwise remain unnoticed. The paper titled "EDNet-20: An Eye Disease Classification Approach combining CNN and Hybrid Attention" is based on solid reasoning to overcome challenges in medical imaging and ocular illnesses. The development of deep learning algorithms has opened up a promising avenue for automating fundus image interpretation. This advancement has the potential to enable the efficient, accurate, and speedy categorization of eye illnesses on a large scale. Convolutional neural networks (CNNs), a kind of deep learning algorithm, are very effective for image analysis jobs due to their remarkable capability to extract intricate features and patterns from extensive datasets. To create reliable classification systems for identifying and categorizing different eye illnesses, CNN models are trained using annotated fundus image datasets.

The primary objective of the work is to eradicate the potential for incorrect diagnosis and treatments by developing automated classification systems using deep learning techniques that can efficiently analyze large volumes of fundus pictures. Implementing this would enable streamlined screening, thereby enhancing the precision and dependability of identifying ocular conditions.

1.3.1 Research Questions

The researchers in the study titled "EDNet-20: An Eye Disease Classification Approach combining CNN and Hybrid Attention" formulated a series of inquiries to guide their research. The inquiry explored the outcomes of deep learning models, the practical implications of automated sickness identification utilizing fundus images, and the classification technique.
Key questions for further research:

1. What methods may be used to enhance the accuracy and efficiency of classifying eye illnesses in fundus photos using deep learning?

The primary purpose of deep learning in image classification via knowledge acquisition from fundus photographs is the focus of the investigation. The primary objective of the study is to investigate the mechanisms of deep learning in order to enhance the efficiency and accuracy of classifying eye diseases in fundus images.

2. Regarding the diagnosis of eye illnesses, how do different kinds of convolutional neural networks perform?

The primary focus of the investigation is on how deep learning models in image analysis to aid in the categorization of eye illnesses. The objective of this study is to enhance the proposed model by evaluating its performance in comparison to other models.

3. What is the relationship between the model's generalizability and the most important variables?

The investigation explores the fundamental factors that impact overall performance, including image quality, demographic information, and data volume. The primary objective of the research is to enhance performance by discovering the underlying factors that contribute to it.

4. How can one differentiate between different eye illnesses based on the fundus image? What are the key elements of this image?

The query aims to differentiate between different eye illnesses by analyzing fundus image attributes in order to provide valuable information. The objective of this study is to enhance the diagnosis of eye diseases by examining the underlying anatomical elements associated with these problems.

5. What is the comparative effect of the recommended model on the diagnosis and classification of eye illnesses, as compared to more traditional approaches?

The investigation explores the possible impacts of using the study's suggested framework to enhance the efficacy of eye treatment as compared to traditional methods. The objective of the study is to ascertain the enhancements made to treatment procedures upon the implementation of the model.

6. How can we use the triumphs and failures of different categorization systems to develop a diagnostic tool that is more dependable and accurate?

The objective of the investigation is to analyze the advantages and disadvantages of various methods for categorizing eye diseases. The aim is to develop diagnostic tools by using the gathered information.

7. How can we use this newly discovered knowledge to improve patient care and generate better healthcare results, and what are the implications of the research for the growing field of medical imaging?

The topic under discussion is the most effective method of using the knowledge gained from medical imaging research in real-life scenarios.

The thesis is structured based on a series of research inquiries that together direct the examination of the intricacies involved in classifying eye diseases using deep learning methods.

1.4 Objectives

The predicted outcomes of the research "EDNet-20: An Eye Disease Classification Approach combining CNN and Hybrid Attention" could involve a variety of discoveries and notable achievements.
The expected outcomes include the following:

1. Creation of a disease classification system.

The objective is to create deep learning models that can precisely detect and classify different eye diseases using fundus photographs with a high level of accuracy. The system will use cutting-edge deep learning techniques, including advanced models such as VGG19, MobileNetV2, InceptionV3, and DenseNet201, in conjunction with the suggested model. The primary objective of the method is to get the most accurate results in classifying various ocular diseases.

2. An assessment and comparison review of deep learning models.

The selected deep learning models will be deployed to conduct evaluations and analyses in order to get relevant insights. The research will provide an analysis of performance indicators, accuracy, and efficiency models, as well as reveal valuable settings.

3. Efficient dataset generation.

Various data preparation approaches will be investigated to improve performance. The techniques of resizing, grayscale conversion, smoothing, sharpening, histogram equalization, and data augmentation will be investigated in order to generate an optimum dataset.

4. Examine the significant features of the fundus photograph.

The project will use image-processing methods such as image filtering, feature extraction, and picture segmentation to gain knowledge about the anatomical structures associated with different illnesses. This part aims to optimize deep learning models.

5. Incorporation into the medical treatment process.

Integrating the model into the current processes helps optimize the diagnosing process. The key factors to consider when evaluating the performance of an automated system are the precision of its classification, the efficiency of its speed, the reliability of its model, its capacity to handle larger workloads, and its cost-effectiveness compared to conventional techniques.

6. Utilize the advantages of various classification methods and verify the outcomes.

In order to create a more reliable and precise diagnostic tool, it is crucial to use the advantages of categorization methods and address their shortcomings.

7. Participate in educational discussions.

The focus of the research serves to add to the academic literature on the intersection of deep learning and eye problems.

1.5 Scope and Limitations

The objective of the project is to classify prevalent eye conditions such as diabetic retinopathy, macular scar, and myopia by using high-resolution retinal photographs obtained from both public and commercial medical databases. In addition, the research also investigates the classification of less common eye diseases to assess the strength and adaptability of the classification model. The development and evaluation of classification systems include the use of machine learning methods such as Convolutional Neural Networks (CNNs) and Deep Learning models. Evaluation conducted using conventional measures such as accuracy, precision, recall, and F1-score. Nevertheless, the variability in picture quality and resolution has a substantial impact on the effectiveness of classification algorithms. Furthermore, the dataset exhibit an imbalance, resulting in models that are biased. Elaborate models need significant computing resources and lengthier training durations. The incorporation of novel classification methods into clinical processes may encounter opposition as a result of alterations in traditional practices.

1.6 Report Layout

Chapter 1: Introduction

The first section of the report is the introduction. The book provides an introduction to the research, highlighting the significance of categorizing and identifying eye illnesses, the use of deep convolutional neural networks (CNNs), and the need to compare different approaches for detection and classification. This document outlines the research questions, objectives, and overall framework of the study.

Chapter 2: Background

The background chapter thoroughly explores relevant literature and past investigations. This part offers a thorough understanding of the theoretical foundations, research methods, and technical advancements related to the categorization and diagnosis of eye diseases, as well as deep convolutional neural networks (CNNs). It establishes the foundation for the current inquiry, identifying specific areas where there is a shortage of existing information.

Chapter 3: Research Methodology

The research methodology chapter provides a comprehensive explanation of the specific methods used to conduct the study. This document provides a comprehensive summary of the study strategy, data-collecting methods, model structure, and the rationale behind the selection of certain methodologies. Furthermore, this chapter discusses ethical concerns and recognizes any constraints that may impact the research.

Chapter 4: Implementation

This chapter represents the implications of models, model optimization to create a robust classifier. The chapter summarizes the implications of all models training, model testing and evaluation.

Chapter 5: Result and Analysis

This chapter presents the empirical findings obtained by using transfer learning and deep convolutional neural networks (CNNs) to categorize eye disorders. The study includes thorough evaluations of the effectiveness of several detection and segmentation techniques, along with visual illustrations and statistical measures to verify the results.

Chapter 6: Impact on society, Environment and Sustainability

The chapter on social effect explores the broader ramifications of the study results on healthcare and medical diagnostics. The text explores the potential benefits of advanced techniques for categorizing eye illnesses, including improved patient care, greater treatment results, and overall improvement in public health. The chapter also examines potential technological developments and their wider social ramifications.

Chapter 7: Conclusion and Future Research

This last chapter consolidates the whole study by providing a concise overview of the main discoveries, restating the conclusions drawn from the research, and suggesting practical uses and potential areas for further investigation. The text explores the consequences for future investigation, offering a guide for those who are interested in building upon the existing work.

The structured layout ensures a logical flow of information, guiding the reader through the study process, from the introduction to the wider implications and suggestions. This approach facilitates a comprehensive comprehension of the research process and its potential impact on both the scholarly and practical implementations of deep learning in the classification of eye diseases.

1.7 Summary

Worldwide, 2.2 billion people suffer from an eye illness of some kind. This includes central serous chorioretinopathy, glaucoma, myopia, retinitis pigmentosa, retinal detachment, and diabetic retinopathy. Timely identification and management of these ailments are crucial in averting permanent visual impairment. While the classification of many illnesses is still difficult, deep learning approaches have shown promise in differentiating between healthy and diseased retinal fundus pictures. To evaluate eye diseases, medical imaging techniques including Fundus Fluorescein Angiography, Color Fundus Photography, and Optical Coherence Tomography may be used. Deep learning, which is a consequence of advances in processing power and neural networks, may be used by ophthalmologists to evaluate data, segment pictures, diagnose illnesses autonomously, and perhaps forecast outcomes. By focusing on target areas, the attention process improves the classification accuracy of CNN models.

The effort uses color fundus pictures to classify eye disorders with the goal of improving medical treatments and reducing blindness and visual impairments. Two are used in the research to enhance CNN models' performance. The study presents EDNet-20, an automated CNN-based model that is further improved by classifying color fundus pictures and improving the model. The investigation's main goal is to develop a trustworthy convolutional neural network model for the diagnosis of eye disorders in order to have an impact on the worldwide eye care sector. The goal of the project is to avoid incorrect diagnosis and treatments by developing automated classification systems that use deep learning methods.

CHAPTER 2
Literature Review

2.1 Overview

In order to fully understand the fundamental ideas and approaches used in this study on "EDNet-20: An Eye Disease Classification Approach combining CNN and Hybrid Attention," it is essential to set foundations by introducing basic concepts and defining important terms. Transfer learning, a key principle in artificial intelligence, plays a major role in this work by allowing the model to use information acquired from unrelated tasks to enhance the classification of eye diseases. Deep Convolutional Neural Networks (CNNs) are used to effectively capture complex patterns in medical imaging, making them integral for image interpretation. Identification is the process of finding eye diseases in medical photographs, whereas segmentation entails accurately identifying the spatial boundaries of the damaged areas. This study employs these terminologies in an organized way to investigate the relationships among transfer learning, deep convolutional neural networks (CNNs), and the comparative examination of detection and classification methodologies. It is crucial for readers to have a shared awareness of these initial details in order to understand the complexities of the research and appreciate the sophisticated methods used in diagnosing eye diseases.

2.2 Related Work

A number of prominent studies have shown the effectiveness of CNNs in accurately identifying and diagnosing certain eye disorders, including diabetic retinopathy, glaucoma, and age-related macular degeneration.

Classification and Detection

In this study, Srivastava et al. [36] investigate the use of artificial intelligence in the area of ophthalmology. This subject is particularly suited for computer algorithms because of the availability of digital images and quantifiable data. Their main areas of concentration include the retina, cornea, anterior region, and pediatric ophthalmology. They provide very accurate disease segmentations, including a computer-based segmentation of the optic nerve for glaucoma scoring.

Mayya et al. [37] used convolutional neural networks and multiple preprocessing techniques to identify COD in eye fundus images. They evaluated the ResNeXt50 model, which achieved an accuracy of 86.06% and an F1 score of 89.53%, with the optimal outcomes obtained using RGB RoI cropped images.

Li et al. [38] conducted a study to compare automatic and traditional systems in detecting diabetic retinopathy, age-related macular degeneration, and glaucomatous optic neuropathy. They used an independent dataset from China and compared diagnostic consistency between different experience levels and non-physician graders. The study found that the deep learning system had higher sensitivity (97.8% vs. 91.1%) but lower specificity (92.5% vs. 99.6%) than general ophthalmologists.

Zang et al. [39] developed an automated diagnostic system that can classify ailments by utilizing extensive structural OCT and OCTA data. The semi-sequential 3-dimensional convolutional neural networks of the framework can be employed to diagnose diabetic retinopathy and age-related macular degeneration. Non-DR or AMD data is employed in the second phase of the glaucoma diagnostic procedure. The framework's AUC was greater than 0.9 for each illness that was assessed, suggesting that it performed exceptionally well.

Song et al. [40] evaluated the diagnostic capabilities of a robotically aligned optical coherence tomography (RAOCT) system combined with a deep learning model for identifying referable posterior segment pathology in OCT images from emergency department patients. The model, RobOCTNet, was trained and validated to differentiate between OCT images requiring ophthalmology consultation and those not. The model achieved an AUC of 1.00 in internal testing and 0.91 in external validation, with a sensitivity of 95% and a specificity of 76%.

Choudhary et al. [41] created an artificial intelligence model that can automatically identify different stages of retinal illness. The OCT-scanned retina pictures are classified into four groups using a VGG-19 architecture consisting of 19 layers. The model attained a classification accuracy of 99.17%, with specificities and sensitivity rates of 0.995 and 0.99, respectively. The authors propose that the model has the potential to be modified for the treatment of several additional retinal illnesses.

In their study, Saranya et al. [42] devised a methodology to promptly diagnose diabetic retinopathy (DR) by detecting red lesions in retinal pictures. The model employs pre-processed pictures to minimize unwanted disturbances and improve the distinction between different elements, and it utilizes the UNet architecture for the purpose of semantic segmentation. A sophisticated convolutional layer architecture is included for the purpose of pixel-level class labeling. The performance of the model was evaluated using four datasets. The IDRiD dataset achieved a specificity of 99%, a sensitivity of 89%, and an accuracy of 95.65%. The STARE dataset achieved a specificity of 96.65%, a sensitivity of 95.23%, and an accuracy of 96%.

Kaur et al. [43] developed a method for classifying retinal images using a three-step process. They used CLAHE and DNCNN neural networks for image enhancement, applied morphological and K-means algorithms for segmentation, and then inputted the segmented images into their EyeNet model, built on the ResNet-18 architecture. The model achieved an impressive 99.76% accuracy.

Karlin et al. [44] developed a deep-learning classifier to identify thyroid eye disease using external photographs. They compared their model's performance with that of domain experts using two datasets. An ensemble model was constructed using five neural networks trained on different subsets of Stein Eye Institute data. The final prediction was combined, resulting in an overall accuracy of 89.2% on the test set and 86% on the survey set.

Sun et al. [45] developed an end-to-end deep learning model for classifying UWFI diseases using a limited dataset. They used high-performing convolutional neural networks (CNNs) like EfficientNet-B7, DenseNet, and ResNet-101, and initialized parameters from pre-trained models from ImageNet. The model used the normal eye as a classification reference and included an additional self-attentive layer to enhance feature dimension balance. EfficientNet-B7 was chosen for the deep learning model due to its superior performance on the internal test set, achieving an average AUC of 0.9880.

Attention Mechanism

Wang et al. [46] addressed two primary obstacles in the classification of fundus images: the classification of multiple labels and the restrictions on the availability and quality of data. The researchers used a problem transformation-based methodology and utilized various CNN models. They integrated a CNN MBConv block and a transformer into a unified network architecture. The researchers developed a complex neural network called MBSaNet, which utilizes convolutional blocks and self-attention modules to detect various fundus illnesses. The model attained an Area Under the Curve (AUC) of 0.891 on the test set from an external source and 0.878 on the test set from the same source as the model.

Huang et al. [47] introduced the Global Attention Block (GAB) for use in feed-forward convolutional neural networks (CNNs). The GAB creates an attention map, which is used to compute adaptive feature weights. This module is easily integrated into any CNN, improving classification performance. They developed a lightweight model called GABNet, which serves as a feature extraction network after data preprocessing. GABNet outperformed the Xception and EfficientNetV2B3 algorithms without transfer learning, achieving a 3.7% increase in classification accuracy. The model achieved 95.4% in the limited model and 96.5% in the complete model.

Dinpajhouh et al. [48] developed Attention-DenseNet, a model for detecting and assessing diabetic retinopathy using a pre-trained convolutional neural network. The model incorporates an attention mechanism, prioritizing significant regions more effectively than other CNN models. The APTOS 2019 dataset and fivefold cross-validation were used to evaluate the model's performance, achieving an overall accuracy of 98.44% for detection and 83.69% for severity grading.

Lu et al. [49] created an automated multilabel classification algorithm to detect fundus illnesses from color fundus pictures. The model, using a convolutional neural network, has the capability to accurately categorize both normal circumstances and seven prevalent forms of fundus illnesses. A method of attention was included to augment the process of extracting features. The model demonstrated a high level of accuracy in multi-disease classification, with 98.64% and 94.27% accuracy on the processed training and validation sets, respectively. This impressive performance was obtained despite the use of a tiny training dataset.

Ma et al. [50] developed a U-Net algorithm with a supervised attention mechanism for retinal vessel segmentation to improve precision and extract details of small vessels and edges. They integrated three sub-modules: the DFM encoding module, the CSE decoding module, and the SFM supervised fusion module to optimize the foundational network. The network's performance was evaluated using three publicly available datasets, with the DRIVE dataset achieving a high accuracy of 0.9570 with the combination of Base+CSE+SFM, while STARE and CHASE_DB1 datasets achieved the highest accuracy of 0.9660 and 0.9667 with the proposed model (DFM+CSE+SFM).

2.3 Comparison Between Existing Works

The study "EDNet-20: An Eye Disease Classification Approach combining CNN and Hybrid Attention" evaluates the effectiveness of transfer learning and deep Convolutional Neural Networks (CNNs) in detecting and classifying eye diseases. It compares these methods with traditional methods and advanced classification techniques. The research provides valuable insights into the strengths and weaknesses of each approach, offering a comprehensive understanding of their impacts on diagnostic procedures. The study also emphasizes the importance of a comprehensive approach to medical diagnostics, integrating both conventional and modern techniques. The results enhance our understanding of eye disease detection and classification and offer practical suggestions for improving current diagnostic methods. This comparative analysis will form the basis for future chapters, laying the groundwork for discussions on societal impact, recommendations, and implications for future research.

Table 2.3.1: Comparative Analysis of Related Works

Study	**Image Type**	**Model**	**Accuracy**
Srivastava et al. (2023) [36]	Fundus Image	Convolutional Neural Network (CNN)	The study emphasizes AI applications in the area of ophthalmology and presents various disease segmentation approaches from other studies
Mayya et al. (2023) [37]	Fundus Image	Convolutional Neural Network (CNN)	Accuracy 86.06% and F1 score 89.53%
Z. Li et al. (2023) [38]	Fundus Photographs	Convolutional Neural Network (CNN)	The comparison between DLS and Ophthalmologists demonstrates sensitivity (97.8% vs 99.6%) and specificity (92.5% vs 99.6%)
Zang et al. (2023) [39]	Optical coherence tomography angiography (OCTA) and optical coherence tomography (OCT)	Convolutional Neural Network (CNN)	The framework demonstrated robust performance, with the AUCs of the ROC curves for the test dataset exceeding 0.9 for each disease evaluated in the study.
Song et al. (2023) [40]	Optical coherence tomography (OCT)	Convolutional Neural Network (CNN)	During validation, RobOCTNet achieved an

			AUC of 91%, sensitivity of 95%, and specificity of 76%
Choudhary et al. (2023) [41]	Optical coherence tomography (OCT)	Convolutional Neural Network (CNN)	The customized VGG19 achieved an accuracy of 99.17% with a specificity and sensitivity of 99.5% and 99%, respectively.
Saranya et al. (2023) [42]	Color Fundus Image	Convolutional Neural Network (CNN)	The result divided into two parts, the IDRiD dataset red lesion detection system achieved an accuracy of 95.65%, and the STARE dataset DR classification system got an accuracy of 96%.
Kaur et al. (2023) [43]	Color Fundus Image	Convolutional Neural Network (CNN)	The proposed EyeNet model achieved an accuracy of 99.76%
Karlin et al. (2023) [44]	Photographs	Ensemble Neural Network	The model achieved an accuracy of 89.2% and 86% on the test set and survey set respectively
Sun et al. (2023) [45]	Ultra-widefield fundus images	Convolutional Neural Network (CNN)	The EfficientNet-B7 network achieved an average AUC of 98.8%
Wang et al. (2023) [46]	Fundus Image	Convolutional Neural Network (CNN)	The proposed MBSaNet achieved an AUC value of 89.1% on the off-site test set and 87.8% on the on-site test set.
Huang et al. (2023) [47]	Optical coherence tomography (OCT)	Convolutional Neural Network (CNN)	The GABNet model achieved an accuracy of 95.4% in the limited model and 96.5% in the complete model
Dinpajhouh et al. (2023) [48]	Color Fundus Image	Convolutional Neural Network (CNN)	The proposed model Attention-DenseNet demonstrates an overall accuracy of 98.44% for the detection and 83.69% for the grading task.
Lu et al. (2023) [49]	Color Fundus Image	Convolutional Neural Network (CNN)	The proposed model achieved an accuracy of 98.64% on the training

			set and 94.27% on the validation set.
Ma et al. (2024) [50]	Digital Retinal Images and Color Fundus Image	Convolutional Neural Network (CNN)	The STARE and CHASE_DB1 datasets attained an accuracy of 96.6% and 96.67%, respectively, with the proposed model.

2.4 Open Issues

The research focuses on the classification of eye diseases using fundus images using deep learning techniques. Retinal diseases pose a global health challenge, necessitating precise and prompt detection for effective treatment. The research aims to overcome the limitations of traditional methods, which often fall short in terms of sensitivity, specificity, and diagnostic accuracy. The research explores the use of transfer learning, an emerging concept in artificial intelligence, to improve the classification and detection of eye diseases. Transfer learning enables models to utilize knowledge gained from different tasks, addressing the complexities and variations found in retinal images. Combining deep Convolutional Neural Networks (CNNs) with attention mechanisms further enhances the model's effectiveness in extracting detailed patterns and features from complex visual data. The research also includes a comparative analysis of detection and classification methods, examining their strengths, weaknesses, and potential synergies. The research seeks to provide insights that surpass traditional classification techniques, refine current diagnostic methodologies, and improve the impact of eye diseases on public health.

2.5 Summary

This study explores the classification of eye diseases using fundus images using deep learning. Transfer learning enhances classification by using information from unrelated tasks. Deep Convolutional Neural Networks capture complex patterns in medical imaging. The study investigates the relationships between these concepts and compares detection and classification methodologies. Understanding these details is crucial for understanding the complexities of eye disease diagnosis. The research explores the use of transfer learning in artificial intelligence to improve the classification and detection of eye diseases using fundus images. It aims to overcome traditional methods' limitations in sensitivity, specificity, and diagnostic accuracy. The study also compares detection and classification methods, aiming to provide insights that improve public health and enhance the understanding of eye diseases.

CHAPTER 3

Methodology

3.1 Overview

The study "EDNet-20: An Eye Disease Classification Approach Combining CNN and Hybrid Attention" explores the use of advanced AI methods for medical diagnostics, specifically in classifying eye diseases. The research aims to address the challenges of conventional eye disease classification techniques by incorporating transfer learning and deep Convolutional Neural Networks (CNNs) with attention mechanisms. Transfer learning enhances the classification of eye diseases, while deep convolutional neural networks (CNNs) are used to discern detailed patterns from medical images, improving diagnostic precision and sophistication. The comparative analysis of different methods for detecting and categorizing abnormalities in fundus images helps to understand their strengths and weaknesses, enabling the creation of more reliable and efficient diagnostic instruments. The research combines artificial intelligence, medical imaging, and ophthalmology to advance eye disease classification using innovative approaches like transfer learning and deep convolutional neural networks.

The CNN base model was trained using a batch size of 16, a learning rate of 0.001, and the Adam optimizer. The entire dataset has undergone training for 100 epochs, with the loss function being specified as Categorical Cross-Entropy. This investigation was conducted on a machine with an Intel Core i5-10400F CPU, NVidia GeForce GTX 1080 GPU, 16 GB RAM, and a Samsung 980 Pro 500GB PCIe 4.0 M.2 NVMe SSD for storage. The study focused on eye disease classification architectures, including the proposed EDNet-20 as well as established models like VGG16, VGG19, DenseNet201, DenseNet169, ResNet50, Xception MobileNet, and MobileNetV2.

3.2 Proposed Methodology

Various steps are performed to develop an effective CNN-based classification model for eye disease. Figure 3.2.1 depicts the whole process, illuminating the numerous processes necessary to create a successful computer-aided classification system that might assist in proper disease diagnosis.

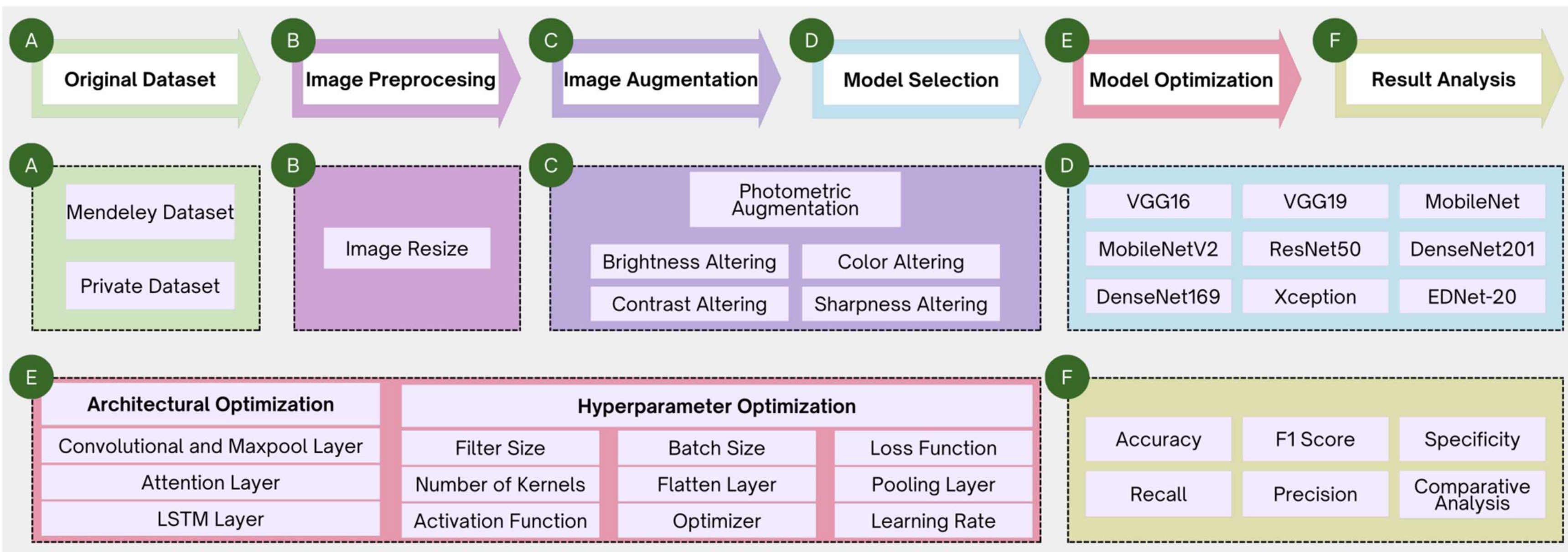

Figure 3.2.1: Methodology of the Work

Figure 3.2.1 Illustrate an overall approach to classifying eye disease based on two datasets is as follows: Frame (A) shows that the two datasets (i.e., The Mendeley and Private dataset) are collected. (B) Using image pre-processing steps, an enhanced pre-processed dataset is generated. (C) The pre-processed dataset is augmented using photometric augmentation techniques. (D) A baseline CNN model has been developed. (E) A robust model is developed by architectural and hyper parameter optimization. (F) The proposed model results are analyzed using performance analysis.

The research conducted many tests using two datasets that are specifically connected to eye problems. Each dataset underwent several image processing procedures to improve the quality of the pictures and eliminate any imperfections. In order to tackle the problems of data imbalance and shortage, several data augmentation techniques were used, such as photometric, geometric, and elastic deformation approaches. Out all these options, photometric augmentation was shown to be the most effective. The supplemented dataset was then divided into three parts: 70% for training, 20% for testing, and 10% for validation. These portions were then used as input for a baseline CNN model. Model optimization was conducted to get the most favorable performance. EDNet-20 was created by making adjustments to the layer architecture and hyper parameters of the baseline CNN model, using this optimization as a basis. Eight advanced transfer learning models, namely VGG16, VGG19, DenseNet201, DenseNet169, ResNet50, Xception MobileNet, and MobileNetV2, were trained and assessed on the datasets. Their accuracy was then compared to that of the proposed model. An analysis was conducted on the performance of the EDNet-20 model using different metrics on the two datasets. An k-fold cross-validation test was performed to assess the presence of overfitting.

3.2.1 Dataset Description

The proposed model has been tested on two datasets. This section provides pertinent details about these datasets and insights into the comprehensive testing methodology employed to evaluate the model's performance. The specifics of the datasets are outlined below.

A. Mendeley Dataset

This research analyzed 5229 images of healthy and affected eye images from a Mendeley dataset [51]. The dataset contains ten different classes of images, but one of them are not considered in this study due to the very small number of images. The images were in JPG format and had a size of 2004×1690 pixels. Figure 3.2.1.1 shows the Mendeley dataset in pictorial form for different classes. The dataset description is summarized in Table 3.2.1.

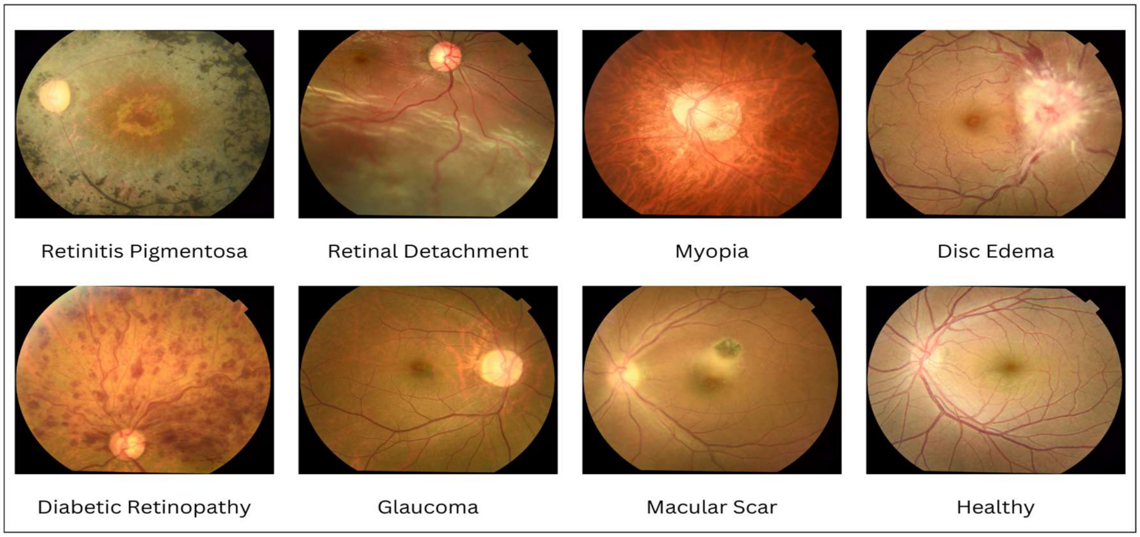

Figure 3.2.1.1: Random images from the Mendeley dataset

B. Private Dataset

The private dataset containing 1320 images of healthy and affected eye images was collected from Anawara Hamida Eye Hospital in Faridpur and BNS Zahrul Haque Eye Hospital in Faridpur district with the help of the hospital authorities. The images were graded by a skilled doctor. The dataset was captured by Topcon digital fundus machines (TL-211) which are linked to a separate Nikon DSLR camera (D90) with a maximum resolution of 4288x2848. Figure 3.2.1.2 shows the Mendeley dataset in pictorial form for different classes. The dataset description is summarized in Table 3.2.1.

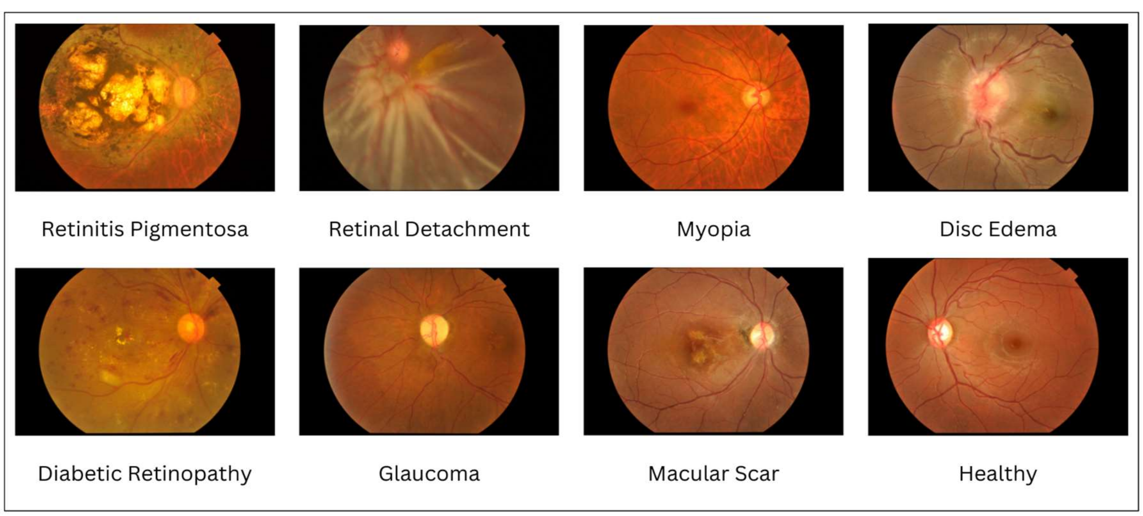

Figure 3.2.1.2: Random images from the Private dataset

Table 3.2.1: Datasets description

Name	Mendeley Dataset	Private Dataset
Total Number of Images	5229	1320
Average Dimension	2004×1690	4288x2848
Color Grading	RGB	RGB
Data Format	JPG	JPG
Classes	**Number of Images**	
Retinitis Pigmentosa	139	47
Retinal Detachment	125	92
Myopia	500	198
Disc Edema	127	86
Diabetic Retinopathy	1509	281
Glaucoma	1349	315
Macular Scar	444	211
Healthy	1024	65
Central Serous Chorioretinopathy	101	6

3.2.2 Image Pre-processing

The initial stage in employing deep learning networks with images involves pre-processing the image data. Pre-processing images is considered a crucial initial step before inputting them into a neural network, and this process is essential for attaining satisfactory accuracy and reducing the model's computational time [52]. Image pre-processing enhances quality by reducing noise and eliminating unnecessary artifacts, ensuring clearer and more accurate images [53]. Preprocessing steps enhance image quality, and these techniques extract valuable information from images [54]. Developing an effective neural network-based classifier for eye diseases involves several crucial steps. This study utilized two color fundus image datasets for all experiments. For each dataset, similar image processing techniques were applied. Initially, the images were resized to 224 x 224 pixels. The image pre-processing is demonstrated in Figure 3.2.2.1.

A. Image Resize

Image resizing is crucial for deep learning models, as smaller images lead to faster training and greater computational efficiency. Additionally, consistent image size is often required for many deep-learning architectures [55]. Resizing an image during pre-processing involves changing its dimensions, either increasing or decreasing its size. This process might be employed to modify the image's aspect ratio or to fit it into a specific space. In this research, the images have been resized to 224 x 224 pixels, resulting in a different aspect ratio. An output of the resized image is shown in Figure 3.2.2.1.

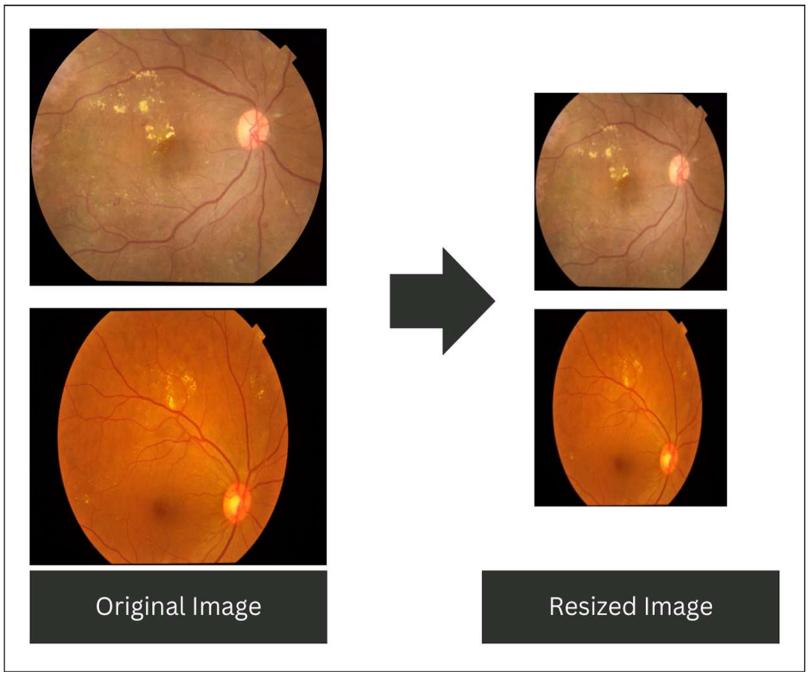

Figure 3.2.2.1: Pre-processed Image

3.2.3 Data Augmentation

Data augmentation is crucial in the field of deep learning models, particularly in the development of models for visual object recognition. Data augmentation is an effective method for addressing the problem of inadequate training data. Data augmentation refers to the practice of using modified real-world samples to expand the size of training or testing datasets [56]. This method bolsters deep learning models by enlarging the training dataset with new images generated from the original ones, thereby improving the model's ability to generalize and mitigate overfitting. The effectiveness of ML algorithms often increases with more data, even if it's of lower quality, as long as the model can extract useful information from the dataset [57]. With limited datasets, harsh regularization or simpler architectures are needed, but data augmentation techniques have shown great success in image analysis and classification [58]. Deep learning models require a large dataset to improve performance, prevent overfitting, and facilitate advanced feature extraction [59, 60]. An additional concern is the preservation of picture quality, which is of utmost importance in the medical domain to preserve critical image characteristics [61]. As a result, the dataset was expanded using photometric augmentation techniques.

3.2.3.1 Photometric Augmentation

Photometric augmentation is a method of image processing used to alter or improve the visual characteristics of a digital image [62]. Mathematical operations on visual data adjust the contrast, sharpness, color, and brightness. Photometric augmentation modifies pixel brightness, sharpness, contrast, and color while maintaining the original geometry of the image [64]. Photometric image augmentation alters image pixel content while maintaining spatial structure, focusing on adjusting visual attributes like color, brightness, sharpness, and contrast [64]. This technique enhances image quality and excellence by reducing visual noise and improving readability. Additionally, it can create multiple versions of the same image with different characteristics, enabling users to quickly select their preferred variant. In this study, four photometric augmentation methods were applied to increase images in the dataset.

A. Brightness Altering

Brightness refers to the overall lightness or darkness in an image, adjusted in photometric data augmentation to simulate diverse lighting conditions and enhance dataset diversity [65]. Equation (1) has been incorporated to enhance image quantity and adjust brightness levels.

$$Brightness\ (x) = Source\ (x) + factor \qquad (1)$$

The brightness of the image at position (x) is determined by adding a factor to the brightness of the source image at the same position. When the value of the factor is greater than 1, the image will be brighter, and when it is less than 1, the image will be darker.

B. Contrast Altering

Contrast refers to the difference in brightness between pixels within a region of interest (ROI) and those in the background of an image. When contrast increases, light areas become lighter and dark areas darker. In photometric augmentation, adjusting contrast involves modifying the difference between the bright and dark parts of an image. This alteration helps vary the details, thereby expanding the range of training data available for machine learning models [65]. Equation (2) has been incorporated to enhance image quantity and adjust contrast levels.

$$Contrast\ (x) = Source\ (x) + factor \qquad (2)$$

The contrast of x is defined as the source value of x with a specific factor added to it. When the value of the factor is greater than 1, the contrast will increase, and when it is less than 1, the contrast will decrease.

C. Sharpness Altering

Sharpness in photography depends on how well a system preserves details, which is crucial for photo quality. Tone zone boundaries are key factors affecting sharpness. Adjusting overall brightness or darkness, termed brightness alteration in photometric augmentation, diversifies datasets and simulates varied lighting conditions for machine-learning tasks [65]. Equation (3) has been incorporated to enhance image quantity and adjust sharpness levels.

$$Sharpness\ (x) = Source\ (x) + factor \qquad (3)$$

Sharper edges result from applying a sharpening factor above 1, while a factor below 1 softens and blurs the image.

D. Color Altering

Altering the image's color levels is an effective way to enhance it. When performing photometric augmentation, adjusting and enhancing color attributes helps deep learning models better adapt to changes in color and lighting, thereby improving their capability [66]. Equation (4) has been incorporated to enhance image quantity and perform color alteration.

$$Color\ (x) = Source\ (x) + factor \qquad (4)$$

When the value of the factor is greater than 1, the color intensifies, and when it is less than 1, it becomes weakened.

Table 3.2.3.1: Detailed Information on Image counts in the original and augmented dataset

Class	Number of Images			
	Without Augmentation		With Augmentation	
	Mendeley dataset	Private dataset	Mendeley dataset	Private dataset
Retinitis Pigmentosa	139	47	695	235
Retinal Detachment	125	92	625	460
Myopia	500	198	2500	990
Disc Edema	127	86	635	430
Diabetic Retinopathy	1509	281	3000	1405
Glaucoma	1349	315	2700	1575
Macular Scar	444	211	2050	1055
Healthy	1024	65	2600	325
Central Serous Chorioretinopathy	101	06	505	30

Table 3.2.3.1 explaining the description of both datasets including image classes with their actual number and number after augmentation.

3.3 Hardware and Software Requirements

The research project "EDNet-20: An Eye Disease Classification Approach Combining CNN and Hybrid Attention" requires a substantial amount of resources in the areas of software, hardware, data gathering, model validation, ethical consideration, documentation, and reporting.

Need in Hardware

Access to a high-powered GPU (NVidia GeForce GTX 1080) for use in deep learning model training; other high-performance computing resources are available.

Storage and Servers: Huge backup solutions and PCIe 4.0 M.2 NVMe SSDs (Samsung 980 Pro 500GB) are available.
Desktop Computers: PCs with powerful processing capabilities used for data preparation, model building, and evaluation.

Needed Programs and Files
Windows 11 is the OS.
Programs for Deep Learning: TensorFlow and PyTorch; for Convolutional Neural Network (CNN) Model Development, Training, and Evaluation.
Python provides an adaptable and widely used environment for running deep learning algorithms and analyzing data.
Data analysis tools and libraries: NumPy, OpenCV, and Pandas, among others.
Data visualization frameworks such as Matplotlib and Seaborn

Gathering and Preparing Data
Collaborating with healthcare organizations to get datasets, including Mendeley and clinical datasets.
Labels for eye disease categorization have been attached to the datasets.
Data preprocessing includes tasks such as dividing the data into test, validation, and training sets and enhancing and normalizing images.

Training and Evaluation of Models
To facilitate transfer learning, make use of pre-trained models such as ResNet50, Xception MobileNet, DenseNet201, DenseNet169, VGG19, and MobileNetV2.
Criteria for Assessment: Precision, Accuracy, Recall, F1-score, and Specificity are the detection metrics.
The generalizability of the model may be checked by doing k-fold cross-validation.

Factors related to ethics and regulations
Get the go light from the appropriate ethical committees.
Ensuring data protection policies are in place to preserve patient information is crucial for data privacy.
Whenever possible, be sure to get people's informed permission.

Reporting and Documentation
Documentation: Keep detailed records of all procedures, programs, and outcomes.
Writing: Get your reports and articles ready to be published.

Software Archive: When working with code, use a version control system such as Git (GitHub).

Thoroughly attending to these implementation needs will allow the research project on deep learning for eye disease categorization to be carried out, producing results that are both robust and useful to clinical practice.

3.4 Project Management and Financial Analysis

Both the planning and finance stages of a research project are crucial. The efficacy of a study and its impact on the progression of knowledge depends on the researchers' adeptness in effectively overseeing the project's activities and financial resources. Create a complete project plan outlining each stage, including data collection, model selection, training, and evaluation. Include specific due dates for each step. Identify the essential resources for the project and allocate them appropriately. Identify any risks and uncertainties that might impact the final result of the project. Achieving success in a project requires the implementation of a well-defined strategy and the continuous monitoring of progress according to that plan. The study's total budget plan would include expenses such as personnel expenditures, software and hardware purchases, data collection, and managerial expenses. By integrating project management and finance ideas, the project may enhance its efficiency, accountability, and financial sustainability, resulting in increased value creation and a more impactful outcome.

The projected budget for the project varies from BDT 5,800 to BDT 10,500, encompassing crucial elements such as expert consultations (BDT 2,000-3,000) with healthcare professionals, necessary software and tools (BDT 1,000-2,000) for implementing models, and data collection and processing (BDT 1,000-2,000) from local hospitals. In addition, expenditures related to documentation and report writing, ranging from BDT 1,000 to BDT 2,000, will be incurred to guarantee thorough reporting of research results. Furthermore, incidental costs, ranging from BDT 800 to BDT 1,500, will be allocated to cover unanticipated expenses such as travel and communication. The project intends to successfully fulfill its goals and contribute to the progress of analytics in healthcare and improvement of patient outcomes by maintaining a clear and organized financial plan. Periodic progress evaluations and financial audits will be carried out to ensure that the project remains on schedule and within the allocated budget shown in table 3.4.1.

Table 3.4.1: Budget Estimation

S No.	Sector	Projected Cost (BDT)
1	Visit Professionals and Healthcare	2000-3000
2	Software and Tools	1000-2000
3	Data Collection and Processing	1000-2000
4	Documentations	1000-2000
5	Others	800-1500
	Total Projected	5800-10500

3.5 Summary

The proposed methodology for eye disease classification uses advanced deep learning techniques, starting with dataset collection and image preprocessing. Convolutional Neural Networks and other deep learning models will be evaluated for their effectiveness in classifying eye diseases. Cross-validation techniques will ensure models generalize well to unseen data, and the most effective approach will be identified.

CHAPTER 4

Implementation

4.1 Overview

The process of developing, training, and evaluating classification models is explained in the implementation section of the eye disease categorization study. It offers a comprehensive understanding of research objectives by including the process of data collection through to execution. Crucial processes include data collection and preparation, model development, training and validation, and ultimately, clinical integration and deployment.

4.2 Train Models

The Model Training section outlines the methods and processes used to develop and refine deep learning models for various eye diseases. It covers model architecture, training procedures, and evaluation metrics. The model architecture determines their effectiveness, and the dataset is divided into training, validation, and test sets for evaluation. Parameters like learning rate, batch size, and epoch number are adjusted to optimize performance. Loss functions and optimization algorithms are employed to minimize the difference between predicted and actual labels.

4.2.1 Transfer Learning Models

This study evaluates the performance of several Transfer Learning models when trained on the same datasets, while also considering the training duration, to determine the effectiveness of the proposed strategy.

I. VGG16

VGG16 is notable for its consistent use of 2x2 max-pooling layers with a stride of 2, and it favors 3x3 convolution layers with a stride of 1 throughout its architecture. This uniform arrangement of convolution and pooling layers extends across the entire model. It also includes two fully connected (FC) layers and a SoftMax layer for final output. The "16" in VGG16 refers to the total number of layers, including weights [67, 68].

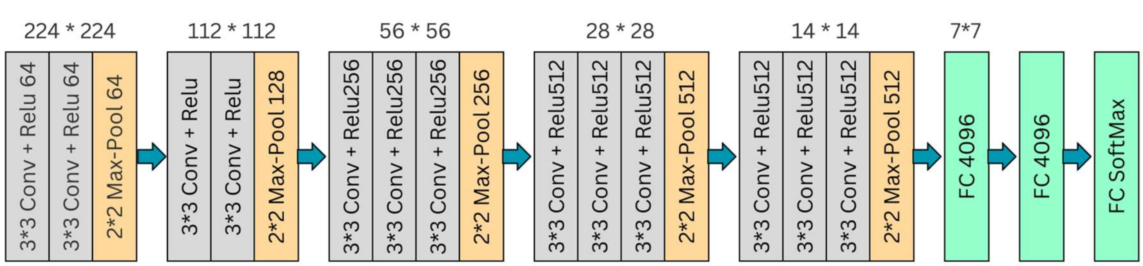

Figure 4.2.1.1: Architecture of VGG16

II. VGG19

The VGG architecture was initially designed for image recognition tasks. In VGG19, it employs 19 weight layers with a reduced 3 × 3 convolutional filter size. This model achieved top ranks in the ILSVR (ImageNet) competition of 2014, securing both first and second places. It operates with a fixed input image size of 224 × 224 and is trained on the extensive ImageNet dataset, comprising millions of images [68].

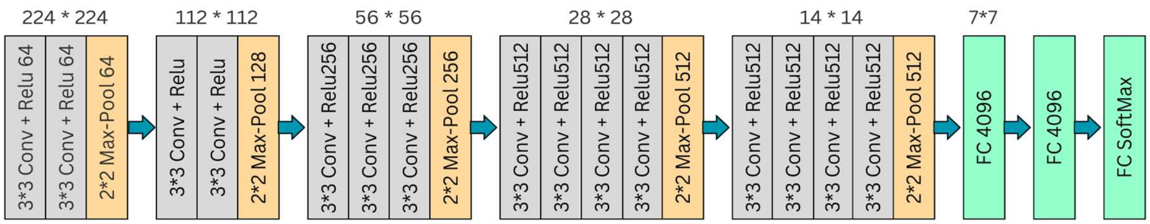

Figure 4.2.1.2: Architecture of VGG19

III. DenseNet201

DenseNet-201 is a convolutional neural network based on DenseNet, featuring 201 layers. DenseNet models are characterized by dense connections between layers, where each layer gets input from all previous layers. This design promotes feature reuse, reduces parameter counts, and improves gradient flow, enhancing both training efficiency and performance. DenseNet-201 aims for high accuracy in image classification, typically trained on large datasets such as ImageNet [68, 69].

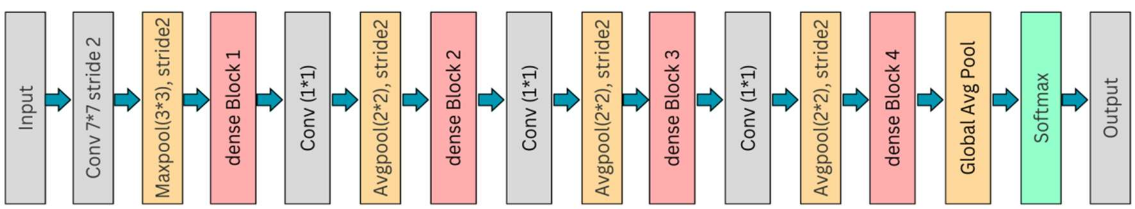

Figure 4.2.1.3: Architecture of DenseNet201

IV. DenseNet169

DenseNet-169 is an advanced convolutional neural network with 169 layers. It builds on the DenseNet model by incorporating dense connectivity, where each layer gets input from all preceding layers. This design promotes efficient feature reuse, reduces parameters, and improves gradient flow. These benefits enhance training effectiveness and performance, making DenseNet-169 ideal for high-accuracy tasks like image classification on large datasets such as ImageNet [68, 69].

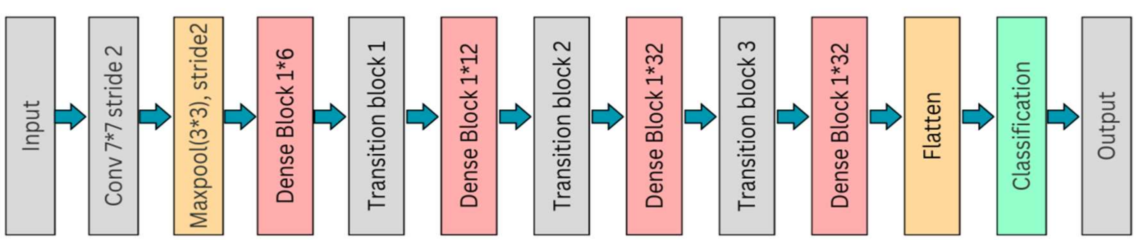

Figure 4.2.1.4: Architecture of DenseNet169

V. MobileNet

MobileNet is a CNN architecture made for mobile and embedded systems, focusing on efficiency. It uses depthwise separable convolutions to lower parameters and computation, yet maintains accuracy. Its lightweight design suits tasks like real-time object detection and image classification on devices with limited hardware [70].

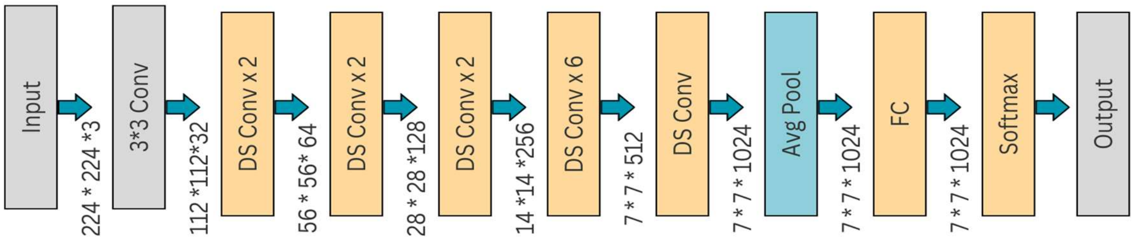

Figure 4.2.1.5: Architecture of MobileNet

VI. MobileNet V2

MobileNet v2 is an improved version of the MobileNet architecture, further optimizing performance for mobile and embedded devices. It introduces inverted residuals and linear bottlenecks to enhance efficiency while maintaining accuracy. MobileNet v2 is designed to be even more lightweight and suitable for tasks requiring real-time processing on hardware-constrained devices like smartphones and IoT [71].

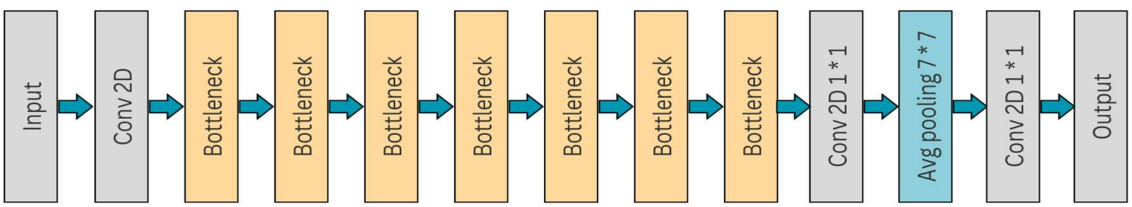

Figure 4.2.1.6: Architecture of MobileNetV2

VII. ResNet50

ResNet-50, a 50-layer convolutional neural network, is a member of the ResNet family known for its deep structure and residual connections. These connections address the vanishing gradient issue by providing direct paths for gradient flow during training. ResNet-50 is extensively applied in computer vision tasks such as image classification and feature extraction from large datasets like ImageNet, offering efficient training and high performance in deep learning tasks [72].

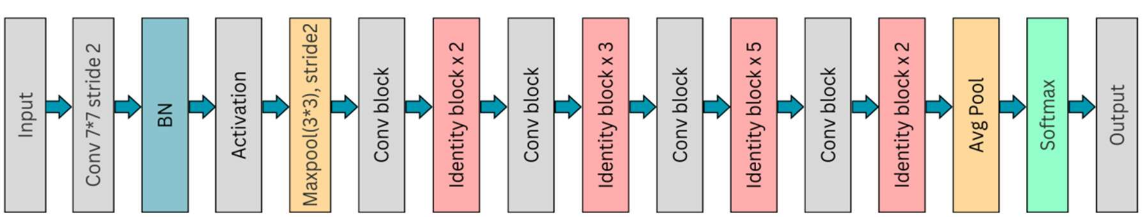

Figure 4.2.1.7: Architecture of ResNet50

VIII. Xception

Instead of the usual convolutional layers, Xception improves convolutional neural networks by adding depthwise separable convolutions to each node. This strategy is based on MobileNet's research. This strategy has the potential to improve its ability to identify complex patterns in data by streamlining the procedure and minimizing the number of parameters. Xception's optimizations considerably enhance computer vision applications, such as photo classification and object detection [73].

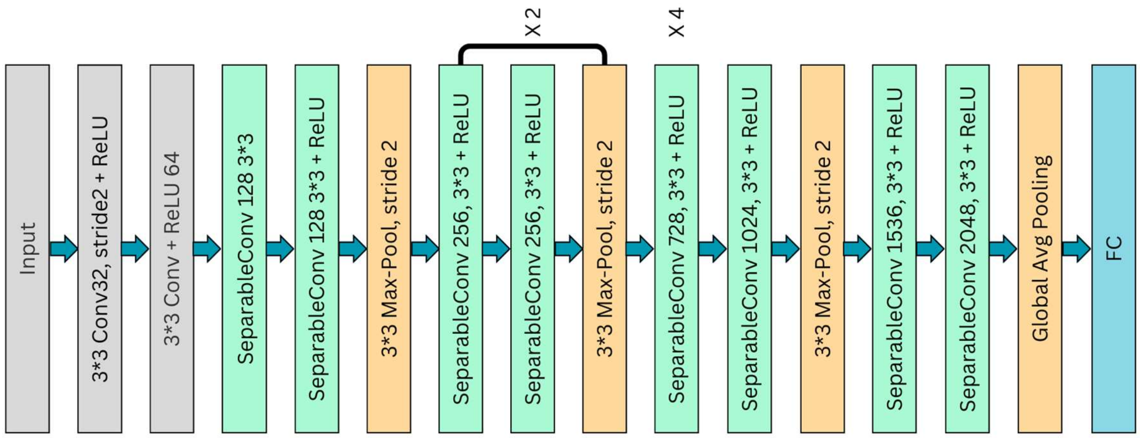

Figure 4.2.1.8: Architecture of Xception

4.2.2 Hybrid Attention Mechanism Block

The significance of the Attention Mechanism in deep learning is widely recognized [74]. However, traditional standalone Attention Mechanisms face limitations, such as difficulty in effectively capturing valuable features or information. To address these challenges, this research employs a Hybrid Attention Mechanism Block. The Hybrid Attention Mechanism block comprises two elements: (A) Channel Self Attention Block and (B) Spatial Self Attention Block. The Hybrid Attention Module consists of three parallel Channel Self Attention Blocks, known as Selective Kernel Networks (SKNets) [75, 76] and Spatial Attention Mechanism (SAM) [76]. Figure 4.2.2.1 shows the Hybrid Attention Mechanism Block architecture.

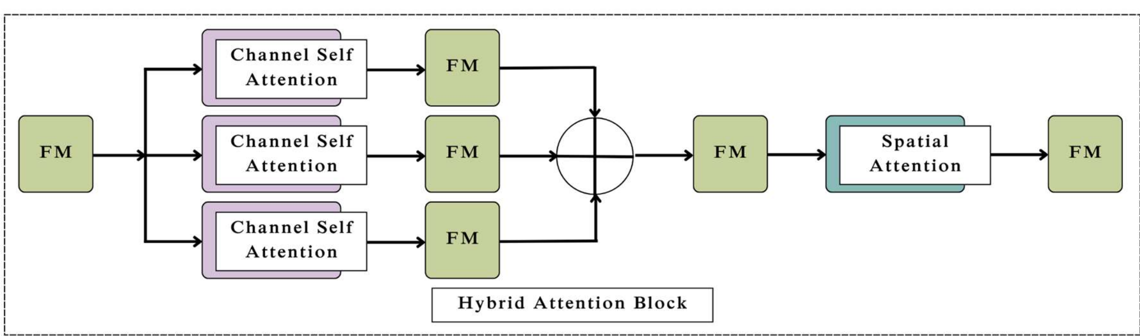

Figure 4.2.2.1: Hybrid Attention Block Architecture.

A. Channel Self-Attention Block

Channel-wise attention performs better by adjusting the importance of different channels within a feature map, allowing it to extract crucial channel-specific information more effectively [77]. Self-attention is a mechanism within a sequence that connects different positions to encode data, using importance scores to enhance the modeling of long-range dependencies [78]. Figure 4.2.2.2 shows the Channel Self Attention Mechanism Block architecture.

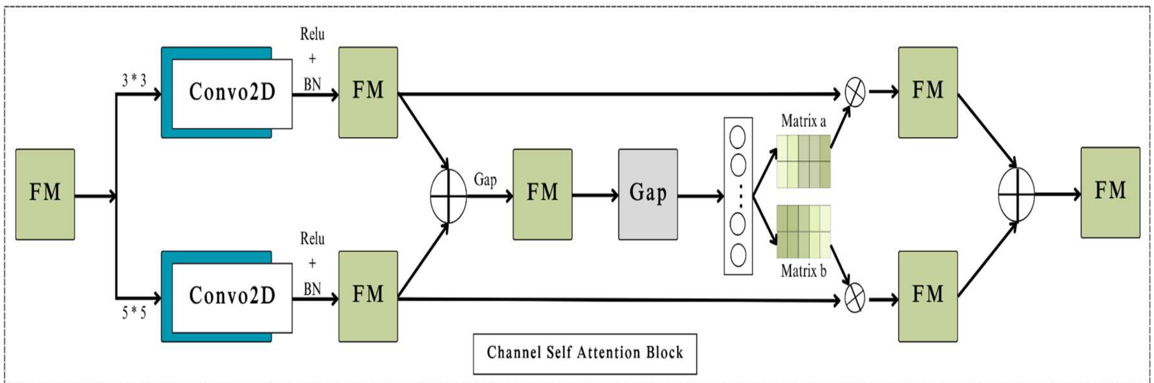

Figure 4.2.2.2 Channel Self Attention Block Architecture.

B. Spatial Self-Attention Block

Spatial attention functions as an adaptive mechanism for selecting specific regions in space to focus on and deciding where to direct attention [79]. Channel attention is focused on voxel interactions and task relationships in a small area, while spatial attention is focused on integrating information about where things are in space on a larger scale, creating complex links between local features that are given different amounts of weight [80]. Figure 4.2.2.3 shows the Spatial Attention Mechanism Block architecture.

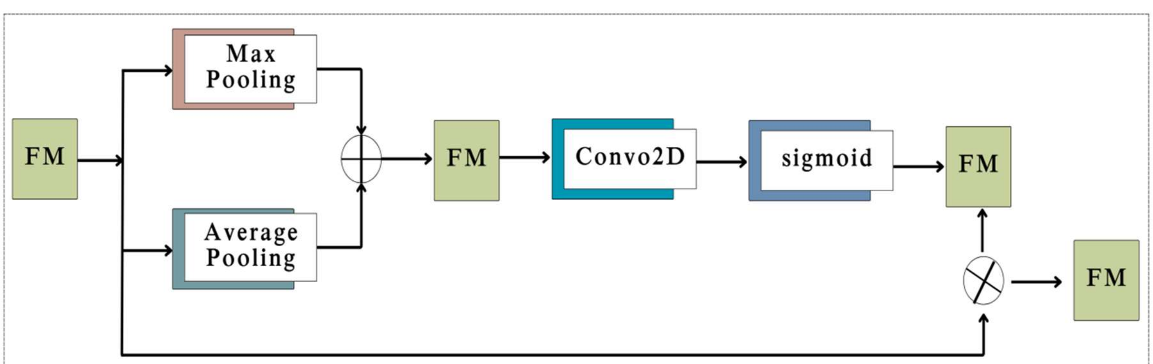

Figure 4.2.2.3 Spatial Self-Attention Block Architecture.

4.2.3 Base CNN Model

A Base CNN (Convolutional Neural Network) model was initially proposed, suitable for use in the deep learning process. Figure 4.2.3.1 shows a baseline CNN model architecture.

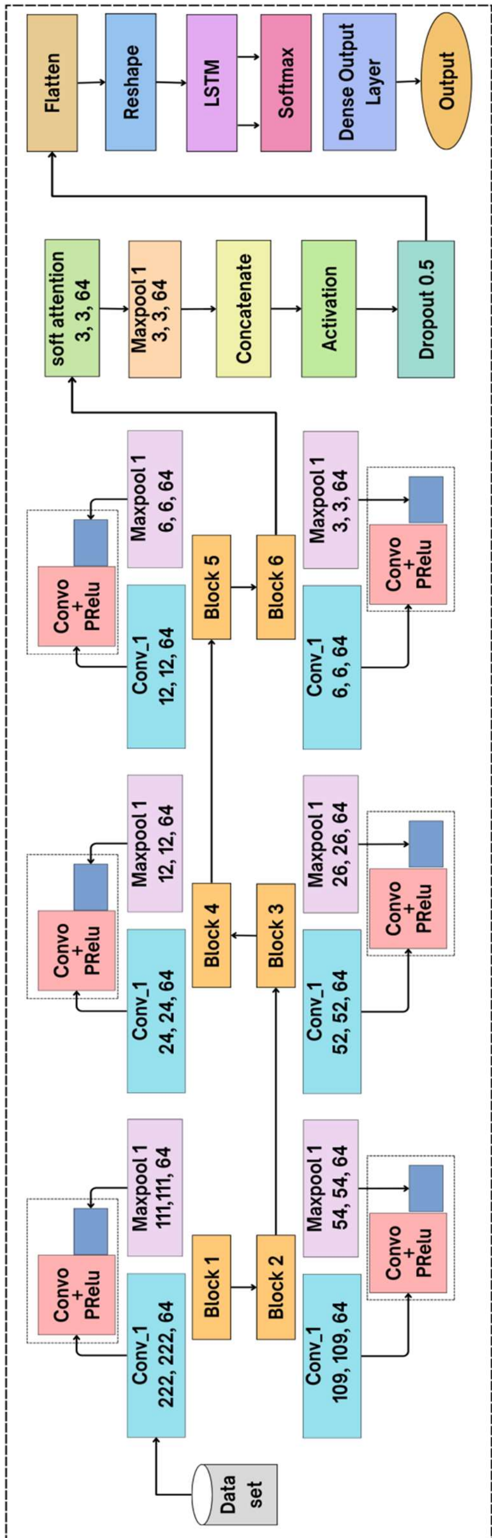

Figure 4.2.3.1: Baseline CNN model architecture.

The base CNN model consists of six blocks, each containing a convolutional layer followed by a max-pooling layer. The convolutional layer is the most crucial component in CNN architecture, as it extracts features from the input data [81].

$$conv\ (x, y) = (a \times b)[x, y] \sum i\ \sum j\ \ a[i, j] b[x - i, y - j] \qquad (5)$$

Where *conv(x,y)* is the output of the following layer, *a* is the input image, the kernel is represented as *b*, and the size of the kernel is defined by *i,j*, and × is the convolution operation.

The initial configuration includes a convolutional kernel size of 5 x 5 and a dropout of 0.7. Each block utilizes 64 convolutional kernels. The convolutional layer generates feature maps, which show the output of the previous layer after applying the filter.
After convolution, a pooling layer picks out the most important values from the hidden units. Its goal is to reduce the number of features while keeping important data and getting rid of unnecessary ones by lowering the resolution in the layer below it [82].

The base model uses PReLU as its activation function, and softmax is employed as the final layer activation function. With 5 x 5 filters and the rectified linear unit (PReLU) as its activation function, each convolutional layer effectively pulls out information. PRelu activation can fix problems with regular activation functions, making it easier for deep convolutional neural networks to fit models and find images [83].

4.2.4 Model Optimization

Model optimization involves enhancing the performance of a machine learning model by adjusting several parameters to get superior outcomes. Within the realm of deep learning, specifically in the context of Convolutional Neural Networks (CNNs), several tactics and approaches are used.

4.2.5 Architectural Optimization

Creating an optimal model by modifying the architecture and hyperparameters of a base CNN model involves several crucial steps. These adjustments are designed to enhance the model's performance for specific tasks, such as classifying fundus images in this scenario. In the process of architectural optimization, adjustments are made to the convolutional layers, maxpool layers, attention module, and LSTM layer to develop the most efficient and optimal model architecture. A convolution layer extracts features from input data in CNNs, while a pooling layer reduces dimensionality by extracting patches from feature maps [84]. Pooling layers down-sample feature maps, preserving essential information while reducing computational complexity and ensuring spatial invariance [85, 86]. Various configurations of convolutional and maxpool layers result in different levels of accuracy. The adjustment of these layers aims to achieve the optimal model that offers the highest accuracy. Attention has significantly improved computer vision systems and deep neural network understanding by emphasizing relationships among elements within sequences through self-attention mechanisms [87, 88]. It enables models to prioritize critical regions, such as the Region of Interest (ROI), to achieve peak accuracy and efficiency [89]. This study used several attention mechanisms, such as soft attention, channel attention, spatial attention, and a new hybrid attention module, to find and choose the most efficient model design. After identifying the ideal attention module, the suitability of LSTM was evaluated to define its design. This was done by utilizing its capacity to handle both spatial and sequential information, which ultimately improved the predictive capabilities of the model [90, 91, 92].

4.2.6 Hyperparameter Optimization

Following the identification of the most appropriate model architecture, this study conducted a hyperparameter optimization to refine and optimize the model's parameters. The size of a model's convolutional kernel affects its ability to capture features. Smaller kernels improve detail and efficiency, whereas bigger kernels capture wider features but increase parameter complexity, which may reduce model performance [93]. ReLU has become the dominant activation function in deep learning because of its simplicity and improved performance. This is important for ensuring that the complexity of the neural network matches the complexity of the dataset to prevent overfitting or under-convergence during training [94, 95]. It is often recommended to use smaller batch sizes and moderate learning rates to improve convergence and achieve greater generalization during the training of a network [96, 97, 98]. By making changes to the flattened layer, the network's ability to connect convolutional features to fully connected layers is improved, which leads to better design and better model performance [99, 100]. Changing the loss function and optimizer may have a significant influence on both the training process and the model's ability to achieve an optimum solution [101].

4.2.7 Proposed EDNet-20 Model

EDNet-20 has been developed through architectural and hyperparameter optimization within the Bayes CNN framework. Figure 4.2.7.1 depicts the EDNet-20 model architecture.

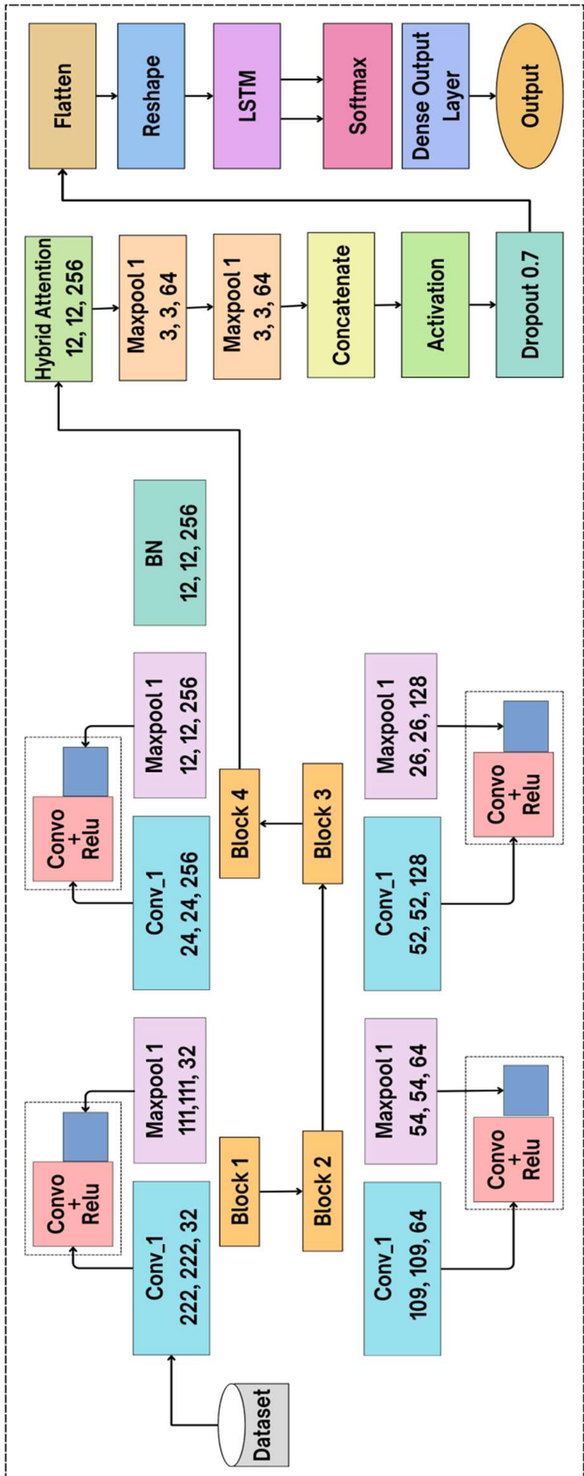

Figure 4.2.7.1: Proposed EDNet-20 model architecture.

The proposed EDNet-20 model consists of four blocks, each containing a total of twenty layers. In order to reduce the dimensions of the network's feature maps and capture significant information, each block starts with a convolutional layer succeeded by a max-pooling layer. In order to enhance the stability of the network's learning process, a batch normalization layer is included after the max-pooling layer in the fourth block. ReLU is used as the activation function for each convolutional layer, with a kernel size of 3 x 3. ReLU is often used as the activation function in Convolutional Neural Networks (CNNs) due to its reduced computing expense. It converts all input values into non-negative integers. The rectified linear unit (ReLU) activation function is often considered the most common in the field of deep learning [102]. The first four blocks consist of kernels in the following order: 32, 64, 128, and 256. The optimization process utilizes Adam with a batch size of 32.

The model is trained using a learning rate of 0.001. The Hybrid Attention module consists of two maxpool layers and is combined with the fourth block. The concatenate layer is used to integrate several signals. The activation layer is tasked with converting the feature maps and producing the state vectors for the model. To enhance the model's performance and capacity to generalize, it may be enhanced by including a regularization layer with a rate of 0.5. The subsequent action involves using the flatten and reshape layers to seamlessly merge the output of the model into a unified vector. Subsequently, using deep neural networks with an LSTM layer is the subsequent course of action to rectify faults. The ultimate result is generated by standardizing and arranging the model's output via the use of softmax and dense layers.

4.3 Model Evaluation

Model assessment is an essential stage in research, with the goal of assessing the effectiveness and dependability of the models created for categorizing different eye disorders. This section provides an overview of the fundamental principles of metrics used to assess the effectiveness of these models.

4.3.1 Evaluation Metrics

In order to evaluate the performance of the proposed model, many metrics have been assessed. A true positive (TP) is when the model accurately detects a positive instance. A true negative (TN) occurs when the model accurately detects and classifies a situation as negative. False positives (FP) and false negatives (FN) occur when the model makes erroneous predictions of positive and negative instances, respectively. Accuracy refers to the proportion of accurate predictions produced by the model, expressed as a percentage.

Accuracy

Measures the overall correctness of the model. It is the ratio of correctly predicted observations (both true positives and true negatives) to the total observations.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision

Measures the accuracy of positive predictions. It is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$precision = \frac{TP}{TP+FP} \quad (7)$$

Recall

Measures the fraction of actual positives that are correctly identified. It is the ratio of correctly predicted positive observations to all observations in the actual class.

$$recall = \frac{TP}{TP+F} \quad (8)$$

Specificity

Measures the accuracy of negative predictions. It is the ratio of correctly predicted negative observations to the total predicted negative observations.

$$specificity = \frac{TN}{TN+FP} \quad (9)$$

F1 Score

The harmonic mean of precision and recall. It provides a single metric that balances both precision and recall.

$$f1\ score = 2 \times \frac{precision \times recall}{precision+rec} \quad (10)$$

4.4 Summary

The implementation phase focused on developing deep learning models, specifically Convolutional Neural Networks (CNNs) for image classification tasks. Various architectures like VGG16, VGG19, DenseNet201, DenseNet169, ResNet50, Xception MobileNet, and MobileNetV2, were trained and tested to determine the most suitable model. The models were trained on two datasets, with each image annotated according to the specific eye disease it represented. A stratified k-fold cross-validation approach was used to ensure robustness and generalizability. Hyperparameters were tuned, and techniques like dropout and batch normalization were applied to prevent overfitting. The trained models were evaluated using performance metrics like accuracy, precision, recall, and F1-score. The best-performing models demonstrated high accuracy and robustness.

CHAPTER 5
Result and Analysis

5.1 Overview

The Results and Analysis section presents the research findings. It evaluates developed models, their performance metrics, and implications. The section includes a comparative analysis of different algorithms, hyperparameter optimization efforts, common patterns in misclassified images, cross-validation experiments, and performance evaluation on independent test sets to ensure model robustness and generalizability.

5.2 Experimental Results

The experimental results section evaluates eye disease classification models, highlighting performance metrics and outcomes from diverse dataset testing.

5.2.1 Performance Analysis of the Proposed Model

After completing the model optimization on the base CNN model, the robust final EDnet-20 model has been generated, and that classification performance is considerably enhanced. Tables 5.2.1.1 and 5.2.1.2 present the statistical analysis for the EDNet-20 model, such as precision, recall, and f1 score.

Table 5.2.1.1: The performance of EDNet-20 model on the Mendeley dataset

Dataset	Class	Precision	Recall	F1-score	Overall Accuracy
Mendeley	CSCR	73.28	95.05	82.76	**94.26**
	DR	97.65	93.51	95.53	
	DE	75.32	93.70	83.51	
	GC	98.07	94.37	96.18	
	MS	90.62	95.72	93.10	
	MP	93.16	95.40	94.27	
	RD	81.69	92.80	86.89	
	RP	78.82	96.40	86.73	
	NM	96.78	93.85	95.29	

Table 5.2.1.2: The performance of the proposed model on the private dataset

Dataset	Class	Precision	Recall	F1-score	Overall Accuracy
Private	CSCR	23.53	66.67	34.78	**91.80**
	DR	98.88	94.66	96.73	
	DE	91.95	93.02	92.49	
	GC	98.34	95.18	96.73	
	MS	96.59	93.84	95.19	
	MP	93.94	93.94	93.94	
	RD	92.47	93.48	92.97	
	RP	75.41	97.87	85.19	
	NM	89.47	91.89	90.67	

Tables 5.2.1.1 and 5.2.1.2 present a comprehensive assessment of the performance of the proposed model, EDNet-20, across two distinct datasets: Mendeley and Private. The model consistently delivers outstanding results across various classes within these datasets, reflected in high accuracy, recall, and F1-score values that approach or reach one. This indicates a strong ability to accurately identify samples for each dataset and class. Additionally, the test accuracy spans from 91.80% to 94.26%, showcasing the model's robustness and reliability. In the below figure, 5.2.1.1 shows the confusion matrix of the proposed model in different modality datasets.

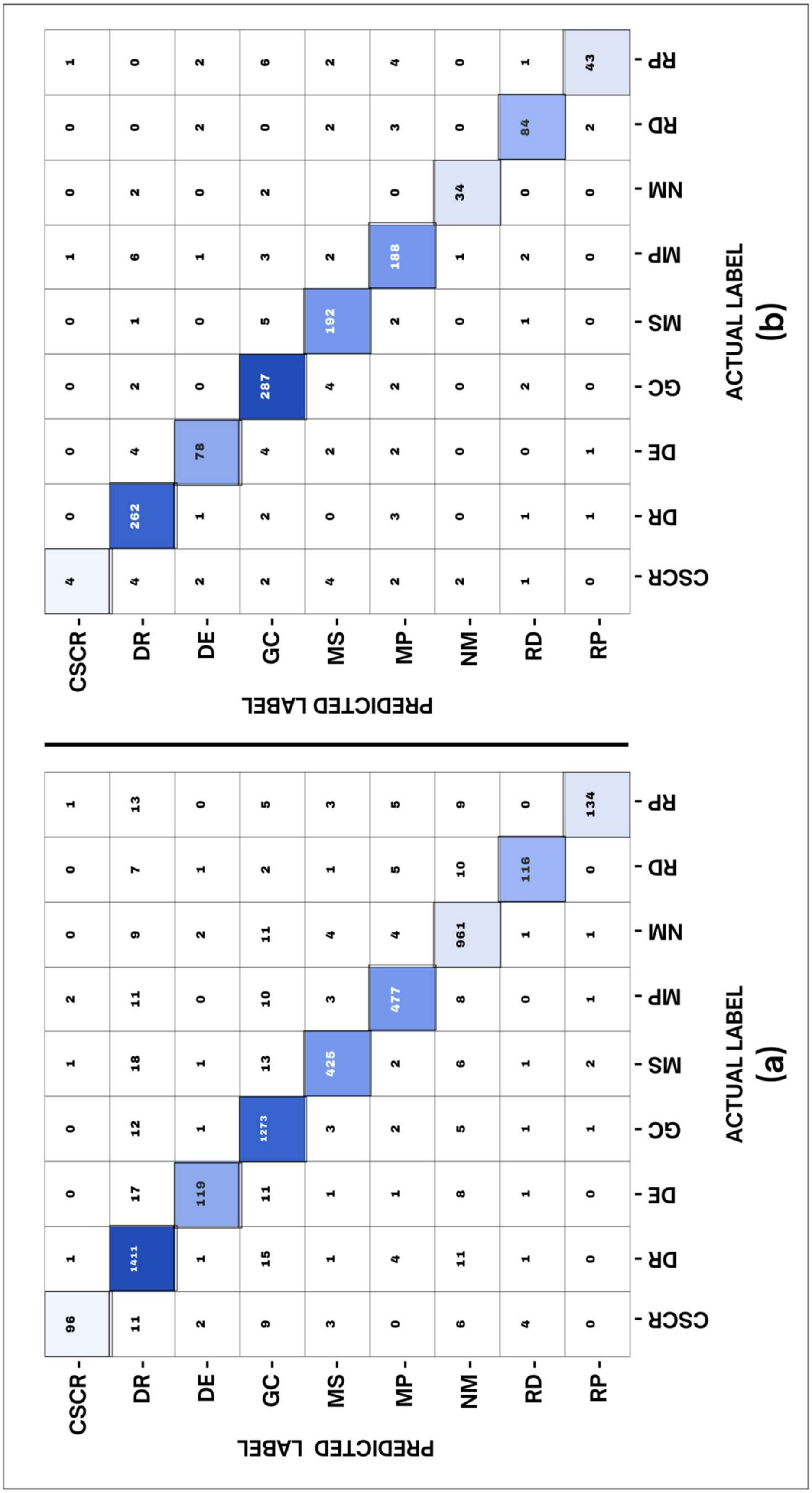

Figure 5.2.1.1: Confusion Matrix of EDNet-20

Figure 5.2.1.1 shows the confusion matrix of datasets (a) Mendeley and (b) Private dataset of the proposed model. The performance of an algorithm may be seen by the specific table layout. While the x-axis values represent the predicted label of the test images and the y-axis values represent the actual label of the test images, the diagonal values denote the True Positives.

5.2.2 Analysis of Accuracy and Loss

In deep learning, accuracy and loss for training, testing, and validation datasets are vital indicators of a model's performance and its ability to generalize. Training Accuracy reflects the percentage of correctly predicted instances in the training dataset, while Training Loss measures how accurately the model predicts the training data, showing the difference between predicted and actual values. Validation and test accuracy and loss operate similarly in evaluating the model's performance on new data. Table 5.2.2.1 shows the performance of the proposed model.

Table 5.2.2.1: Values of Accuracy and Loss

Dataset	Train Accuracy	Train Loss	Test Accuracy	Test Loss	Validation Accuracy	Validation Loss
Mendeley	99.64	0.0559	94.26	3.0970	96.06	2.4560
Private	99.64	0.0559	91.80	4.0140	96.06	2.4560

Figure 5.2.2.1 Visual Representation of Accuracy and Loss

Figure 5.2.2.1 illustrates the accuracy and loss of training, testing, and validation for two datasets: Mendeley (a) and Private (b). Analysis of the EDNet-20 model's performance reveals consistent accuracy between training and validation across both datasets. However, variations are observed in the model's testing accuracy and loss: achieving 94.26% accuracy with a loss of 3.0970 on the Mendeley dataset and 91.80% accuracy with a loss of 4.0140 on the private dataset.

5.2.3 K-Fold Cross Validation

During cross-validation, the dataset is divided into K folds in order to assess the model's performance using fresh data. The research used K-fold cross-validation to evaluate the suggested model. This validation approach is performed by using both training and test datasets, where each fold acts as a distinct data segment for every k training and validation iteration. This methodology facilitates the comprehension of how randomness and variable bias impact accuracy, which is measured by the difference between the actual and projected accuracy. Figure 5.2.3.1 shows K Fold Cross Validation accuracy.

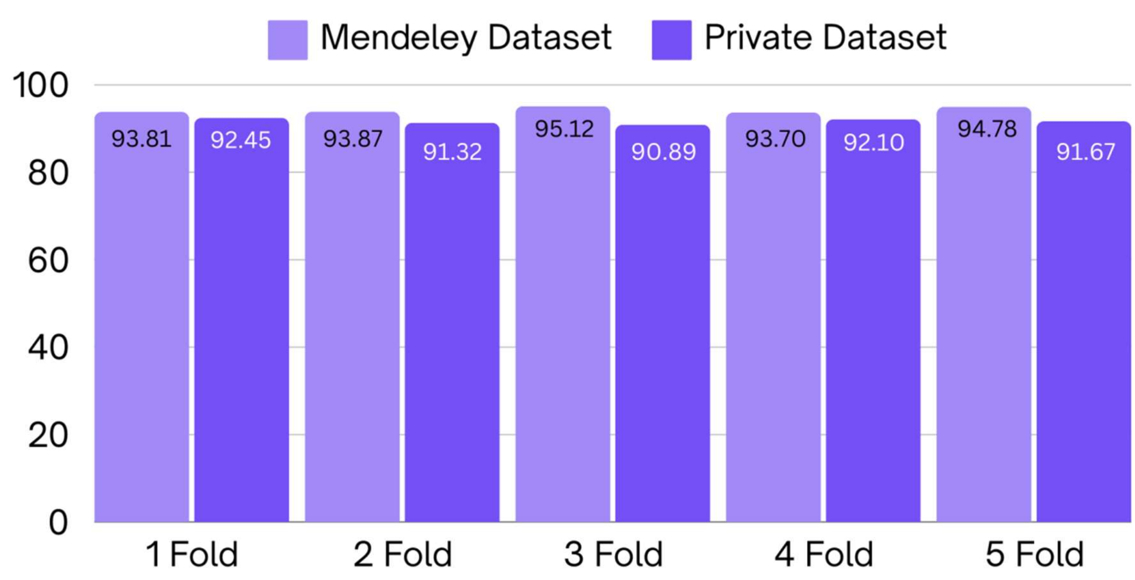

Figure 5.2.3.1: K Fold Cross Validation

This study investigates the effects of using different K-fold cross-validation numbers, specifically 1-fold, 2-fold, 3-fold, 4-fold, and 5-fold values. As a result, the testing accuracy fluctuated among 93.81%, 93.87%, 95.12%, 93.70%, and 94.78% on the Mendeley dataset and 92.45%, 91.32%, 90.89%, 92.10%, and 91.61% on the private dataset. Given the notable increase in accuracy observed with 5-fold cross-validation in the study, this method seems to offer the most reliable estimate of the model's performance.

5.3 Comparative Analysis

This section assesses and compares the efficacy of several deep learning models in the detection and classification of eye illnesses. It emphasizes their advantages and disadvantages to provide a thorough insight into their clinical usefulness.

5.3.1 Comparison with Attention-based Models

The proposed model is tested against the baseline CNN and three other attention-based models to evaluate its performance. The comparison utilizes several metrics, including accuracy, precision, recall, specificity, and F1-score. Although attention-based models are adept at highlighting relevant sections of the input data and often attain high accuracy due to their advanced architecture, the proposed model seeks to enhance these outcomes through optimization techniques. The proposed model demonstrated promising testing accuracy, potentially outperforming the attention-based models in the evaluated scenarios, showcasing its effectiveness and potential benefits in real-world applications.

A. Table 5.3.1.1 represents the comparison between the proposed model and attention-based models on the Mendeley dataset.

Table 5.3.1.1: Comparison between the proposed model and attention-based models on Mendeley dataset.

Model Configuration	Accuracy	Precision	Recall	Specificity	F1 Score
Base CNN	72.52±4.65	72.17±4.48	72.85±4.18	72.95±3.52	71.83±4.98
CNN + Soft Attention	79.91±3.54	81.95±3.45	81.92±3.75	79.96±3.44	81.93±3.25
CNN + Channel Self Attention	90.14±2.12	90.14±2.23	90.94±2.27	90.10±2.56	90.14±2.68
CNN + Spatial Self Attention	90.21±1.45	90.38±1.41	90.21±1.51	90.59±1.33	90.24±1.30
Proposed EDNet-20 (CNN + Hybrid Attention)	**94.26±0.31**	**95.16±0.29**	**94.25±0.08**	**99.29±0.19**	**94.58±0.23**

B. Table 5.3.1.2 represents the comparison between the proposed model and attention-based models on the private dataset.

Table 5.3.1.2: Comparison between the proposed model and attention-based models on Private dataset.

Model Configuration	Accuracy	Precision	Recall	Specificity	F1 Score
Base CNN	75.34±3.89	70.94±5.13	74.12±6.07	71.68±4.83	73.56±6.21
CNN + Soft Attention	78.67±4.23	83.24±4.57	82.58±4.92	80.72±4.65	82.47±4.11
CNN + Channel Self Attention	90.27±3.10	91.38±3.17	91.72±3.45	89.96±3.64	90.48±3.92
CNN + Spatial Self Attention	89.87±2.34	91.52±2.19	90.96±2.44	91.18±2.27	90.75±2.12
Proposed EDNet-20 (CNN + Hybrid Attention)	**91.80±0.24**	**93.03±0.26**	**91.80±0.11**	**98.98±0.25**	**92.27±0.24**

Tables 5.3.1.1 and 5.3.1.2 provide a comprehensive summary of the performance of various models on the Mendeley dataset. The models tested include the Base CNN, CNN + Soft Attention, CNN + Channel Self Attention, CNN + Spatial Self Attention, and the Proposed EDNet-20 (CNN + Hybrid Attention). These models were trained on both Mendeley and private datasets. The accuracy results indicate the effectiveness of each model on these datasets. EDNet-20 consistently surpassed the other models in terms of accuracy, precision, recall, specificity, and F1 score. Notably, CNN + Spatial Self Attention also performed well, consistently achieving accuracy rates above 90% across all datasets. In contrast, CNN + Channel Self Attention showed variable performance, while Base CNN and CNN + Soft Attention had lower accuracies. Overall, the findings demonstrate that EDNet-20 consistently outperforms the other models, achieving an exceptional accuracy of 94.26% and 91.80% on the Mendeley and private datasets, respectively. These results highlight EDNet-20's potential as a highly effective model for eye disease classification.

5.3.2 Comparison with Transfer Learning Models

The proposed model undergoes testing against eight advanced transfer learning CNN models to assess its performance. Evaluation involves metrics such as accuracy, precision, recall, specificity, and F1-score. While transfer learning CNN models are skilled at identifying pertinent aspects of input data and frequently achieve high accuracy because of their sophisticated architecture, the proposed model aims to improve upon these results through optimization techniques. The proposed model has shown promising testing accuracy, potentially surpassing the transfer learning CNN models in the studied scenarios, demonstrating its efficacy and potential advantages in practical applications.

A. **Table 5.3.2.1 shows the comparative analysis of classification results for transfer learning CNN models with the proposed EDNet-20 model on the Mendeley dataset.**

Table 5.3.2.1: Comparison between the proposed model and transfer learning CNN models.

Method	VGG 16	VGG 19	Mobile Net	Mobile NetV2	ResNet 50	DenseNet 201	DenseNet 169	Xcepti on	**EDNet -20**
Accuracy	66.33 ±3.82	81.21 ±4.12	71.92 ±4.15	75.12 ±2.87	78.62 ±3.21	78.43 ±2.49	81.27 ±2.15	81.79 ±2.44	**94.26 ±0.31**
Precision	69.45 ±3.15	80.04 ±3.89	76.03 ±4.72	74.65 ±2.99	78.81 ±3.15	78.16 ±1.97	81.58 ±2.39	81.24 ±1.91	**95.16 ±0.29**
Recall	67.89 ±3.75	81.97 ±4.05	72.37 ±4.35	75.03 ±2.74	78.61 ±3.20	77.98 ±1.94	81.94 ±2.68	81.18 ±1.87	**94.25 ±0.08**
Specificity	69.26 ±3.49	79.86 ±3.67	73.14 ±4.09	75.45 ±3.12	77.04 ±3.92	76.58 ±2.65	80.92 ±2.86	81.17 ±2.74	**99.29 ±0.19**
F1 Score	70.04 ±3.97	80.55 ±3.98	74.26 ±4.01	75.78 ±3.05	78.97 ±3.46	78.09 ±2.05	81.59 ±2.39	81.18 ±1.87	**94.58 ±0.23**

B. **Table 5.3.2.2 shows the comparative analysis of classification results for transfer learning CNN models with the proposed EDNet-20 model on the Private dataset.**

Table 5.3.2.2: Comparison between the proposed model and transfer learning CNN models.

Method	VGG 16	VGG 19	Mobile Net	Mobile NetV2	ResNet 50	DenseNet 201	DenseNet 169	Xcepti on	**EDNet -20**
Accuracy	63.48 ±3.21	77.43 ±3.78	70.28 ±3.92	72.65 ±3.34	75.84 ±3.78	74.95 ±2.87	80.12 ± 2.55	79.22 ±2.21	**91.80 ±0.24**
Precision	66.72 ±2.87	76.12 ±3.42	77.19 ±4.38	71.68 ±3.21	82.29 ±3.49	80.18 ±2.37	83.16 ±2.20	78.32 ±2.96	**93.03 ±0.26**
Recall	64.91 ±3.45	77.56 ±3.57	71.02 ±3.98	76.05 ±3.06	81.12 ±3.74	76.04 ±2.31	78.26 ±3.05	82.45 ±2.34	**91.80 ±0.11**
Specificity	66.18 ±3.08	75.96 ±3.21	74.26 ±3.78	78.57 ±3.56	73.92 ±4.37	79.13 ±2.48	82.13 ±3.17	78.12 ±3.12	**98.98 ±0.25**
F1 Score	66.92 ±3.52	76.87 ±3.51	72.59 ±3.64	77.80 ±3.33	76.05 ±4.01	72.93 ±3.05	76.54 ±2.98	84.20 ±2.21	**92.27 ±0.24**

Tables 5.3.2.1 and 5.3.2.2 provide a comprehensive summary of the performance of multiple models on the Mendeley and private datasets. The models tested include VGG16, VGG19, DenseNet201, DenseNet169, MobileNet, MobileNet V2, ResNet50,

Xception, and EDNet-20, all trained on both datasets. Accuracy results indicate the models' performance, with EDNet-20 consistently outperforming others across all metrics. DenseNet169, Xception, and VGG19 also consistently achieved accuracy rates above 80%. In contrast, models like MobileNet V2, DenseNet201, and ResNet50 showed varying performance, and VGG16 and MobileNet demonstrated lower accuracies. Overall, these findings underscore EDNet-20's exceptional performance, consistently achieving high accuracy of 94.26% and 91.80% on the Mendeley and private datasets, respectively. These results highlight EDNet-20's potential for classifying eye diseases.

5.4 Summary

Figure 5.4.1 presents the Transfer Learning CNN model test accuracy comparison with EDNet-20 on the Mendeley dataset.

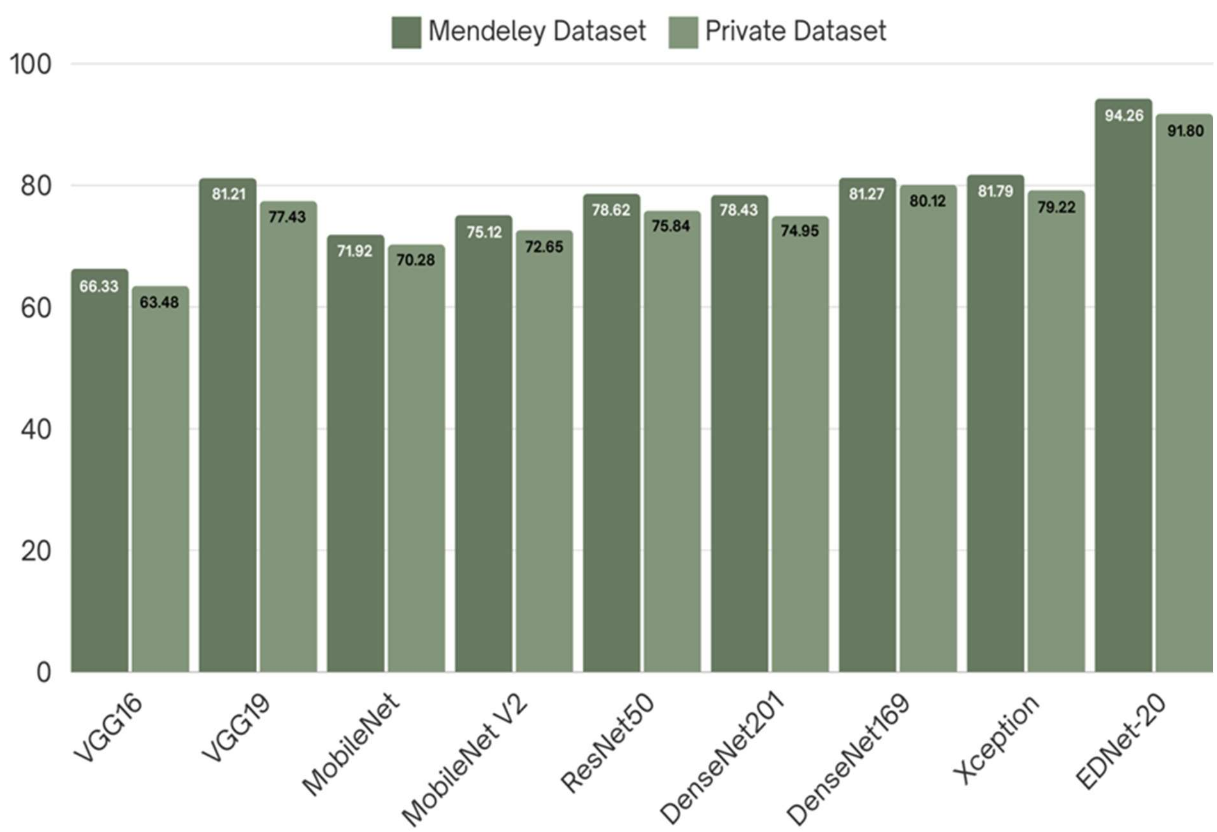

Figure 5.4.1 Transfer Learning CNN model comparison

Figure 5.4.1 presents the Transfer Learning CNN model test accuracy comparison with EDNet-20 on the Private dataset.

This in-depth research investigates the effectiveness of deep convolutional neural networks (CNNs) in categorizing nine different types of fundus images. Utilizing a substantial dataset comprising 6555 original photos from two sources, the study employs advanced augmentation techniques to create five images from each original photograph. Focusing on classifying eye diseases, the research evaluates the performance of a baseline CNN model, and various transfer learning CNN models, and introduces a novel model. The study rigorously tests eight sophisticated transfer learning CNN models, including VGG16, VGG19, DenseNet201, DenseNet169, MobileNet, MobileNet V2, ResNet50, and Xception, alongside the proposed EDNet-20, across the nine fundus image categories. Notably, EDNet-20 emerges as the top performer, achieving an impressive accuracy of 94.26% and 91.80%, as illustrated in figure 5.4.1.

CHAPTER 6
Impact on Society, Environment and Sustainability

6.1 Impact on Life

Machine learning and artificial intelligence are being used in eye disease categorization research to enhance patient treatment and provide social and economic advantages. Precise and timely identification of ocular conditions, such as diabetic retinopathy or glaucoma, may halt the advancement to critical stages, such as loss of vision. Advanced categorization systems have the ability to customize treatment programs for particular patients, resulting in enhanced and individualized care. Timely identification of ocular conditions may avert visual impairment, maintain autonomy, and alleviate emotional and psychological distress in individuals who may encounter anxiety and despair as a result of deteriorating eyesight.

6.2 Impact on Society and Environment

Impact on Society

The research "EDNet-20: An Eye Disease Classification Approach Combining CNN and Hybrid Attention" employs sophisticated artificial intelligence methodologies, including transfer learning and deep Convolutional Neural Networks (CNNs), to enhance medical diagnosis. This methodology improves the accuracy and efficiency of classifying eye ailments, empowers healthcare professionals to make wiser decisions, and results in improved clinical results and societal advantages.

Eye diseases provide substantial obstacles to healthcare systems globally, and the integration of more accurate diagnostic tools may enhance the quality of patient care by minimizing incorrect diagnoses and expediting treatment procedures. EDNet-20's enhanced classification algorithms facilitate the creation of customized treatment regimens for particular eye diseases, resulting in enhanced patient results and satisfaction.

The research also supports the progress of telemedicine technology, which may enable diagnosis remotely in rural areas with limited availability of eye care practitioners. Precise categorization of illnesses may minimize useless interventions and hospitalizations, enhancing the handling of ocular-related problems in society.

The finding has the potential to transform precision medicine in the identification and treatment of eye diseases, improving the accessibility and cost of healthcare, especially in disadvantaged areas with limited medical resources. Emerging technology may be used to provide sophisticated diagnostic tools to regions lacking access to traditional healthcare facilities.

Impact on Environment

The "EDNet-20: An Eye Disease Classification Approach Combining CNN and Hybrid Attention" research initiative aims to improve medical diagnostics by integrating transfer learning and deep convolutional neural networks. These technologies require intensive computational processes, often requiring energy-intensive GPUs and hardware. However, advancements in hardware efficiency, cloud computing strategies, and sustainable practices can mitigate environmental concerns. Optimizing algorithms for energy efficiency and adopting renewable energy sources in data centers can reduce the carbon footprint associated with such research. This aligns technological progress with eco-friendly practices, ensuring healthcare innovations uphold both medical excellence and environmental responsibility.

6.3 Ethical Aspects

The "EDNet-20: An Eye Disease Classification Approach Combining CNN and Hybrid Attention" study adheres to strict ethical standards, ensuring patient privacy and confidentiality. It obtains informed consent where necessary, respecting individual rights. The study's commitment to transparency promotes ethical practices in information dissemination, supports reproducibility, and invites scrutiny. This ethical framework highlights the dedication to societal well-being and the responsible application of advanced medical technologies; ensuring advancements in diagnostics contribute to enhanced patient care.

6.4 Sustainability Plan

The sustainability strategy for "EDNet-20: An Eye Disease Classification Approach Combining CNN and Hybrid Attention" emphasizes minimizing its environmental footprint while maximizing societal benefits. This includes optimizing computational efficiency through improved algorithms and adopting energy-efficient practices throughout model training and deployment. Additionally, the study leverages cloud computing methodologies to reduce overall energy consumption. The commitment to environmentally responsible data management and storage further underscores these efforts. These initiatives illustrate a commitment to integrating green computing principles into technological advancements, ensuring that advances in medical diagnostics align with environmental sustainability. Ongoing refinement of these sustainable practices is essential for fostering responsible and environmentally conscious scientific innovation.

6.5 Summary

The research used machine learning and artificial intelligence to enhance patient care and provide societal and economic advantages. This technology has the potential to detect and cure eye diseases, therefore avoiding vision loss and minimizing mental suffering. Additionally, it facilitates the development of telemedicine technology, allowing for remote diagnosis in distant regions. This discovery has the potential to revolutionize precision medicine, enhancing the accessibility and affordability of healthcare, particularly in underprivileged regions. The research highlights the need of being transparent, following ethical principles, and using cloud computing to decrease energy use. The emphasis on environmentally appropriate data handling and storage highlights the incorporation of green computing ideas into medical diagnostics.

CHAPTER 7
Conclusion and Future Work

The study aims to improve the classification and treatment of eye diseases using the medical imaging technique CFP. Early-stage diagnosis is challenging due to time-consuming screening and factors like noise, artifacts, and equipment supply disruptions. There is a global shortage of medical equipment, radiologists, and experts capable of interpreting screening data, especially in rural areas and developing countries. Computer-aided detection (CAD) technologies can help reduce radiologists' workloads by evaluating eye fundus images across different modalities. Recent advancements have led to the development of AI solutions for the automated identification of eye diseases based on these imaging modalities. The research begins with image pre-processing and employs augmentation techniques to increase the dataset size. Model optimization techniques are applied to create a robust, fine-tuned CNN model (EDNet-20) for classifying fundus images into different classes. EDNet-20 excels in distinguishing between diseased and normal fundus images, with its accuracy surpassing comparable works. Various statistical analyses, including K-fold cross-validation, are conducted to ensure the model's reliability and robustness.

7.1 Conclusion

The study aims to create a reliable model for classifying eye diseases using advanced techniques. Two datasets were used to assess the model's effectiveness. Using diverse image processing methods, the quality and quantity of eye disease lesion images were improved, and overfitting was prevented. EDNet-20, a robust model based on the Convolutional Neural Network (CNN) architecture, achieved an accuracy of 94.26% on the Mendeley dataset and 91.80% on a proprietary dataset. Comparative analysis showed that EDNet-20 consistently surpassed other models in terms of accuracy and reliability, indicating its potential for practical applications in medical diagnostics. The research emphasizes the importance of integrating advanced image processing and deep learning methods to enhance diagnostic precision. EDNet-20's achievements demonstrate how deep learning strategies can improve medical diagnostics, patient outcomes, and healthcare operations. This research sets a groundbreaking standard for eye disease classification and lays the groundwork for future advancements in medical imaging and artificial intelligence research.

7.2 Limitations and Future Work

The research used two small datasets for deep learning, highlighting the need for more comprehensive and diverse datasets to improve the robustness and accuracy of eye disease classification models. The small size, low quality, and minimal parameters of the datasets led to a lengthy and time-consuming addressing procedure, highlighting the need for substantial computational effort to achieve meaningful results. Future research aims to explore new constraints associated with eye diseases through advanced medical imaging techniques, enhancing the model's precision and potentially uncovering novel insights into the early detection and treatment of eye diseases. This approach could significantly contribute to medical diagnostics, leading to better patient outcomes and more efficient healthcare solutions.

References:

[1] Li, S., Chen, L., & Fu, Y. (2023). Nanotechnology-based ocular drug delivery systems: recent advances and future prospects. *Journal of Nanobiotechnology*, *21*(1), 232.

[2] Wang, C., & Pang, Y. (2023). Nano-based eye drop: Topical and noninvasive therapy for ocular diseases. *Advanced Drug Delivery Reviews*, *194*, 114721.

[3] World Health Organization, available at << https://www.who.int/publications/i/item/9789241516570>>, last accessed on 08/06/2024 at 9 PM.

[4] Bourne, R., Steinmetz, J. D., Flaxman, S., Briant, P. S., Taylor, H. R., Resnikoff, S., ... & Tareque, M. I. (2021). Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *The Lancet global health*, *9*(2), e130-e143.

[5] Bourne, R. R., Jonas, J. B., Bron, A. M., Cicinelli, M. V., Das, A., Flaxman, S. R., ... & Resnikoff, S. (2018). Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe in 2015: magnitude, temporal trends and projections. *British Journal of Ophthalmology*, *102*(5), 575-585.

[6] Al-Fahdawi, S., Al-Waisy, A. S., Zeebaree, D. Q., Qahwaji, R., Natiq, H., Mohammed, M. A., ... & Deveci, M. (2024). Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. *Information Fusion*, *102*, 102059.

[7] Pardue, M. T., & Allen, R. S. (2018). Neuroprotective strategies for retinal disease. *Progress in retinal and eye research*, *65*, 50-76.

[8] Vaziri, K., Schwartz, S. G., Kishor, K. S., & Flynn Jr, H. W. (2016). Tamponade in the surgical management of retinal detachment. *Clinical Ophthalmology*, 471-476.

[9] Yan, Y., & Liao, Y. J. (2021). Updates on ophthalmic imaging features of optic disc drusen, papilledema, and optic disc edema. *Current Opinion in Neurology*, *34*(1), 108-115.

[10] Khalid, S., Akram, M. U., Hassan, T., Nasim, A., & Jameel, A. (2017). Fully automated robust system to detect retinal edema, central serous chorioretinopathy, and age related macular degeneration from optical coherence tomography images. *BioMed research international*, *2017*.

[11] Leasher, J. L., Bourne, R. R., Flaxman, S. R., Jonas, J. B., Keeffe, J., Naidoo, K., ... & Vision Loss Expert Group of the Global Burden of Disease Study. (2016). Global estimates on the number of people blind or visually impaired by diabetic retinopathy: a meta-analysis from 1990 to 2010. *Diabetes care*, *39*(9), 1643-1649.

[12] Ou, X., Gao, L., Quan, X., Zhang, H., Yang, J., & Li, W. (2022). BFENet: A two-stream interaction CNN method for multi-label ophthalmic diseases classification with bilateral fundus images. *Computer Methods and Programs in Biomedicine*, *219*, 106739.

[13] Li, J. Q., Welchowski, T., Schmid, M., Mauschitz, M. M., Holz, F. G., & Finger, R. P. (2020). Prevalence and incidence of age-related macular degeneration in Europe: a systematic review and meta-analysis. *British Journal of Ophthalmology*, *104*(8), 1077-1084.

[14] Holden, B. A., Fricke, T. R., Wilson, D. A., Jong, M., Naidoo, K. S., Sankaridurg, P., ... & Resnikoff, S. (2016). Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*, *123*(5), 1036-1042.

[15] Verbakel, S. K., van Huet, R. A., Boon, C. J., den Hollander, A. I., Collin, R. W., Klaver, C. C., ... & Klevering, B. J. (2018). Non-syndromic retinitis pigmentosa. *Progress in retinal and eye research*, *66*, 157-186.

[16] Xu, Z. Y., Azuara-Blanco, A., Kadonosono, K., Murray, T., Natarajan, S., Sii, S., ... & CORDS Study Group. (2021). New classification for the reporting of complications in retinal detachment surgical trials. *JAMA ophthalmology*, *139*(8), 857-864.

[17] Naing, S. L., & Aimmanee, P. (2024). Automated optic disk segmentation for optic disk edema classification using factorized gradient vector flow. *Scientific Reports*, *14*(1), 371.

[18] Rim, T. H., Kim, H. S., Kwak, J., Lee, J. S., Kim, D. W., & Kim, S. S. (2018). Association of corticosteroid use with incidence of central serous chorioretinopathy in South Korea. *JAMA ophthalmology*, *136*(10), 1164-1169.

[19] Onugwu, A. L., Nwagwu, C. S., Onugwu, O. S., Echezona, A. C., Agbo, C. P., Ihim, S. A., ... & Khutoryanskiy, V. V. (2023). Nanotechnology based drug delivery systems for the treatment of anterior segment eye diseases. *Journal of Controlled Release*, *354*, 465-488.

[20] Shamsan, A., Senan, E. M., & Shatnawi, H. S. A. (2023). Automatic classification of colour fundus images for prediction eye disease types based on hybrid features. *Diagnostics*, *13*(10), 1706.

[21] Muthukannan, P. (2022). Optimized convolution neural network based multiple eye disease detection. *Computers in Biology and Medicine*, *146*, 105648.

[22] Sarki, R., Ahmed, K., Wang, H., Zhang, Y., & Wang, K. (2021). Convolutional neural network for multi-class classification of diabetic eye disease. *EAI Endorsed Transactions on Scalable Information Systems*, *9*(4).

[23] Gargeya, R., & Leng, T. (2017). Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*, *124*(7), 962-969.

[24] Sarki, R., Ahmed, K., Wang, H., Zhang, Y., Ma, J., & Wang, K. (2021). Image preprocessing in classification and identification of diabetic eye diseases. *Data Science and Engineering*, *6*(4), 455-471.

[25] Puneet, Kumar, R., & Gupta, M. (2022). Optical coherence tomography image based eye disease detection using deep convolutional neural network. *Health Information Science and Systems*, *10*(1), 13.

[26] Alsaih, K., Lemaitre, G., Rastgoo, M., Massich, J., Sidibé, D., & Meriaudeau, F. (2017). Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *Biomedical engineering online*, *16*, 1-12.

[27] Tariq, H., Rashid, M., Javed, A., Zafar, E., Alotaibi, S. S., & Zia, M. Y. I. (2021). Performance analysis of deep-neural-network-based automatic diagnosis of diabetic retinopathy. *Sensors*, *22*(1), 205.

[28] Burlina, P. M., Joshi, N., Pekala, M., Pacheco, K. D., Freund, D. E., & Bressler, N. M. (2017). Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA ophthalmology*, *135*(11), 1170-1176.

[29] Moraru, A. D., Costin, D., Moraru, R. L., & Branisteanu, D. C. (2020). Artificial intelligence and deep learning in ophthalmology-present and future. *Experimental and Therapeutic Medicine*, *20*(4), 3469-3473.

[30] Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B. S., Waldstein, S. M., & Bogunović, H. (2018). Artificial intelligence in retina. *Progress in retinal and eye research*, *67*, 1-29.

[31] Zhen, Y., Chen, H., Zhang, X., Meng, X., Zhang, J., & Pu, J. (2020). Assessment of central serous chorioretinopathy depicted on color fundus photographs using deep learning. *Retina*, *40*(8), 1558-1564.

[32] Abbasi, S., & Bedeer, E. (2022, June). Deep learning-based list sphere decoding for Faster-than-Nyquist (FTN) signaling detection. In *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)* (pp. 1-6). IEEE.

[33] Abbasi, S., & Bedeer, E. (2023). Low complexity classification approach for Faster-than-Nyquist (FTN) signaling detection. *IEEE Communications Letters*, *27*(3), 876-880.

[34] Gu, R., Wang, G., Song, T., Huang, R., Aertsen, M., Deprest, J., ... & Zhang, S. (2020). CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation. *IEEE transactions on medical imaging*, *40*(2), 699-711.

[35] Eom, H., Lee, D., Han, S., Hariyani, Y. S., Lim, Y., Sohn, I., ... & Park, C. (2020). End-to-end deep learning architecture for continuous blood pressure estimation using attention mechanism. *Sensors*, *20*(8), 2338.

[36] Srivastava, O., Tennant, M., Grewal, P., Rubin, U., & Seamone, M. (2023). Artificial intelligence and machine learning in ophthalmology: A review. *Indian Journal of Ophthalmology*, *71*(1), 11-17.
[37] Mayya, V., Kulkarni, U., Surya, D. K., & Acharya, U. R. (2023). An empirical study of preprocessing techniques with convolutional neural networks for accurate detection of chronic ocular diseases using fundus images. *Applied Intelligence*, *53*(2), 1548-1566.
[38] Li, Z., Guo, X., Zhang, J., Liu, X., Chang, R., & He, M. (2023). Using deep learning models to detect ophthalmic diseases: A comparative study. *Frontiers in Medicine*, *10*, 1115032.
[39] Zang, P., Hormel, T. T., Hwang, T. S., Bailey, S. T., Huang, D., & Jia, Y. (2023). Deep-learning–aided diagnosis of diabetic retinopathy, age-related macular degeneration, and glaucoma based on structural and angiographic OCT. *Ophthalmology Science*, *3*(1), 100245.
[40] Song, A., Lusk, J. B., Roh, K. M., Hsu, S. T., Valikodath, N. G., Lad, E. M., ... & Kuo, A. N. (2024). RobOCTNet: Robotics and Deep Learning for Referable Posterior Segment Pathology Detection in an Emergency Department Population. *Translational Vision Science & Technology*, *13*(3), 12-12.
[41] Choudhary, A., Ahlawat, S., Urooj, S., Pathak, N., Lay-Ekuakille, A., & Sharma, N. (2023, January). A deep learning-based framework for retinal disease classification. In *Healthcare* (Vol. 11, No. 2, p. 212). MDPI.
[42] Saranya, P., Pranati, R., & Patro, S. S. (2023). Detection and classification of red lesions from retinal images for diabetic retinopathy detection using deep learning models. *Multimedia Tools and Applications*, *82*(25), 39327-39347.
[43] Kaur, M., & Kamra, A. (2023). Detection of retinal abnormalities in fundus image using transfer learning networks. *Soft Computing*, *27*(6), 3411-3425.
[44] Karlin, J., Gai, L., LaPierre, N., Danesh, K., Farajzadeh, J., Palileo, B., ... & Rootman, D. (2023). Ensemble neural network model for detecting thyroid eye disease using external photographs. *British Journal of Ophthalmology*, *107*(11), 1722-1729.
[45] Sun, G., Wang, X., Xu, L., Li, C., Wang, W., Yi, Z., ... & Chen, C. (2023). Deep learning for the detection of multiple fundus diseases using ultra-widefield images. *Ophthalmology and Therapy*, *12*(2), 895-907.
[46] Wang, K., Xu, C., Li, G., Zhang, Y., Zheng, Y., & Sun, C. (2023). Combining convolutional neural networks and self-attention for fundus diseases identification. *Scientific Reports*, *13*(1), 76.
[47] Huang, X., Ai, Z., Wang, H., She, C., Feng, J., Wei, Q., ... & Zeng, F. (2023). GABNet: Global attention block for retinal OCT disease classification. *Frontiers in Neuroscience*, *17*, 1143422.

[48] Dinpajhouh, M., & Seyyedsalehi, S. A. (2023). Automated detecting and severity grading of diabetic retinopathy using transfer learning and attention mechanism. *Neural Computing and Applications*, *35*(33), 23959-23971.

[49] Lu, Z., Miao, J., Dong, J., Zhu, S., Wu, P., Wang, X., & Feng, J. (2023). Automatic multilabel classification of multiple fundus diseases based on convolutional neural network with squeeze-and-Excitation attention. *Translational Vision Science & Technology*, *12*(1), 22-22.

[50] Ma, Z., & Li, X. (2024). An improved supervised and attention mechanism-based U-Net algorithm for retinal vessel segmentation. *Computers in Biology and Medicine*, *168*, 107770.

[51] Mendeley, available at << https://data.mendeley.com/datasets/s9bfhswzjb/1 >>, last accessed on 03-05-2024 at 11 PM.

[52]Montaha, S., Azam, S., Rafid, A. K. M. R. H., Ghosh, P., Hasan, M. Z., Jonkman, M., & De Boer, F. (2021). BreastNet18: a high accuracy fine-tuned VGG16 model evaluated using ablation study for diagnosing breast cancer from enhanced mammography images. *Biology*, *10*(12), 1347.

[53]Tariq, M., Iqbal, S., Ayesha, H., Abbas, I., Ahmad, K. T., & Niazi, M. F. K. (2021). Medical image based breast cancer diagnosis: State of the art and future directions. *Expert Systems with Applications*, *167*, 114095.

[54] Bama, K. A., & Priyadharsini, S. S. (2022). Groundnut leaves and their disease, deficiency, and toxicity classification using a machine learning approach. In *Cognitive Systems and Signal Processing in Image Processing* (pp. 253-275). Academic Press.

[55] Mohammad, A., Utso, M. Z., Habib, S. B., & Das, A. K. (2021). Predicting Retinal Diseases using Efficient Image Processing and Convolutional Neural Network (CNN). *Journal of Engineering Advancements*, *2*(04), 221-227.

[56] Wang, X., Wang, K., & Lian, S. (2019). A survey on face data augmentation. *arXiv preprint arXiv:1904.11685*.

[57] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*.

[58] Bjerrum, E. J. (2017). SMILES enumeration as data augmentation for neural network modeling of molecules. *arXiv preprint arXiv:1703.07076*.

[59] Khatun, T., Razzak, A., Islam, M. S., & Uddin, M. S. (2023). An extensive real-world in field tomato image dataset involving maturity classification and recognition of fresh and defect tomatoes. *Data in Brief*, *51*, 109688.

[60] Khatun, T., Nirob, M. A. S., Bishshash, P., Akter, M., & Uddin, M. S. (2024). A comprehensive dragon fruit image dataset for detecting the maturity and quality grading of dragon fruit. *Data in Brief*, *52*, 109936.

[61] Hussain, Z., Gimenez, F., Yi, D., & Rubin, D. (2017). Differential data augmentation techniques for medical imaging classification tasks. In *AMIA annual symposium proceedings* (Vol. 2017, p. 979). American Medical Informatics Association.
[62] Sivamani, S., Chon, S. I., Lee, D. H., & Park, J. H. (2022). Importance of Adaptive Photometric Augmentation for Different Convolutional Neural Network. *Computers, Materials & Continua*, *72*(3).
[63] Taylor, L., & Nitschke, G. (2018, November). Improving deep learning with generic data augmentation. In *2018 IEEE symposium series on computational intelligence (SSCI)* (pp. 1542-1547). IEEE.
[64] Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, *16*, 100258.
[65]https://medium.com/lunit/photometric-data-augmentation-in-projection-radiography-bed3ae9f55c3
[66] Alomar, K., Aysel, H. I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. *Journal of Imaging*, *9*(2), 46.
[67] Yang, H., Ni, J., Gao, J., Han, Z., & Luan, T. (2021). A novel method for peanut variety identification and classification by Improved VGG16. Scientific Reports, 11(1), 15756.
[68] Subramanian, N., Elharrouss, O., Al-Maadeed, S., & Chowdhury, M. (2022). A review of deep learning-based detection methods for COVID-19. *Computers in Biology and Medicine*, *143*, 105233.
[69] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
[70] Phiphiphatphaisit, S., & Surinta, O. (2020, March). Food image classification with improved MobileNet architecture and data augmentation. In Proceedings of the 3rd International Conference on Information Science and Systems (pp. 51-56).
[71] Tragoudaras, A., Stoikos, P., Fanaras, K., Tziouvaras, A., Floros, G., Dimitriou, G., ... & Stamoulis, G. (2022). Design space exploration of a sparse mobilenetv2 using high-level synthesis and sparse matrix techniques on FPGAs. Sensors, 22(12), 4318.
[72] Ji, Q., Huang, J., He, W., & Sun, Y. (2019). Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images. Algorithms, 12(3), 51.
[73] Moataz, L., Salama, G. I., & Abd Elazeem, M. H. (2021, December). Skin cancer diseases classification using deep convolutional neural network with transfer learning model. In Journal of physics: conference series (Vol. 2128, No. 1, p. 012013). IOP Publishing.

[74] Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, *452*, 48-62.
[75] Li, X., Wang, W., Hu, X., & Yang, J. (2019). Selective kernel networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 510-519).
[76] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).
[77] Tao, W., Li, C., Song, R., Cheng, J., Liu, Y., Wan, F., & Chen, X. (2020). EEG-based emotion recognition via channel-wise attention and self attention. *IEEE Transactions on Affective Computing*, *14*(1), 382-393.
[78] Hu, D. (2020). An introductory survey on attention mechanisms in NLP problems. In *Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys) Volume 2* (pp. 432-448). Springer International Publishing.
[79] Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., ... & Hu, S. M. (2022). Attention mechanisms in computer vision: A survey. *Computational visual media*, *8*(3), 331-368.
[80] Zhou, P., Cao, Y., Li, M., Ma, Y., Chen, C., Gan, X., ... & Chen, C. (2022). HCCANet: histopathological image grading of colorectal cancer using CNN based on multichannel fusion attention mechanism. *Scientific reports*, *12*(1), 15103.
[81] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, *8*, 1-74.
[82] Jie, H. J., & Wanda, P. (2020). RunPool: A dynamic pooling layer for convolution neural network. *International Journal of Computational Intelligence Systems*, *13*(1), 66-76.
[83] QingJie, W., & WenBin, W. (2017, June). Research on image retrieval using deep convolutional neural network combining L1 regularization and PRelu activation function. In *IOP Conference Series: Earth and Environmental Science* (Vol. 69, No. 1, p. 012156). IOP Publishing.
[84] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, *9*, 611-629.
[85] Gholamalinezhad, H., & Khosravi, H. (2020). Pooling methods in deep neural networks, a review. *arXiv preprint arXiv:2009.07485*.
[86] Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation*, *11*(3), 52.

[87] Hafiz, A. M., Parah, S. A., & Bhat, R. U. A. (2021). Attention mechanisms and deep learning for machine vision: A survey of the state of the art. *arXiv preprint arXiv:2106.07550*.
[88] Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in vision: A survey. *ACM computing surveys (CSUR)*, *54*(10s), 1-41.
[89] Hore, P., & Chatterjee, S. (2019). A comprehensive guide to attention mechanism in deep learning for everyone. *American Express*.
[90] Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.
[91] Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. *Journal of machine learning research*, *3*(Aug), 115-143.
[92] Mehtab, S., Sen, J., & Dutta, A. (2021). Stock price prediction using machine learning and LSTM-based deep learning models. In *Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, October 14–17, 2020, Revised Selected Papers 2* (pp. 88-106). Springer Singapore.
[93] Han, S., Meng, Z., Li, Z., O'Reilly, J., Cai, J., Wang, X., & Tong, Y. (2018). Optimizing filter size in convolutional neural networks for facial action unit recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5070-5078).
[94] Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, *503*, 92-108.
[95] Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Sciences*, *10*(5), 1897.
[96] Kandel, I., & Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT express*, *6*(4), 312-315.
[97] Rathore, P. S., Dadich, N., Jha, A., & Pradhan, D. (2018). Effect of learning rate on neural network and convolutional neural network. *Int. J. Eng. Res. Technol*, *6*(17), 1-8.
[98] Wilson, D. R., & Martinez, T. R. (2001, July). The need for small learning rates on large problems. In *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222)* (Vol. 1, pp. 115-119). IEEE.
[99] Stephen, O., Sain, M., Maduh, U. J., & Jeong, D. U. (2019). An efficient deep learning approach to pneumonia classification in healthcare. *Journal of healthcare engineering*, *2019*(1), 4180949.

[100] Chen, F. and Tsou, J.Y., 2022. Assessing the effects of convolutional neural network architectural factors on model performance for remote sensing image classification: An in-depth investigation. International Journal of Applied Earth Observation and Geoinformation, 112, p.102865.
[101] Desai, C. (2020). Comparative analysis of optimizers in deep neural networks. *International Journal of Innovative Science and Research Technology*, *5*(10), 959-962.
[102] He, J., Li, L., Xu, J., & Zheng, C. (2018). ReLU deep neural networks and linear finite elements. *arXiv preprint arXiv:1807.03973*.

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title: EDNET-20: AN EYE DISEASE CLASSIFICATION APPROACH COMBINING CNN AND HYBRID ATTENTION

Student ID: 201-15-14266

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Integrate recently gained and previously acquired knowledge to identify an **eye disease classification** problem for the Final Year Design Project (FYDP)	**PO1**
CO2	Analyze different aspects of the goals in designing a solution for this FYDP	**PO2**
CO3	Explore diverse problem domains through a literature review, delineate the issues, and establish this goals for the FYDP	**PO4**
CO4	Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP	**PO11**
Phase -II		
CO5	Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP	**PO3**
CO6	Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP	**PO5**
CO7	Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding.	**PO6**
CO8	Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks.	**PO7**
CO9	Implement ethical principles and adhere to professional standards and norms in this FYDP	**PO8**

CO10	Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP.	**PO9**
CO11	Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP.	**PO10**
CO12	Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	This project demonstrates **fundamental engineering (K3)** principles by employing deep neural networks, data augmentation techniques, and diverse CNN architectures for image processing and classification tasks. The project demonstrates **specialist knowledge (K4)** by conducting transfer learning with advanced CNN architectures like VGG16, VGG19, DenseNet201, DenseNet169, ResNet50, Xception MobileNet, and MobileNetV2, enhancing eye disease classification accuracy in color fundus images, crucial for computer-aided diagnosis.	**Page no:** [19-24] **Section:** [3.2.1] **Page no:** [29-32] **Section:** [4.2.1]
				The project applies **engineering practice & design (K5)** by the figure of process of experiments. The project addresses **engineering practice & technology (K6)** by	**Page no:** [17] **Section:**

				employing CNN model.	[3.2] **Page no:** [34] **Section:** [4.2.3]
				This project ensures to **K8 (Research Literature)** by synthesizing insights from recent studies, to advance eye disease classification using deep learning, showcasing a comprehensive understanding of current methodologies.	**Page no:** [10-16] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	This project addresses **EP-2** by recognizing the hurdles in eye disease classification, including limitations of traditional methods and the complexities of integrating transfer learning and deep CNNs. Through comparative analysis, it confronts challenges in understanding spatial distributions, offering insights for refining diagnostic methodologies.	**Page no:** [16] **Section:** [2.4]
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	This project addresses **EP-3** by meticulously comparing experimental outcomes, highlighting Transfer Learning as the chosen significant solution for enhancing eye disease classification amidst multiple potential approaches.	**Page no:** [41-48] **Section:** [5.2, 5.3]

4.	**EP4: Familiarity of Issues**	Yes	CO8	This project's interdisciplinary approach extends beyond computer science and engineering, impacting medical diagnostics in eye disease, contributing to advancements in public health and healthcare practices which indicate **EP-4**.	**Page no:** [10-16, 49] **Section:** [2.2, 5.4]
5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A
6.	**EP6: Extends of stakeholders involved and conflicting requirements**	No	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	This project's comprehensive approach addresses high-level problems by integrating various components across data collection, statistical analysis, and proposed methodology, ensuring a holistic solution to complex challenges in medical diagnostics which ensures **EP-7**.	**Page no:** [17-25] **Section:** [3.2, 3.2.1, 3.2.2, 3.2.3]

Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:

SN	EA Definition	Attainment	CO	Justification	References

1.	**EA1: Range of resources**	Yes	CO11	This project utilizes diverse resources such as high-performance computing infrastructure, GPUs, deep learning frameworks, annotated datasets, and ethical considerations to ensure systematic research and contribute to advancements in eye disease classification through transfer learning and deep CNNs.	**Page no:** [23-25] **Section:** [3.3]
2.	**EA2: Level of interaction**	No		N/A	N/A
3.	**EA3: Innovation**	Yes		This project comes with a new CNN model enhanced by hybrid attention mechanism to improve the effectiveness of the CNN model in case of image classification.	**Page no:** [36-38] **Section:** [4.2.7]
4.	**EA4: Consequences for society and the environment**	Yes		This project contributes to society by improving healthcare through advanced eye disease classification methods, while also promoting environmental sustainability by employing efficient computational resources and adhering to ethical guidelines for patient data privacy.	**Page no:** [51-52] **Section:** [6.1, 6.2, 6.3]

5.	**EA-5: Familiarity**	Yes		This project expands upon existing research by examining a novel approach in eye disease classification through transfer learning and deep CNNs, demonstrated through preliminary terminologies and a comprehensive comparative analysis, offering new insights into the field.	**Page no:** [14-15] **Section:** [2.1, 2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	This project addresses **CO4** by integrating effective project management and financial oversight, ensuring meticulous planning, resource allocation, and budget estimation for optimal resource utilization throughout the research lifecycle.	**Page no:** [27] **Section:** [3.4]
2	CO9	Yes	The project demonstrates adherence to ethical principles by prioritizing patient privacy, obtaining informed consent, and transparently documenting the research process, ensuring responsible knowledge dissemination and societal well-being through the ethical application of advanced healthcare technologies which comply with **CO9**.	**Page no:** [52] **Section:** [6.3]
3	CO10	No	N/A	N/A
4	CO12	Yes	The project's dedication to continuous learning **(CO12)** and adaptation within the dynamic technological landscape is reflected in its comprehensive data collection, rigorous statistical analysis, meticulous methodology development, and thorough experimental results and analysis, showcasing a commitment to staying updated and refining techniques to address modern challenges.	**Page no:** [17-22, 41-45] **Section:** [3.2, 5.2]

EDNet20 Version 1

ORIGINALITY REPORT

14%	9%	9%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
2	"Advanced Computing and Intelligent Technologies", Springer Science and Business Media LLC, 2024 Publication	1%
3	www.ncbi.nlm.nih.gov Internet Source	1%
4	www.mdpi.com Internet Source	<1%
5	Sidratul Montaha, Sami Azam, Abul Kalam Muhammad Rakibul Haque Rafid, Pronab Ghosh et al. "BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images", Biology, 2021 Publication	<1%
6	Teixeira, Tiago Oliveira. "Image Classification of OCT Scans Using Machine Learning",	<1%