KULLIYYAH OF INFORMATION AND COMMUNICATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

FINAL YEAR PROGRESS REPORT

PROJECT ID

1205 D

PROJECT TITLE

OCT-BASED DIAGNOSIS OF RETINAL DISEASES USING DEEP LEARNING

STUDENT

FADWA RAMADAN ALI HASSAN (2024334)

SUPERVISOR

DR. SHARYAR WANI

JUNE 2024

SEMESTER 2 2023/2024

FINAL YEAR PROGRESS REPORT

PROJECT ID

1205 D

PROJECT TITLE

OCT-BASED DIAGNOSIS OF RETINAL DISEASES USING DEEP LEARNING

PROJECT CATEGORY

DEVELOPMENT

by

FADWA RAMADAN ALI HASSAN (2024334)

SUPERVISED BY

DR. SHARYAR WANI

In partial fulfillment of the requirement for the

Bachelor of Computer Science

Kuliyyah of Information and Communication Technology

International Islamic University Malaysia

ACKNOWLEDGEMENTS

All gratitude and praise to Allah SWT, the Most Beneficent, and Most Merciful, for granting me the knowledge, strength, and perseverance to complete this project. Allah's blessings and guidance have been the foundation of my journey.

To my supervisor, I would like to extend my sincere gratitude .Dr. Sharyar Wani for all his help, encouragement, and support during this project. His knowledge and insights were crucial to my development. I deeply value his willingness, openness to my inquiries, and constant encouragement, all of which enabled me to overcome obstacles.

I am also incredibly grateful to my family, whose unwavering support and faith in me have been a main source of strength. Their words of encouragement and support throughout my studies have enabled me to pursue my ambitions and reach this milestone. Their dedication and sacrifices inspire me to strive for excellence in all aspects of my life.

Furthermore, I would like to extend my gratitude to my colleagues, and friends who provided me with encouragement and support throughout this journey. Their words of kindness and belief in me helped me stay motivated and focused on my goals.

May Allah SWT grant them all the best in this world and the hereafter.

ABSTRACT

This project explores the application of deep learning in retinal disease diagnosis using Optical Coherence Tomography (OCT) images. The goal is to develop a reliable and automated system for classifying four common retinal conditions: choroidal neovascularization (CNV), diabetic macular edema (DME), drusen, and normal. Two convolutional neural network (CNN) models were investigated: a basic CNN architecture and a pre-trained ResNet50 model with various train-test ratio. The training dataset consisted of a curated collection of labelled OCT images. Both models were evaluated using accuracy, precision, and recall metrics on a held-out test set. The ResNet50 model achieved superior accuracy (96.69%) compared to the basic CNN model (91.16%). Both models demonstrated high precision and recall, For the CNN model, it achieved a precision of 91.70% and a recall of 91.16%, indicating a low rate of false positives (8.30%). Similarly, the ResNet-50 model achieved an even higher precision (96.85%) and recall (96.69%). This work demonstrates the potential of deep learning for accurate and automated diagnosis of retinal diseases, potentially enhancing diagnostic capabilities and enabling earlier intervention for patient care. Furthermore, to facilitate user interaction and real-time image analysis, a mobile application has been developed for deployment. Future work will focus on exploring more conditions and further deep learning architectures, deploying, and integrating the developed model into a user-friendly platform for healthcare application.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

OCT	Optical Coherence Tomography
CNNs	Convolutional Neural Networks
DME	Diabetic Macular Edema
CNV	Choroidal Neovascularization
RPE	Retinal Pigment Epithelium
AMD	Age-Related Macular Degeneration
FPOA	Flower Pollination Optimization Algorithm
ANU-Net	Attention-based Nested U-Net
FD-CNN	Fully Dense Fusion Convolutional Neural Network
D-SVM	Deep Support Vector Machine
D-KNN	Deep K-Nearest Neighbor
LIME	Local Interpretable Model-Agnostic Explanations
UCSD	University of California, San Diego
Duke OCT	Duke Optical Coherence Tomography
PolyLoss	Polynomial Loss
Grad-CAM	Gradient-weighted Class Activation Mapping
SRT	Selective Retina Therapy
HOG	Histogram of Oriented Gradients
k-NN	k-Nearest Neighbor
DOR	Diagnostic Odds Ratio

CHAPTER ONE

INTRODUCTION

1.1 Project Overview

Blindness affects millions worldwide, primarily due to delayed diagnoses of retinal conditions. In ophthalmology, optical coherence tomography (OCT) is a vital tool, it offers high-resolution detailed images of the retina. These images are frequently used for diagnosing retinal conditions such as choroidal neovascularization (CNV), Drusen and diabetic macular edema (DME). Accurate and timely diagnosis of all of these conditions is crucial for proper treatment. However, interpreting OCT scans can be time-consuming for healthcare professionals and difficult for patients.

OCT images are generated by using light waves to produce cross-sectional images of the eye's retina, providing valuable features for diagnosis. for example, DME's OCT images include increased retinal thickness, particularly in the central part of the retina (Figure 1) [1], due to fluid accumulation. In contrast, OCT images show Abnormal blood vessel growth beneath the retina, originating from the choroid layer and are often diagnosed as CNV. On the other hand, Drusen's OCT images exhibit deposits that accumulate beneath the retinal pigment epithelium (RPE).

Despite the benefits of OCT, the manual interpretation of these images can be prone to human error and requires significant expertise and time. The complexity of retinal diseases combined with the shortage of ophthalmologists underlines the demand for accurate, efficient and automatic analyzing methods. This is a tricky task, and the existing approaches are not very effective in developing efficient systems that could analyze medical images with high accuracy, but deep learning, a subset of artificial intelligence, could serve as a way out of this problem.

The primary objective of this project is to employ the power of deep learning to develop a rapid and accurate diagnostic tool for the early detection and management of retinal conditions. This project tackles this challenge by developing a deep learning application for the precise diagnosis and classification of critical retinal diseases, including CNV, DME, Drusen, and normal condition, using retinal OCT images. By employing a dataset containing more than 84000 images, which encompass various retinal pathologies, the deep learning architecture is trained on diverse retinal conditions, incorporating transfer learning for improved performance.

The potential of having such a tool is important. Early and accurate detection of retinal diseases can prevent vision loss and improve patient outcomes. The use of automated diagnostic tools can benefit ophthalmologists by offering a second opinion and decreasing the list of tasks. Additionally, such tools can facilitate screening programs in underserved areas where access to specialized healthcare professionals is limited.

In summary, this project aims to harness deep learning techniques to enhance the diagnostic process in ophthalmology. By automating the analysis of OCT images, the project seeks to provide a reliable, efficient, and accessible diagnostic tool that can contribute to the early detection and management of retinal diseases, ultimately improving patient care and outcomes.

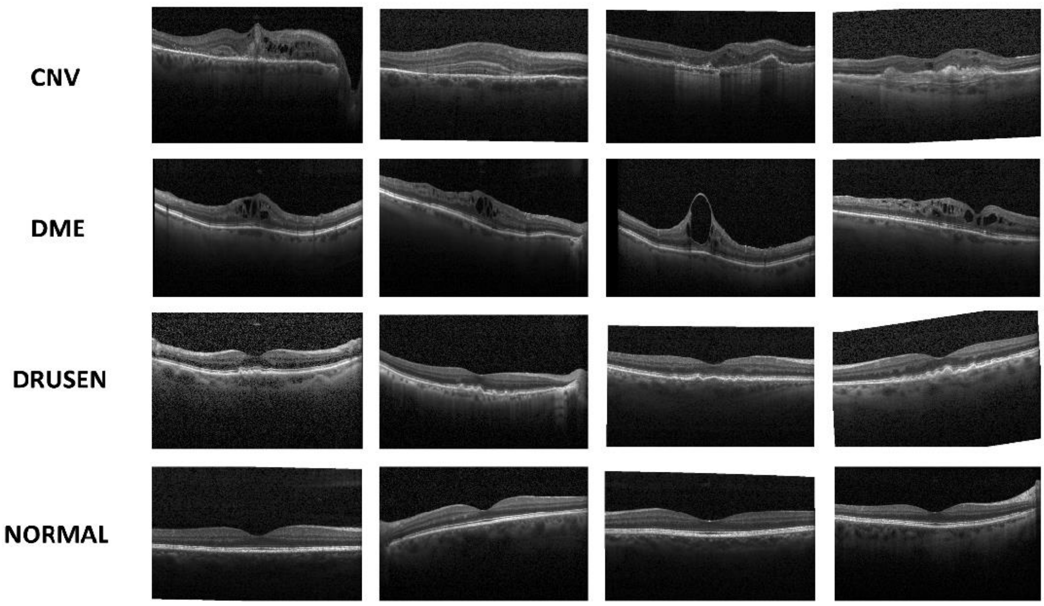

Figure 1: OCT image samples

1.2 Problem Statement

Retinal pathologies pose a significant challenge in ophthalmology, demanding accurate and timely diagnosis for effective management. OCT is a crucial imaging technique, generating high-resolution scans for detailed retinal assessment. However, the interpretation of OCT images is a complex and time-consuming task for healthcare professionals. This challenge is further exacerbated by the need for swift and precise diagnoses in cases of CNV, DME, and Drusen. The current manual grading process, involving multiple tiers of graders, introduces delays and potential human errors. Additionally, the absence of an automated system contributes to the overwhelming workload of healthcare providers, hindering their ability to provide timely interventions and tailored treatment plans. This project aims to address this problem by developing a convolutional neural network (CNN) based automated classification system for retinal conditions, enhancing efficiency, and accuracy, and ultimately improving patient care in the field of ophthalmology.

1.3 Project Objectives

This project focuses on implementing a simplified form of IoT-like communication and machine learning as a solution to the problem statements as stated previously, the main objectives are as follows:

1. To implement deep learning algorithms that accurately identify and detect the condition of the retinal.
2. To develop a real-time classification system using a Convolutional Neural Network (CNN) model.
3. To design a mobile application capable of analyzing and interpreting retinal OCT images.

1.4 Significance of Project

This project is crucial because of its ability to make changes within the ophthalmology and healthcare fields. The project offers solutions related to eye issues by developing a mobile app with a model that classifies retina OCT scans into CNV, Drusen, DME, and Normal OCT retina. An automatic classification system driven by CNN decreases diagnostic time and possibilities for mistakes. and that helps in improving the effectiveness of service delivery by providing faster and more accurate diagnoses. Moreover, this project illustrates how deep learning can be useful in terms of public health. and it's used for diagnostic purposes within the globe of medical technologies. Finally, the importance that lies on this project depends on the potential to transform the way that physicians diagnose the various cases of retinal diseases, leading medical practitioners into modern times.

1.5 Project Schedule

Process	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
Finalizing Title														
Literature review														
Chapter 1														
Data Collection														
Methodology														
Designing & implementing														
Testing														
Showcase														
Report Completion														

Table1: Gantt Chart for FYP 1

Process	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
Consultation														
Dataset Improvement														
Model Enhancing														
Application Development														
System Integration														
System Testing														
Report Writing														
Showcase														
Report Completion														

Table2: Gantt Chart for FYP 2

CHAPTER TWO

REVIEW OF PREVIOUS WORK

2.1 Literature Review

The evolution of deep learning in ophthalmology, particularly for Optical Coherence Tomography (OCT) image analysis, has opened new avenues in diagnosing retinal diseases. Reviewing previous works provides a guide through the way other scholars conducted their research. This literature review aims to investigate a different and unique method for employing deep learning for categorizing OCT images. Thus, in this chapter, 10 different articles have been reviewed, those studies discuss the topic of retinal diagnosis using Deep Learning and OCT images.

A compelling Deep Learning solution for detecting and analyzing AMD in OCT images is presented by Gilakara Muni Nagamani & Sudhakar (2023). Recognizing the crucial role of early AMD diagnosis in preserving vision, their method automates both disease classification and lesion segmentation. As a first step, the researchers applied preprocessing on the OCT images to remove the natural curvature of the retina and crop out the diagnostic regions of interest. Subsequently, they used a deep learning model called SqueezeNet to categorise retina OCT images into the following groups: normal, early AMD (drusen), and late AMD (choroidal neovascularization). To further improve performance, settings of the CNN model are adjusted using a flower pollination optimization algorithm (FPOA). Researchers also propose ANU-Net, a powerful model for precisely identifying AMD lesions within retinal layers. Its performance on the UCSD dataset is impressive, achieving near-perfect accuracy (98.7%), sensitivity (99.7%), and specificity (99.8%) in distinguishing normal, early, and late AMD cases. Furthermore, through a combination of a CNN classification and nested U-Net segmentation, the presented model by the researchers can distinguish between normal and diseased retinal scans, while also locating abnormalities. The high precision highlights this system's potential to aid preliminary diagnosis and early intervention by clinicians in ophthalmic practice. (Gilakara Muni Nagamani & Sudhakar, 2023)

While Gilakara & Sudhakar focused on a novel CNN and U-Net combination for AMD classification and segmentation, Kayadibi & Güraksın (2023) explored a different path by employing a hybrid Deep Learning classifier for retinopathy diagnosis. However, noise in OCT images can cause an issue. To tackle it and to improve image quality, the authors implemented a double preprocessing which comprises a speckle reduction filter. The researchers' approach tackles feature extraction from OCT images through a fully dense fusion convolutional neural network (FD-CNN). This technique is specially trained to efficiently identify features crucial for disease classification. They further refined diagnosis by implementing a hybrid strategy, employing Deep Support Vector Machines (D-SVM) and Deep K-Nearest Neighbors (D-KNN) with their established decision boundaries to handle diverse image qualities and disease presentations. This combination potentially builds a more robust and versatile model for real-world use. Additionally, they delve deeper into the classification process by generating heat maps with Local Interpretable Model-Agnostic Explanations (LIME), pinpointing the vital regions of the retinal layer used for decision-making. The model applied has a remarkable performance of the proposed D-SVM model on two public datasets, namely Duke and UCSD and OCT. With a 99.6% accuracy rate on the UCSD dataset and a 97.5% accuracy rate on the Duke dataset, the D-SVM model demonstrated how well the suggested approach can classify retinal diseases. (Kayadibi & Güraksın, 2023)

Further exploring avenues beyond CNNs and transfer learning, Li et al (2023) introduce a promising model utilizing Swin Transformers. According to their researchers, "Swin Transformer V2 exhibits excellent performance in many computer vision tasks" (p.2), particularly for accurate OCT image classification, as misdiagnosis can have serious consequences. Using Swin Transformer V2, the authors suggest a multi-foveal disease classification model that includes network training, data enhancement, and preprocessing operations. To improve model interpretability during testing, specific network parameter tuning, the use of PolyLoss as a loss function, and visualisation techniques like Grad-CAM heatmap and confusion matrix were implemented. The researchers additionally discussed comparing various loss functions and continuously optimising network parameters to find the best network model for classifying multiple fundus diseases. The presented findings show how accurately the model classified OCT images. As the model's functionality is improved by the PolyLoss function, retinal OCT image classification performance improves. (Li et al., 2023)

Demonstrating the continued potential of hybrid approaches, Saifur Rahman Shatil & Mir (2023) propose a novel method combining CNNs with transfer learning. That assists in improving accuracy. Focusing on four key retinal conditions (NORMAL, DME, DRUSEN, and CNV), the authors used "Labeled (OCT) and Chest X-Ray Images for Classification" dataset (83,484 training images, 968 testing images). The researcher's work represents the first use of pre-trained models such as DenseNet169 and ResNet152V2 for the classification of retinal diseases using transfer learning. Additionally, they developed a new CNN model achieving a remarkable accuracy of 98.34%. Through preprocessing techniques and meticulous training workflows, their approach outperforms existing models, demonstrating the potential of transfer learning for enhancing classification accuracy in OCT image analysis. (Saifur Rahman Shatil & Mir, 2023)

Highlighting the importance of feature selection in optimizing classification accuracy Khan et al (2023) present a unique approach that leverages ant colony optimization using hybrid and transfer learning approaches, the researchers take further strides by introducing a novel method combining pre-trained models, ant colony optimization for feature selection, and multiple classifiers to achieve impressive accuracy in OCT image classification. This achieved superior accuracy. By looking at more types of eye issues, their work included diseases like branch retinal vein occlusion and AMD. The researchers use three trained models which are (DenseNet-201, InceptionV3 and ResNet-50) to fine-tune them for the best feature extraction. This is done by using transfer learning. ACO is very important in choosing the most different features from what was taken out. This will improve how well we can classify or sort things. Significantly, adding ACO improves the total accuracy to 99.1% surpassing previous approaches%, compared with just 97.4% without it. This beat old ways in correctness and shows how well ACO works to pick the best features. Using two classifiers KNN and SVM, was found to give the best results with Cubic SVM and DenseNet-201. In addition, this shows how flexible this method can be. Furthermore, recorded times for the method show that it's possible to use it in real-life situations. The researchers did some important work. It opens the way for quicker and more correct ways to find retinal diseases using OCT pictures by computer alone. (Khan et al., 2023)

Furthering the quest for interpretable and accurate retinal disease categorization using OCT images, He et al (2023) introduce a novel Swin-Poly Transformer network. This approach attempts to overcome the drawbacks of manual image interpretation to improve the accuracy of diagnosis. Besides this, the researchers apply PolyLoss for lossy data optimization and that of Score-CAM for improving image interpretability. Those optimization methods enhance a comprehensive appreciation of how the model makes decisions. and they permit an in-depth examination of the rationale behind the decision made by the model. Thus, an important factor in their study is due to the nature of data collected for OCT 2017 and OCT C8 from their OCT images. The used methods by researchers have achieved remarkable accuracy (99.80%) and an Area Under Curve (AUC) of 99.99%, outperforming both convolutional neural network and Vision Transformer (ViT) approaches. Their study has established an important standard for combining deep learning models with medical imaging analysis, setting the path for more innovations in this critical. (He et al., 2023)

Zbinden et al (2016) propose a novel approach to overcoming a key challenge in Selective Retina Therapy (SRT), Properly measuring treatment results in real time. Unlike regular methods that hurt nearby tissue, SRT gives a more focused way. But it's hard to tell how well SRT works because we can't see images in real time. The researchers came up with a smart answer. It uses OCT to see changes in the retina during treatment quickly. They combine this new way of taking pictures with a teaching method using feature extraction and Random Forest classification to effectively measure the results of SRT. Es. Their way is especially accurate - 95% for pig eyes outside the body and human patients. This shows it could be used in real medical situations. This ability, which works well on human patients with the ex vivo trained algorithm, is very important for creating strong medical tools. (Zbinden et al., 2016)

Venkatraman & Sumathi (2019) point out a system studies 114 OCT pictures from different tools. By using steps like removing noise and separating areas, they got important features out of it. Most notably, they use Histogram of Oriented Gradients (HOG) features and a k-NN classifier. This method helps them to accurately tell normal pictures of the eye from those showing diseases related to fluid buildup. This way works very well. It gets 89.29% right and a special number called DOR at 2.0, both being better than old systems that only work in some areas and show it is tough to beat. The researcher's work is a big step forward in making OCT images for retinal problems faster and easier to analyze. Their system gives a hopeful way to give true and detailed tests. This can help patients have better results earlier on, leading to early actions for health. (Venkatraman & Sumathi, 2019)

In another study by Hasan et al (2021), they investigate the effectiveness of using deep learning to separate retinal illnesses in OCT images. In addition, they show how useful OCT is in eye care for finding and treating diseases. Furthermore, they mention works done by Tal and Nugroho, who used models to divide Retina images. These comparisons also show that deep learning is better at pulling out features from OCT images. Hasan and his team show how well deep learning can be used to classify retinal diseases correctly and automatically from OCT pictures. (Hasan et al. , 2021)

None Mohammad Talebzadeh et al., 2024 propose a novel deep triplet network incorporating a conditional loss function tailored for limited data samples. Inspired by VGG16, the foundational architecture processes triplets of anchor, positive and negative samples. The triplet network facilitates embedding samples of the same class in proximity while maximizing the distance between different classes in the embedding space. Experiments utilize an open-access OCT dataset spanning retinal abnormalities categorized into CNV, DME, Drusen, and Normal classes. Simulation results on the public OCT dataset demonstrate the proposed model's superior performance over state-of-the-art techniques like DenseNet , InceptionV3 , ResNet152 , and ResNet50 in terms of accuracy for retinal disease classification, as shown in Table 3. The ablation study in Table 3 validates the efficacy of the conditional triplet loss with penalty and reward terms, achieving 89.4% recall, outperforming the basic triplet loss (78.62%) and Siamese network (72.28%). (None Mohammad Talebzadeh et al., 2024)

Method	Model	Accuracy
Islam et al. [15]	DenseNet	92.64%
Kermany et al. [5]	InceptionV3	90.37%
Kim et al. [6]	Resnet152	91.86%
Li et al. [7]	ResNet50	91.48%
Our proposed	VGG16	92.81 %

Table 3: Comparison of model's accuracy with state-of-the-art techniques for Retinal diseases classification

Leandro et al., 2023 discussed utilizing deep learning models, specifically VGG-16, ResNet-50, and Inception-v3, to analyze OCTscans for detecting retinal abnormalities. These models demonstrated high accuracy levels ranging from 93% to 99% in classifying normal retinas, high myopia, and other retinal diseases. The use of Grad-CAM visualization techniques provided insights into the decision-making process of convolutional neural networks, enhancing interpretability and supporting clinical decision-making in ophthalmology. (Leandro et al., 2023)

Elkholy & Marzouk, 2024 utilized image processing techniques, transfer learning, and fine tuning to improve the accuracy of the models, achieving high classification rates for diseases. The proposed methodologies showcase significant advancements in leveraging AI for medical imaging analysis in ophthalmology. The proposed model achieved an accuracy of about 94% initially, which increased to approximately 97% after fine-tuning. It showed high efficiency in predicting four types of eye diseases (Elkholy & Marzouk, 2024)

Schlegl et al., 2018 presents a fully automated method to detect and quantify intraretinal cystoid fluid (IRC) and subretinal fluid (SRF) in macular OCT volume scans of various exudative macular diseases. The methodology involved using OCT scans from 1200 eyes with AMD, DME, and RVO, imaged with Spectralis and Cirrus devices. Results showed high accuracy in detecting IRF and SRF in all diseases, with reliable results even in DME where SRF occurred less frequently. Automated quantification of fluid was possible in all OCT scans, demonstrating the robustness and sensitivity of the method. (Schlegl et al., 2017)

Goutam et al., 2022 explored the application of deep learning strategies in the diagnosis of retinal diseases, focusing on glaucoma, age-related macular degeneration, diabetic retinopathy. The researchers covered the use of different deep learning models, including CNN and RNN, for analyzing fundus images to detect diseases like cataract, AMD, ROP, and others. The paper emphasizes the significance of dataset quality, image pre-processing techniques, and model architecture in achieving accurate disease diagnosis. Future research directions include weakly supervised learning, fundus image synthesis, lightweight network design, improving generalization, implementing federated learning, and smartphone-based diagnosis. The authors highlighted the potential for interpretable AI models and simultaneous diagnosis of multiple retinal diseases. Various advancements in deep learning techniques, such as transfer learning, attention mechanisms, and multimodal approaches, have shown promising results in improving the segmentation and diagnosis of retinal diseases. (Goutam et al., 2022)

2.2 Table of Summarized Studies on Retinal Disease Classification using OCT Images

Study Title	Authors	Year	Methods	Dataset	Main Findings	Key Contribution
Automated classification of age-related macular degeneration from optical coherence tomography images using deep learning approach.	Gilakara & Sudhakar	2023	SqueezeNet CNN + ANU-Net segmentation	UCSD public dataset	98.7% accuracy, 99.7% sensitivity, and 99.8% specificity in 3-way AMD classification (normal, early, late)	Novel combination of CNN and U-Net for classification and segmentation
An Explainable Fully Dense Fusion Neural Network with Deep Support Vector Machine for Retinal Disease Determination.	Kayadibi & Güraksın	2023	FD-CNN feature extraction + D-SVM & D-KNN classification	UCSD & Duke public datasets	99.6% accuracy on UCSD, 97.5% on Duke	Hybrid deep learning with interpretable heatmaps
Multi-Fundus Diseases Classification Using Retinal Optical Coherence Tomography Images with Swin	Li et al	2023	Swin Transformer V2	OCT2017 & OCT-C8 datasets	Achieved good performance on both datasets, highlighting model effectiveness	Utilizes Swin Transformer V2 for multi-disease classification

Transformer V2						
Retinal OCT Image Classification Based on CNN and Transfer Learning.	Saifur Rahman Shatil & Mir	2023	DenseNet169 & ResNet152V2 + new CNN	Labeled OCT and Chest X-Ray dataset	98.34% accuracy, outperforming existing models	First application of DenseNet169 and ResNet152V2 for retinal disease classification
Optical Coherence Tomography Image Classification Using Hybrid Deep Learning and Ant Colony Optimization.	Khan, Pin, Aziz, Jung, & Nam	2023	Pre-trained models + ACO feature selection + KNN & SVM classification	Soonchunhyang University Bucheon Hospital dataset	99.1% accuracy with ACO, 97.4% without, outperforming existing models	Hybrid deep learning with ACO for feature selection and multi-classifier approach
An interpretable transformer network for the retinal disease classification using optical coherence tomography.	Yan et al.	2023	Swin-Poly Transformer framework + PolyLoss + Score-CAM	OCT 2017 & OCT C8 datasets	99.80% accuracy and AUC of 99.99%, outperforming CNN and ViT	Interpretable transformer network with confidence score maps
Automatic assessment of time-resolved OCT images for selective retina therapy.	Song et al.	2016	Time-resolved OCT + feature extraction + classification (Random Forest)	Ex vivo porcine eyes & in vivo human patients	95% accuracy in both ex vivo and in vivo conditions	Non-infiltrating time-resolved OCT for real-time SRT outcome assessment
Classification of Retinal Disorders Based on Fluid Patterns in OCT Images.	Venkatraman & Sumathi	2017	Preprocessing + HOG feature extraction + k-NN classification	114 OCT images from TOPCONN and ZEISS equipment	89.29% accuracy and DOR of 2.0	Automated system for classifying retinal disorders based on fluid patterns
Enhanced Deep Learning Model for Classification of Retinal Optical Coherence Tomography Images.	Mahendran & Murugan	2023	Modified ResNet 50 + forest algorithm	Dataset with various retinal disorders	High accuracy, precision, and recall across various categories	Enhanced deep learning model with improved classification performance
Classification Performance Analysis of Retinal OCT Image using Handcrafted and Deep Learning	Islam et al.	2021	Handcrafted and deep learning features + SVM	Dataset with NORMAL, DRUSEN, DME, and CNV classes	98.34% accuracy with proposed CNN model	Comparison of handcrafted and deep learning features for SVM classification

Feature with Support Vector Machine.						
Deep learning-based retinal abnormality detection from OCT images with limited data	None Mohammad Talebzadeh et al.	2024	Deep triplet network with conditional triplet loss function, Incorporated penalty term for underperforming triplets, Reward term for optimal triplet cases, Foundational architecture inspired by VGG16	Open-access OCT dataset categorized into CNV, DME, Drusen, and Normal classes	Conditional triplet loss with penalty/reward achieved 89.4% recall, better than basic triplet loss (78.62%) and Siamese network (72.28%)	incorporate a conditional loss function tailored for limited data samples of retinal OCT images to enhance classification accuracy and mitigate overfitting issues.
OCT-based deep learning models for the identification of retinal degeneration signs	Leandro et al.	2023	Utilized VGG-16 convolutional neural network architecture for image analysis and classification.	Labeled images were preprocessed to create 9 predictive binary models to identify retinal degeneration signs.	Successfully identified retinal degeneration signs using OCT-based deep learning models.	Demonstrated the efficacy of deep learning models in identifying retinal degeneration signs through OCT imaging.
Deep learning-based classification of eye diseases	Elkholy & Marzouk	2024	Multi-modal approaches for cataract case identification	Publicly available dataset from Kaggle	Achieved an accuracy of approximately 97% after fine-tuning	Efficient prediction of four types of eye diseases
Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning	Schlegl et al.,	2017	utilized a fully automated AI method to detect and quantify intraretinal cystoid fluid (IRC) and subretinal fluid (SRF) in macular OCT volume scans of various exudative macular diseases	1200 OCT volume scans from eyes affected by neovascular age-related macular degeneration (AMD), (DME), and macular edema secondary to retinal vein occlusion (RVO).	high accuracy in detecting and quantifying IRC and SRF in all diseases, with reliable results even in DME where SRF occurred less frequently	ighly robust and sensitive automated method for detecting and quantifying macular fluid in routine OCT images of various exudative macular diseases.
A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Image	Goutam et al.	2022	deep learning strategies for retinal disease diagnosis using fundus images, covering datasets, evaluation metrics, image pre-processing techniques, and	Various publicly available fundus image datasets are consolidated, including IDRiD, Messidor, DRIVE, Kaggle, EyePACS,	highlights the importance of dataset quality, image pre-processing techniques, and model architecture in achieving accurate disease diagnosis.	offers a comprehensive overview of deep learning strategies in retinal disease diagnosis, outlines future research directions, and provides insights into

			commonly used models.			the application of deep learning models for diagnosing various retinal diseases using fundus images

Table 4: Summarized Previous Studies

CHAPTER THREE

METHODOLOGY

3.1 Design

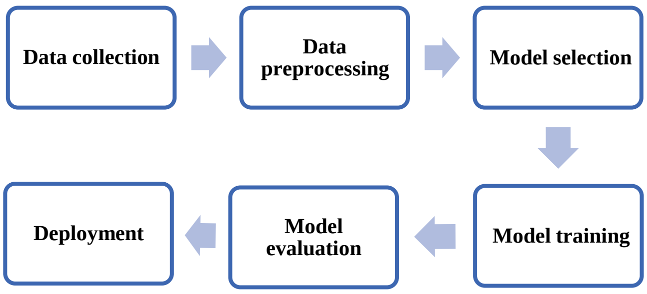

Figure 2: Methodology process

1. Data Collection

The dataset used in this project, "Retinal OCT Images (optical coherence tomography)" from Kaggle, contains 84,495 images categorized into NORMAL, CNV, DME, and DRUSEN. The images are split into two sets (Train, Test) for training, and testing purposes. Retrospectively, between July 2013 and March 2017, these images were collected from adult patients at several hospitals, providing a diverse range of OCT data for retinal disease analysis. Choosing this dataset was driven by its well-organized structure, balanced representation of target diseases, and high-quality image resolution.

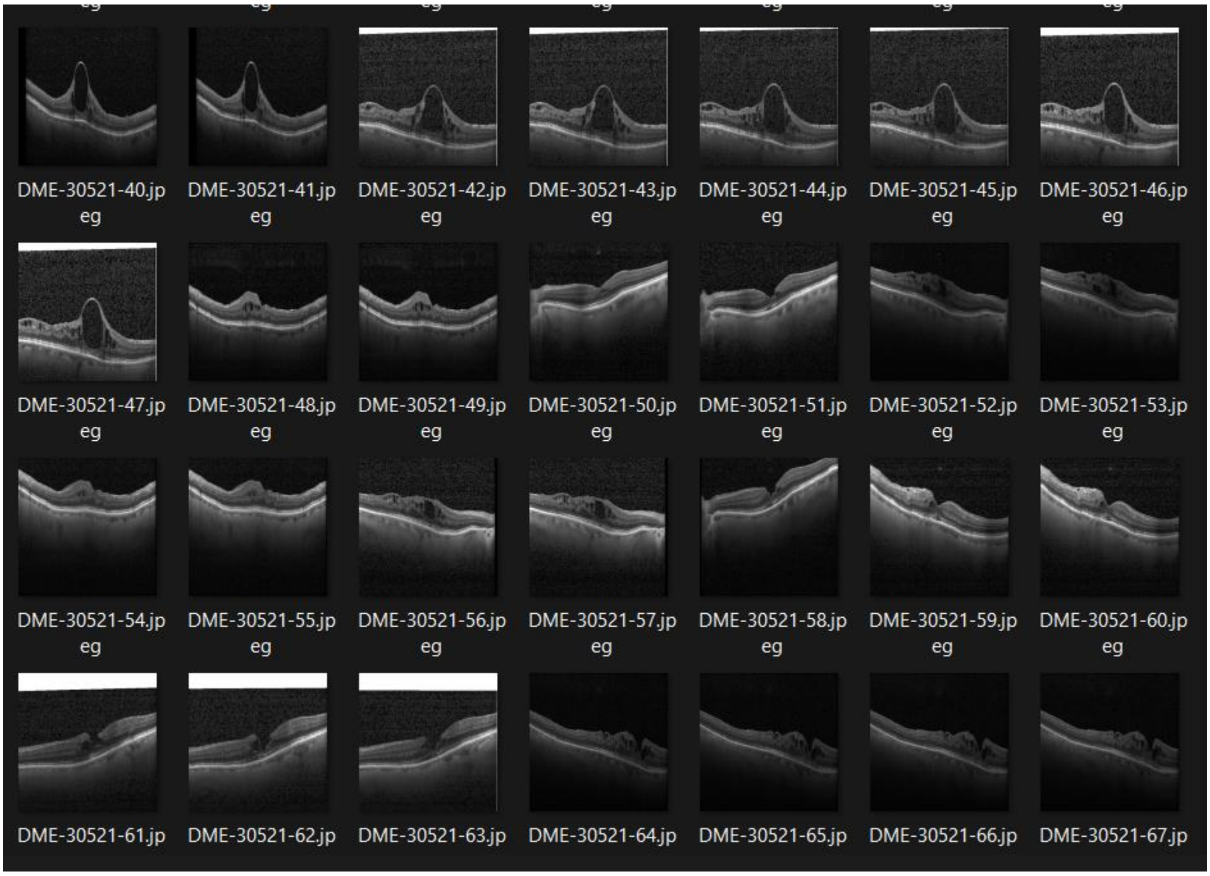

Figure 3: Sample of The Original Images

2. Data Preprocessing

To ensure effective model training and consistent format, the original OCT images in the "Retinal OCT Images" dataset were pre-processed with standardized resizing and normalization techniques. The following steps were performed:

1. Image Filtering: Image files were scanned for valid extensions (.jpg, .jpeg, .bmp, .png). Images with incompatible formats were removed to avoid loading errors. In addition, Potential issues like corrupted files or invalid paths were identified and addressed through exception handling mechanisms.

2. Data Normalization: All image pixel intensities were scaled to a common range (0-1) by dividing by the maximum intensity value (255). This helps improve model convergence and prevents issues with numeric range differences.

3. Batch Creation: To safeguard against bias and guarantee the model's ability to perform well on unseen data, the preprocessed data was strategically divided into

distinct training, test sets. This approach involved creating approximately 2618 batches, each containing 32 images.

3. **Model Selection**

This project employs two model architectures for automated diagnosis of four retinal diseases (CNV, DME, DRUSEN, and NORMAL): Convolutional Neural Network (CNN) and ResNet-50. Both are powerful deep learning methods commonly used in image recognition tasks.

The CNN model uses a multi-layered filter structure to extract important features from the OCT images. It includes Convolutional layers which are used for pattern recognition and extract the feature map, and pooling layers which are used to reduce dimensionality, dropout layers for regularization, and dense layers for final classification. The choice of this architecture is because of its suitability for multi-class classification tasks. The model's design, including the number of layers (13 layers), filter sizes, and dropout rates, is aimed at optimizing the available data. The model complexity is intentionally kept manageable for efficient training and inference on standard hardware configurations.

Figure 4: CNN Layers

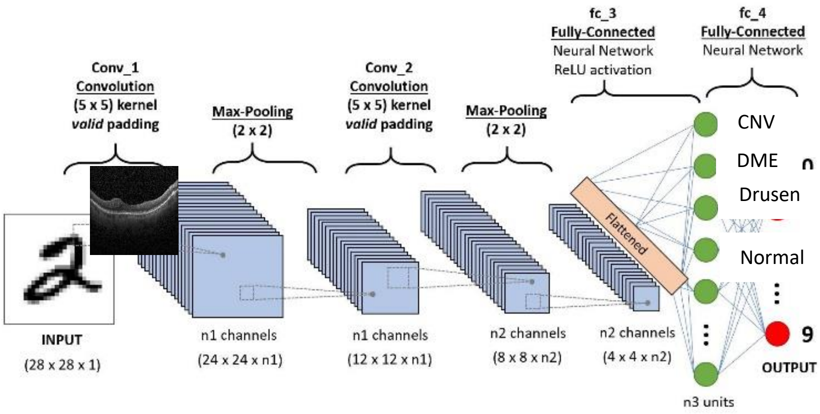

Figure 5: CNN Architecture

On the other hand, ResNet50, which is a type of CNN model with 50 layers has a special architecture that allows for deeper networks without sacrificing performance. By introducing skip connections that permit information to bypass certain layers, making it easier for the network to learn. ResNet50 allows for a deeper network (175 layers 50 of them are convolutional layers) compared to the standard CNN. This can potentially capture more complex patterns and relationships in the OCT images, leading to improved classification accuracy for retinal diseases. ResNet50 also has the benefit of skip connections, which reduces the effect of the vanishing gradient issue, and this may occur in deep networks and impair efficient learning. ResNet50 has a proven track record of high accuracy in various image recognition tasks, including medical image analysis.

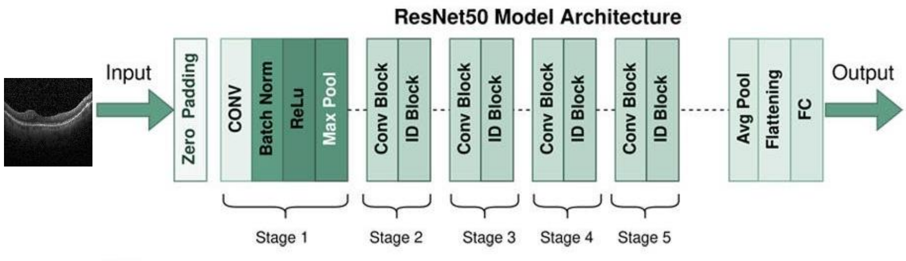

Figure 6: ResNet50 Architecture

4. Model Training

4.1 CNN Model Architecture:

- Convolutional Layers: Employs three convolutional layers with increasing filter counts (32, 64, 128) and kernel sizes of 3x3 to extract a hierarchy of features from the OCT images. A max pooling layer with a 2x2 window comes after each convolutional layer to lessen dimensionality and capture the most noticeable features. The increasing filter counts in the convolutional layers allow the model to learn increasingly complex features at each stage.
- Dropout Layers: Strategically incorporates dropout layers after each max pooling layer (with 0.3 dropout rate) and after the flatten layer (with 0.5 dropout rate) to prevent

overfitting and enhance model generalization. The flatten layer's goal is to take the multidimensional feature maps and turn them into a one-dimensional vector so that the dense layers can use them.

- Dense Layers: After squeezing the features from the convolutional layers, the model feeds them into a dense layer with 512 neurons. This layer, using a 'relu' activation, helps the model learn non-linear patterns between features. To make sure the model doesn't memorize the training data too closely, there's another dropout layer after this.
- Output Layer: comprising four neurons and powered by the 'SoftMax' activation function, serves as the model's final decision-maker. It assigns probability scores to each of the four disease categories, enabling the model to confidently identify the most probable diagnosis.

4.2 ResNet50 Model Architecture:

- Pretrained Base: Leverages a pre-trained ResNet50 model, excluding the top layers, as a strong foundation for feature extraction. Initializes the model with weights trained on the ImageNet data set, which is a massive collection of images, allowing it to benefit from prior knowledge of visual patterns. Using a pre-trained model have meany benefits, such as faster convergence and potentially better performance due to the transfer of knowledge from ImageNet.
- Global Average strikes a balance between capturing essential information and reducing computational complexity. This layer efficiently condenses each feature map into a single, representative value, streamlining subsequent classification tasks while potentially enhancing the model's ability to generalize to new data.
- Dense Layers: Adds a dense layer containing 512 neurons and 'relu' activation to enable further learning of complex relationships between features.in addition, it incorporates a dropout layer with a 0.5 rate to combat overfitting.

- Output Layer: Employs a final output layer containing four neurons and 'SoftMax' activation, like the CNN model, to produce class probabilities for the four disease categories.

5. **Model Evaluation**

The performance of models was evaluated on the test set using four key metrics:

- Accuracy: This measures the overall proportion of images where the model correctly predicted the disease category (CNV, Drusen, DME, Normal).
- Precision: This focuses on the positive predictive value. It represents the proportion of images predicted to have a specific disease that actually belong to that class (i.e., how many true positives are among predicted positives).
- Recall: This emphasizes the ability of the model to identify all relevant cases. It represents the proportion of images with a specific disease that the model correctly classified (i.e., how many true positives out of all actual positives).
- Loss: This quantifies the difference between the model's predictions and the actual labels. Lower loss values indicate better model performance.

6. **Deployment**

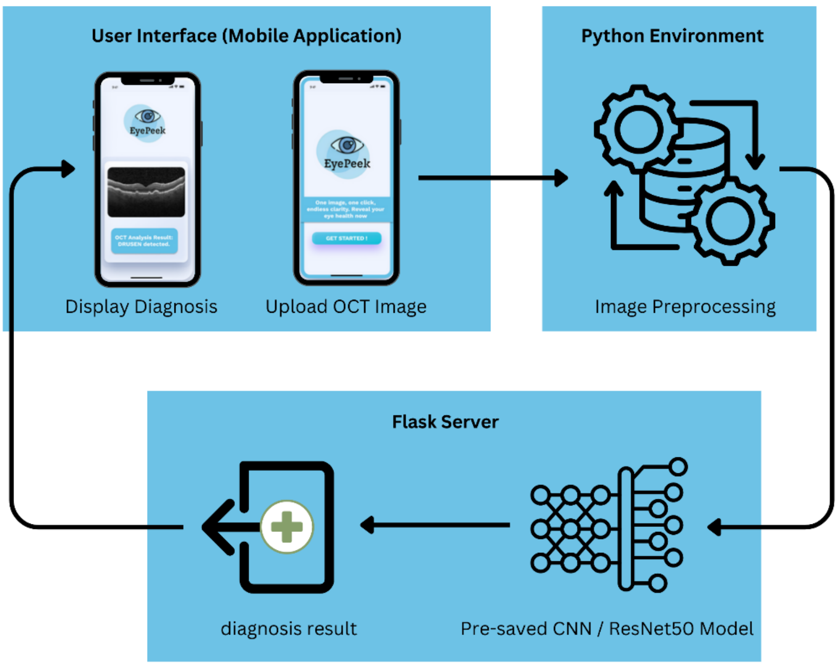

Figure 7: System Development

Deployment involves creating a system that contains three environments, the user interface, Python, and the Flask server environment. Each environment has its own functionalities and helps in building and connecting the system.

- **User Interface Environment:** mobile application, that allows users to upload OCT images and receive diagnostic results, using Flutter and Dart.
- **Python Environment**: this environment creates the Flask server and preprocessing steps.
- **Flask server:** in this environment, the image is received, and the preprocessing is applied, then the pre saved models (CNN/ResNet50) predict the diagnosis, once the result is output, it will be sent back to the User interface environment.

Flask Implementation Details:
The Flask application utilizes several key functionalities:

- Image Uploading: It is an application that aims to enable the users to input OCT images via a graphical user interface.
- Base64 Encoding/Decoding: Most of the images that are uploaded are converted to Base64 for transmission between the browser and the server.
- Image Preprocessing: After receiving the image data, resizing and normalization are performed on it to match the input data format required by the model.
- Model Prediction: The pre-trained CNN8020 model is then employed to classify the uploaded image in terms of the disease class as either CNV, DME, Drusen or Normal.
- Result Response: The predicted class is then passed back to the user interface, and the diagnosis is shown on the application screen.

Benefits of Flask Deployment:

- Accessibility: Using the built Flask application, the model can be easily accessed and used by a user without necessarily having to understand some concepts in machine learning.
- Real-time Classification: From this, the users are able to get their diagnoses in real time, making the process faster and more efficient.
- Scalability: The developed Flask application can also be expanded and enhanced to

cater for more end users.

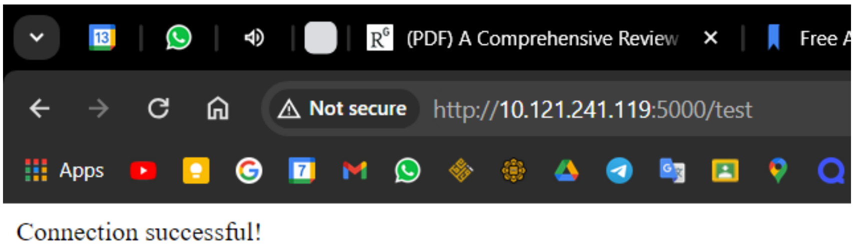

Figure 8: Successful working Flask server

3.2 Implementation

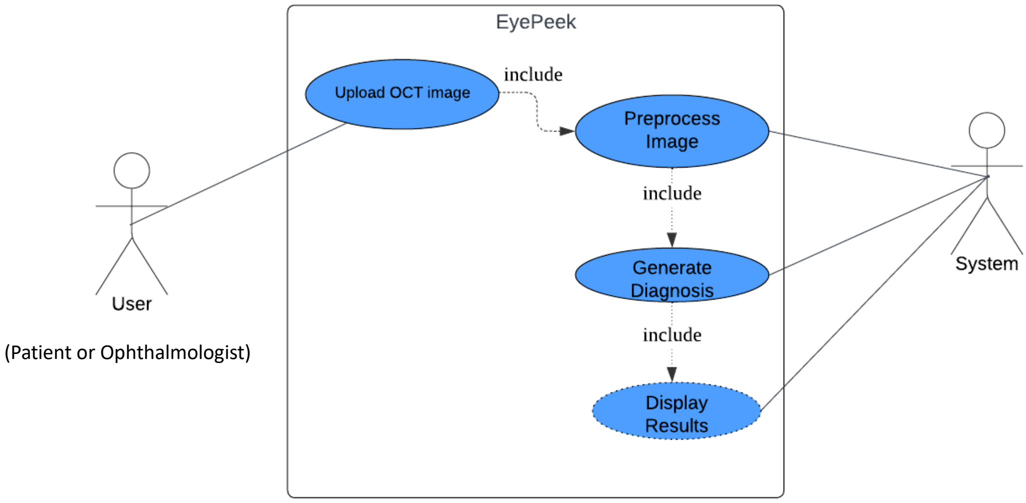

Figure 9: Use Case Diagram

Actors:

- User: An ophthalmologist or a patient who interacts with the mobile application to upload OCT images and receive diagnostic results.
- System: The OCT image diagnosis system, including the User Interface, Python Environment, and Flask Server.

Use Cases:

1. Upload Image:
   - Actor: User
   - Description: User launches the mobile app, selects an OCT image from their

device, and uploads it to the system.

2. Preprocess Image:
   - Actor: System (Python Environment)
   - Description: System receives the uploaded image, create necessary preprocessing steps (e.g., resizing, normalization), and prepares it for analysis.
3. Generate Diagnosis:
   - Actor: System (Flask Server)
   - Description: System passes the preprocessed image to the deep learning model, which analyzes it and predicts the most likely disease category (CNV, DME, DRUSEN, or NORMAL).
4. Display Results:
   - Actor: System (User Interface)
   - Description: System receives the predicted diagnosis from the model, formats it for user understanding, and displays it clearly within the mobile app, along with any relevant additional information or guidance.

Figure 10: Application interface

CHAPTER FOUR

ANALYSIS OF RESULTS AND OUTCOMES

4.1 System Evaluation

The application is built using the Flutter framework; it functions as the first step analysis of OCT retinal images. Users can interact with the application in the following ways:

- Image Selection: OCT retinal images can be chosen by the user directly on their device without the need for specialized software using file pickers.
- Image Display: after the image is selected users are able to change the uploaded image or proceed with the classification process.
- Automated Analysis: Once an image is chosen, the application relies on a pre-trained deep learning model to assess the image for possible eye conditions.
- Analysis Results: The selected image is well presented to the users within the application and the class which has been predicted as CNV, Drusen, DME or Normal, is prominently shown on the screen.
- Disease Definition: Based on the predicted class, the program gives brief information about the type of eye problem that may be related to it.
- "Learn More" Buttons: some information link called “Learn more” are shown below the results based on context. By clicking on a button, the educational resource related to the predicted class is opened through the web browser of the device, giving users an opportunity to explore the condition in detail.

4.2 Model Evaluation and Performance

Base Model Evaluation involves assessing the trained model's performance on unseen data to ensure its generalizability and reliability. By using early stopping on the test set to prevent overfitting. The performance of the CNN and ResNet50 models are summarized in the table below:

	CNN Model		ResNet50 Model	
Train-Test split	80:20	70:30	80:20	70:30
Accuracy	0.9116	0.9052	0.9668	0.9377
Loss	0.3268	0.3798	0.0955	0.1197
Precision	0.9170	0.9123	0.9685	0.9397
Recall	0.9116	0.9052	0.9669	0.9377

Table 5: The performance of the CNN and ResNet50 models

ResNet50 80:20 achieved the highest accuracy of 96.68%, while CNN 80:20 followed closely with 91.16% accuracy. The gap in accuracy between these two models is 5.52%. Training with a larger portion of the data (80% in this case) led to slightly better performance compared to a 70:30 train-test split. This is evident in both CNN and ResNet models. Overall, the results suggest that the ResNet50 architecture, especially when trained with a larger training set, is more effective in diagnosing retinal diseases from OCT images compared to the CNN models implemented.

The decision of selecting CNN architecture (CNN8020) and training it on 80% of the data set was a deliberate process aimed at reducing overfitting. Overfitting happens when a model gets overly tuned to the training data and as a result has poor performance on new data. Hence, to avoid overfitting of the model and enhance its ability to classify images not seen during training.

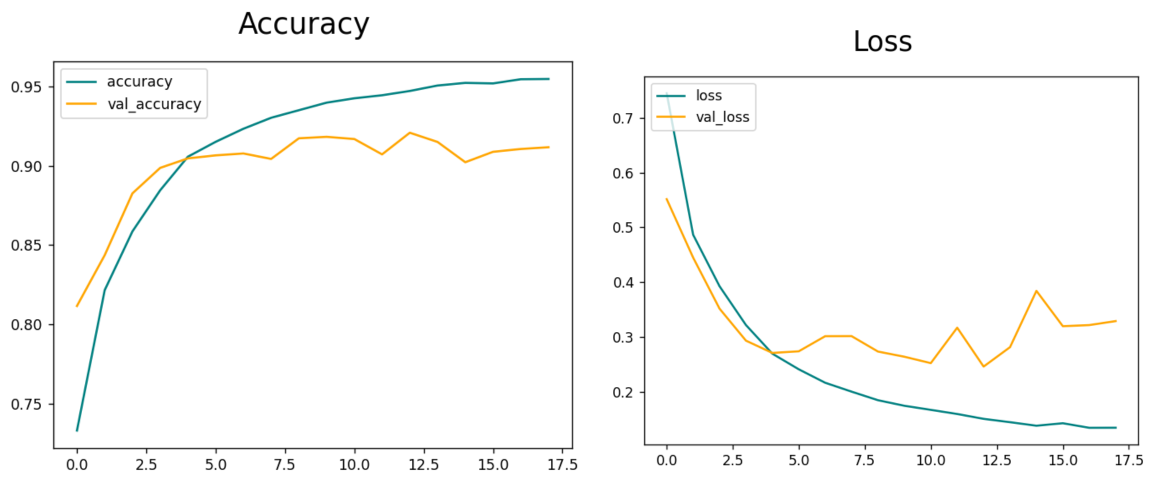

Figure 11: The Loss and Accuracy Plot of The CNN 80:20 Model

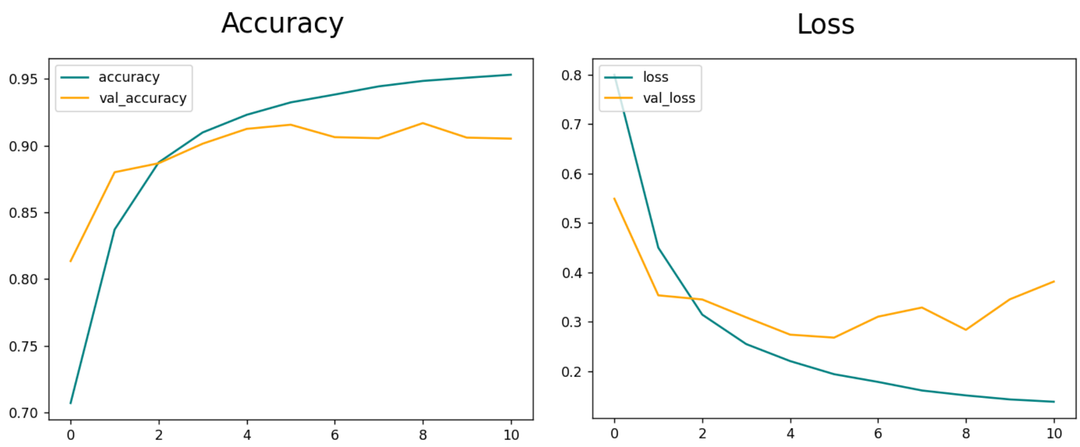

Figure 12: The Loss and Accuracy Plot of The CNN 70:30 Model

4.3 User Acceptance Testing (System Validation)

To evaluate the usability and functionality of the mobile application, a two-pronged approach was implemented:

- Ophthalmology Professionals: An online UAT session was conducted via Google Meet on June 8, 2024, with three participants. Two participants were ophthalmology students, and the third was a practicing ophthalmologist. These individuals were briefed on the project's objectives and background before testing the application. Positive Feedback on Design: The first participant appreciated the application's design, highlighting that "the design and logo give a professional medical feel, which is really good."
- Non-Medical Users: To assess the application's accessibility for users with average eye knowledge, ten participants participated in a separate UAT. The results of their functionality testing are presented in a table5.

	Condition				Accuracy
Person	CNV	DRUSEN	DME	NORMAL	
A	✓	×	✓	✓	75%
B	✓	✓	✓	✓	100%
C	✓	✓	✓	✓	100%
D	✓	✓	✓	✓	100%
E	✓	×	✓	✓	75%
F	✓	✓	✓	×	75%
G	✓	✓	✓	✓	100%
H	✓	✓	✓	✓	100%
I	✓	✓	✓	✓	100%
J	✓	✓	✓	✓	100%
Sum Accuracy	100%	80%	100%	90%	92.5%

Table 6: Summary of User Functionality Testing

CHAPTER FIVE

CONCLUSION AND FUTURE WORK

The exploration of deep learning for the automated classification of retinal diseases using OCT images has demonstrated promising potential. The comparative analysis of CNN and ResNet-50 architectures highlighted the superior performance of the ResNet-50 model, achieving an impressive accuracy of 96.68%. This underscores the efficacy of deep learning techniques in the field of ophthalmology, providing a robust tool for diagnosing conditions like CNV, DME, and DRUSEN with high precision and recall rates.

The methodology employed in this project encompassed a comprehensive workflow, from data collection and preprocessing to model training, evaluation, and deployment. The use of the "Retinal OCT Images" dataset from Kaggle ensured a diverse and high-quality set of images, facilitating the training of effective models. Preprocessing steps, including resizing and normalization, were critical in preparing the data for efficient and accurate model training.

The deployment of the trained models through a mobile application using Flutter, Python, and Flask showcased the practical application of this technology. This deployment not only provided real-time diagnostic capabilities but also enhanced accessibility for both medical professionals and patients. The user acceptance testing conducted with ophthalmology professionals and non-medical users further validated the application's usability and effectiveness, with an overall accuracy of 92.5% in user functionality testing.

Despite the promising results, there are limitations that need to be addressed. The dataset, while diverse, requires further expansion to include more varied demographics and conditions to ensure the generalizability of the models. Additionally, ongoing refinement of the models and the incorporation of explainability techniques are necessary to enhance their transparency and trustworthiness in clinical settings. Future research should also focus on clinical trials to validate the models' performance in real-world scenarios.

In Brief, this project demonstrates the potential of deep learning in revolutionizing retinal disease diagnosis. By providing fast, accurate, and accessible diagnostic tools, deep learning can significantly improve patient care in ophthalmology. This project serves as a foundational step towards the development of advanced deep learning-based diagnostic tools, paving the way for more efficient and accurate diagnosis and ultimately better patient outcomes in the field of retinal disease management.

REFERENCES

1. Ibrahim, M. R., Fathalla, K. M., & M. Youssef, S. (2020). HyCAD-OCT: A Hybrid Computer-Aided Diagnosis of Retinopathy by Optical Coherence Tomography Integrating Machine Learning and Feature Maps Localization. Applied Sciences, 10(14), 4716. https://doi.org/10.3390/app10144716.
2. Gilakara Muni Nagamani, & Sudhakar, T. (2023). Automated classification of age-related macular degeneration from optical coherence tomography images using deep learning approach. IAES International Journal of Artificial Intelligence, 12(4), 2011–2011. https://doi.org/10.11591/ijai.v12.i4.pp2011-2021.
3. Saifur Rahman Shatil, & Mir. (2023). Retinal OCT Image Classification Based on CNN and Transfer Learning. Lecture Notes in Networks and Systems, 433–444. https://doi.org/10.1007/978-3-031-27524-1_41.
4. Kayadibi, İ., & Güraksın, G. E. (2023). An Explainable Fully Dense Fusion Neural Network with Deep Support Vector Machine for Retinal Disease Determination. International Journal of Computational Intelligence Systems, 16(1). https://doi.org/10.1007/s44196-023-00210-z.
5. Li, Z., Han, Y., & Yang, X. (2023). Multi-Fundus Diseases Classification Using Retinal Optical Coherence Tomography Images with Swin Transformer V2. Journal of Imaging, 9(10), 203–203. https://doi.org/10.3390/jimaging9100203.
6. He, J., Wang, J., Han, Z., Ma, J., Wang, C., & Qi, M. (2023). An interpretable transformer network for the retinal disease classification using optical coherence tomography. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-30853-z.
7. Khan, A., Pin, K., Aziz, A., Jung Woo Han, & Nam, Y. (2023). Optical Coherence Tomography Image Classification Using Hybrid Deep Learning and Ant Colony Optimization. Sensors, 23(15), 6706–6706. https://doi.org/10.3390/s23156706.
8. Zbinden, S., Şerife Seda Kucur, Steiner, P., Wolf, S., & Sznitman, R. (2016). Automatic assessment of time-resolved OCT images for selective retina therapy. International Journal of Computer Assisted Radiology and Surgery, 11(6), 863–871. https://doi.org/10.1007/s11548-016-1383-6.
9. Venkatraman, K., & Sumathi, M. (2019). Classification of Retinal Disorders Based on Fluid Patterns in OCT Images. Indian Journal of Public Health Research & Development, 10(7), 1367. https://doi.org/10.5958/0976-5506.2019.01779.0.
10. Hassan, E., Samir Elmougy, Mai Ramadan Ibraheem, M. Shamim Hossain, Khalid Al-

Mutib, Ghoneim, A., AlQahtani, S. A., & Talaat, F. M. (2023). Enhanced Deep Learning Model for Classification of Retinal Optical Coherence Tomography Images. Sensors, 23(12), 5393–5393. https://doi.org/10.3390/s23125393.

11. Hasan, M. J., Alom, M. S., & Fatema, U. (2021, May 28). Classification Performance Analysis of Retinal OCT Image using Handcrafted and Deep Learning Feature with. . . ResearchGate. https://doi.org/10.13140/RG.2.2.36346.82888/2.
12. serej, arash dehghanian. (2022, December 26). ResNet-50. Medium. https://medium.com/@arashserej/resnet-50-83b3ff33be7d
13. Ratan, P. (2020, October 28). Convolutional Neural Network Made Easy for Data Scientists. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/
14. None Mohammad Talebzadeh, None Abolfazl Sodagartojgi, None Zahra Moslemi, None Sara Sedighi, None Behzad Kazemi, & None Faezeh Akbari. (2024). Deep learning-based retinal abnormality detection from OCT images with limited data. World Journal of Advanced Research and Reviews, 21(3), 690–698. https://doi.org/10.30574/wjarr.2024.21.3.0716
15. Leandro, I., Lorenzo, B., Aleksandar, M., Dario, M., Rosa, G., Agostino, A., & Daniele, T. (2023). OCT-based deep-learning models for the identification of retinal key signs. Scientific Reports, 13(1), 14628. https://doi.org/10.1038/s41598-023-41362-4
16. Elkholy, M., & Marzouk, M. A. (2024). Deep learning-based classification of eye diseases using Convolutional Neural Network for OCT images. Frontiers in Computer Science, 5. https://doi.org/10.3389/fcomp.2023.1252295
17. Schlegl, T., Waldstein, S. M., Bogunovic, H., Endstraßer, F., Sadeghipour, A., Philip, A.-M., Podkowinski, D., Gerendas, B. S., Langs, G., & Schmidt-Erfurth, U. (2018). Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning. Ophthalmology, 125(4), 549–558. https://doi.org/10.1016/j.ophtha.2017.10.031
18. Goutam, B., Hashmi, M. F., Geem, Z. W., & Bokde, N. D. (2022). A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Images. IEEE Access, 10, 57796–57823. https://doi.org/10.1109/access.2022.3178372

APPENDICES

- **Results for the CNN 80:20 Model.**

```
Accuracy: 0.9116174693595719
Precision: 0.9170438181690977
Recall: 0.9116174693595719
Loss: 0.32679665
```

- **Results for the CNN 70:30 Model.**

```
Accuracy: 0.9052879912538121
Precision: 0.9122877793824054
Recall: 0.9052879912538121
Loss: 0.37980217
```

- **Results for the ResNet50 80:20 Model.**

```
Accuracy: 0.9668921324314703
Precision: 0.9684727329846726
Recall: 0.9668921324314703
```

- **Results for the ResNet50 70:30 Model.**

```
Accuracy: 0.9376879586190304
Precision: 0.9397074312842725
Recall: 0.9376879586190304
```

- **Application functions demonstration**

https://drive.google.com/file/d/14636ydChrGNkks7ysr_5yT5XTdU-dlcj/view?usp=sharing

- **Code repository**

https://github.com/fody-ram/FYP2

Project Logbook

Name	**Fadwa Ramadan Ali Hassan**	**Project Id**	**1205 D**
Matrix No	**2024334**	**Supervisor**	**DR. SHARYAR WANI**

No	Date	Meeting Discussion	Target met	Supervisor Signed
1	07/03/2024	Progress update	Yes	
2	14/03/2024	Optimizing the model	Yes	
3	28/03/2024	Creating more models	Yes	
4	11/04/202	Project Deployment	Yes	
5	25/04/2024	System integration	Yes	
6	09/05/2024	System validation	Yes	
7	23/05/2024	Finalize technical report	Yes	
8	30/05/2024	Presentation preparation	Yes	