An Efficient Deep Learning-Based Multi-Classification of Ocular Toxoplasmosis and its Secondary Complications from Fundus Images

Submitted by

AVISHEK ROY SPARSHO
20301269
JOYEETA SAHA
20301087
MD. MUSFIQUR RAHMAN SHOVON
20301332
MARZUK AHAMED
20301169
ANONNA AMIN BIVA
20301382

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

Department of Computer Science and Engineering
Brac University
December 2024

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

AVISHEK ROY SPARSHO

20301269

JOYEETA SAHA

20301087

MD. MUSFIQUR RAHMAN SHOVON

20301332

MARZUK AHAMED

20301169

ANONNA AMIN BIVA

20301382

Approval

The thesis titled "An Efficient Deep Learning-Based Multi-Classification of Ocular Toxoplasmosis and its Secondary Complications from Fundus Images" submitted by

1. AVISHEK ROY SPARSHO (20301269)
2. JOYEETA SAHA (20301087)
3. MD. MUSFIQUR RAHMAN SHOVON (20301332)
4. MARZUK AHAMED (20301169)
5. ANONNA AMIN BIVA (20301382)

Of Fall, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on December 02, 2024.

Examining Committee:

Supervisor:
(Member)

Dr. Md. Ashraful Alam, PhD

Associate Professor
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi, PhD

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

Abstract

Ocular Toxoplasmosis is considered a leading cause of visual impairment when not diagnosed and treated correctly on time. Its complications pose a challenge for accurate diagnosis and treatment. This study represents an efficient deep-learning approach for the multi-classification of Ocular Toxoplasmosis and its secondary complications using image datasets. In order to obtain a better fit of the model and further improve the data set, we incorporated its secondary complications. By carefully and accurately collecting a diverse set of data, we significantly enriched our custom dataset and enhanced its potential for insightful analysis. Deep learning techniques help us to develop an accurate multi-classification system for ocular toxoplasmosis and its associated complications. In this research, we utilized multiple CNN models and Transformer models, all of which demonstrated excellent accuracy. Additionally, we introduced two hybrid models by combining CNN and Transformer architectures. Finally, we also implemented robust ensemble architecture's two methods by combining the best-performing models: VGG19, ViT, and ResNet50. Furthermore, we integrated the XAI (GRAD-CAM) with the best accuracy 98% and F1 score 97% achiever Hybrid2(VGG19+ViT) model to enhance the understanding of our classification process and provide more transparency in the model's decision-making.

Keywords: Ocular Toxoplasmosis, Deep learning, Multi-classification, Fundus images, Ophthalmology, Visual impairment

Acknowledgments

First and foremost, all praise is due to Almighty Allah for His blessings which have allowed us to complete our thesis without any major interruptions.

We express our deepest gratitude to our Supervisor, [Dr. Md. Ashraful Alam], for his invaluable support and guidance throughout our work. His assistance was indispensable whenever we faced challenges.

We also extend our sincere thanks to our Research Assistant. His insightful guidance and constructive feedback have been instrumental in advancing our research and refining our work.

Finally, we are profoundly grateful to our parents for their unwavering support and encouragement. Their prayers and support have been instrumental in bringing us to the brink of our graduation.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the most common eye infections which is known as ocular toxoplasmosis is caused by the single-celled parasite Toxoplasma gondii. Typically, the parasite infects the eyes, resulting in inflammation and damage, particularly to the choroid and retinal layers. While toxoplasmosis is a disease that can affect anyone at any time, location affects how frequently it affects the eyes; warm, unsanitary places tend to have greater incidence. The main symptoms of ocular toxoplasmosis include focal necrotizing retinitis, which is often accompanied by granulomatous anterior uveitis and vitritis. Occasionally, it may manifest as papillitis. Retinal perivasculitis, glaucoma, cataracts, posterior synechiae, cystoid macular edema, and chorioretinal vascular anastomoses are examples of secondary effects. Typically present in the second decade of life, toxoplasmic retinochoroiditis has a 79% recurrence rate after five years [1]. The morphological variants of the retinal lesions include punctate inner and outer retinal lesions, as well as huge destructive lesions. Complications and a higher risk of sight loss are linked to large destructive lesions, which frequently require treatment. Whereas outer retinal lesions usually recur gradually, punctate interior lesions may be harmless. Severe unilateral papillitis is the rare but inevitable outcome of toxoplasmic papillitis. Cataracts, secondary glaucoma, choroidal neovascularization, and chronic iridocyclitis are among the complications. Macular involvement and the length of active episodes affect the visual prognosis; some patients have irreversible sight loss. In the US, it is the most prevalent retinal infection that can seriously impair eyesight. To estimate the burden of Toxoplasma gondii infection and ocular toxoplasmosis, we analyzed data from population-based research, outbreaks, and the U.S. census. Their estimates show that each year in the United States, 4,839 (range = 2,150–7,527) people get symptomatic ocular toxoplasmosis, 21,505 people have ocular lesions (both symptomatic and asymptomatic), and 1,075,242 people are infected with T. gondii [7]. The burden of ocular illness in the US is significantly increased by toxoplasmosis.Despite not exhibiting any symptoms, up to half of the world's population is infected with T. gondii.According to Holland's research, people with T. gondii infection had an approximate 2% chance of developing an eye ailment [7]. It is estimated that 21,505 people would get ocular lesions annually if this rate were applied to the yearly estimate of 1,075,242 T. gondii infections [7]. According to the British Columbia epidemic,people (median estimate of 4,839) would yearly acquire symptomatic ocular toxoplasmosis, with rates of 0.2–0.7% for symptomatic retinitis among T. gondii-infected individuals[7]. According to these estimates, over 21,000 people in the US had T. gondii ocular lesions in 2009, and over 4,800 had symptomatic ocular toxoplasmosis[7].This assumes that people develop eye lesions within a year of infection. Notably, larger annual incidence rates than predicted may arise from ocular toxoplasmosis recurrence in the years that follow. The ratio of symptomatic to asymptomatic disease is known to be influenced by particular outbreak-specific characteristics, and the global geographic heterogeneity in ocular toxoplasmosis rates is partly explained by differences in T. gondii strains and infection prevalence. In essence, the percentage of women who do not have a history of immunity before becoming pregnant and the frequency of Toxoplasma exposure during pregnancy determine the incidence of congenital toxoplasmosis. In the US, estimates of congenital infection have varied from one in 3000 to one in 10,000 live births[11]. Although T. gondii cysts can infect a variety of species, the main way that they are transmitted to humans is by eating undercooked or raw meat. Another frequent mode of transmission for T. gondii is the consumption of food, water, or soil tainted with oocysts excreted in the feces of infected cats, the final host. Thus, it is evident that ocular toxoplasmosis, with its vast prevalence, potential for severe vision impairment, and difficulties with diagnosis and treatment, is becoming a global problem. Millions of people worldwide are impacted by the high incidence of Toxoplasma gondii infection, which emphasizes the urgency. Due to their limits in sensitivity and specificity, traditional diagnostic techniques including clinical examinations and serological testing might result in wrong or delayed diagnosis, which can damage vision. The intricate nature of treatment, combining antibiotics and anti-inflammatory medicines, further complicates the problem, as outcomes vary across individuals, and recurrent bouts add to long-term difficulties. In response to these problems, Deep Learning presents intriguing options for upgrading the approach to ocular toxoplasmosis. Leveraging Deep Learning for image processing, medical images, such as fundus photos or retinal scans, can be interpreted more precisely, enabling early and accurate diagnosis. Predictive analytics, applying machine learning algorithms, can assess the likelihood of ocular toxoplasmosis based on numerous criteria, including demographic data and clinical history. This facilitates the early detection of individuals who pose a risk and expedites the diagnostic procedure. Furthermore, through the analysis of large datasets containing patient outcomes and reactions to various therapeutic approaches and facilitates the creation of customized treatment regimens.So the purpose of this research article is to use Deep Learning algorithms to provide a more customized and successful treatment approach, which may reduce the variation in treatment outcomes seen with conventional methods.

1.1 Research Problem

The primary focus of the research problem is the difficulties in promptly diagnosing ocular toxoplasmosis, which is a serious issue because it can cause permanent vision loss. Because of its widespread effects on public health and the difficulties associated with both diagnosis and treatment, ocular toxoplasmosis has become a major worldwide problem. The causal parasite, Toxoplasma gondii, affects millions of people globally, underscoring the necessity of efficient management techniques. Sensitivity and specificity issues with traditional diagnostic techniques, such serological testing and clinical examination, can result in erroneous or delayed diagnosis, which can impair vision. The intricate nature of treatment, combining antibiotics and anti-inflammatory medicines, further complicates the problem, as outcomes vary across individuals, and recurrent bouts add to long-term difficulties. The goal of the study is to investigate the factors that hinder the early diagnosis of ocular toxoplasmosis. These factors include awareness, symptom variability, and the limits of current diagnostic tools. In general, a number of variables, including the disease's stage, the experience of the healthcare provider, and the diagnostic techniques used, might affect how accurately medical reports diagnose ocular toxoplasmosis. Diagnosing ocular toxoplasmosis can be difficult, particularly in the early stages, and the sensitivity and specificity of conventional diagnostic techniques, such as clinical examination and serological testing, may be compromised. For the early detection of ocular toxoplasmosis, the use of deep learning and image processing techniques presents significant benefits over conventional medical interventions. Through the application of these technologies, medical imaging' complex patterns and minute abnormalities can be methodically examined, leading to the early identification of the disease. After being trained on large datasets, deep learning outperforms traditional medical exams in terms of accuracy and sensitivity. Potential human biases are mitigated by the impartiality and consistency inherent in deep learning and image processing techniques, guaranteeing a consistent manner of interpreting medical images. These technologies speed up detection processes that are essential for prompt interventions in ocular illnesses through automated screening. This efficiency leads to a quicker diagnostic process as well as less labor for healthcare providers, allowing for a more targeted approach to treatment plans. A thorough understanding of ocular toxoplasmosis is made possible by the integrative powers of deep learning models, which cover a variety of data sources and patient-specific factors. Furthermore, the versatility and ongoing learning capacity of deep learning models make them well-suited for dynamic medical scenarios, further reinforcing their position in enhancing early diagnosis and intervention options for Ocular Toxoplasmosis.

1.2 Research Objectives

The main goal of this research is to address the need for more accurate and timely detection of ocular toxoplasmosis. Diagnosing conventional ocular toxoplasmosis is based on a simple clinical examination. However, a typical cases benefit from ocular fluid testing, such as a polymerase chain reaction for parasite DNA or evaluating intraocular antibody production for exact etiology determination. Due to the high seroprevalence in communities, serological testing for T. gondii antibodies is usually ineffective [17]. In many cases, this conventional method of detection cannot give the proper accuracy due to the delay between symptoms and local antibody production. The object of this paper is to analyze the deep learning methods of ocular toxoplasmosis detection for higher accuracy and sensitivity. Fundus Photography (FP) makes use of optical cameras to monitor, diagnose, and plan therapy for retinal diseases[5]. Our custom dataset contains various retinal fundus images, which were collected to test and enhance the deep learning diagnostic process. We applied different pre-processing techniques to optimize the dataset. Our goal is to compare the effectiveness of deep learning models including hybrid models and ensemble models with conventional medical methods. By doing so, we aim to observe higher accuracy rates with deep learning, advancing the diagnostic process for ocular toxoplasmosis. Additionally, this study explores the transferability of these models across different demographic groups, considering factors such as age and gender, to ensure the models' consistency and effectiveness in diverse patient populations. Ultimately, the research aims to enhance deep learning methods in the field of ophthalmology, contributing to future advancements and more efficient diagnostic techniques. Additionally, this study also incorporates explainable AI (XAI) techniques to provide clear, interpretable insights into the decision-making process.

Chapter 2

Related Work

Since ocular toxoplasmosis became a prominent ocular pathology, a great deal of study has been conducted to determine the pathophysiology, epidemiology, and best practices for diagnosis and treatment. In order to present a thorough overview of the current status of research on ocular toxoplasmosis, this literature review aims to synthesize existing knowledge. If not properly and quickly detected, this parasite infection—caused by the protozoan Toxoplasma gondii—poses a serious risk to the health of the eyes, resulting in visual impairment and other consequences. This study attempts to provide insight into the changing field of ocular toxoplasmosis research, with an emphasis on recent developments, diagnostic techniques, and treatment approaches.This review aims to make a contribution to the ongoing efforts to improve clinical outcomes and our understanding of this difficult-to-treat ocular condition by looking at the most recent literature, finding knowledge gaps, and investigating novel treatments.

In the paper authored by Syed Rehan Shah, Salman Qadri et.al, the authors compare the performance of various deep learning architectures-Inception V3, VGG16, VGG19, CNN, and ResNet50-for the early detection of rice disease. The dataset contained about 2000 images from Kaggle :1200 of diseased leaves and 800 of healthy leaves,resized to 224×224 pixels. It was further split into 70:30 for training and testing , with preprocessing techniques like augmentation and brightness enhancement .ResNet 50 was chosen with the connection-skipping structure addressing the gradient vanishing problem. This network comprises of 18 convolving networks and featured transfer learning that imported weights from ImageNet for pre-training. Other augmentation processes such as the alterations of angle and zoom were applied.Experiments were performed in Python 3.7 and Tensorflow using a Tesla T4 GPU with performance parameters including binary cross-entropy loss, overall accuracy and F1 score. The web application is made using Gradio allowing users to interact with the model by uploading images and exhibiting the text results. Several models including ResNet 50 were compared and the results obtained from the experiment indicate that ResNet 50 was the optimally performing model for classifying images containing rice leaf blast disease.[33]

The paper authored by Payel Patra and Tripty Singh, highlights the detection of Diabetic Retinopathy (DR) by using an Improved ResNet 50-InceptionV3 and hybrid DiabRetNet Structures.The frameworks are based on Convolutional Neural Networks (CNNs).The proposed method constructs its DRD framework within three stages, namely; preprocessing fundus images through normalization and augmenta-

tion, feeding these preprocessed images into different architectures of convolutional neural network for feature vector extraction, and classifying these images according to their diabetic retinopathy severity levels from no DR to proliferative diabetic retinopathy. Different architectures of CNNs, including ResNet50, Inception V3, VGG-19, DenseNet-121, and MobileNetV2, are used under this work, improving the accuracy up to 93.79% - an improvement of about 7% over the previously published methods. This demonstrates how effective deep learning techniques are for early detection and diagnosis of Diabetic Retinopathy, which finally would improve patient outcomes.[27]
The paper reviews the use of deep learning techniques, specifically convolutional neural networks (CNNs), combined with transfer learning for image classification in the agricultural domain, particularly for detecting rice diseases. It highlights the effectiveness of pre-trained models like ResNet50, Inception V3, VGG16, and VGG19, emphasizing their ability to improve predictive performance when adapted to specific datasets. Foundational CNN architectures, such as those proposed by Szegedy et al. (2016) and He et al. (2015), are discussed, showcasing their advancements in handling large-scale image recognition tasks. Previous studies explored feature-based machine learning methods and simpler CNN architectures for rice disease detection, achieving moderate success but suffering from issues like overfitting, computational inefficiency, and limited adaptability to diverse disease patterns. Challenges such as the complexity of early disease symptoms, reliance on extensive labeled datasets, and computational resource constraints are identified as significant hurdles.[34]
The paper reviews the use of ensemble learning methods with deep convolutional neural networks (CNNs) for image classification,which highlighting their ability to improve predictive performance by combining multiple models. It also discusses foundational techniques like bagging and boosting, as well as advanced approaches such as stacking and the Super Learner, which uses cross-validation to optimally weight base models. Previous successes with CNN ensembles, such as those by Krizhevsky et al. (2012) and He et al. (2015), relied on unweighted averaging, leaving a gap in exploring more adaptive methods. Here are several hallenges such as overconfidence in predictions, sensitivity to weak learners, and the need for diverse architectures are identified as critical issues. The paper addresses these by systematically comparing ensemble methods, including the Super Learner and Bayes Optimal Classifier, demonstrating their potential to outperform traditional approaches while providing robustness against poorly performing models.[15]
The combination of deep learning and fuzzy logic techniques is examined in the paper "Design of Deep Ensemble Classifier with Fuzzy Decision Method for Biomedical Image Classification" by Abhishek Das, Saumendra Kumar Mohapatra, and Mihir Narayan Mohanty in order to enhance the classification of biomedical images. Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) are the four deep learning models that the authors use to categorize brain cancers using MRI scans and detect COVID-19 in chest X-rays. To handle uncertainty and enhance decision-making in classification tasks, the suggested method combines the outputs of various models using a fuzzy min–max (FMM) classifier as a meta learner. Along with highlighting the benefits of ensemble learning—especially for heterogeneous ensembles—the research also shows how effective the suggested approach is in terms of high accuracy, exceeding conventional classifiers. The findings demonstrate how important

ensemble learning and fuzzy logic are for improving biological image processing and diagnosis. [25]
In their 2024 study, "Ensemble Learning Optimized Medical Image Classification With Deep Convolutional Neural Networks," Müller, Soto-Rey, and Kramer investigate how ensemble techniques—such as Augmenting, Stacking, and Bagging—affect deep learning-based medical image processing. Stacking yields the most performance rise (up to 13% F1-score increase), followed by bagging (11%) and augmentation (4%), according to their analysis of four datasets (CHEMIST, COVID, ISIC, and DRD). They discover that unweighted averaging and other basic strategies perform better than more sophisticated ones like Support Vector Machines. The paper emphasizes how ensemble learning can increase model robustness, decrease overfitting, and improve clinical medical image categorization performance. [26]
Iqball and Wani (2023) present an ensemble approach using Deep Convolutional Neural Networks for medical image classification, particularly focused on Covid-19 and pneumonia detection from chest X-rays. By assigning weights based on individual model accuracy, the ensemble method improves upon traditional techniques by emphasizing high-performing models like ResNet101, InceptionV3, and MobileNetV2. Tested on two X-ray datasets, the approach demonstrates superior classification accuracy for normal, Covid-19, and pneumonia cases, showcasing improved generalization across data distributions. This method offers a promising direction for enhancing medical image classification models.[32]
The paper "An Ensemble Model for the Diagnosis of Brain Tumors through MRIs" by Ghafourian et al. presents the method of combining machine learning and data mining techniques for tumor detection. Each image is pre-processed to eliminate its background region and identify brain tissue, followed by SSO-based segmentation and SVD feature extraction to enhance accuracy and speed. An ensemble of Naïve Bayes, SVM, and KNN classifiers is used, with majority voting determining the final result. Tested on the BRATS 2014 and BTD20 datasets, the model achieves 98.61% and 99.13% accuracy, respectively. The method improves diagnosis efficiency, although training time is longer. Future work could explore other classifiers and parallel processing.This research shows significant improvements over previous efforts, confirming the effectiveness of ensemble learning and segmentation techniques in medical image classification. [31]
A study on the reliability of Deep Learning (DL) models for the automatic diagnosis of ocular toxoplasmosis (OT) from fundus images was carried out by Parra R, Ojeda V, Vázquez Noguera JL, García-Torres M, Mello-Román JC, Villalba C, Facon J, Divina F, Cardozo O, Castillo VE, and Matto IC. Even while DL models are useful, building trust in the medical community depends on how easily they can be interpreted. Physicians are more inclined to accept interpretable forecasts, hence the suggested methodology adds a fresh way to evaluate DL models: assessing the similarity between a DL model's decision criteria and an ophthalmologist's. The study, which made use of three DL architectures, found that model complexity, predictive power, and trust score are inversely correlated. The study not only supplies an open dataset of annotated eye fundus images for OT diagnosis but also introduces a domain-specific way to evaluate prediction models based on trust. Future developments might compare trust scores with conventional ML models, compare with other attribution techniques, and conduct user surveys with ophthalmologists. These developments could offer important insights into the medical community's

acceptance of DL model [23].
Matthew T.S. Tennant, MD, Uriel Rubin, MD, Faraz Oloumi, PhD, and Parampal S. Grewal, MD, wrote the review. Investigating deep learning's (DL) increasing influence in ophthalmology is FRCSC. For example, digital photos, optical coherence tomography, and visual fields are just a few of the diagnostic modalities where DL shows great promise as an emerging technology. Notably, DL techniques have been helpful in identifying numerous eye diseases, including cataracts, glaucoma, age-related macular degeneration, and diabetic retinopathy. With a discussion of their current uses in ophthalmic care and predictions for future developments, the study highlights the quick progress of DL approaches. The essay recognizes the constraints and ambiguities surrounding the best uses of deep learning in ophthalmology, despite its transformative promise. By relieving doctors of tedious duties and allowing a more concentrated approach to patient care, it is anticipated that as DL technology advances, it will blend smoothly into ophthalmic care. In the end, DL has the potential to improve physician-patient interactions and maximize ophthalmology treatments, both surgical and medicinal [16].
Authors Yanling Ouyang, Fuqiang Li, Qing Shao, Florian M. Heussen, Pearse A. Keane, Nicole Stübiger, Srinivas R. Sadda, and Uwe Pleyer have conducted a study with the goal of characterizing the clinical manifestation of subretinal fluid (SRF) in the posterior pole in eyes that have active ocular toxoplasmosis (OT) using spectral domain optical coherence tomography (SD-OCT). 45.5% of eyes with usual active OT and 51.3% of eyes with active OT showed signs of SRF, according to the study. The SRF's total volume was 0.47 mm3 and its mean maximum height was 161.0 µm. After receiving standard anti-toxoplasmosis therapy, SRF showed 100% absorption in situations where it was linked to active retinal necrosis. SRF clearance took 33.8 days on average, with a mean clearance rate of 0.0128 mm3/day.According to the results, when evaluated using SD-OCT, patients with active OT frequently have SRF, also known as serous retinal detachment. Most SRF linked to retinal necrosis responded well to standard therapy; however, even for smaller-sized SRF, cases with concurrent cystoid macular edema (CME) or choroidal neovascularization (CNV) showed less favorable responses or remained refractory to standard or combined intravitreal treatments [12].
Bambang Krismono Triwijoyo, Boy Subirosa Sabarguna, Widodo Budiharto, and Edi Abdurachman present a deep learning method for classifying eye illnesses using color fundus images in their work. When it comes to the diagnosis of eye illnesses, retinal pictures are particularly important. This is especially true for conditions like diabetes and high blood pressure, where microvascular changes are symptomatic of issues. While periodic ophthalmoscopy is the standard for disease screening, the paucity of ophthalmologists provides a problem for broad screening. The research investigates automatic analysis with a deep learning technique to assist ophthalmologists in diagnosing eye disorders. It does this by utilizing digital fundus cameras. Using the STARE dataset, which consists of scaled photographs classified into 15 kinds of eye illnesses, the chapter describes the use of convolutional neural networks (CNN) as classifiers for retinal images. The results of the studies show that input images with dimensions of 61×71 and 31×35 pixels yield the maximum training accuracy, while 31×35 pixels, the lowest resolution, yields the highest test accuracy of 80.93% [19].
"Artificial Intelligence and Imaging Processing in Optical Coherence Tomography

and Digital Images in Uveitis," a paper by María Abellanas et al., examines the major advances in computer vision—which include both AI and classical algorithms—in terms of improving the analysis and interpretation of OCT images in retinal research. With a special emphasis on OCT pictures, this review seeks to briefly summarize current advancements in the application of computer vision to uveitis imaging processing. Although significant development is still needed, computer vision technology is evolving to help with uveitis diagnosis and prognosis.It is possible to create standardized and objective results for upcoming clinical studies by automating image processing for uveitis patients. This work is critical for developing better understanding of the disease and how it progresses, as well as for designing treatment trials. Enhancing diagnostic skills, streamlining analyses, and adding to a more thorough understanding of uveitis through the integration of AI and imaging processing in uveitis research is a promising option that could eventually lead to better patient care and results [24].

When there is proof to justify such distinctions, the study addresses the benefits of SS-OCT over spectral-domain OCT. In conclusion, the review highlights the advancements in technology made possible by SS-OCT in retinal imaging, with its incorporation into clinical practice increasing our capacity for diagnosis. The authors believe that continued The paper authored by Jayesh Vira, Alessandro Marchese, Rohan Bir Singh, and Aniruddha Agarwa, titled "Swept-source optical coherence tomography imaging of the retinochoroidal and beyond," provides a comprehensive review of the advancements ushered in by swept-source optical coherence tomography (SS-OCT) imaging. The ability to quickly and accurately get high-resolution images of different layers' precise patho-anatomy has completely transformed retinochoroidal imaging technology. The review digs into the fundamentals and workings of SS-OCT, addressing its applications across varied illnesses such as age-related macular degeneration, diabetic retinopathy, pachychoroid spectrum diseases, and inflammatory vitreoretinal pathologies. A literature review is conducted for each disease, supported by illustrative pictures, and emphasizing the usefulness of SS-OCT and optical coherence tomography angiography. Promising revolutionary implications in patient care and improving our understanding of disease pathogenesis, quick, high-resolution imaging advances may soon allow real-time volumetric information of the entire retina and choroid [14].

This review of ocular toxoplasmosis, which focuses mainly on the disease's postnatally acquired form, was written by Dimitrios Kalogeropoulos, Hercules Sakkas, Bashar Mohammed, Georgios Vartholomatos, Konstantinos Malamos, Sreekanth Sreekantam, Panagiotis Kanavaros, and Chris Kalogeropoulos. Most commonly, retinochoroiditis is the symptom of ocular toxoplasmosis, which is the cause of posterior uveitis; Toxoplasmosis serologic testing is a useful tool for accurate diagnosis, which is based on the recognition of clinical symptoms. A conclusive diagnostic is provided by the detection of parasite DNA in the vitreous or aqueous material. There is growing evidence to support alternate therapies, such as intravitreal antibiotics, in addition to systemic corticosteroids, which are now the mainstay of current treatment. The report underlines current breakthroughs in diagnostic and treatment procedures, stressing their impact on avoiding or minimizing vision loss in persons with ocular toxoplasmosis. As a means of improving the management of this intricate clinical entity, it ends by emphasizing the need for additional study to deepen our understanding of the epidemiology, etiology, diagnosis, and treatment

[21].
Jolanda D.F. de Groot-Mijnes, Jose G. Montoya, and Justus G. Garweg write a paper on the difficulties in diagnosing toxoplasmic retinochoroiditis. They highlight the disease's localized character and the possibility that it won't always cause a noticeable systemic immune response. Diagnostic exploration is necessary for uncommon patients, however diagnosis is facilitated in conventional clinical presentations. In certain cases, the low DNA burden requires vitreous sampling for confirmation, even though the aqueous humor analysis may detect particular antibodies or parasite DNA. Individual differences in the time at which a particular antibody begins to produce can provide challenges for laboratory confirmation and can lead to false negatives. The limits of existing diagnostic methods are acknowledged by the authors, who suggest a therapeutically customized diagnostic methodology for atypical cases. The insufficiencies of depending exclusively on serological testing are brought to light, as well as the difficulty in identifying ocular toxoplasmosis due to seroprevalence patterns. The diagnostic use of IgM detection is restricted, and Toxoplasma-specific IgG, although prevalent, has low specificity. Given variables such as recent infection or reactivated latent diseases, a sophisticated interpretation of serological results is essential. The work highlights the usefulness of laboratory testing in clinical diagnosis, especially in atypical cases, and it recognizes the continued difficulties in comprehending parts of humoral immunity and the function that strain variability of parasites plays in the evolution of the disease. To better understand the pathophysiology of ocular toxoplasmosis and address these complexities, the authors emphasize the necessity for ongoing study [9].
The paper authored by Maenz M, Schlüter D, et al., the authors address the past, present and new aspects of ocular toxoplasmosis.It discusses the prevalence, socio economic impact, regional variations, clinical characteristics, treatment choices, and role of Toxoplasma gondii in ocular toxoplasmosis (OT).T. gondii, an internal protozoan parasite, goes through morphological and metabolic stages before releasing infectious sporozoites from felid oocysts.Although there is clinical heterogeneity, necrotizing retinochoroiditis is the normal presentation of OT.Historical and recent progress in establishing animal models for OT is discussed.Parasite entrance, disease initiation, self-limitation, recurrence, diagnostics, and parasite features are among the criteria used to classify animal OT models.Despite advancements, unresolved topics in ocular toxoplasmosis (OT) research, such as pathogenesis, clinic, and treatment, underscore the need for further exploration in future studies [8].
The paper authored by Stanford MR, Gilbert RE.The authors discussed the current evidence of treating ocular toxoplasmosis.Toxoplasmic retinochoroiditis is typically treated with pyrimethamine and sulfonamides, which are frequently combined with corticosteroids. However, there is insufficient evidence for antibiotic efficacy in recurring instances, with available research being restricted, weak, and old, demanding thorough randomized trials.Treatment is required for immunosuppressed patients, however preventive congenital infection treatment is debatable. Abandoned surgical alternatives and a focus on prevention highlight the lack of a vaccine or safe cysticidal medication. Genetic research on Toxoplasma gondii strains requires additional investigation. Prospective trials are critical for evaluating antibiotic efficacy, appropriate treatment, and recurrence prevention [6].
In the paper conducted by Yan Tong, Wei Lu, et al, the authors have discussed the application of machine learning in ophthalmic imaging modalities.AI and ma-

chine learning are transforming ophthalmology by analyzing enormous information and offering high-quality, tailored diagnosis in ocular illnesses. AI trends in ocular imaging show a tremendous increase in techniques used for medical imaging, with MRI and CT. Fundus Photography (FP) uses optical cameras to monitor, diagnose, and arrange treatment for retinal illnesses. Innovations such as ultra-wide field scanning laser ophthalmoscopy broaden the scope of FP, and FDA-approved AI grading algorithms like IDx-DR enable real-world applications.AI uses deep CNNs like GoogleNet to detect early diabetic retinopathy (DR).Optical Coherence Tomography (OCT) delivers non-invasive retinal morphological data.AI helps with segmentation and classification tasks in OCT scans.AI's overall impact in ocular imaging represents a shift from conceptual to objective medical imaging approaches [18].

The research conducted by Butler NJ,Furtadp JM et al, the authors talk about the clinical features,pathology and management of ocular toxoplasmosis II.Ocular toxoplasmosis causes recurrent posterior uveitis, which can be detected clinically or by ocular fluid testing. Treatment, which typically includes pyrimethamine, sulfadiazine, and corticosteroids, presents toxicity issues with no cure. Atypical symptoms are severe, particularly in immunocompromised patients, and serological testing has limited diagnostic value due to its high seroprevalence. Congenital toxoplasmosis can have significant systemic effects, but immunocompetent instances are usually asymptomatic. T. gondii infection of the eyes results in coagulative retinal necrosis and granulomatous choroiditis in fetuses, immunocompromised patients, and adults. H&E, immunohistochemistry, electron microscopy, and PCR are used to detect parasites that live in necrotic areas and near blood vessels. Clinical diagnosis involves distinctive findings for typical cases, while ocular fluid testing, especially PCR, is valuable for atypical cases, offering high sensitivity and specificity targeting the toxoplasmic B1 gene; limitations include false-positives and false-negatives, and combining PCR with the Goldmann-Witmer coefficient (GWC) .Serological testing has a limited role due to common seropositivity, with IgM detection in pregnancy indicating [10].

The paper authored by Rodrigo Parra, Verena Ojeda et.al, the authors uphold the deep learning(DL) tool for ocular toxoplasmosis diagnosis. Three DL architectures-random-weight CNN, VGG16 pretrained on ImageNet, and Resnet18 pretrained on ImageNet , have been tested.Ophthalmologists detect ocular toxoplasmosis (OT) using fundus images during eye exams, however machine learning, particularly Deep Learning (DL) and Convolutional Neural Networks (CNNs), has shown promise in evaluating retinal images for ocular diagnosis.In model training, CNNs use weight-sharing to lower parameters, enabling deeper networks with fewer parameters. All models undergo data augmentation, which includes random cropping and flips. Transfer learning, specifically for VGG16 and Resnet18, entails fine-tuning for ocular toxoplasmosis (OT) classification, addressing limited annotated data in the domain, and pretraining on a bigger dataset. Dataset split into training (70%), validation (10%), and test (20%) sets for model fitting, hyperparameter tuning, and final evaluation. Two experiments were performed. In the first, VGG and Resnet surpassed the CNN in accuracy, sensitivity, and specificity, with VGG unexpectedly outperforming Resnet, contrary to ImageNet results. In the second experiment, a proposed trust score calculated using successfully predicted images demonstrated that models that excelled in predictive metrics received lower trust scores.The findings indicate

that complicated models can attain higher sensitivity and specificity [22].
In this paper by Astrid M Tenter , Anja R Heckeroth et al., the authors have discussed toxoplasmosis caused by Toxoplasma gondii and various routes of its transmission. According to recent research, T. gondii has at least two clonal lineages. The transmission route includes vertical transmission during pregnancy and horizontal transmission through ingestion of oocysts, tissue cysts, or tachyzoites from various sources such as meat, offal, blood products, tissue transplants, and unpasteurized milk.Epidemiological importance of these routes varies and remains incompletely understood.Toxoplasma isolates from humans and animals belong to the same species. Prenatal infections occur in a range of 1 to 120 per 10,000 births, indicating a diverse incidence. In women of childbearing age, seroprevalence varies widely from 4 to 85%. Public health organizations, such as the World Health Organization, emphasize the need for accurate epidemiological data on T. gondii [3].
A benchmark study is proposed in the paper "Benchmarking Deep Learning Frameworks for Automated Diagnosis of Ocular Toxoplasmosis: A Comprehensive Approach to Classification and Segmentation," written by Syed Samiul Alam, Samiul Based Shuvo, Shams Nafisa Ali, Fardeen Ahmed, Arbil Chakma, and Yeong Min Jang. The paper discusses the difficulties in diagnosing OT. The present diagnostic techniques for OT, which is caused by T. gondii, are difficult, expensive, and require specialized experts. OT frequently causes visual issues. For the purpose of OT identification from fundus images, the research assesses the performance of current pre-trained networks utilizing transfer learning. Furthermore, for lesion segmentation, the study evaluates the effectiveness of segmentation networks based on transfer learning.The objective is to provide guidance to future researchers in the development of a deep learning (DL)-based, automated, user-friendly, and accurate diagnostic procedure that is affordable. The article performs a thorough examination of different feature extraction methods for lesion segmentation and OT categorization. MobileNetV2 outperforms other pre-trained models in terms of accuracy, recall, and F1 score when it comes to classification tasks. Other models that are assessed include VGG16, InceptionV3, ResNet50, and DenseNet121. For segmentation, U-Net design is adopted, and MobileNetV2/U-Net outperforms other architectures, particularly achieving superior results with the Jaccard loss function during training. This study offers insights into the best application of DL approaches and serves as a useful resource for the creation of effective and precise automated diagnostic tools for occupational therapy [28].
The study undertaken by Nancy Roizen and colleagues focuses on the impact of visual impairment resulting from toxoplasmic chorioretinitis on cognitive performance in children with congenital toxoplasmosis. Based on an extensive assessment of 64 kids, the study finds that kids with vision impairment, especially if they have macular involvement, perform worse on timed tasks that require them to distinguish between tiny intersecting lines. Notably, these kids' verbal scores are still lower than those of kids with normal vision even with compensating stronger verbal skills, suggesting that visual impairment has a more widespread impact on cognitive tests. The results of the study provide important new information about the complex relationship between cognitive and ocular function in congenital toxoplasmosis, pointing to the need for specialized compensatory intervention measures to improve learning and performance in children with these kinds of visual abnormalities. A basis for measuring cognitive function separate from visual impairment is laid by the identifi-

cation of certain patterns of difficulty, which may lead to more focused interventions for this population [2].

To find out how frequently people with congenital toxoplasmosis who received treatment during their first year of life developed new chorioretinal lesions, Laura Phan and colleagues carried out a prospective longitudinal research. The 132-child group was followed for an average of 10.8 years while receiving treatment with leucovorin, sulfadiazine, and pyrimethamine. Remarkably, only 31% of the assessed kids acquired new chorioretinal lesions; 14% of them acquired central lesions, and 25% of them acquired newly discovered peripheral lesions. The study's results show a marked difference from earlier reports on children who were either untreated or received limited treatment, when new lesions were much more common.Notably, 41 percent of those who developed new lesions had events at or after the age of 10, emphasizing the value of long-term monitoring well into the second decade of life in order to obtain a thorough evaluation of treatment effectiveness. This study emphasizes the value of prolonged monitoring to better understand the long-term impact of early therapies and provides important insights into the lower frequency of new chorioretinal lesions in instances of congenital toxoplasmosis treated during early life [4].

Carlos Cifuentes-González and colleagues conducted a comprehensive systematic review and meta-analysis, seeking to analyze risk variables for recurrences, visual impairment, and blindness in patients with ocular toxoplasmosis (OT). OT's trajectory was determined by analyzing 72 investigations conducted globally. Using data from 4,200 OT patients in various geographic locations, the meta-analysis found that 49% of cases had recurrences, with South America having a greater frequency than Europe. There were reports of visual impairment in 35% of eyes and blindness in 20% of eyes, with comparable prevalence patterns. Lesion locations close to the macula or the optic nerve, together with many recurrences, were noteworthy risk factors for blindness.Preventive therapy with Trimethoprim/Sulfamethoxazole showed noteworthy protective effects, which is noteworthy. OT patients who have certain risk factors may benefit from preventive therapy, as the review identifies clinical, environmental, and parasite-related aspects that affect the prognosis of the condition. In order to effectively manage ocular toxoplasmosis, physicians and researchers can benefit greatly from the thorough study that highlights the significance of customized therapies based on individual risk profiles [29].

In this work, Fatemeh Moradi et al. examine the effects of curcumin (CR) on behaviors resembling anxiety and depression that mice develop after a persistent infection with Toxoplasma gondii. Five sets of rats were used in the experiment, and behavioral testing supports the findings that chronic T. gondii infection causes symptoms resembling depression and anxiety. Significant antidepressant effects were observed with the administration of curcumin, particularly at dosages of 40 and 80 mg/kg. In the study, which explores the mechanisms behind these effects, proinflammatory cytokines and oxidative stress indicators in the hippocampus of infected mice are modulated by curcumin. Based on T. gondii-induced affective disorders, curcumin's antidepressant activities may be attributed to its capacity to control cytokine networks and oxidative stress.The psychological effects of a long-term Toxoplasma gondii infection are highlighted by this study, which also provides new insights into the treatment of affective disorders linked to parasite infections. Curcumin is one such therapeutic medication [30].

Chapter 3

Dataset and Pre-processing

3.1 Dataset Collection

In this research, we focus on detecting ocular toxoplasmosis using retinal fundus images and we also have considered its secondary complications to enrich our dataset and provide early detection and create awareness for the patients. After determining diabetes and cataract as common secondary conditions, we were able to collect three datasets which included images of active, inactive and healthy ocular toxoplasmosis as well as datasets for the secondary complications (diabetes and cataracts).
In our ocular toxoplasmosis dataset, we initially had three classes: Active OT, Inactive OT and Healthy. To strengthen the dataset, we merged the Active OT and Inactive OT classes into a single Ocular Toxoplasmosis class while keeping the Healthy class unchanged. The diabetes and cataract datasets were divided into True and False classes, with False cases included in the Healthy class. This resulted in a final custom dataset with four key classes:

1. **Healthy**
2. **Ocular Toxoplasmosis**
3. **Secondary Complications - Diabetes**
4. **Secondary Complications - Cataract**

3.2 Data Preprocessing

The initial dataset comprised 15,692 fundus images. Careful pre-processing method is used to ensure a robust and clean dataset which is suitable for accurate model training. To achieve this, we implemented several pre-processing techniques aimed at enhancing image quality and eliminating noisy or low-quality samples.

3.2.1 Normalization

Image normalization is a crucial preprocessing step in medical image processing, designed to adjust the intensity values of an image to a common scale, typically within the range [0, 255], using methods like cv2.normalize(). By applying a linear transformation to the pixel values, the minimum intensity is mapped to 0 and the maximum to 255, standardizing images with varying brightness and contrast. This uniformity across the dataset reduces variability caused by different imaging conditions, allowing models to better generalize across diverse samples. As a result, normalization helps mitigate biases introduced by varying illumination or contrast, ensuring that all images contribute equally to subsequent analysis and model training.

Laplacian and Thresholding

The performance of Deep learning models depends on the image quality where enhancing image quality is crucial for improving the accuracy of models. Laplanian model can enhance the image details, improve feature detection and reduce noise of images. Besides Thresholding is a technique used to segment an by converting the image into a binary format. In this process the pixels are set to a specific value depending on their intensity relative to a given threshold. The context of thresholding can enhance the contrast, simplify the image structure and improve model performance. In our data preprocessing pipeline, we applied a Laplacian filter with a value of 50 and thresholding with a value of 100 to enhance the quality of fundus images and improve model performance. Laplacian filter is a second order derivative filter which is used to sharpen the images. Complex changes can be easily found through this filter as well as this enhancement not only improve the model performance but also reduce the background noise and highlight the more prominent features. The combination of these techniques allowed for better feature extraction, noise reduction and improved classification accuracy which are essential for medical image analysis and model generalization across diverse datasets.

Grad Scale

Grad scale was used to measure the sharpness of images, ensuring that only those with clear and well-defined features were included in the dataset. This method is essential in highlighting critical structures which are crucial for accurate diagnosis and classification in medical imaging tasks.

Contrast Ratio Adjustments

Adjusting the contrast ratio improves the visibility of essential features by increasing the contrast between different regions of images. This step is essential for enhancing subtle patterns in the images. Detecting subtle patterns makes the model much easier to explore and identify dataset correlations.

3.2.2 Noise reduction

Noise reduction is an essential preprocessing step in image processing, aimed at smoothing images and reducing unwanted pixel variations that can interfere with subsequent analysis. A common method for this is cv2.GaussianBlur(), which applies a convolution operation using a Gaussian kernel. The Gaussian kernel is a matrix of weights that averages pixel values based on their proximity to the center pixel, effectively reducing noise by smoothing the image. This process helps eliminate minor variations while preserving important features. Reducing noise is particularly significant for tasks like edge detection and segmentation, where cleaner images lead to more accurate feature extraction and improved overall performance.

3.2.3 Scaling

Scaling in image processing refers to the process of resizing images to a different size or resolution while maintaining the aspect ratio or modifying it as needed. This is commonly done to ensure consistency in image dimensions when preparing data for machine learning models, especially in cases where models require uniform input sizes. In our research, we employed various deep learning models, such as VGG19, ResNet50, Inception V3, Swin Transformer, Vision Transformer, Hybrid and Ensemble model each of which requires input images of uniform size of 224*224 for optimal performance. To ensure compatibility with these models, we converted all images in our dataset to a standard resolution of 224*224 pixels.

3.2.4 Segmentation Mask Processing

The ‘transforms.Compose()‘ method from ‘torchvision.transforms‘ is used to create a sequence of transformations that are applied sequentially to images or segmentation masks. This method allows you to define a pipeline of preprocessing steps, such as resizing, normalization and format conversion, which are essential for preparing the data for deep learning models. For segmentation tasks, this can include operations like resizing the mask to a fixed size (224x224) and converting it to a tensor format, making it compatible with PyTorch models.
In the context of segmentation masks, these transformations ensure that the masks are standardized in terms of dimensions and format, which is crucial for deep learning tasks such as semantic segmentation. Accurate preprocessing of segmentation masks is important because it ensures that the masks align with the corresponding input images and are ready for model training. Properly formatted masks improve the model’s ability to learn from the data, leading to better performance in identifying and segmenting regions of interest.

3.2.5 Blood vessel extraction

The **frangi()** method from the **skimage.filters** module is specifically designed to enhance and extract tubular structures, such as blood vessels, from images. It works by analyzing the Hessian matrix of the image, which captures the second-order local structure information. The Frangi filter enhances regions that resemble vessels by applying a series of filters with varying scales and orientations, making it highly effective in highlighting elongated, tube-like structures such as blood vessels.
In medical image processing, particularly for tasks like retinal vessel analysis or angiography, the Frangi filter is valuable for isolating blood vessel networks from segmentation masks. By enhancing these critical features, it aids in precise vessel detection, which is important for diagnosing vascular diseases and assessing overall blood vessel health.

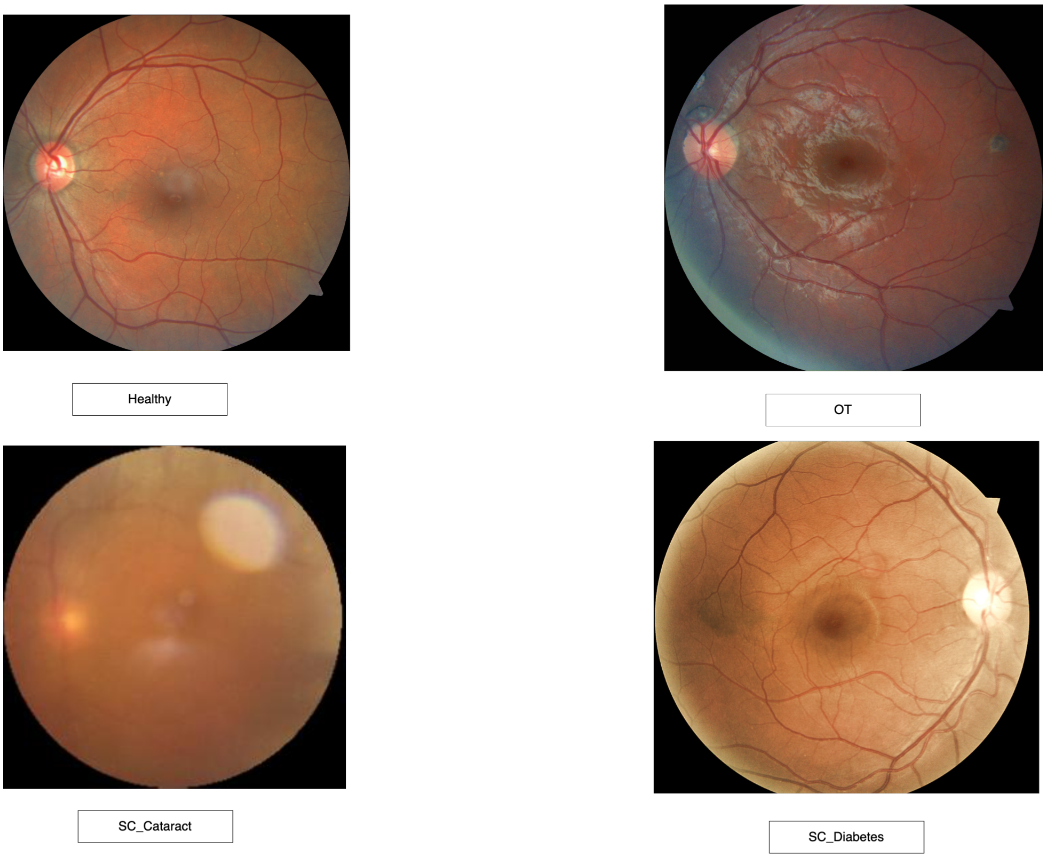

Figure 3.1: Images Before Pre-processing

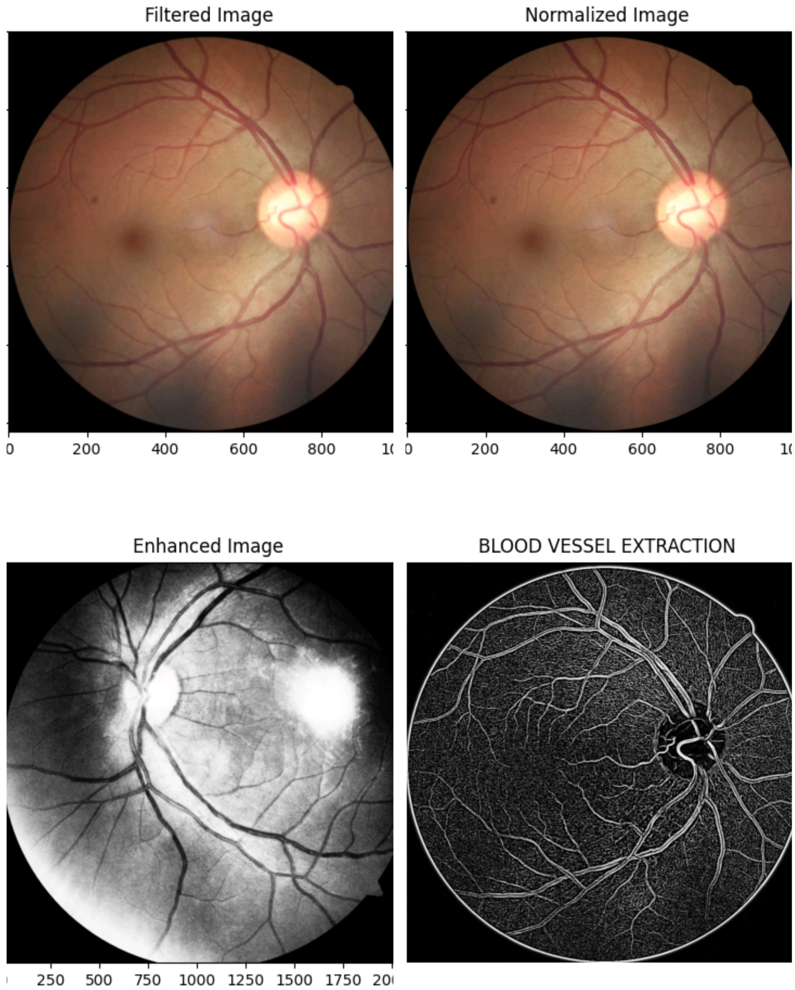

Figure 3.2: Pre-processing Steps

These initiatives greatly enhanced the quality of the dataset and making it more balanced and consistent. During the pre-processing time, we deducted the noisy images.

Chapter 4

Research Methodology

4.1 Our Methodology

The workflow diagram in Figure 4.1 gives an overview of each step we have done to train and evaluate our models.

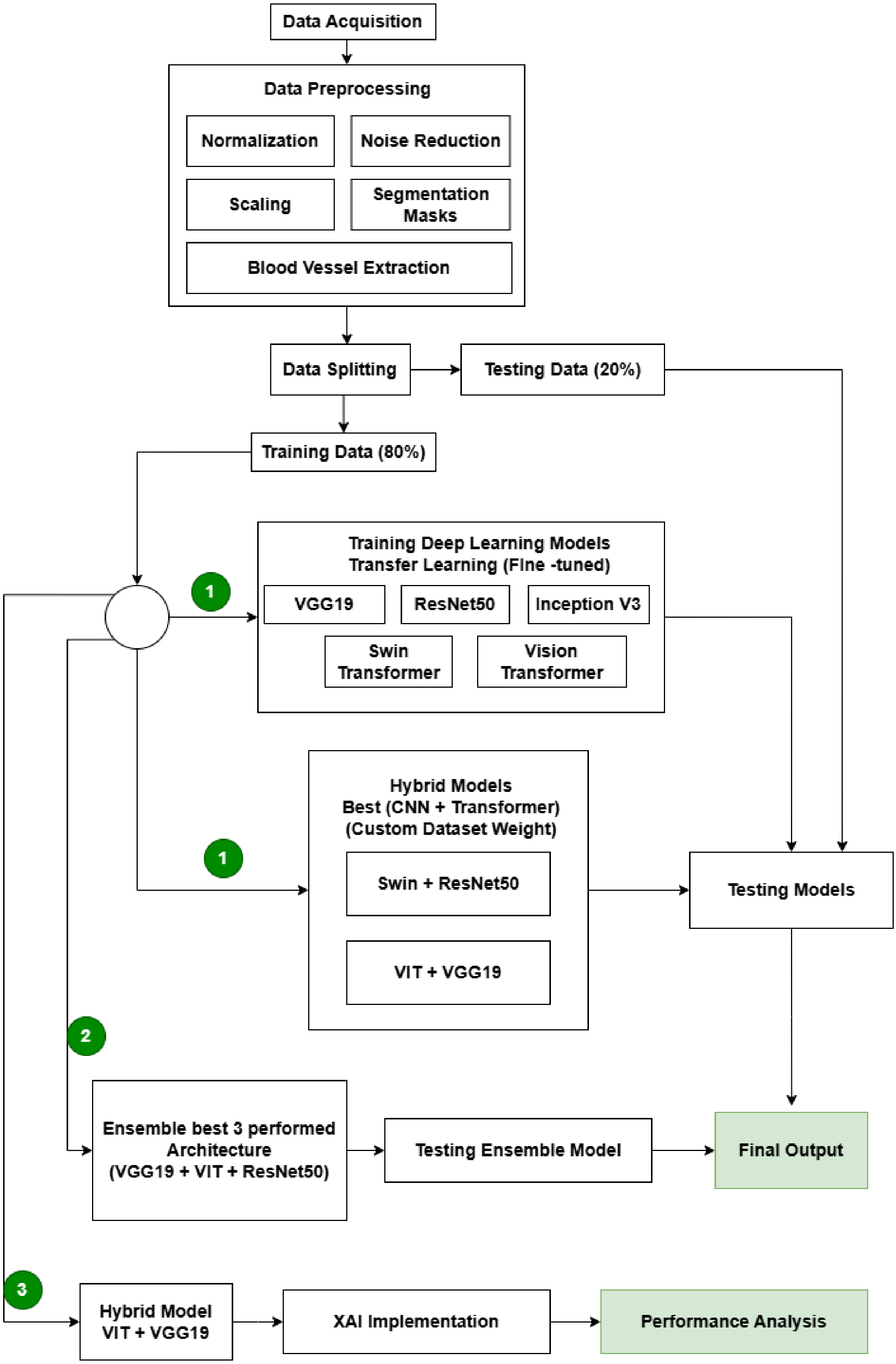

Figure 4.1: WorkFlow Diagram of the Proposed Model

4.2 Data Splitting

The above dataset was divided into two portions. Here 80% for training and 20% for testing. The majority of the data is in the training set which is used to train machine learning models and also helps to learn patterns and features from images. On the other hand, the testing set consists of the remaining 20% data which is used to assess the models performance. Moreover, it evaluates how it converts to new and unseen data in a good way. This approach also helps to ensure that the model doesn't overfit and it can make accurate predictions on the data which hasn't been encountered during the training period.

Class	Train Data	Test Data	Total Data
Healthy	966	241	1207
Ocular Toxoplasmosis	224	56	280
Diabetes	10658	2665	13323
Cataract	700	175	875

Table 4.1: Class Distribution

4.3 Training Parameters

The model was trained by using key parameters like a learning rate of 0.01 and a batch size of 32. In the training session the model used a total of 15 epochs. For classification, the Adam optimizer was chosen which combines the strength of momentum and RMSProp and makes it fitting for adaptive learning. Moreover, we have used the ReLU activation function in the hidden layers for introducing non-linearity and the sigmoid activation function in the output layer. For improving the performance and reducing the unnecessary computation, we have created early stopping which terminates training once minimal loss is gained and helps to prevent overfitting and preserve computational resources.

4.4 Deep Learning Models for Classification

Deep learning has revolutionized the field of machine learning, particularly in the domain of classification tasks. Classification involves assigning input data into predefined categories and deep learning models have proven to be exceptionally powerful at handling complex and high-dimensional data. Leveraging multi-layered neural networks, these models are capable of automatically learning and extracting intricate patterns and features from raw data, reducing the need for manual feature engineering. For our multiclass image classification, we have used ResNet50, VGG19, Inception-V3, Swin Transformer, Vision Transformer (ViT) and two custom model Hybrid model1 (ResNet50+ Swin Transformer) and Hybrid model2 (VGG19+ Vision Transformer).

4.4.1 VGG19

VGG19 is a deep convolutional neural network which has 19 layers. Above these 19 layers, it has 16 convolution layers and 3 fully connected layers. This model is efficient for image classification and feature extraction in transfer learning. In image classification tasks it provides very high accuracy. It is characterized by its simplicity and uniform architecture, consisting of small 3x3 convolutional filters throughout the network. Despite its depth, VGG19 is known for capturing intricate details and fine-grained features in image which is effective at capturing fine details in images.

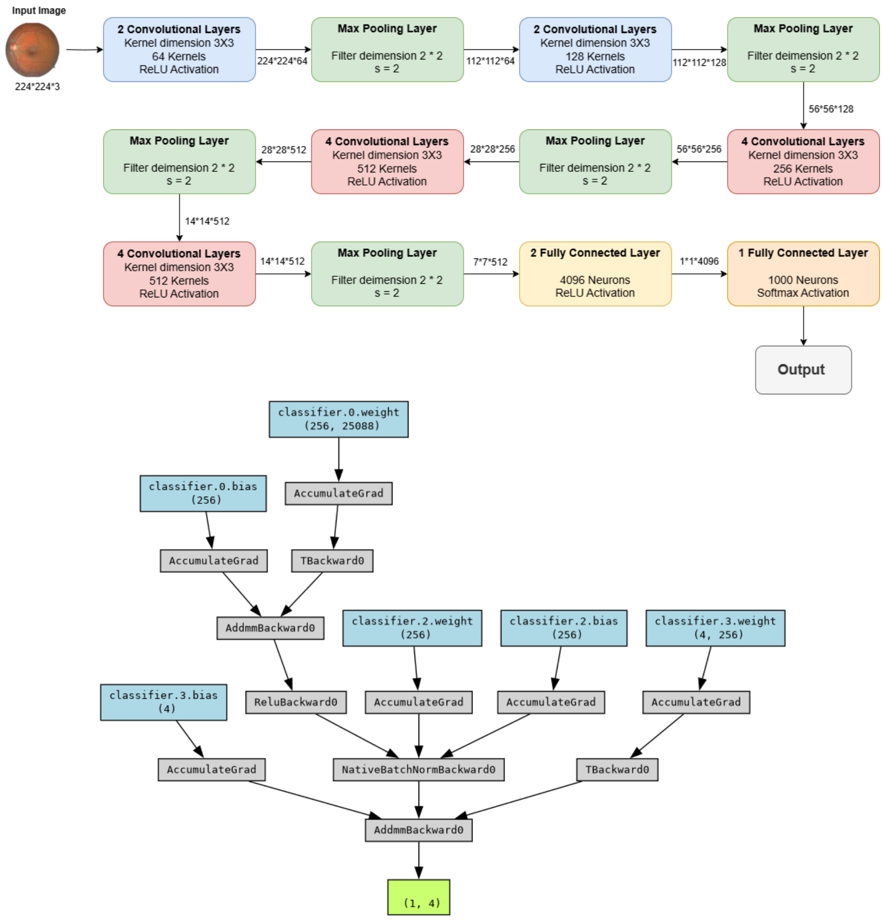

Figure 4.2: VGG19 Model Architecture

4.4.2 ResNet-50

ResNet introduced the key concepts of residual learning with skip connections which can enable a very deep neural network. ResNet has up to 152 layers to make a complex deep learning model. ResNet50 is a powerful and widely used convolutional neural network which has 50 layers. With 50 layers, ResNet50 employs the innovative concept of residual learning through shortcut connections which effectively address the vanishing gradient problem that typically hampers the training of deep networks. Vanishing gradient problem hampers the activity of the model where resent can successfully overcome this problem. These shortcut connections allow gradients to flow more easily through the network and enable the training of much deeper architectures without degradation. As a result, ResNet50 exhibits robust performance across a variety of complex image classification tasks. That's why it is a popular and reliable choice in both academic research and industrial applications. Its ability to maintain high accuracy while being computationally efficient has cemented ResNet50 as a go-to model for many vision-related challenges. [13]

A residual block in ResNet can be expressed as:

$$y = \mathcal{F}(x, \{W_i\}) + x \tag{4.1}$$

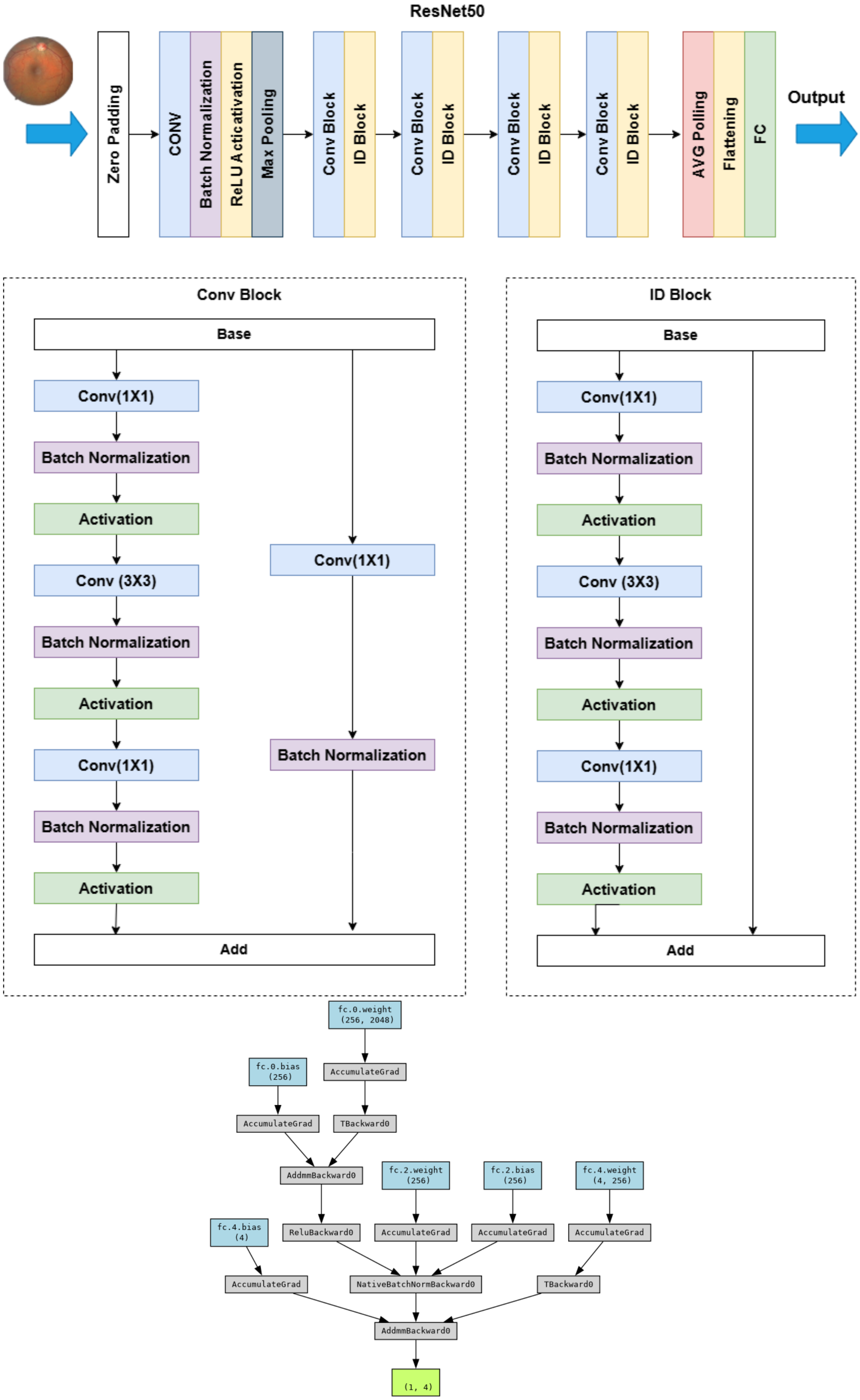

Figure 4.3: ResNet50 Model Architecture

4.4.3 Inception-v3

The Inception-v3 model is part of the inception family of networks. This model employs inception modules that allow the network to capture multi-scale features within the same layer. This model extracts information from images at multiple scales, improving classification performance efficiently when convolutions of different sizes in parallel are used. Inception-v3 model is highly effective for multi-class feature extraction. Its architecture can reduce computational complexity efficiently. By using various datasets, it provides a high accuracy.

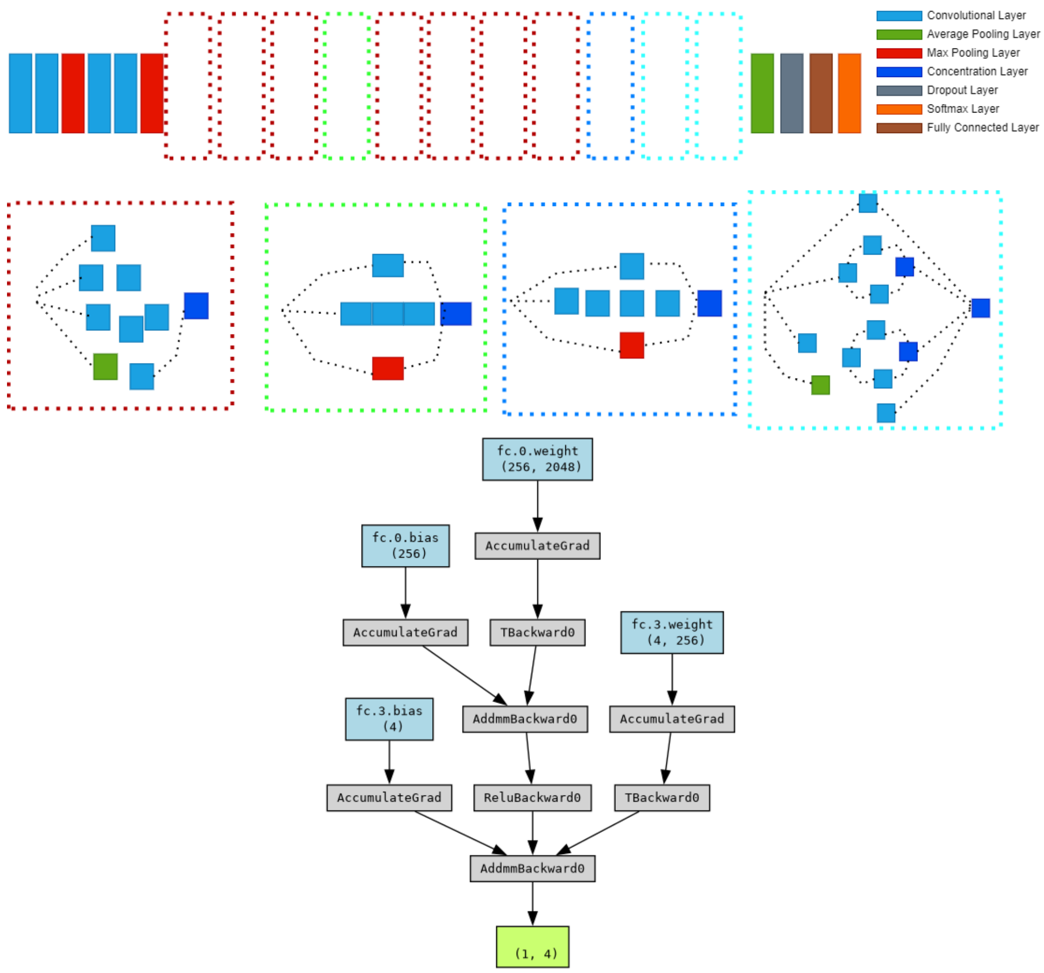

Figure 4.4: Inception-v3 Model Architecture

4.4.4 Swin Transformer

Swin Transformers is an advanced vision transformer with a hierarchical way of processing the image. It overcomes some of the shortcomings of prior Vision Transformer models, such as their quadratic computational cost in relation to input picture size and the absence of hierarchical feature representations. Swin Transformer replaces the normal multi-head self-attention module in a transformer block with a module based on shifted windows while maintaining the other layers. It builds hierarchical feature maps by merging image patches in deeper layers. The Swin Transformer's primary novelty is that it computes self-attention inside small windows rather than throughout the full screen. This local self-attention mechanism

decreases computational cost from quadratic to linear in terms of input picture size, making the Swin Transformer more scalable and efficient for high-resolution image processing applications.
Here, the input picture is divided into non-overlapping local windows or patches using the Swin Transformer. Self-attention procedures are calculated inside each local window, enabling the model to capture spatial dependencies and interactions among the patches within that window. However, the Swin Transformer uses a smart shifting window partitioning method to allow data interaction between windows and represent long-range relationships throughout the entire picture. This method involves moving the window dividers between subsequent Swin Transformer blocks which allows the model to focus on various areas of the picture and develop connections between distinct windows. This change in window partitioning helps the model capture global dependencies while preserving computational efficiency and making it easier for contextual information to spread. A feed-forward network is applied to each window following the shifting window attention and merging processes. This adds non-linearity and helps to further refine the feature representations.In later stages, neighboring patches are merged to lower the spatial resolution of the feature maps.The shifted window partitioning approach, along with the hierarchical merging of windows and patches, allows the Swin Transformer to capture both local and global dependencies within the picture while retaining linear computational cost in relation to the input image size.

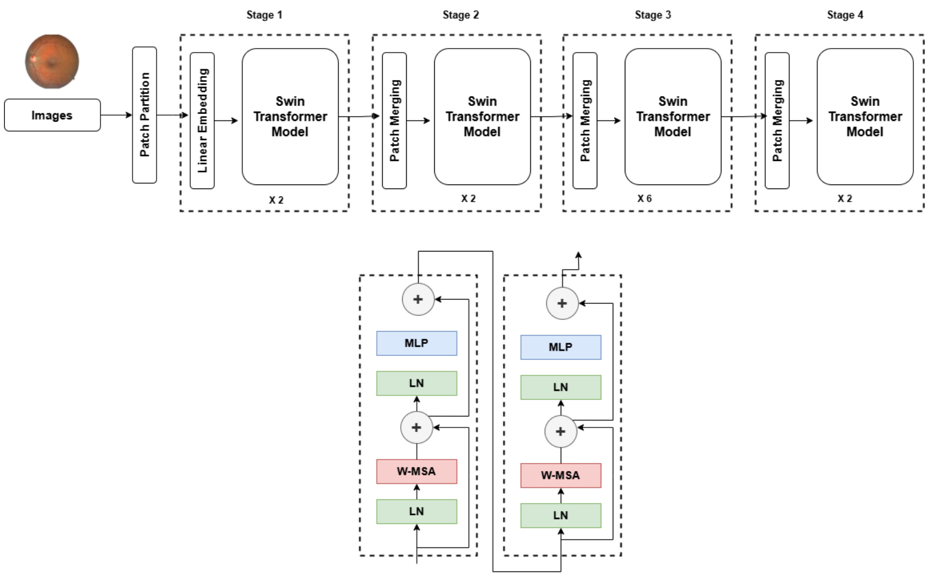

Figure 4.5: Swin Transformer Model Architecture

4.4.5 Vision Transformer (ViT)

The vision transformer is a type of deep-learning model used for image classification. They are inspired by the success of transformer models in the natural language processing. Self-attention is an important part of vision transformer. Self-attention in vision transformation helps the model grasp the relationships between different parts of the image by giving important scores to patches and focusing on the most essential information. This allows the model to better interpret the image and execute various computer vision tasks. Here, the input picture is divided into fixed-size patches which are then linearly embedded into a vector representation. Position embeddings are used to encode spatial information and the resulting sequence of embedded patches is put into a conventional transformer encoder.

$$\textbf{Attention}(Q, K, V) = \textbf{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4.2}$$

The Transformer encoder employs multi-headed self-attention processes which enable the model to detect long-range relationships between picture patches. A specific "classification token" is added to the input sequence and the transformer encoder's final representation is utilized to classify images. The ViT model is fully trained on picture data, learning rich visual representations straight from raw pixels without the need of hand-crafted features [20].

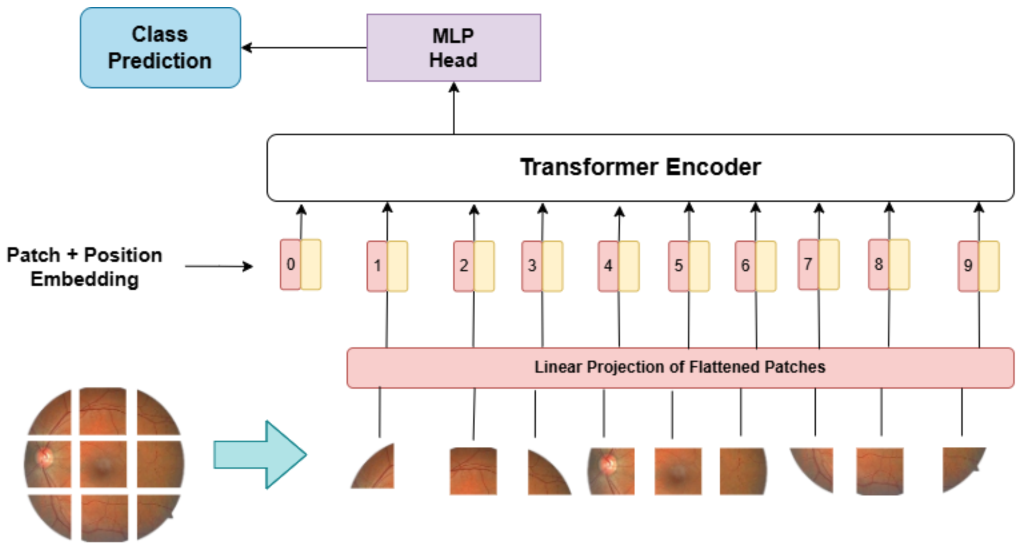

Figure 4.6: Vision Transformer Model Architecture

4.4.6 Swin + ResNet50

In this hybrid model, the Swin Transformer and ResNet50 architectures were combined to create a new, enhanced model. The described network architecture combines two pre-trained models, namely ResNet50 and Swin Transformer to perform image classification task. It commences by processing an input image of size 224x224x3 and incorporates data augmentation methods such as resizing, random flipping, rotation and color adjustment. In the feature extraction phase, ResNet50 employs frozen layers with fine-tuning on Layer 4, resulting in 2048 features, while the Swin Transformer freezes early layers and fine-tunes the last two layers, producing 50,176 features. These extracted features are then merged to create a combined output of 52,224 features. The merged features undergo fully connected layers with ReLU activation and dropout, ultimately leading to a final classification layer. By utilizing pre-trained models and fine-tuning specific layers, the architecture applies transfer learning to enhance performance, particularly for tasks with limited training data.

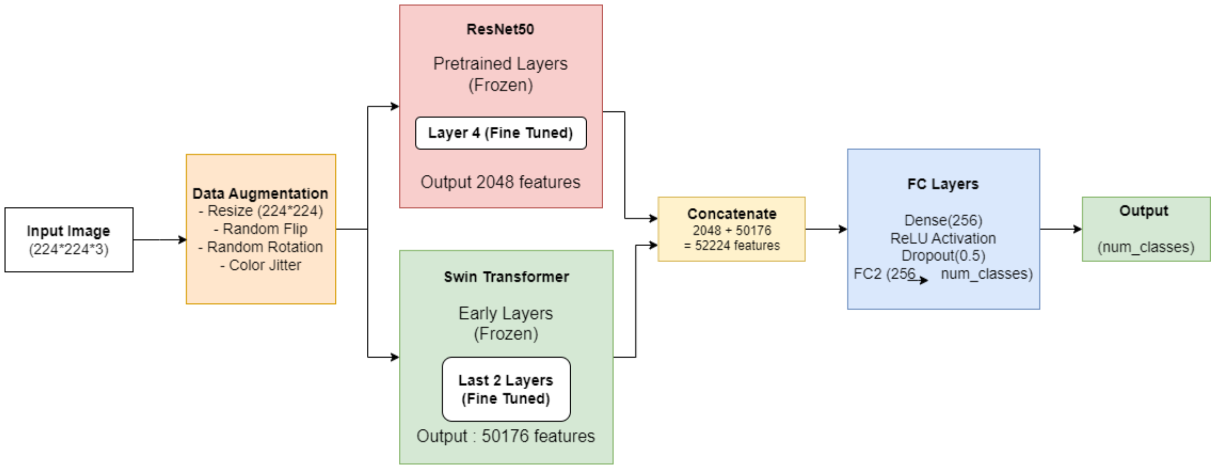

Figure 4.7: Hybrid 1 Model Architecture

4.4.7 VGG19 + Vit

The model for image classification utilizes a hybrid neural network, integrating VGG19 and a Vision Transformer to process a 224x224x3 image simultaneously. The CNN component of VGG19, employs pretrained weights and fine-tunes all layers to extract 25,088 features using convolutional layers, max pooling and ReLU activation. Concurrently, the Vision Transformer generates 768 features through patch embedding, positional encoding and transformer layers, also undergoing full fine-tuning. These combined features amount to 25,856, which then pass through fully connected layers. The process begins with a dense layer of 256 neurons with ReLU activation and a 0.5 dropout for regularization. The Softmax activation is utilized by the last dense layer to generate class probabilities. This design combines the CNN's capability to capture local features with the Vision Transformer's effectiveness in capturing global dependencies, resulting in improved classification performance.

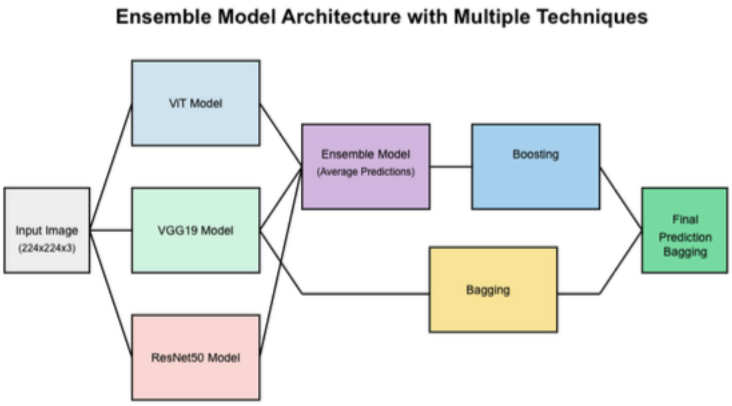

Figure 4.8: Hybrid 2 Model Architecture

4.4.8 Ensemble of VGG19, VIT and RESNET50

Ensemble modeling is such a process of multilayer diverse models which are used to predict the outcomee by using different modeling algorithm and different training datasets. The goal of this modeling is to reduce the generalization error of a prediction and have the possibility of high accuracy than a single classifier. In this model several models are merged and acts as a single model. Here, average voting is also a common ensemble procedure where averaging softmax class probabilities a posterior label is generated for all base learners. For averaging layer modification, in each model the inputs and outputs remain same.

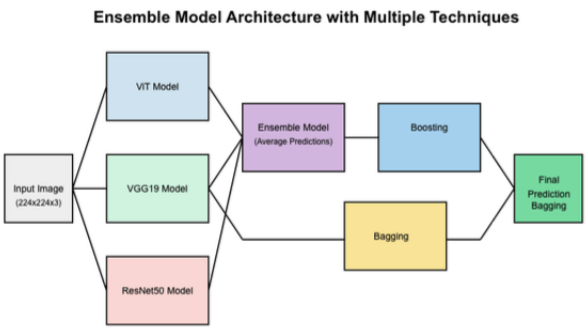

Figure 4.9: Ensemble Model Architecture with Multiple Techniques

Average layer :

In the proposed model, here the output of VGG19, Vision Transformer(VIT), and ResNet50 which wil be taken as inputs for this averaging layer. This layer will take the average value of these output. By averaging the predictions from this architecture, this models goal is to improve overall prediction accuray by averagiing the strengths of both of the CNN and Transformer based models.

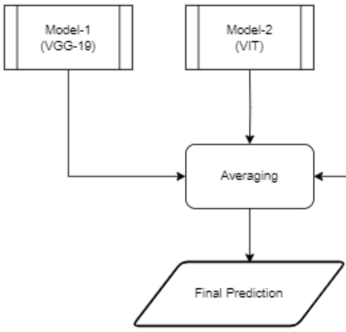

Figure 4.10: Average layers of the Ensemble model

Output layer :

Based on the result of the averaging layer, Ocular Toxoplasmosis patients can be identified easily from fundus images. If the patients probability of the first index is greater than the second index, the patient is Ocular Toxoplasmosis negative, otherwise, the patient is Ocular Toxoplasmosis positive. This process ensures that the model strongly classifies the condition by comparing the output probabilities from the several networks, providing a trustable diagnosis based on the fundus images.

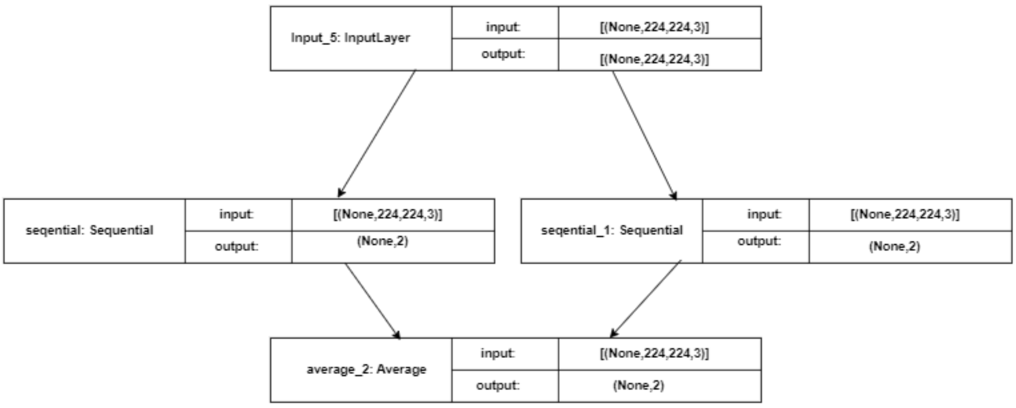

Figure 4.11: Output layers of the Ensemble model

4.4.9 Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence is a technique where a more explainable model is generated with a high level of learning performance [45]. By this technique human can easily understand ,trust and manage the emerging generation of intelligent partners. XAI allows supervised care by adapting deep learning techniques to obtain explainable attributes. It also contains procedures for acquiring more hierarchical, generalizable, and explanatory representations, and other pattern inference tools for asserting an understandable paradigm with any black -box testing model. Client happiness, influencing the multispectral images goal attainment, credibility analysis, and consistency are all areas where XAI thrives. In other terms, Explainable AI (XAI) is artificial intelligence that permits learners to acknowledge the consequences of the procedure. It is a collection of tools and regulatory frameworks that help individuals in comprehending and deciphering machine learning model projections. Feature attributions for model predictions in Auto ML Tables and AI Platform, and visually investigate model behavior using the What-If Tool can also be generated by the help of XAI.

Gradient-weighted Class Activation Mapping (Grad-CAM)

Grad-CAM is a widely used technique for visualizing and interpreting the decisions made by convolutional neural networks (CNNs). It provides insights into which regions of an input image significantly influence a specific class prediction, making it particularly valuable in image classification and object detection tasks.

The process begins with an input image and a pre-trained CNN model that generates class predictions. Grad-CAM computes the gradients of the target class score—representing the model's output for a particular class—relative to the feature maps of the last convolutional layer. This step is crucial as it reveals how variations in the feature maps can affect the score for the target class.

Once the gradients are calculated, they are averaged across the spatial dimensions (height and width) of the feature maps. This averaging yields a set of importance weights for each feature map, highlighting which features are most relevant to the prediction of the target class. Each feature map is then weighted by its corresponding importance weight, emphasizing the features that contribute significantly to the model's decision.

The next step involves summing the weighted feature maps to create a single-channel heatmap. This heatmap visually represents the areas of the input image that are most influential for the class prediction. To enhance the interpretability of the heatmap, a Rectified Linear Unit (ReLU) activation function is applied, which retains only the positive values. This ensures that only the regions contributing positively to the prediction are highlighted, while irrelevant or negative contributions are ignored.

Finally, the generated heatmap is resized to match the dimensions of the original input image. This allows for a clear visualization of the significant regions when superimposed on the original image, providing a comprehensive view of where the model is focusing its attention during classification.

Chapter 5

Results & Analysis

5.1 Model Performance

In machine learning algorithm, the most important thing is to find out the performance of the model. The learning curve is firstly prioritized for this research results and analysis part which confirmed that whether any models have overfit. In addition, accuracy, recall, precision and F1 score is measured for performance of every model.

5.1.1 Performance Analysis with Learning Curve

VGG19

The learning curves for the VGG19 model over 10 epochs show steady improvement. Training accuracy increases from 0.89 to 0.95, while validation accuracy reaches near 0.96, indicating good generalization. Training loss decreases from 0.40 to 0.15, with a similar downward trend in validation loss, showing effective optimization. The alignment between training and validation metrics suggests minimal overfitting and strong model performance.

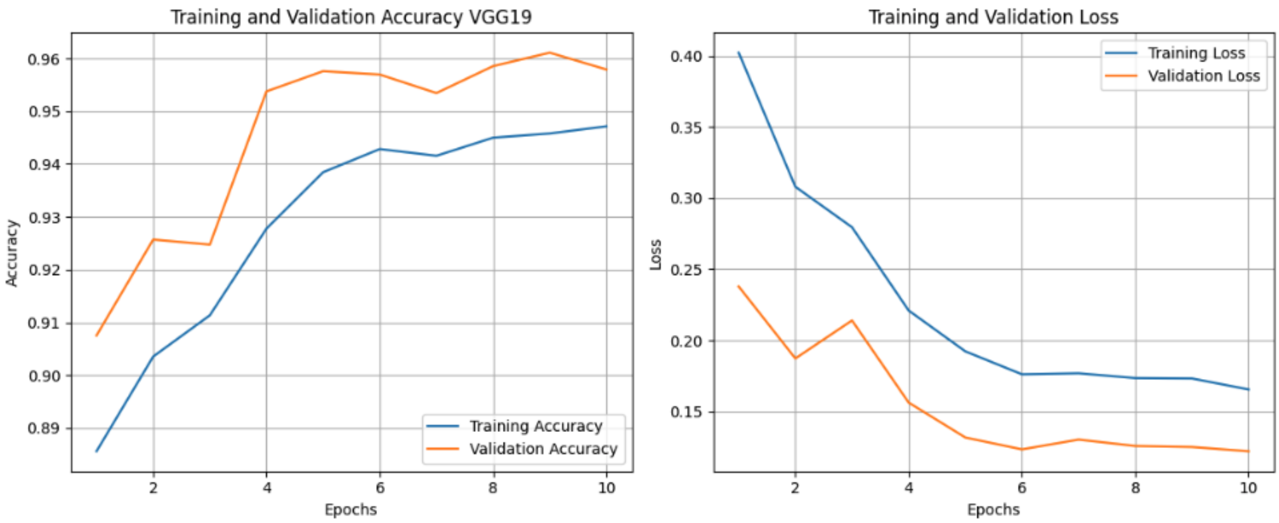

Figure 5.1: VGG 19 Learning Curve

ResNet50

The learning curves for the ResNet50 model over 10 epochs indicate stable progress. Training accuracy steadily increases from 0.89 to 0.92, while validation accuracy stabilizes around 0.93, reflecting good generalization. Training loss decreases significantly from 0.375 to 0.225, and validation loss reduces from 0.275 to 0.188, demonstrating effective optimization. The close alignment between training and validation metrics suggests minimal overfitting and reliable model performance.

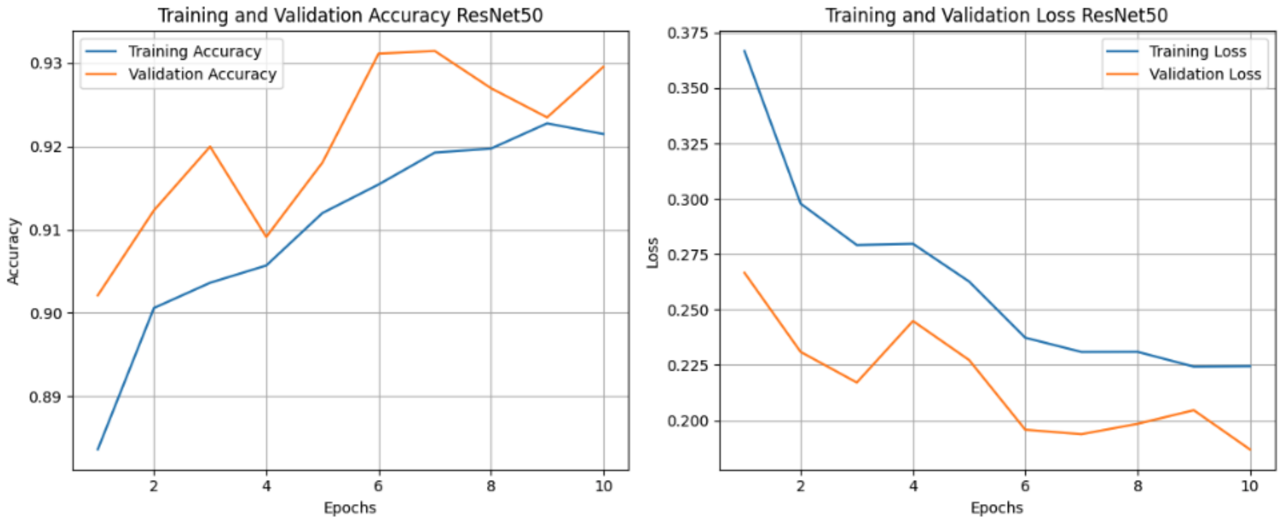

Figure 5.2: ResNet50 Learning Curve

Inception V3

The learning curves for the Inception V3 model over 10 epochs indicate strong performance. Training accuracy improves steadily from 0.88 to 0.91, while validation accuracy reaches 0.92, showing good generalization. Training loss decreases from 0.39 to 0.26, and validation loss reduces consistently from 0.33 to 0.22, reflecting effective optimization. The close alignment between training and validation metrics suggests minimal overfitting and robust model performance.

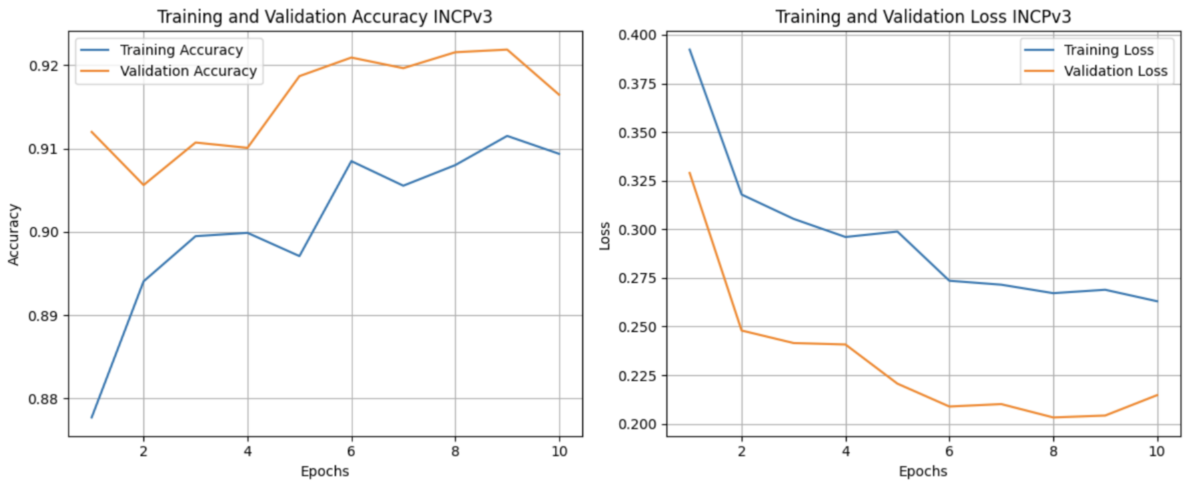

Figure 5.3: Inception V3 Learning Curve

Swin Transformer

In the swin transformer model, validation accuracy was gradually increasing in each epoch. The validation loss is also reduced after 2 epochs which shows a better performance in the validation process. In the training part, the training accuracy and loss curves represent the model that performed better for this dataset. The training and validation curves also demonstrate that there is no overfitting problem in this model. Training and Validation had the same performance proves the models high efficiency.

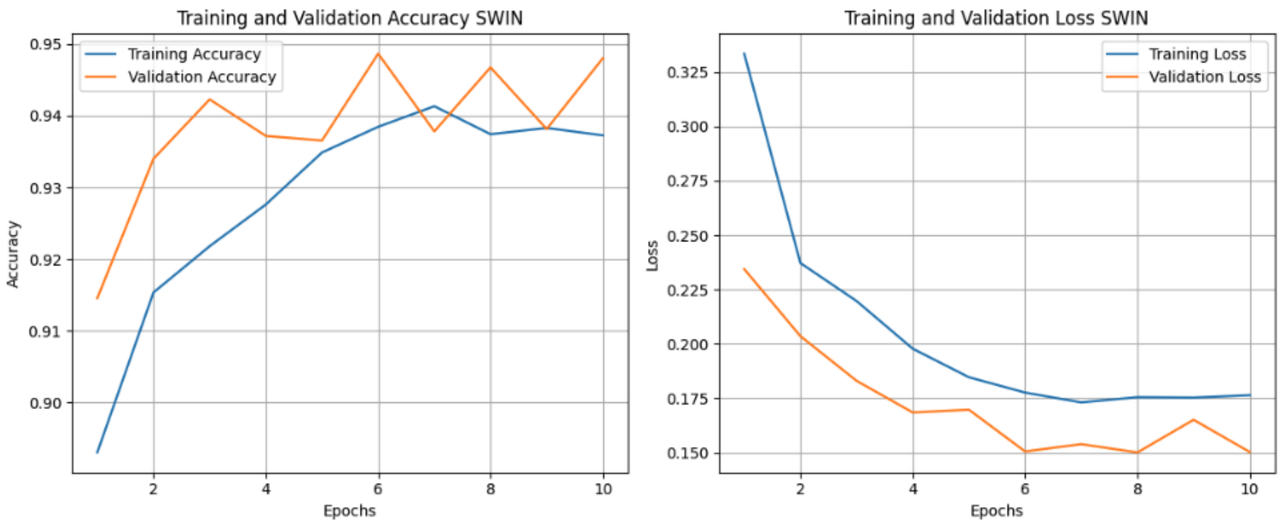

Figure 5.4: Swin Transformer Learning Curve

Vision Transformer

In the above graph 5.5, the accuracy rate of validation increases, and in the case of validation loss, the curve decreases gradually. The loss rate started to decrease from epochs 1 to 6, Here, between epochs 6 and 10, the rate of loss had slight fluctuations. The lowest rate of loss we got was between epochs 6 and 8. Then again, the rate of loss increased a bit between epochs 8 and 10. In the training graph, the rate of loss decreases gradually from epoch 1. No overfitting issue can be identified in the learning curve. This model has also high accuracy and a low loss rate.

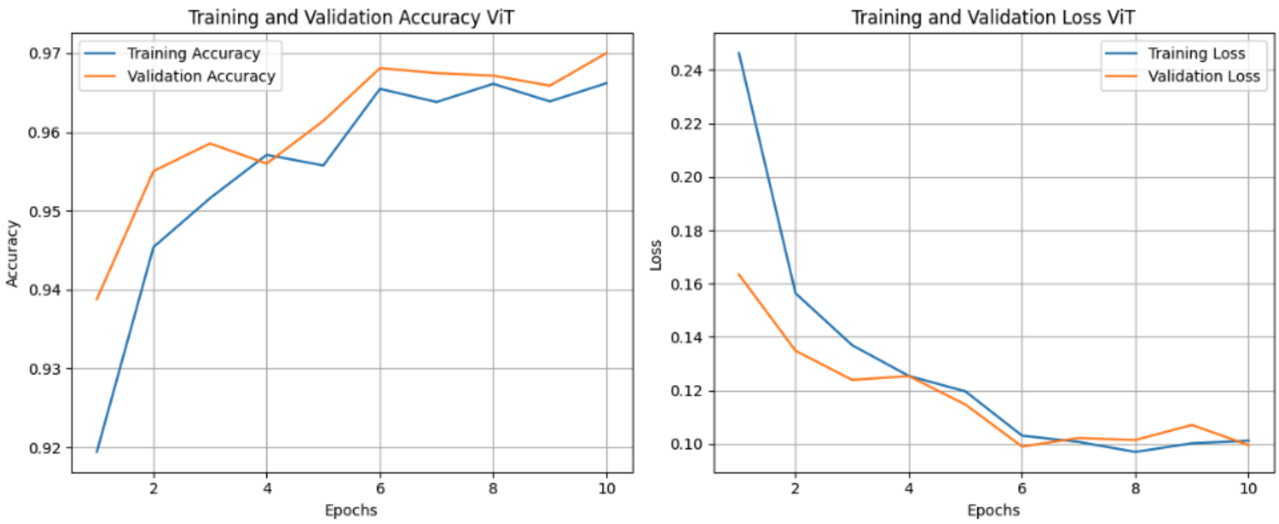

Figure 5.5: Vision Transformer Learning Curve

5.1.2 Hybrid 1 : Swin + ResNet50

In the training and validation graphs (HYBRID1), the training accuracy curve steadily improves, reaching above 90% by the 14th epoch. The validation accuracy follows a similar trend, maintaining a slightly higher performance than the training accuracy. On the loss side, both training and validation loss decrease consistently, with the validation loss stabilizing around 0.20 by the end of the training period. The model shows minimal overfitting, as the accuracy and loss curves for both training and validation remain close.

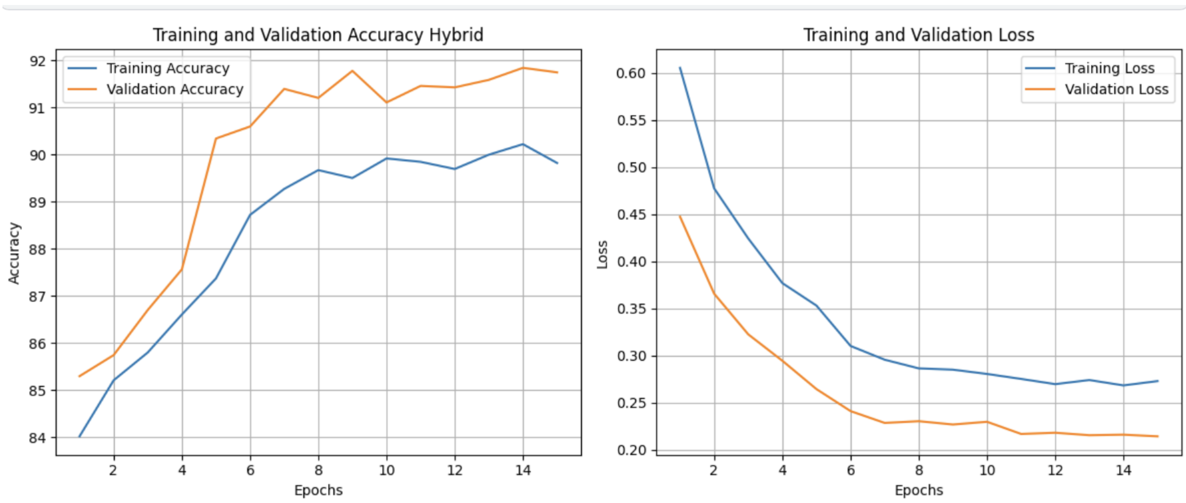

Figure 5.6: SWin + ResNet50 Learning Curve

5.1.3 Hybrid 2 : VGG19 + Vit

In the Hybrid 2 model demonstrates an even stronger performance, with training accuracy approaching closely at 98% approximately at the 10th epoch and validation accuracy tracking like 98% approximately. The loss curves decrease sharply in the first few epochs, with the training and validation losses converging around 125 by the 10th epoch. This model exhibits slightly better overall accuracy and lower loss, suggesting a more efficient training process with minimal overfitting.

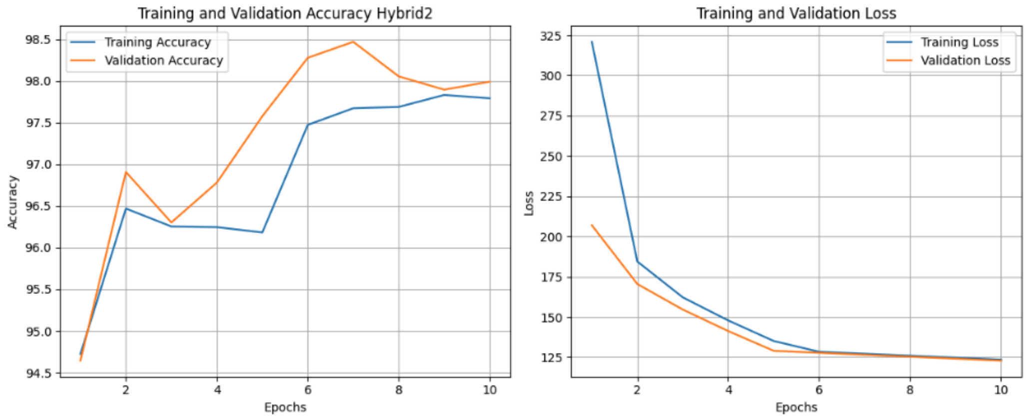

Figure 5.7: VGG19 + Vit Learning Curve

5.1.4 Ensemble Method

A learning curve for an ensemble model illustrates the relationship between training iterations and the model's performance, typically shown as training and validation error rates.After the 10 epoch the error was increasing gradually and the epoch number 12, the loss was maximum after that the loss was decreasing. In the training curve the loss was minimum in the 14th epoch. On the other hand, the validation accuracy remained flat across every epoch.

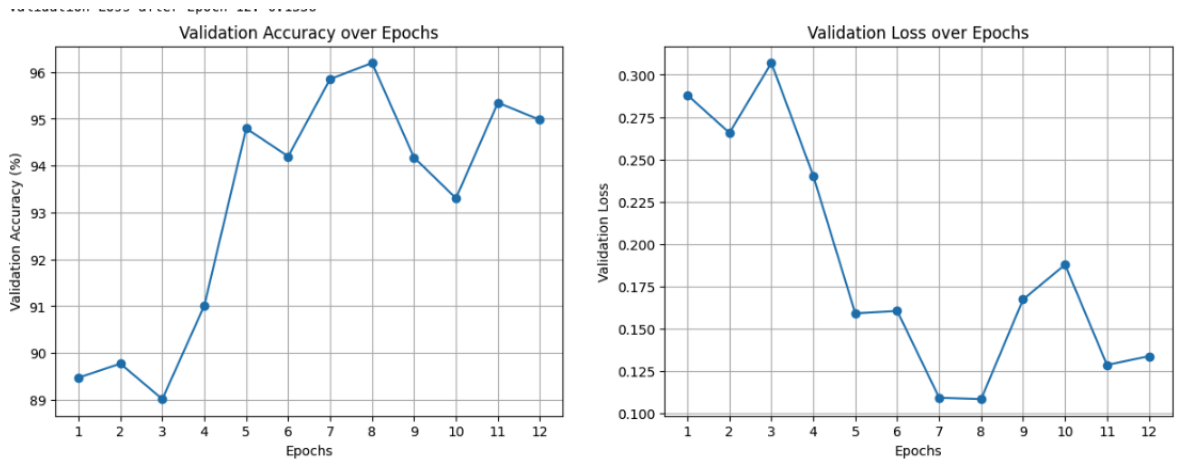

Figure 5.8: Ensemble Bagging Learning Curve

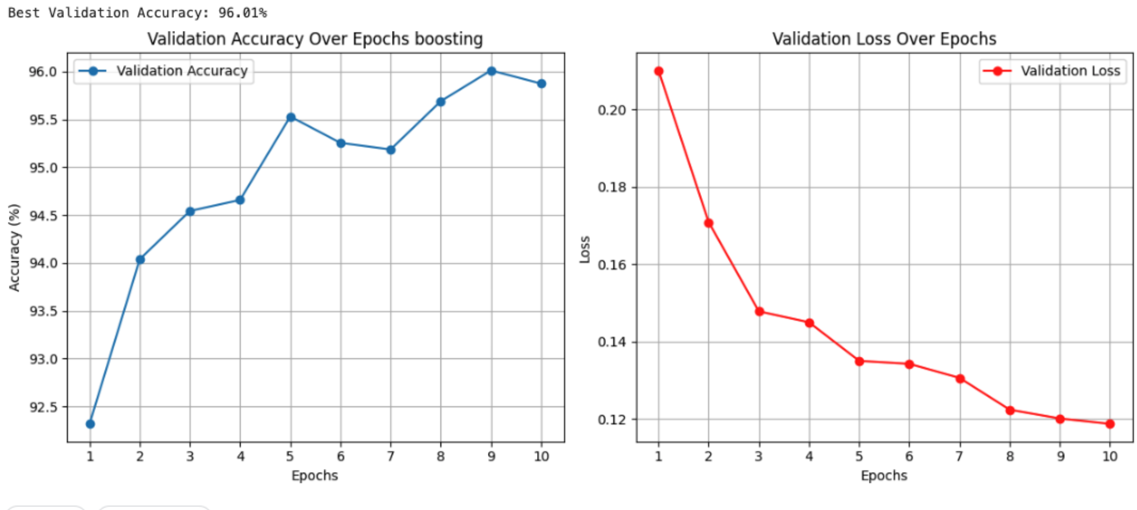

Figure 5.9: Ensemble Boosting Learning Curve

5.1.5 Confusion Matrix

The confusion matrix is a fundamental tool for evaluating the performance of classification models, especially in the context of medical data. Although it is not a performance metric in itself, it provides valuable insights that can be used to derive various performance metrics. In medical applications, the importance of confusion matrix is overemphasized. This allows a detailed analysis of the model's predictions across classes, indicating the number of true positive and true negative predictions, as well as the number of false positives and false negatives. This detailed breakdown is crucial for understanding the performance of the model in the clinical setting, where the cost of misclassification can have a significant impact. Here all the confusion matrices for different models are presented.

VGG19

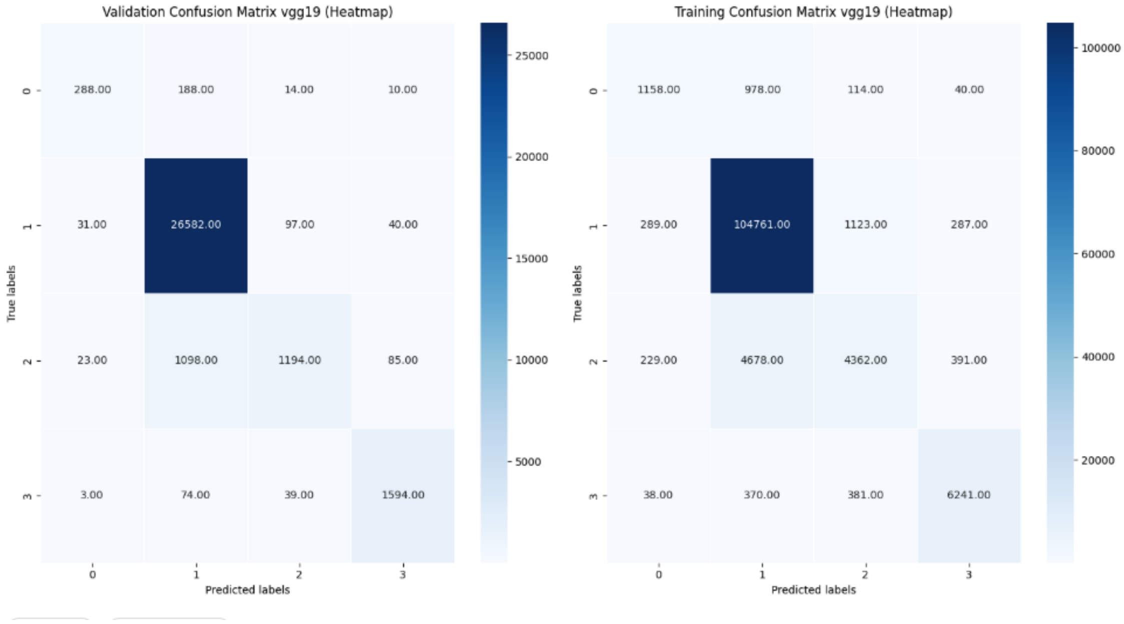

Figure 5.10: Confusion Matrix Heatmap

ResNet50

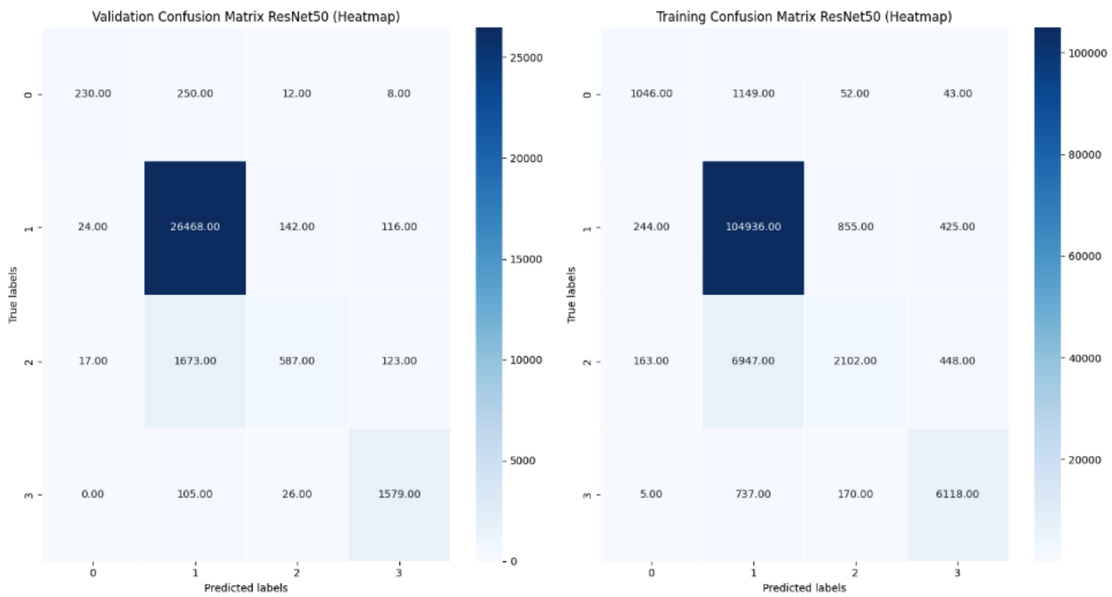

Figure 5.11: Confusion Matrix for ResNet50

Inception V3

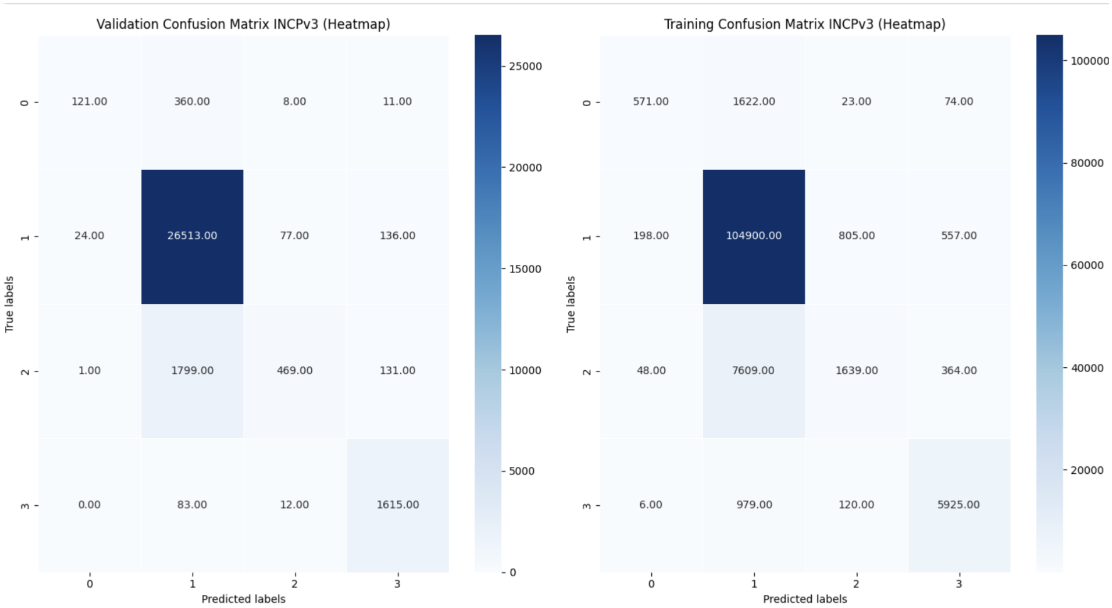

Figure 5.12: Confusion Matrix for Inception V3

Swin Transformer

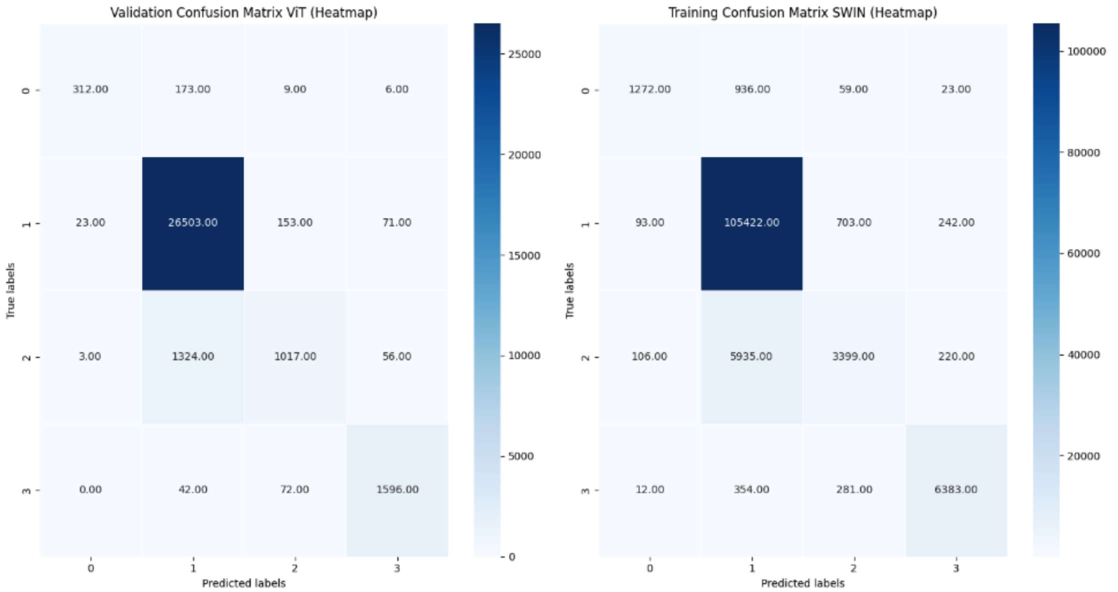

Figure 5.13: Confusion Matrix for Swin Transformer

Vision Transformer

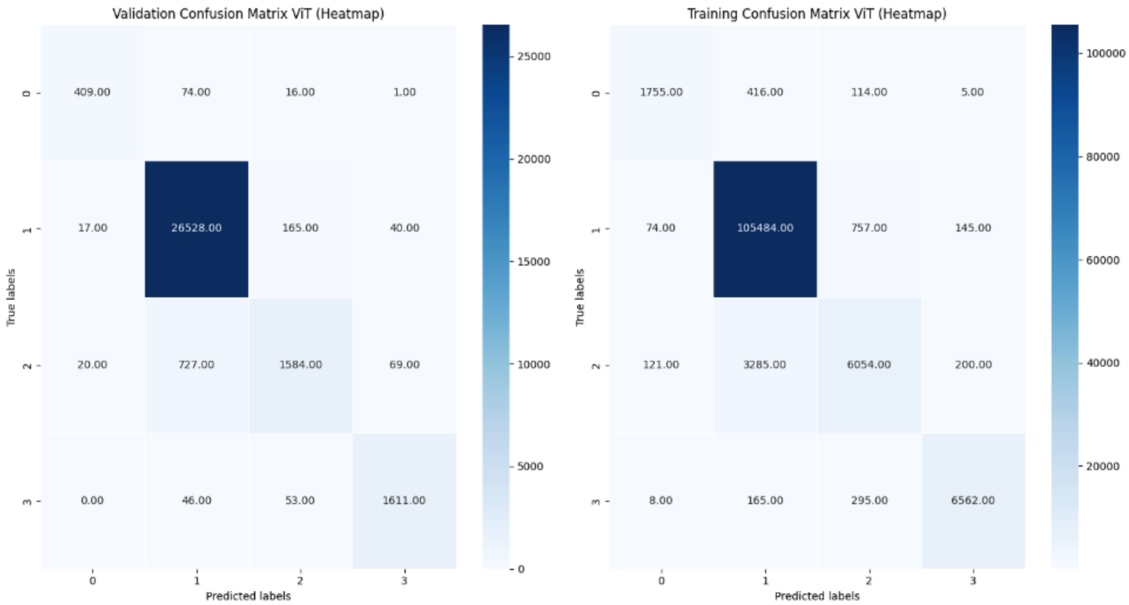

Figure 5.14: Confusion Matrix for Vision Transformer

Hybrid Model 1: Swin + ResNet50

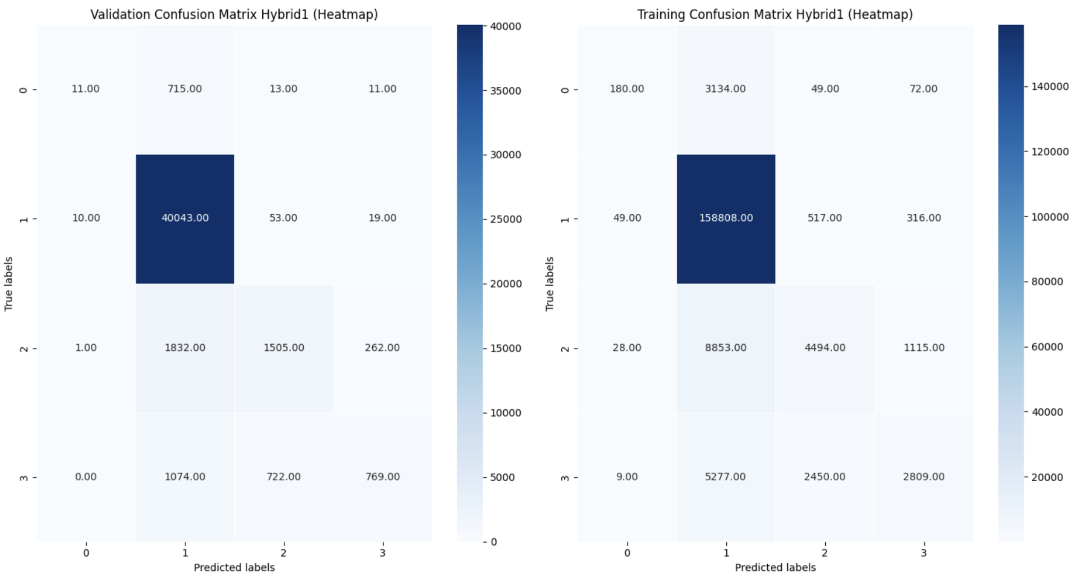

Figure 5.15: Confusion Matrix for Swin + ResNet50

Hybrid Model 2: VGG19 + Vision Transformer

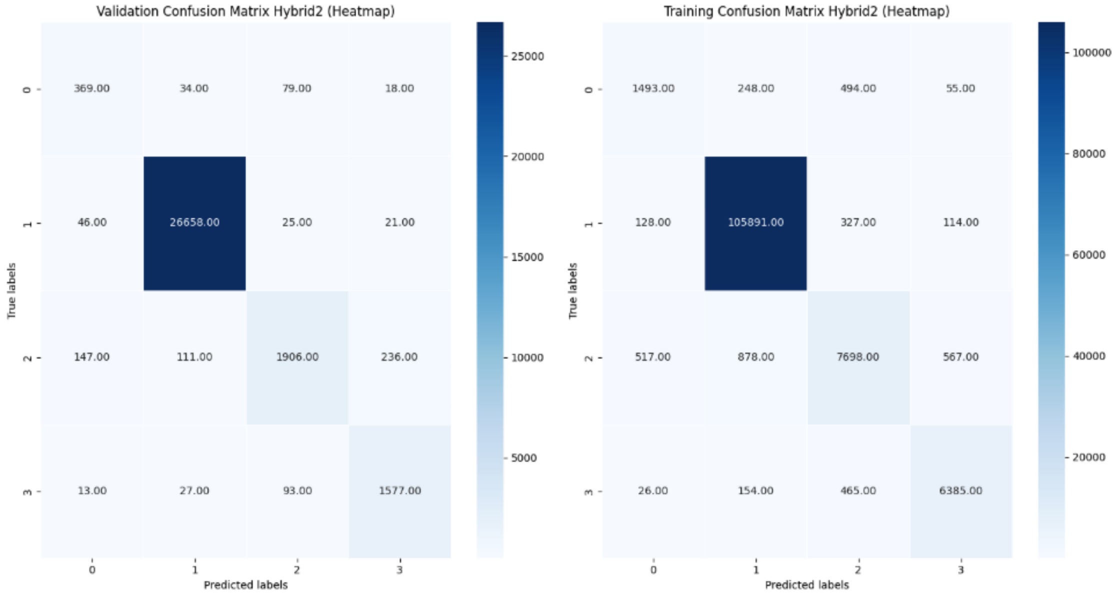

Figure 5.16: Confusion Matrix for VGG19 + Vision Transformer

Confusion Matrix of Ensemble Model

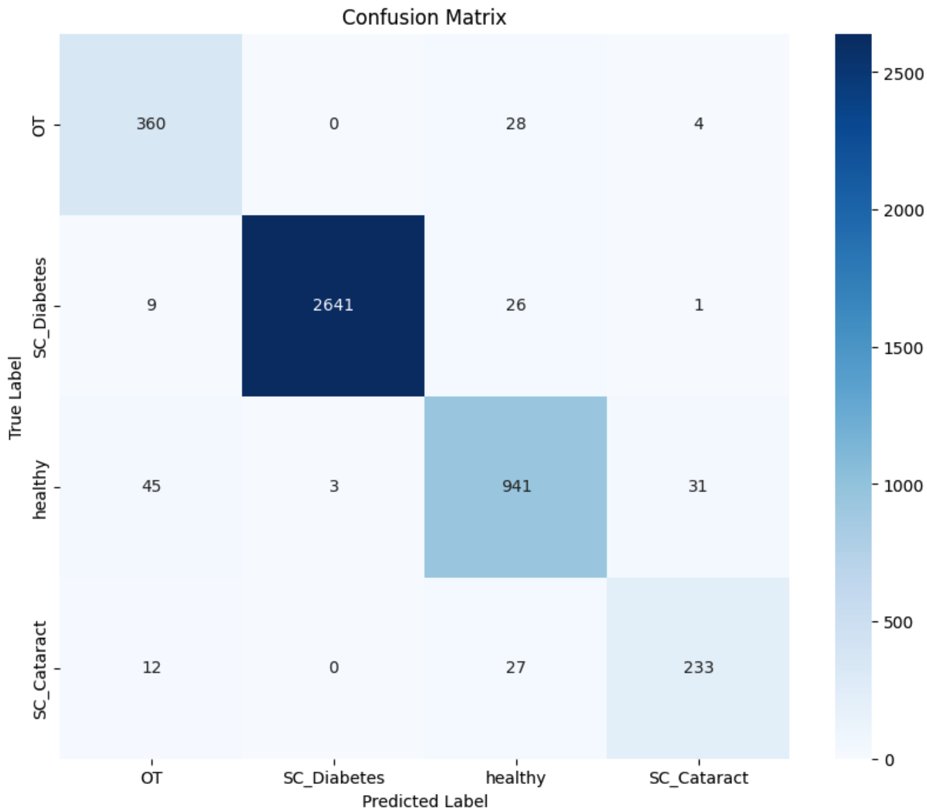

Figure 5.17: Confusion Matrix for Ensemble Bagging

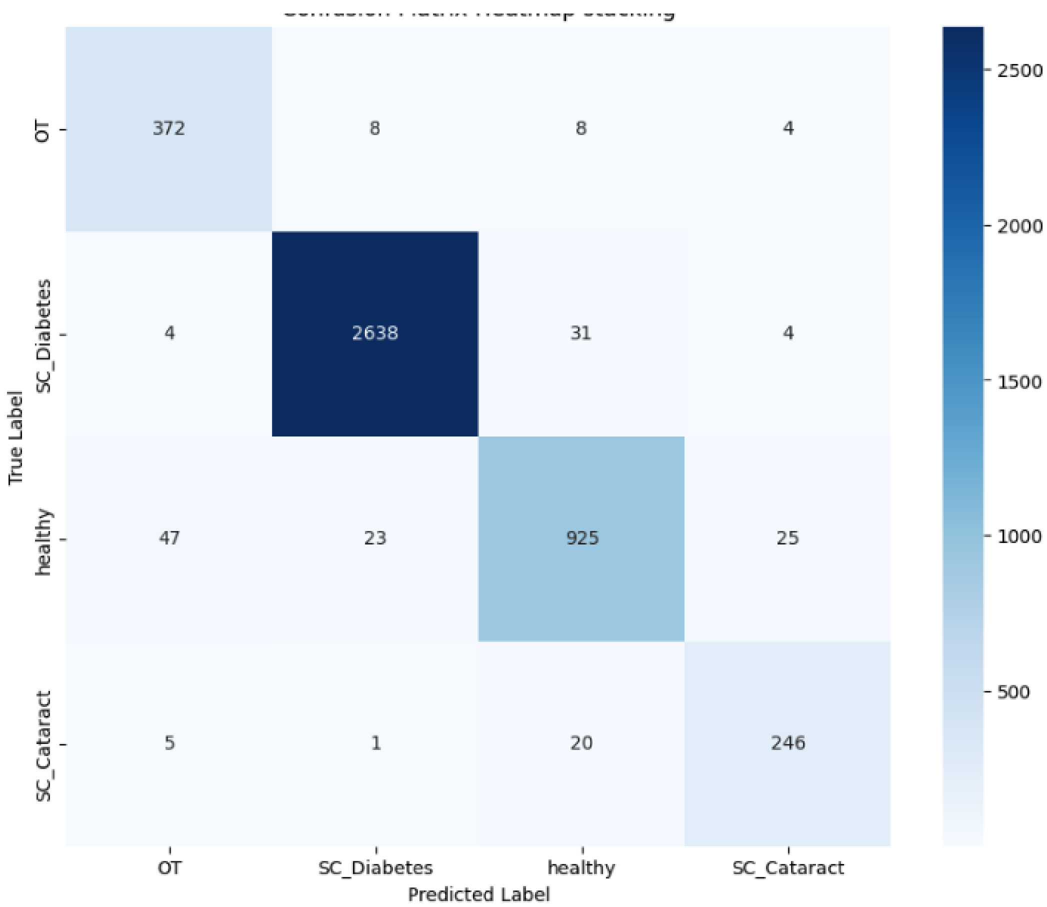

Figure 5.18: Confusion Matrix for Ensemble boosting

5.1.6 ROC Curve Analysis

The ROC (Receiver Operating Characteristic) curve illustrates the diagnostic ability of a classifier across all classification thresholds. The ROC curve shows how well a model can distinguish between different classes. It can find out whether a model is good or bad. ROC curve is mainly used for binary classification but with a proper modification it is usable for multi class classification.

VGG19

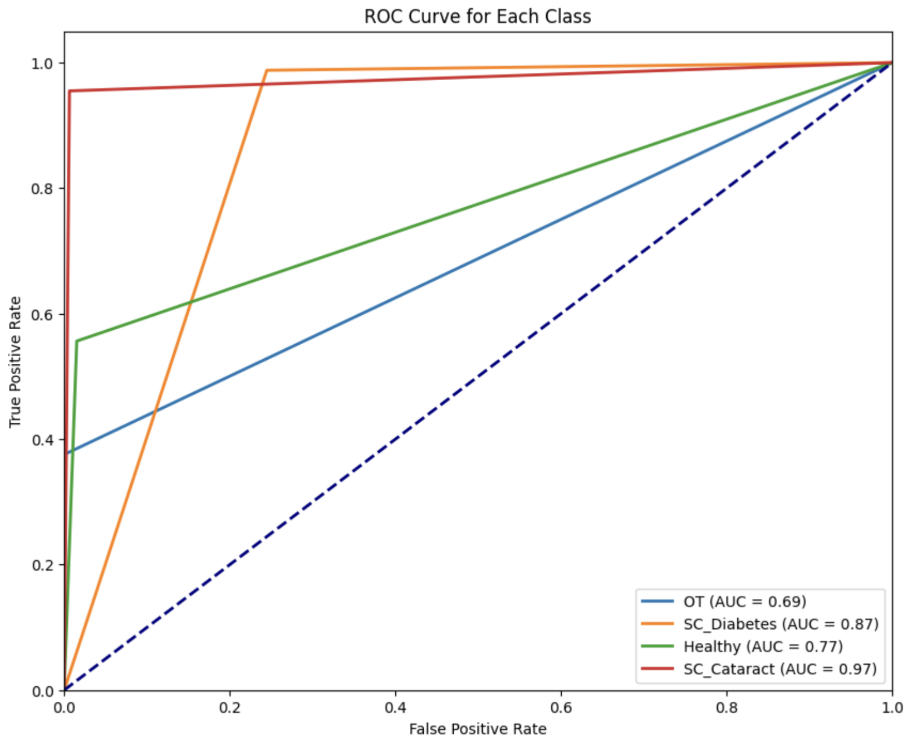

Figure 5.19: ROC Curve for VGG19

ResNet50

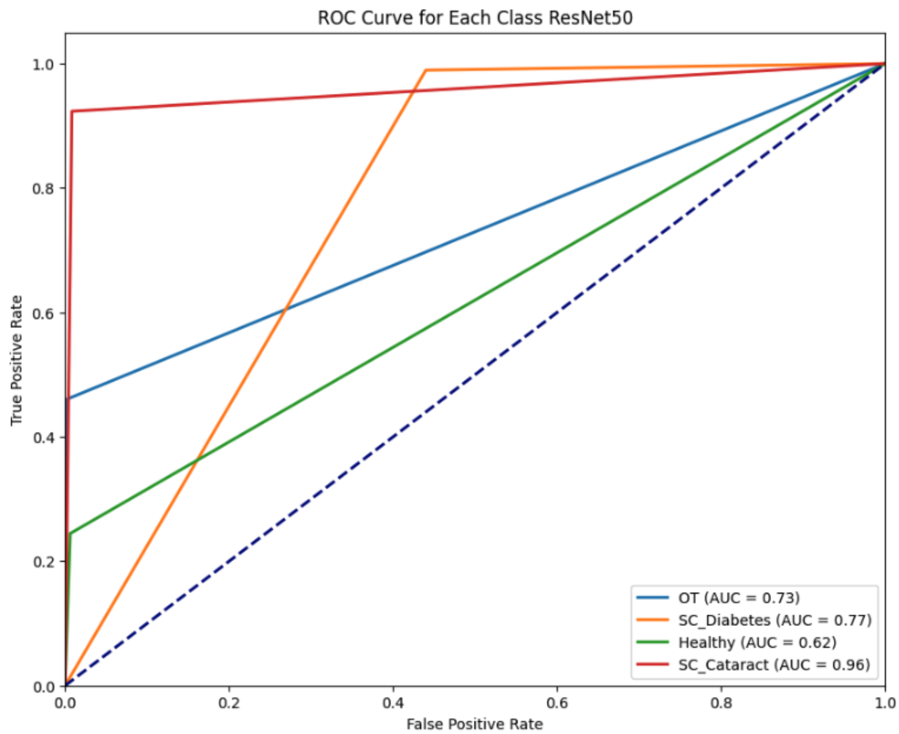

Figure 5.20: ROC Curve for ResNet50

Inception V3

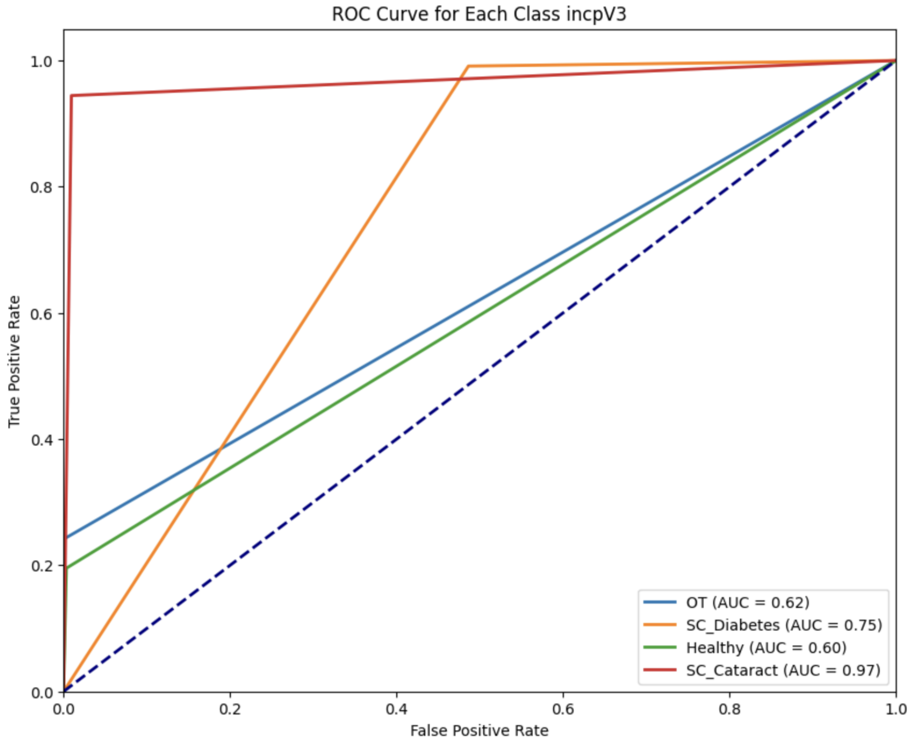

Figure 5.21: ROC Curve for Inception V3

Swin Transformer

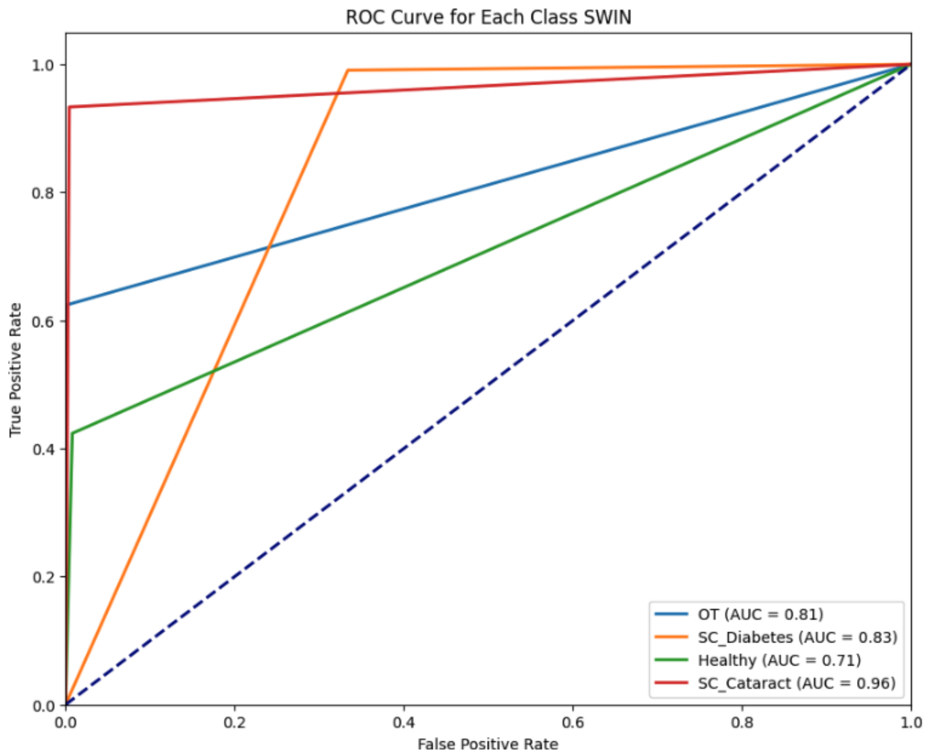

Figure 5.22: ROC Curve for Swin Transformer

Vision Transformer

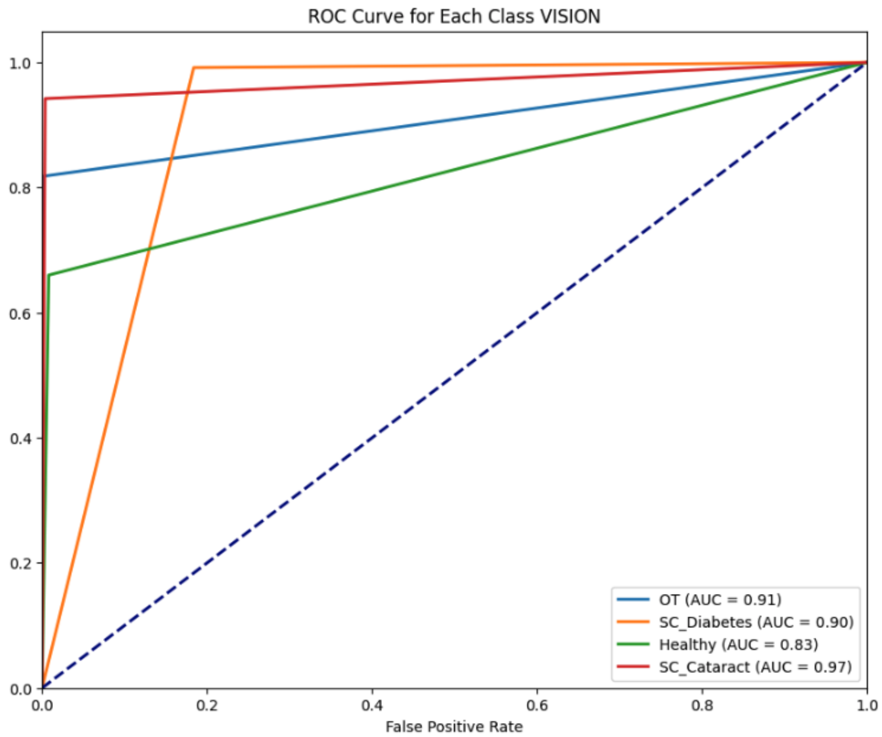

Figure 5.23: ROC Curve for Vision Transformer

Hybrid Model 1: Swin + ResNet50

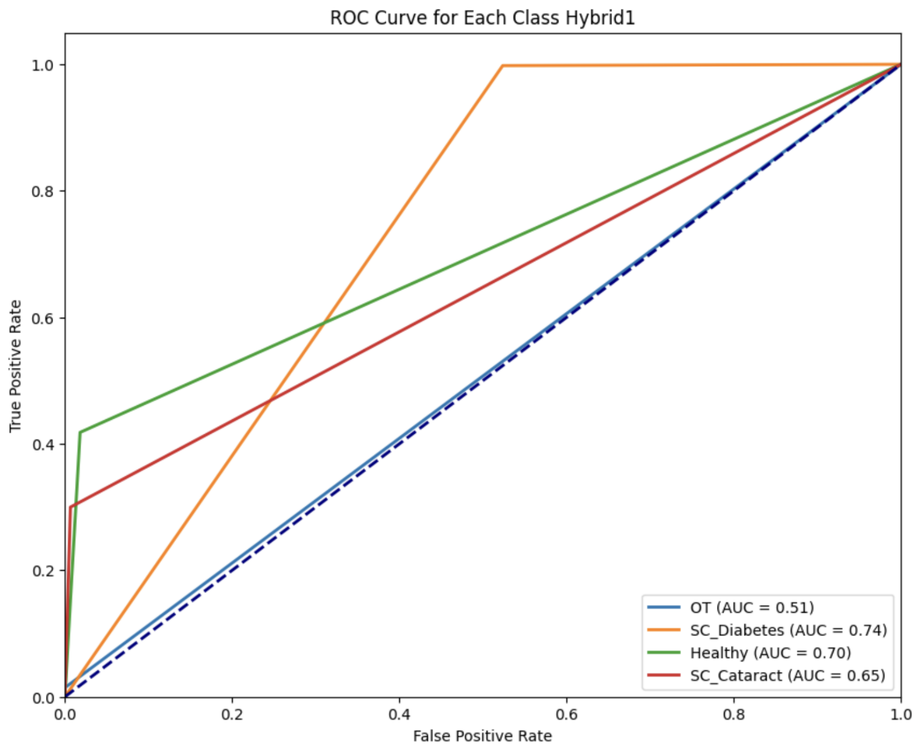

Figure 5.24: ROC Curve for Swin + ResNet50

Hybrid Model 2: VGG19 + Vision Transformer

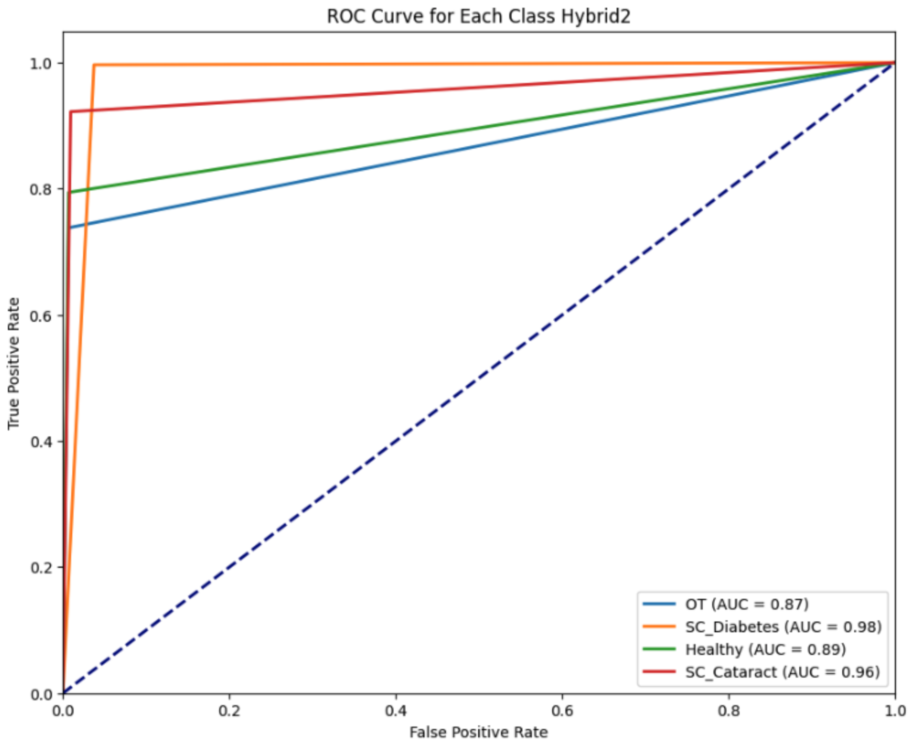

Figure 5.25: ROC Curve for VGG19 + Vision Transformer

Ensembling

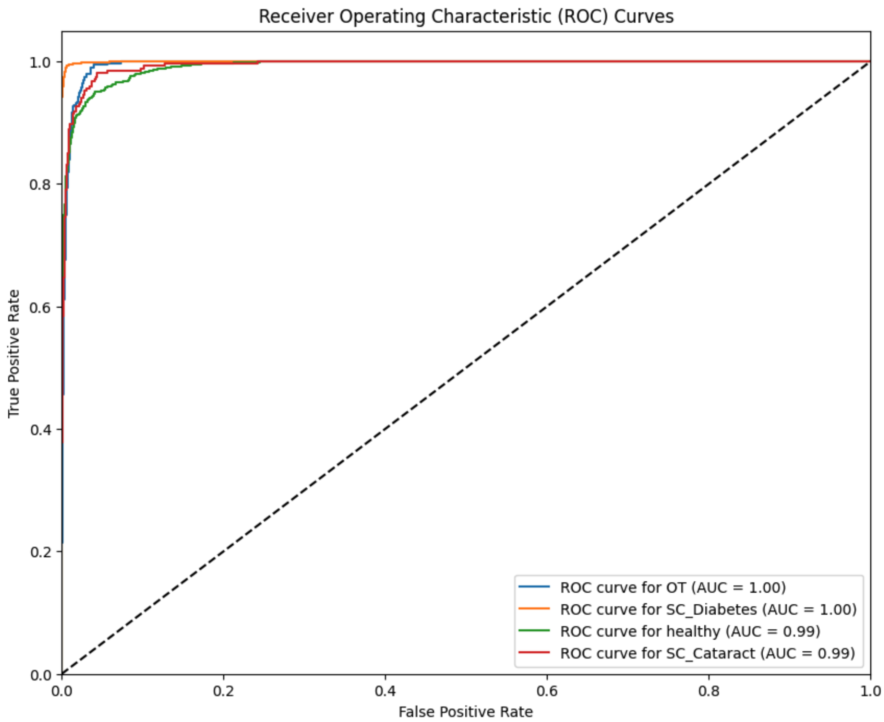

Figure 5.26: ROC Curve for Ensemble

5.2 Model Comparisons

We achieved higher accuracy and f-1 scores from transformer-based models, especially the Vision Transformer and Swin Transformer. Also, we achieved higher accuracy from conventional CNNs like VGG19 with 96.11%. With a 96.62% accuracy and 88.37% f1-score, the Vision Transformer exceeds the Swin Transformer, which comes in second with a 94.87% accuracy and a 79.93% f1-score. In the hybrid models, more potential has been shown such as VGG19+ VIT, which provides the best accuracy of 98.26%, and sensitivity, specificity, precision, and f-1 scores are also higher than others. Traditional models like ResNet50 and Inception V3 perform very well which results in a great percentage of misclassifications. Overall, Hybrid Model 1(VGG19+VIT) shows greater and superior performance whereas other models also show potential.

Architecture	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**F1-Score**
VGG19	96.11%	74.98%	92.33%	89.91%	80.46%
ResNet50	93.14%	88.60%	65.43%	85.19%	80.46%
InceptionV3	92.19%	87.50%	69.73%	85.81%	73.56%
Vision Transformer	96.62%	95.08%	85.29%	92.33%	88.37%
Swin Transformer	94.87%	91.31%	74.29%	90.10%	79.93%
VGG19 + VIT	98.26%	97.29%	99.10%	97.29%	97.29%
Swin + ResNet50	91.84%	89.98%	96.66%	89.98%	89.98%

Table 5.1: Model Performance Metrics

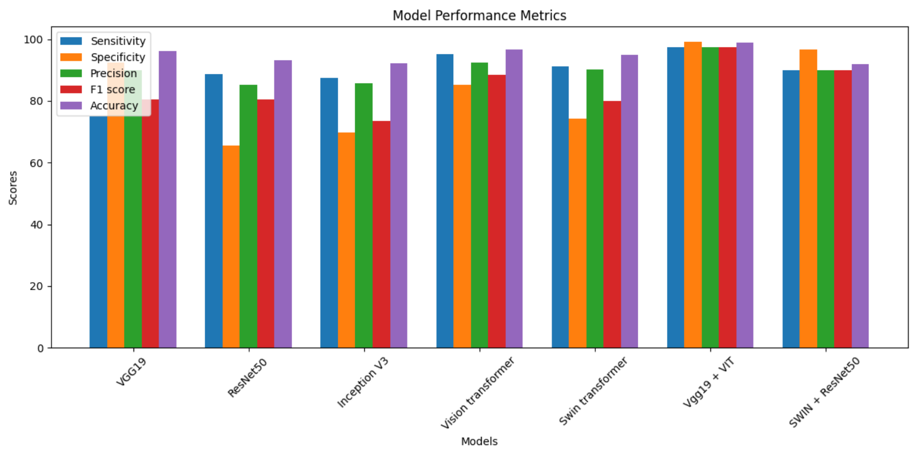

Figure 5.27: Illustration of Comparison Among the Used Architectures Using a Bar-Chart

5.2.1 Ensemble Methods Result Analysis

We have implemented two ensemble techniques - bagging and boosting which are compared in the table. From bagging we have achieved our best accuracy of 96.68% ,recall 95.94%,precision 96.04% and F1 score 95.95% .On the other hand, boosting does well in terms of recall 93.64% and precision 91.97% following with 96.01% of accuracy. Overall, stacking gets the best result while boosting provides a competitive alternate.Based on the best accuracy we implemented XAI model on Ensembling method bagging.

Metric	Precision	ReCall	F1-Score	Specificity	Accuracy
Bagging	96.04%	95.94%	95.98%	98.78%	96.68%
Boosting	91.97%	93.64%	92.75%	93.64%	96.01%

Table 5.2: Ensemble Methods Performance Metrics

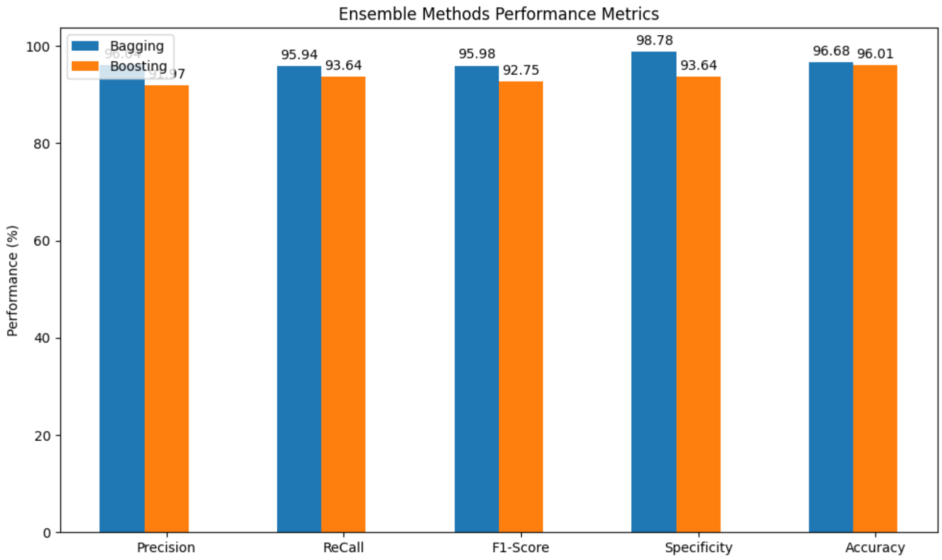

Figure 5.28: Illustration of Comparison between the Ensemble Methods(Bagging & Boosting)

5.3 Explainable Artificial Intelligence

After implementing Explainable AI, we have noticed that there are 3 types of masks represented by different colors. The overlay highlight's the regions of the image that the VGG19 model deemed most relevant for making its classification decision. And here areas of high importance is represented in warm colors (like red and yellow) on top of the original image, while less important areas are depicted in transparent colors.

The GradCAM illustrates how ResNet50 model effectively classifies eye conditions by identifying relevant features with the help of heatmaps it uses color gradient to indicate the importance of different regions. Warmer colors (red, yellow) correspond to higher importance while cooler colors (Blue,Green) indicate lower importance. Grad-CAM masks effectively demonstrate the vision transformer's decision making process .Here its noticeable that Scdiabetes image highlights the central area where the lens is located, it denotes the critical region for the model's diagnosis.

The overlay highlight's the regions of the image that the VGG19 model deemed most relevant for making its classification decision. And here areas of high importance will is represented in warm colors (like red and yellow) on top of the original image, while less important areas are depicted in transparent colors.

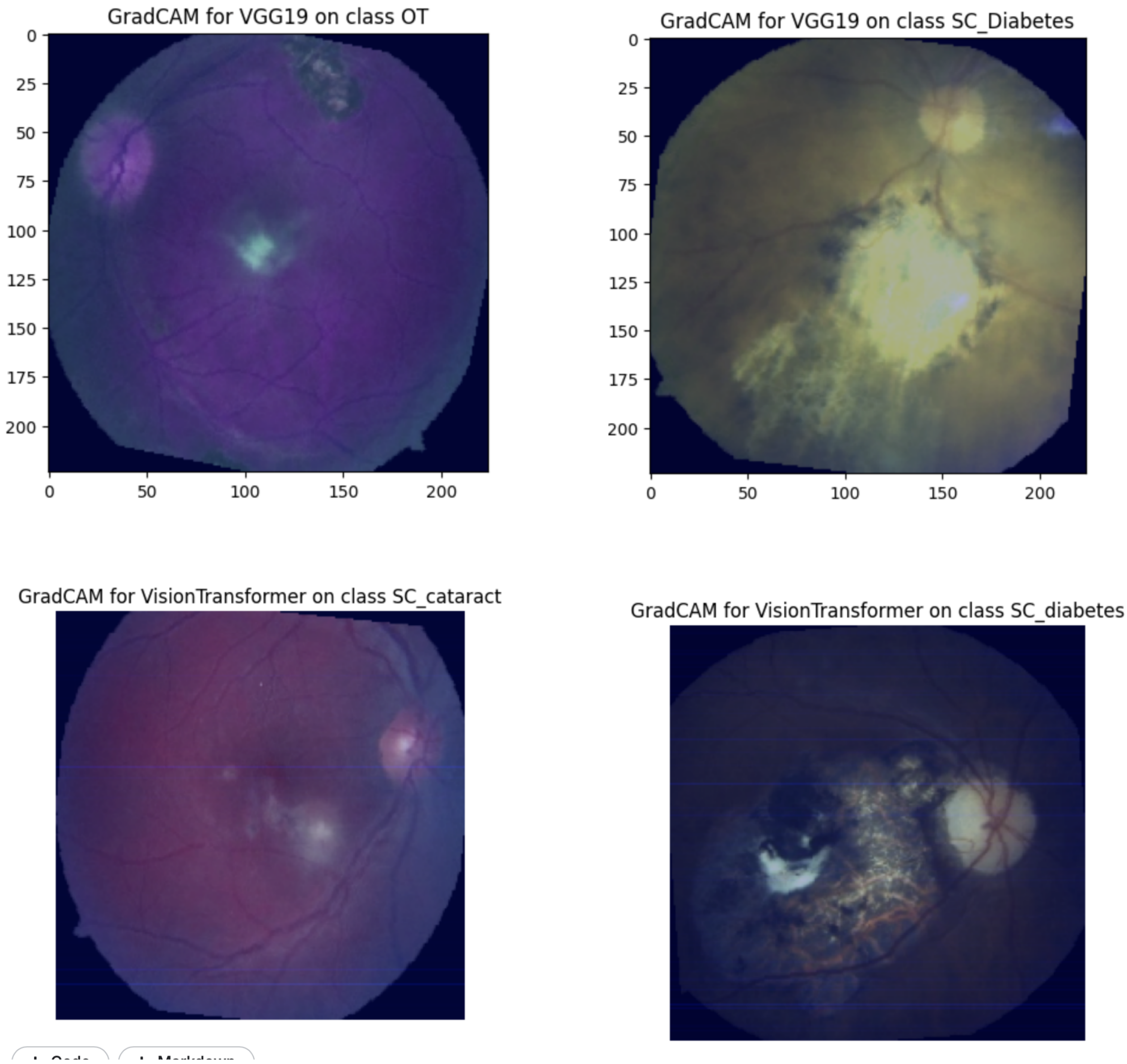

Figure 5.29: XAI Representaion

Grad-CAM masks effectively demonstrate the vision transformer's decision making process .Here its noticeable that Scdiabetes image highlights the central area where the lens is located, it denotes the critical region for the model's diagnosis.

5.4 Challenges and Limitations

Training several deep learning models such as InceptionV3, ResNet50, Vision Transformers for multiclass classification of ocular toxoplasmosis and its secondary complications face several challenges. Also it was challenging to implement hybrid model based on Imagenet weight and transformer model configuration with fine proper fine-tuning. One of the major issues is limited availability of well-organized medical datasets. Ocular toxoplasmosis is a comparably rare condition and its secondary complications can differ frequently due to class imbalance of the dataset. Such imbalance of dataset can also cause the models to clash with the accurate recognition of diminished classes which results in poor abstraction to real world data as well as overfitting issues. With such a small dataset, there is a risk of overfitting where the models normally memorize the specific patterns from the training data instead of learning general aspects which are applicable to unseen facts. Such things can reduce the effectiveness and sharpness of the models in clinical settings. It should need to be solid and vigorous against the variations in patient data, conditions of imaging, presentations of disease. To overcome such challenges, it is important to extend the dataset with a wide range of fundus images that can serve as the diverse demonstrations of ocular toxoplasmosis. Moreover, leading to such dataset include common and extraordinarily rare complications which can improve the models ability to complement the accuracy of classifications. Moreover, class weighting, model tuning, fine-tune dropout , regularization can ease overfitting and grant models to perform a better way with limited data. As deep learning shows boundless commitment in reforming the diagnosis of ocular toxoplasmosis, these drawbacks should be addressed to ensuring the system's accuracy and sharpness in real world medical applications.

Chapter 6

Conclusion

This study presents an innovative and efficient deep-learning approach for the multi-classification of ocular toxoplasmosis and its secondary complications using fundus images. The diagnosis of ocular toxoplasmosis has not received much attention in the field of deep learning. This study addresses the lack of focus on ocular toxoplasmosis in deep learning by proposing an efficient approach for its multiclass-classification and detection using deep learning techniques. The proposed hybrid model integrates multiple deep learning models to improve its capacity to reliably classify diseases and their consequences. Moreover, the suggested technique takes advantage of deep learning methods and hybrid model architecture as well as ensemble algorithm to achieve high classification accuracy and resilient performance in recognizing various manifestations of ocular toxoplasmosis and its secondary complications. It uses a diverse, well-annotated dataset of ocular and its after effect imaging. This method exceeds conventional diagnostic approaches in terms of accuracy, sensitivity, and specificity, resulting in a precise and time-saving diagnostic tool. This deep learning system provides precise, efficient diagnostic capabilities that have the potential to improve the management of ocular toxoplasmosis. It emphasizes deep learning's promise in ophthalmology. The future study intends to increase dataset diversity, provide user-friendly clinical integration tools for early detection, and enhance the model through additional research in a variety of situations to ensure real-world deployment while adjusting to evolving disease patterns.

6.1 Future Work

Future efforts could involve collecting more balanced datasets and exploring other hybrid architectures to improve classification accuracy further. We aim to enrich our dataset with a range of retinal pictures from various populations and areas, together with precise annotations from ophthalmologists to enhance our deep learning studies on ocular toxoplasmosis. We will employ advanced data augmentation approaches to improve the model's resilience and generalization. We want to perform substantial cross-validation and hyper-parameter optimization in order to fine-tune our hybrid models which fuse CNN's with transformer-based architectures. Furthermore, in order to ensure that the interface satisfies clinical usability criteria. We will provide an intuitive interface for real-time disease detection that will allow doctors to upload, evaluate and produce diagnostic reports quickly.Also for framework tensorflow and TPU can be tested. Also more classified classes of disease with enriched dataset could improve significant improvement in classification. Multiple optimizer for instance RMSProp,SGD,SGD with momentum ,AdamW can be implemented to analyze the model performance and changing the rhythm of hyperparameter tuning .To make it more human friendly Stacking method of Ensemble algorithm has potential to become a better option with SHAP XAI explanation.

Bibliography

[1] J. L. Finley, V. V. Joshi, and N. L. Smith, "General Pathology of HIV infection," pp. 733–772, Jan. 2004. DOI: 10.1016/b978-012764051-8/50029-9. [Online]. Available: https://www.sciencedirect.com/topics/medicine-and-dentistry/ocular-toxoplasmosis#:~:text=The%20age%20of%20the%20first,patients%20may%20have%20multiple%20recurrences.

[2] N. Roizen, K. Kasza, T. Karrison, *et al.*, "Impact of visual impairment on measures of cognitive function for children with congenital toxoplasmosis: Implications for compensatory intervention strategies," *Pediatrics*, vol. 118, e379–90, Sep. 2006. DOI: 10.1542/peds.2005-1530.

[3] F. Knapen and P. Overgaauw, "Toxoplasmosis," *The European Journal of Companion Animal Practitioners*, vol. 18, pp. 242–5, Jan. 2008. [Online]. Available: https://www.researchgate.net/publication/233765649_Toxoplasmosis.

[4] L. Phan, K. Kasza, J. Jalbrzikowski, *et al.*, "Longitudinal study of new eye lesions in children with toxoplasmosis who were not treated during the first year of life," en, *Am. J. Ophthalmol.*, vol. 146, no. 3, pp. 375–384, Sep. 2008.

[5] M. R. Stanford and R. E. Gilbert, "Treating ocular toxoplasmosis: Current evidence," en, *Mem. Inst. Oswaldo Cruz*, vol. 104, no. 2, pp. 312–315, Mar. 2009.

[6] M. R. Stanford and R. E. Gilbert, "Treating ocular toxoplasmosis: current evidence," *Memórias do Instituto Oswaldo Cruz*, vol. 104, no. 2, pp. 312–315, Mar. 2009. DOI: 10.1590/s0074-02762009000200027. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/19430659/.

[7] J. L. Jones and G. N. Holland, "Annual burden of ocular toxoplasmosis in the US," en, *Am. J. Trop. Med. Hyg.*, vol. 82, no. 3, pp. 464–465, Mar. 2010.

[8] J. G. Garweg, J. D. F. De Groot-Mijnes, and J. G. Montoya, "Diagnostic approach to ocular toxoplasmosis," *Ocular immunology and inflammation*, vol. 19, no. 4, pp. 255–261, Jul. 2011. DOI: 10.3109/09273948.2011.595872. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/21770803/.

[9] J. G. Garweg, J. D. F. de Groot-Mijnes, and J. G. Montoya, "Diagnostic approach to ocular toxoplasmosis," en, *Ocul. Immunol. Inflamm.*, vol. 19, no. 4, pp. 255–261, Aug. 2011.

[10] N. J. Butler, J. M. Furtado, K. L. Winthrop, and J. R. Smith, "Ocular toxoplasmosis II: clinical features, pathology and management," *Clinical & experimental ophthalmology*, vol. 41, no. 1, pp. 95–108, Sep. 2012. DOI: 10.1111/j.1442-9071.2012.02838.x. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/22712598/.

[11] J. B. McAuley, "Congenital toxoplasmosis," en, *J. Pediatric Infect. Dis. Soc.*, vol. 3 Suppl 1, no. suppl_1, S30–5, Sep. 2014.

[12] Y. Ouyang, F. Li, Q. Shao, *et al.*, "Subretinal Fluid in Eyes with Active Ocular Toxoplasmosis Observed Using Spectral Domain Optical Coherence Tomography," *PloS one*, vol. 10, no. 5, e0127683, May 2015. DOI: 10.1371/journal.pone.0127683. [Online]. Available: https://doi.org/10.1371/journal.pone.0127683.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[14] A. Alibhai, C. Or, and A. Witkin, "Swept source optical coherence tomography: A review," *Current Ophthalmology Reports*, vol. 6, Mar. 2018. DOI: 10.1007/s40135-018-0158-3.

[15] A. B. Cheng Ju and M. van der Laan, "The relative performance of ensemble methods with deep convolutional neural networks for image classification," *Journal of Applied Statistics*, vol. 45, no. 15, pp. 2800–2818, 2018, PMID: 31631918. DOI: 10.1080/02664763.2018.1441383. eprint: https://doi.org/10.1080/02664763.2018.1441383. [Online]. Available: https://doi.org/10.1080/02664763.2018.1441383.

[16] P. S. Grewal, F. Oloumi, U. Rubin, and M. T. Tennant, "Deep learning in ophthalmology: a review," *Canadian journal of ophthalmology*, vol. 53, no. 4, pp. 309–313, Aug. 2018. DOI: 10.1016/j.jcjo.2018.04.019. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30119782/.

[17] Y. Tong, W. Lu, Y. Yu, and Y. Shen, "Application of machine learning in ophthalmic imaging modalities," *Eye and Vision*, vol. 7, Dec. 2020. DOI: 10.1186/s40662-020-00183-6.

[18] Y. Tong, W. Lu, Y. Yu, and Y. Shen, "Application of machine learning in ophthalmic imaging modalities," *Eye and Vision*, vol. 7, Dec. 2020. DOI: 10.1186/s40662-020-00183-6.

[19] B. Triwijoyo, B. Sabarguna, W. Budiharto, and E. Abdurachman, "Deep learning approach for classification of eye diseases based on color fundus images," in Jan. 2020, pp. 25–57, ISBN: 9780128174401. DOI: 10.1016/B978-0-12-817440-1.00002-4.

[20] A. Dosovitskiy, L. Beyer, A. Kolesnikov, *et al.*, *An image is worth 16x16 words: Transformers for image recognition at scale*, 2021. arXiv: 2010.11929 `[cs.CV]`. [Online]. Available: https://arxiv.org/abs/2010.11929.

[21] D. Kalogeropoulos, H. Sakkas, B. Mohammed, *et al.*, "Ocular toxoplasmosis: a review of the current diagnostic and therapeutic approaches," *International ophtalmology/International ophthalmology*, vol. 42, no. 1, pp. 295–321, Aug. 2021. DOI: 10.1007/s10792-021-01994-9. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/34370174/.

[22] R. Parra, V. Ojeda, J. L. V. Noguera, *et al.*, "A Trust-Based Methodology to Evaluate Deep Learning Models for Automatic Diagnosis of Ocular Toxoplasmosis from Fundus Images," *Diagnostics*, vol. 11, no. 11, p. 1951, Oct. 2021. DOI: 10.3390/diagnostics11111951. [Online]. Available: https://www.researchgate.net/publication/355490831_A_Trust-Based_Methodology_to_Evaluate_Deep_Learning_Models_for_Automatic_Diagnosis_of_Ocular_Toxoplasmosis_from_Fundus_Images.

[23] R. Parra, V. Ojeda, J. L. V. Noguera, *et al.*, "Automatic Diagnosis of Ocular Toxoplasmosis from Fundus Images with Residual Neural Networks," May 2021. DOI: 10.3233/shti210143. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/34042728/.

[24] M. Abellanas, M. J. Elena, P. A. Keane, K. Balaskas, D. S. Grewal, and E. Carreño, "Artificial intelligence and imaging processing in optical coherence tomography and digital images in UVeitis," *Ocular immunology and inflammation*, vol. 30, no. 3, pp. 675–681, Apr. 2022. DOI: 10.1080/09273948.2022.2054433. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/35412935/.

[25] A. Das, S. K. Mohapatra, and M. N. Mohanty, "Design of deep ensemble classifier with fuzzy decision method for biomedical image classification," *Applied Soft Computing*, vol. 115, p. 108 178, 2022, ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2021.108178. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494621010309.

[26] D. Müller, I. Soto-Rey, and F. Kramer, "An analysis on ensemble learning optimized medical image classification with deep convolutional neural networks," *IEEE Access*, vol. 10, pp. 66 467–66 480, 2022. DOI: 10.1109/ACCESS.2022.3182399.

[27] P. Patra and T. Singh, "Diabetic retinopathy detection using an improved resnet 50-inceptionv3 and hybrid diabretnet structures," in *2022 OITS International Conference on Information Technology (OCIT)*, 2022, pp. 140–145. DOI: 10.1109/OCIT56763.2022.00036.

[28] S. S. Alam, S. B. Shuvo, S. N. Ali, F. Ahmed, A. Chakma, and Y. M. Jang, *Benchmarking Deep Learning Frameworks for Automated diagnosis of Ocular Toxoplasmosis: A Comprehensive Approach to Classification and Segmentation*, May 2023. [Online]. Available: https://arxiv.org/abs/2305.10975.

[29] C. Cifuentes-González, W. Rojas-Carabali, Á. O. Pérez, *et al.*, "Risk factors for recurrences and visual impairment in patients with ocular toxoplasmosis: A systematic review and meta-analysis," *PloS one*, vol. 18, no. 4, e0283845, Apr. 2023. DOI: 10.1371/journal.pone.0283845. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/37011101/.

[30] "Curcumin ameliorates chronic toxoplasma gondii infection-induced affective disorders through modulation of proinflammatory cytokines and oxidative stress," *National Library Medicine*, no. PMC10008396, Apr. 2023. DOI: 10.22038/IJBMS.2023.68487.14937. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10008396/.

[31] E. Ghafourian, F. Samadifam, H. Fadavian, P. Jerfi Canatalay, A. Tajally, and S. Channumsin, "An ensemble model for the diagnosis of brain tumors through mris," *Diagnostics*, vol. 13, no. 3, 2023, ISSN: 2075-4418. DOI: 10.3390/diagnostics13030561. [Online]. Available: https://www.mdpi.com/2075-4418/13/3/561.

[32] T. Iqball and M. A. Wani, "Weighted ensemble model for image classification," en, *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 557–564, Jan. 2023. DOI: https://doi.org/10.1007/s41870-022-01149-8.

[33] S. R. Shah, S. Qadri, H. Bibi, S. M. W. Shah, M. I. Sharif, and F. Marinello, "Comparing inception v3, VGG 16, VGG 19, CNN, and ResNet 50: A case study on early detection of a rice disease," en, *Agronomy (Basel)*, vol. 13, no. 6, p. 1633, Jun. 2023. DOI: https://doi.org/10.3390/agronomy13061633.

[34] S. R. Shah, S. Qadri, H. Bibi, S. M. W. Shah, M. I. Sharif, and F. Marinello, "Comparing inception v3, VGG 16, VGG 19, CNN, and ResNet 50: A case study on early detection of a rice disease," en, *Agronomy (Basel)*, vol. 13, no. 6, p. 1633, Jun. 2023. DOI: https://doi.org/10.3390/agronomy13061633.