

**MOBILE APP FOR RETINAL DISEASE
CLASSIFICATION USING DEEP LEARNING**

BY

PATTARAKORN INTAPRASIT

**A PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF BACHELOR OF
ENGINEERING IN BIOMEDICAL ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY
LADKRABANG
ACADEMIC YEAR 2023**

SCHOOL OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
PROJECT CERTIFICATE

Project Title	Mobile App for Retinal Disease Classification Using Deep Learning
Student Name	Mr. Pattarakorn Intaraprasit
Degree	Bachelor of Biomedical Engineering
Project Advisor	Dr. May Phu Paing
Academic Years	2023

ABSTRACT

A common retinal illness can have severe consequences for eyesight and eye health. Diabetic retinopathy is an example of a retinal illness that affects the blood vessels in the retina and can result in visual impairment or blindness. Early discovery, adequate treatment, and regular monitoring are critical in controlling these diseases and minimizing their effects on vision. An automatic classification system for detecting retinal illness using optical coherence tomography (OCT) and non-mydriatic camera pictures was developed using the MobileNetV2 deep learning network. With a sample size of 224 pixels, the scans are separated into eleven classes — Age-related macular degeneration (ARMD), branch retinal vein occlusion (BRVO), drusens (DN), diabetic retinopathy (DR), media haze (MH), myopia (MYA), optic disc cupping (ODC), optic disc edema (ODE), optic disc pallor (ODP), tessellation (TSLN), and normal retina (Normal). It is straightforward to convert models from MobileNetV2 into Android mobile apps for point-of-care diagnostics. The model was updated using validation samples, and its accuracy and sensitivity were evaluated. The model's performance was evaluated using measures like accuracy, precision, recall, and F1-score. According to the results of experiments on a large dataset, the suggested MobileNet V2 has an accuracy of 0.9575 for training, 0.9470 for validating, and 0.9630 for testing. Precision, recall, and F1-score values were also obtained, demonstrating high performance in retinal disease categorization.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my advisor, Dr. May Phu Paing, for her continuous guidance and support, which greatly influenced the direction and quality of this study. Also, the author would like to thank the owners of the Retinal Fundus Multi-Disease Image Dataset (RFMiD) 2.0, Panchal, S., Naik, A., Kokare, M., Pachade, S., Naigaonkar, R., Phadnis, P., and Bhave, A., for making the retinal imaging datasets utilized in this investigation available for free. Their efforts in collecting and sharing these datasets have been critical to the advancement of retinal imaging research. Without their dataset, this study would not have been feasible.

Pattarakorn Intaraprasit

TABLE OF CONTENTS

	Page
ABSTRACT	(i)
ACKNOWLEDGEMENTS	(ii)
LIST OF TABLES	(vii)
LIST OF FIGURES	(viii)
LIST OF SYMBOLS/ABBREVIATIONS	(xi)
CHAPTER 1 INTRODUCTION	
1.1 Background and Significance of the Study	1
1.2 Objectives	2
1.3 Scope of the Study	3
1.3.1 Retinal Disease Categories	3
1.3.2 Non-mydriatic Camera Images	3
1.3.3 Data Cleaning and Cross-Validation	3
1.3.4 MobileNetV2 Deep Learning Model	3
1.3.5 Android Mobile Application Integration	3
1.3.6 Technological and Healthcare Implications	4
1.4 Report Outline	4
CHAPTER 2 REVIEW OF THEORY RELATED	5
2.1 Retinal Diseases	5
2.1.1 Retinal Disease Screening	6
2.1.2 Retinal Diseases Class Breakdown	7
2.1.2.1 Age-Related Macular Degeneration (ARMD)	7

2.1.2.2 Branch Retinal Vein Occlusion (BRVO)	9
2.1.2.3 Drusens (DN)	10
2.1.2.4 Diabetic Retinopathy (DR)	11
2.1.2.5 Media Haze (MH)	12
2.1.2.6 Myopia (MYA)	13
2.1.2.7 Optic Disc Cupping (ODC)	14
2.1.2.8 Optic Disc Edema (ODE)	16
2.1.2.9 Optic Disc Pallor (ODP)	17
2.1.2.10 Tessellation (TSLN)	18
2.1.2.11 Normal Retinal	19
2.1.3 Retinal Disease Treatment	19
2.2 Deep Learning	21
2.2.1 Convolutional Neural Network (CNN)	22
2.2.2 MobileNetV2 Architecture	23
2.2.3 Data Augmentations for Biomedical Images	25
2.3 Transfer Learning	26
2.4 Non-Mydriatic Camera Images	27
2.5 Optical Coherence Tomography (OCT) Images	28
2.6 Data Cleaning	30
2.5.1 Isolation Forest	30
2.5.2 Local Outlier Factor (LOF)	32
2.7 Cross Validation	33
2.6.1 K-Fold Cross-Validation	34
2.8 Summary	35
 CHAPTER 3 MATERIALS AND METHODS	36
3.1 Introduction	36
3.2 Dataset	36
3.3 Materials	38
3.3.1 Integrated Development Environment (IDE)	38
3.3.2 Deep Learning Frameworks	39

3.3.3 Hardware Specifications	39
3.3.4 Programming Language	39
3.3.5 Mobile Application Development	40
3.4 Proposed Methodology	41
3.4.1 Parameter Settings	41
3.4.2 Training with Unbalanced Dataset	41
3.4.3 Training with Unbalanced Dataset using Different Class Weight	42
3.4.4 Training with Unbalanced Dataset and using Data Cleaning	42
3.4.5 Training with Data Augmentation Only	43
3.4.6 Training with Data Augmentation before using Data Cleaning	44
3.4.7 Training with Data Cleaning before using Data Augmentation	44
3.5 Creating a Mobile Application using Android Studio	45
3.5.1 Design Application Interface	45
3.6 Summary	46
 CHAPTER 4 EXPERIMENTAL RESULT	47
4.1 Introduction	47
4.2 Evaluation Criteria Equations	47
4.3 Results of Proposed Methodology	48
4.3.1 Result of Training with Unbalanced Dataset	48
4.3.2 Result of Training with Unbalanced Dataset using Different Class Weight	51
4.3.3 Result of Training with Unbalanced Dataset and using Data Cleaning	55
4.3.4 Result of Training with Data Augmentation Only	58
4.3.5 Result of Training with Data Augmentation before using Data Cleaning	61
4.3.6 Result of Training with Data Cleaning before using Data Augmentation	64
4.4 Comparison of All Proposed Methodologies	67
4.5 Mobile Application	68

4.5.1 Use a Testing Image by Taking a Picture	69
4.5.2 Use a Testing Image by Browsing the Gallery	70
4.5.2.1 Prediction for the Rest of the Class using ‘Launch Gallery Method’	73
4.6 Summary	75
 CHAPTER 5 DISCUSSION AND CONCLUSION	76
5.1 Introduction	76
5.2 Discussion	76
5.3 Conclusion	77
5.4 Future Scope	77
 REFERENCES	78

LIST OF TABLES

Tables	Page
3.1 Specifications of dataset	11
3.2 Parameter Settings	22
4.1 Confusion Matrix of MobileNetV2 using Training with Unbalanced Dataset Method	29
4.2 Classification Report on The Testing Set using Training with Unbalanced Dataset Method	35
4.3 Confusion Matrix of MobileNetV2 using Training with Unbalanced Dataset using Different Class Weight Method	36
4.4 Classification Report on The Testing Set using Training with Unbalanced Dataset using Different Class Weight Method	43
4.5 Confusion Matrix of MobileNetV2 using Training with Unbalanced Dataset using Data Cleaning Method	48
4.6 Classification Report on The Testing Set using Training with Unbalanced Dataset using Data Cleaning Method	55
4.7 Confusion Matrix of MobileNetV2 using Training with Data Augmentation Only Method	60
4.8 Classification Report on The Testing Set using Training with Data Augmentation Only Method	60
4.9 Confusion Matrix of MobileNetV2 using Training with Data Augmentation before using Data Cleaning Method	63
4.10 Classification Report on The Testing Set using Training with Data Augmentation before using Data Cleaning Method	63
4.11 Confusion Matrix of MobileNetV2 using Training with Data Cleaning before using Data Augmentation Method	66
4.12 Classification Report on The Testing Set using Training with Data Cleaning before using Data Augmentation Method	66
4.13 Performance Metrics for Different Classification Approaches	67
5.1 Comparison Between the Literature Review and Our Research	76

LIST OF FIGURES

Figures	Page
2.1 Visualization of ARMD images from the RFMiD 2.0 dataset	9
2.2 Visualization of BRVO images from the RFMiD 2.0 dataset	10
2.3 Visualization of DN images from the RFMiD 2.0 dataset	11
2.4 Visualization of DR images from the RFMiD 2.0 dataset	12
2.5 Visualization of MH images from the RFMiD 2.0 dataset	13
2.6 Visualization of MYA images from the RFMiD 2.0 dataset	14
2.7 Visualization of ODC images from the RFMiD 2.0 dataset	15
2.8 Visualization of ODE images from the RFMiD 2.0 dataset	16
2.9 Visualization of ODP images from the RFMiD 2.0 dataset	17
2.10 Visualization of TSLN images from the RFMiD 2.0 dataset	18
2.11 Visualization of normal retinal images from the Kaggle dataset	19
2.12 Deep learning, a subset of machine learning and artificial intelligence	22
2.13 Convolutional Neural Network (CNN) Architecture	23
2.14 MobileNetV2 Architecture	25
2.15 Non-mydriatic Fundus Camera	28
2.16 Optical Coherence Tomography (OCT) Workstation	29
2.17 Isolation Forest Method	31
2.18 Local Outlier Factor (LOF) Method	33
2.19 Visualization of Five-Fold Cross Validation	34
3.1 The procedure for establishing a dataset while obtaining consent from ophthalmologists for the development of a Computer-Aided Disease Diagnosis	36
3.2 TRC-NW300 Non-Mydriatic Retinal Camera	38
3.3 Jupyter Notebook Version 7.0.2 Interface	38
3.4 TensorFlow Logo	39
3.5 Huawei P20 Mobile Phone	40
3.6 Android Studio Interface	40

3.7 Training with Unbalanced Dataset	42
3.8 Training with Unbalanced Dataset using Different Class Weight	42
3.9 Training with Unbalanced and using Data Cleaning	43
3.10 Training with Data Augmentation Only	43
3.11 Training with Data Augmentation before using Data Cleaning	44
3.12 Training with Data Cleaning before using Data Augmentation	45
3.13 Mobile Application Design on Android Studio	46
4.1 The percentage of the images in each class for Training with Unbalanced Dataset Method	48
4.2 Training and Validation Accuracy vs. Epoch using Training with Unbalanced Dataset Method	49
4.3 Training and Validation Loss vs. Epoch using Training with Unbalanced Dataset Method	50
4.4 The percentage of the images in each class for Training with Unbalanced Dataset using Different Class Weight Method	52
4.5 Training and Validation Accuracy vs. Epoch using Training with Unbalanced Dataset using Different Class Weight Method	53
4.6 Training and Validation Loss vs. Epoch using Training with Unbalanced Dataset using Different Class Weight Method	53
4.7 The percentage of the images in each class for Training with Unbalanced Dataset and using Data Cleaning Method	55
4.8 Training and Validation Accuracy vs. Epoch using Training with Unbalanced Dataset using Data Cleaning Method	56
4.9 Training and Validation Loss vs. Epoch using Training with Unbalanced Dataset using Data Cleaning Method	56
4.10 The percentage of the images in each class for Training with Data Augmentation Only Method	58
4.11 Training and Validation Accuracy vs. Epoch using Training with Data Augmentation Only Method	59
4.12 Training and Validation Loss vs. Epoch using Training with Data Augmentation Only Method	59

4.13 The percentage of the images in each class for Training with Data Augmentation before using Data Cleaning Method	61
4.14 Training and Validation Accuracy vs. Epoch using Training with Data Augmentation before using Data Cleaning Method	62
4.15 Training and Validation Loss vs. Epoch using Training with Data Augmentation before using Data Cleaning Method	62
4.16 The percentage of the images in each class for Training with Data Cleaning before using Data Augmentation Method	64
4.17 Training and Validation Accuracy vs. Epoch using Training with Data Cleaning before using Data Augmentation Method	65
4.18 Training and Validation Loss vs. Epoch using Training with Data Cleaning before using Data Augmentation Method	65
4.19 Retinal Diseases Analyzer Mobile Application Icon	68
4.20 Classification of BRVO with a Prediction Score of 87.34%	73
4.21 Classification of DR with a Prediction Score of 94.48%	73
4.22 Classification of DN with a Prediction Score of 32.69%	73
4.23 Classification of MH with a Prediction Score of 89.15%	73
4.24 Classification of MYA with a Prediction Score of 82.25%	74
4.25 Classification of ODC with a Prediction Score of 31.50%	74
4.26 Classification of ODE with a Prediction Score of 90.29%	74
4.27 Classification of ODP with a Prediction Score of 73.12%	74
4.28 Classification of TSLN with a Prediction Score of 45.24%	75

LIST OF SYMBOLS/ABBREVIATIONS

Symbols/Abbreviations	Terms
AI	Artificial Intelligence
ARMD	Age-Related Macular Degeneration
BRVO	Branch Retinal Vein Occlusion
CAD	Computer-aided Diagnostic
CNN	Convolutional Neural Network
CNV	Choroidal Neovascularization
DN	Drusens
DR	Diabetic Retinopathy
ELM	External Limiting Membrane
GCL	Ganglion Cell Layer
INL	Inner Nuclear Layer
IPL	Inner Plexiform Layer
LOF	Local Outlier Factor
MH	Media Haze
ML	Machine Learning
MYA	Myopia
NLP	Natural Language Processing
ODC	Optic Disc Cupping
ODE	Optic Disc Edema
ODP	Optic Disc Pallor
ONL	Outer Nuclear Layer
OPL	Outer Plexiform Layer
RFMiD	Retinal Fundus Multi-Disease Image Dataset
RNFL	Retinal Nerve Fiber Layer
RPE	Retinal Pigment Epithelium
TSLN	Tessellation
VEGF	Vascular Endothelial Growth Factor

CHAPTER 1

INTRODUCTION

The themes of this report will be discussed in this chapter. Some background for the work will be given. Following that, the project's motivations and objectives will be outlined. A project summary will then be provided. Finally, the dissertation will be reviewed chapter by chapter.

Manually analyzing medical pictures is a time-consuming and unreliable traditional way of diagnosing eye diseases. Recent advances in deep learning and artificial intelligence (AI) have enabled the development of computer-aided diagnostic (CAD) systems that can improve the accuracy and speed of detecting retinal diseases [1]. In this research, the focus will be on the development of an automated classification approach for the identification of retinal diseases from non-mydriatic camera images and Optical Coherence Tomography (OCT) images. Additionally, an investigation will be conducted to determine the feasibility of creating Android mobile apps for point-of-care diagnosis using the suggested approach. In brief, the ultimate objective of this research is to embed a deep learning model into an Android smartphone application that can detect retinal illnesses quickly and correctly using retinal scans. The program employs the MobileNetV2 deep learning network to categorize eye disorders from retinal scans in order to improve the speed and accuracy of eye problem diagnosis, ultimately leading to improved patient outcomes.

1.1 Background and Significance of the Study

Numerous illnesses that might have a negative impact on vision and eye health are included in the category of retinal diseases. Age-related macular degeneration (ARMD), branch retinal vein occlusion (BRVO), drusens (DN), diabetic retinopathy (DR), media haze (MH), myopia (MYA), optic disc cupping (ODC), optic disc edema (ODE), optic disc pallor (ODP), and tessellation (TSLN) are among these conditions that stand out due to their prevalence and potential to cause severe visual impairments. The normal retinal class is also included. For appropriate therapies, management plans,

and the avoidance of irreversible vision loss, many retinal illnesses must be accurately and promptly diagnosed.

In the modern healthcare landscape, the integration of advanced technologies has the potential to revolutionize medical diagnostics. Deep learning, a subset of artificial intelligence, has emerged as a transformative tool for the analysis of medical images. It empowers computer-aided diagnostic (CAD) systems to interpret complex visual data with remarkable accuracy, speed, and consistency [2]. This study harnesses the capabilities of the MobileNetV2 architecture to develop an automated classification framework specifically tailored to the identification of a diverse array of retinal diseases from non-mydriatic camera images and OCT images. Furthermore, the incorporation of this model within Android mobile applications enhances the feasibility of point-of-care diagnosis, extending access to efficient and reliable diagnostic solutions.

1.2 Objectives

- Create a robust MobileNetV2-based deep learning model.
- Classify eleven retinal disease categories precisely, including: Age-related macular degeneration, Branch retinal vein occlusion, Drusens, Diabetic retinopathy, Media haze, Myopia, Optic disc cupping, Optic disc edema, Optic disc pallor, Tessellation, and Normal class.
- Incorporate data cleaning methods.
- Employ k-fold cross-validation for enhanced model robustness.
- Do the experiment in six different approaches, including training with unbalanced dataset, training with unbalanced dataset using different class weight, training with unbalanced and using data cleaning, training with data augmentation only, training with data augmentation before using data cleaning, and training with data cleaning before using data augmentation.
- Achieve exceptional accuracy, precision, recall, and F1-score through rigorous training and validation.
- Integrate the model into a mobile android application.
- Demonstrate model superiority over existing methods.

1.3 Scope of the Study

1.3.1 Retinal Disease Categories

The study's primary objective is the accurate classification of ten specific retinal disease categories: age-related macular degeneration, branch retinal vein occlusion, drusens, diabetic retinopathy, media haze, myopia, optic disc cupping, optic disc edema, optic disc pallor, tessellation and normal retinal. The model will be trained to differentiate between these diseases based on patterns and features extracted from retinal images.

1.3.2 Non-mydriatic Camera Images

The dataset used for training, validation, and testing will consist of retinal images captured using a non-mydriatic camera equipped with auto-focus, auto-exposure, and auto-shoot capabilities. The dataset will include images captured after the application of tropicamide at a 0.5% concentration to induce mydriasis, enhancing the quality of fundus images [3]. This choice of dataset limits the study to the analysis of diseases observable through non-mydriatic imaging techniques.

1.3.3 Data Cleaning and Cross-Validation

To augment the resilience and effectiveness of the model, the research will investigate methodologies for managing labels with noise using the isolation forest method to find outlier images and will incorporate k-fold cross-validation techniques throughout the model development phase.

1.3.4 MobileNetV2 Deep Learning Model

The MobileNetV2 deep learning architecture is used as the primary technical methodology in the study to classify diseases. The emphasis will be on perfecting the model to realize remarkable accuracy, precision, recall, and F1-score in recognizing the specified retinal diseases and capabilities to import into mobile devices.

1.3.5 Android Mobile Application Integration

The implementation of the trained deep learning model within an Android mobile application is a pivotal aspect of the study. The mobile application will be designed to provide rapid, accurate point-of-care diagnosis, making use of the model's capabilities for real-time disease classification. The feasibility and effectiveness of the mobile application for enhancing diagnostic accuracy and speed will be assessed.

1.4 Report Outline

The remainder of this report is structured as follows:

The second chapter discusses the current state-of-the-art in retinal diseases, deep learning, convolutional neural network (CNN), non-mydriatic camera images, optical coherence tomography (OCT) images, data cleaning, transfer learning, data augmentation for biomedical images, and k-fold cross-validation.

The third chapter describes the design and implementation of the dataset, materials, methods, including evaluation criteria equations, k-fold cross-validation, and proposed methodology.

The fourth chapter demonstrates the results of all six proposed methodologies from the previous chapter and evaluation metrics, including training accuracy, validating accuracy, testing accuracy, precision, recall, and F1-score.

The report's fifth chapter, which reviews the work done and makes discussion of important aspects of it, brings it to an end. Finally, future work is discussed with an emphasis on clinical settings leading to better patient outcomes.

CHAPTER 2

REVIEW OF THEORY RELATED

This chapter examines the current advancements in utilizing deep learning for the analysis and design phases of the project, specifically focusing on retinal diseases. Eight objectives were served by the investigation: Firstly, the discovery of retinal diseases, including screening, class breakdown, and treatment (section 2.1); secondly, an explanation of the deep learning algorithms is provided (section 2.2); thirdly, an investigation into transfer learning is carried out (section 2.3); fourthly, a study of non-mydriatic camera images is undertaken (section 2.4); fifthly, the illustration of optical coherence tomography (OCT) images is presented (section 2.5); sixthly, the interpretation of the data cleaning is provided (section 2.6); seventhly, an explanation of k-fold cross-validation is given (section 2.7); and lastly, section 2.8 wraps up the entire chapter.

2.1 Retinal Diseases

Retinal diseases comprise a varied array of medical disorders that impact the fragile tissue situated at the rear of the eye, recognized as the retina. The retina assumes a crucial function in visual processes, as it is accountable for capturing light and transforming it into neural impulses subsequently relayed to the brain, facilitating our comprehension of the surrounding environment. Regrettably, the retina is vulnerable to assorted ailments capable of substantially diminishing visual capability and, in certain instances, culminating in irreversible blindness [4].

A prevalent group of retinal ailments is age-related macular degeneration (ARMD), a gradual ailment primarily impacting the central segment of the retina, which is called the macula, responsible for precise vision [5]. ARMD can result in hazy or distorted central eyesight, complicating activities like reading and facial recognition. Another common concern is diabetic retinopathy (DR), which arises when diabetes is inadequately managed over time, causing damage to the retinal blood vessels. This damage can give rise to leakage, hemorrhaging, and the formation of blood vessels, all of which can influence an individual's capacity to perceive clearly.

Retinal detachment is a worrisome circumstance wherein the retina becomes detached from the underlying tissue. This occurrence can be triggered by elements such as physical injury, advancing age, or pre-existing ocular conditions. It is imperative to manage these ailments promptly, given that they have the potential to result in enduring vision impairment. One such hereditary condition called retinitis pigmentosa gradually damages the cells of the retina, leading to symptoms like night blindness and tunnel vision.

The progress in imaging and diagnostic technologies has made it possible to detect diseases at an earlier stage and with greater precision. Techniques like optical coherence tomography (OCT) and fundus photography help capture images of the layers and blood vessels in the retina, assisting doctors in diagnosing these conditions and monitoring their progression [6]. Moreover, recent advancements in intelligent deep learning algorithms have shown encouraging outcomes by automating the analysis of retinal images. This enables more reliable identification of abnormalities.

2.1.1 Retinal Disease Screening

The screening of retinal diseases holds a significant function in promptly identifying and taking proactive measures against a range of ocular conditions that might result in diminished or lost vision. This procedure encompasses a methodical assessment of the retina through specialized imaging methodologies, with the purpose of detecting any irregularities or indications of ailments. The primary objective of retinal disease screening is to identify potential problems during their initial phases, thus enabling timely interventions and therapies to avert the progression of vision decline [7].

Progress in medical technology has notably improved the efficiency of screening for retinal diseases. Optical coherence tomography (OCT) and fundus photography stand out as extensively employed imaging techniques that furnish elaborate and high-detail visuals of the retina's strata, vascular network, and structural soundness. These visuals grant eye care specialists the capacity to meticulously assess retinal well-being and detect nuanced alterations that could hint at the existence of ailments like diabetic retinopathy (DR), and similar conditions [8].

In recent times, the fusion of artificial intelligence, or, in short, AI, and deep learning algorithms has revolutionized the identification of retinal diseases. AI-driven systems exhibit remarkable rapidity and accuracy in scrutinizing retinal images, swiftly pinpointing irregularities, and uncovering potential issues. This expedites the screening process while diminishing the chance of human inaccuracies. Extensive collections of retinal images can be employed to educate AI algorithms, which subsequently grasp how to discern patterns associated with diverse ailments. This capability empowers them to formulate precise prognoses and assist medical practitioners in rendering well-founded judgements.

Retinal disease screening holds immense promise for improving the overall quality of eye care. By catching retinal diseases in their earliest stages, healthcare providers can initiate timely treatments, lifestyle modifications, or interventions to slow down the progression of the condition. This proactive approach not only enhances the chances of preserving visual acuity but also reduces the economic and social burdens associated with advanced-stage retinal diseases. As technology continues to evolve, retinal disease screening is likely to become even more precise, accessible, and integral to maintaining optimal eye health for individuals around the world.

2.1.2 Retinal Diseases Class Breakdown

In the discipline of ophthalmology, the classification and categorization of retinal illnesses are crucial since they allow for precise diagnosis and individualized treatment plans. There is a wide range of illnesses within the category of retinal diseases, each with unique traits and effects on vision. This section digs into the complex breakdown of various retinal conditions, illuminating their prevalence, clinical importance, and effects on the health of patients.

2.1.2.1 Age-Related Macular Degeneration (ARMD)

ARMD stands as a significant retinal disorder, casting its impact primarily on the aging population. With the aging demographic, the prevalence of ARMD has gained prominence, making it a critical concern in the field of ophthalmology. ARMD is a developing degenerative ailment impacting the macula responsible for acute visual

acuity. This disorder presents in two clear variations: "dry" and "wet" ARMD. The depiction of ARMD visuals is provided in Figure 2.1.

Dry ARMD, also termed non-neovascular or atrophic ARMD, represents the more prevalent type, constituting around 80% of ARMD instances [9]. This variant is marked by the buildup of drusen, which are lipid- and protein-based yellow deposits beneath the retinal pigment epithelium (RPE) [9]. With time, these deposits can lead to retinal thinning and degeneration, culminating in gradual central vision decline. Despite the relatively slower advancement of dry ARMD compared to its wet counterpart, its prevalence and potential influence on an individual's quality of life underscore its notable clinical significance.

Wet ARMD, also denoted as neovascular or exudative ARMD, constitutes the remaining 20% of cases, often leading to more pronounced vision impairment [10]. This variation is distinguished by the development of aberrant blood vessels, referred to as choroidal neovascularization (CNV), beneath the macula. These vessels are susceptible to leakage and hemorrhaging, resulting in the buildup of fluid and blood within the retina. This leakage can give rise to abrupt and severe central vision reduction, transpiring swiftly and exerting a significant impact on the patient's capacity to undertake daily activities like reading, driving, and facial recognition.

The significance of ARMD from a clinical perspective arises not only due to its prevalence but also due to its capability to induce irreversible vision decline. Being a primary contributor to blindness in those aged 50 and above, ARMD places a substantial load on both impacted individuals and the healthcare infrastructure [11]. Furthermore, the ramifications of the ailment extend beyond visual impairment; those with ARMD could encounter diminished self-sufficiency, depressive episodes, and a decline in their quality of life.

The impact of ARMD on patients' well-being extends widely, underscoring the requirement for precise diagnosis, prompt intervention, and efficient management methods. Grasping the unique attributes of both dry and wet ARMD, alongside their occurrence and clinical relevance, holds utmost importance in devising strategies for timely identification and personalized treatment strategies. This section's examination of ARMD brings attention to the crucial necessity of tackling this intricate retinal

ailment to alleviate its repercussions and improve patients' visual health and overall welfare.

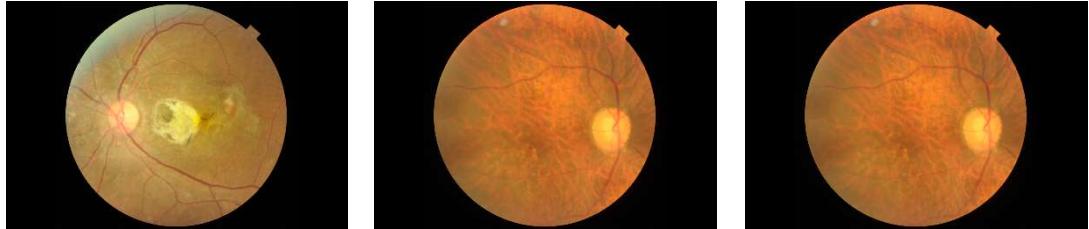


Figure 2.1 Visualization of ARMD images from the RFMiD 2.0 dataset [3].

2.1.2.2 Branch Retinal Vein Occlusion (BRVO)

BRVO stands as a noteworthy retinal vascular ailment of considerable clinical significance in the domain of ophthalmology. This disorder emerges due to the blockage of a retinal vein, disrupting typical blood circulation and giving rise to an array of visual hindrances. BRVO is categorized into two primary forms: those linked with macular edema and those not associated with macular edema. It predominantly impacts individuals across different age ranges, underscoring the importance of holistic comprehension and effective management strategies. BRVO visuals are presented in Figure 2.2.

The frequency of BRVO has attracted attention for its influence on visual well-being and life quality. BRVO ranks among the most prevalent retinal vascular disorders, with a growing occurrence in the elderly population [12]. Factors that elevate risk comprise hypertension, diabetes, atherosclerosis, and glaucoma. This ailment frequently causes retinal bleeding, inflammation, and impaired blood delivery to distinct retinal sectors, culminating in visual disruptions spanning from slight haziness to substantial central vision reduction.

The significance of BRVO from a clinical perspective rests in its capacity to induce permanent vision decline when untreated. The variant associated with macular edema, marked by swelling in the retina's central region called the macula, can give rise to notable distortion and loss of central vision. This impairment can considerably impede one's capacity to participate in everyday tasks. Furthermore, non-macular edema linked to BRVO can result in defects in the peripheral visual field and visual unease.

Apart from evident indications of vision reduction, BRVO exerts adverse consequences on patients' well-being. Individuals afflicted with BRVO might encounter a lowered quality of life, restricted autonomy, and emotional distress stemming from the constraints posed by their impaired eyesight. Consequently, it remains vital to promptly and accurately diagnose patients while administering appropriate interventions to mitigate BRVO's unfavorable repercussions on their health.



Figure 2.2 Visualization of BRVO images from the RFMiD 2.0 dataset [3].

2.1.2.3 Drusens (DN)

Drusens (DN) represent a significant aspect of retinal health, warranting careful consideration due to their prevalence, clinical implications, and effects on patients' vision. Drusens are characterized by the accumulation of yellowish deposits beneath the retinal pigment epithelium, forming small, distinct lesions on the retina's surface [13]. This condition is particularly prevalent in individuals over the age of 50, making it a noteworthy concern within the context of age-related eye diseases. The visualization of DN images is shown in Figure 2.3.

The prevalence of drusens holds significance due to their association with several retinal conditions, including age-related macular degeneration (ARMD). Drusens can be categorized into different types based on their size and appearance, ranging from hard, small drusens to larger, soft drusens [14]. While small drusens are more common and often considered benign, the presence of larger, soft drusens is recognized as a risk factor for the development of ARMD, which can result in central vision loss.

The clinical importance of drusens lies in their potential to signify the presence of underlying retinal pathology and the risk of disease progression. Soft drusens, in

particular, can lead to the breakdown of the retinal pigment epithelium, or, in short, RPE, compromising retinal function and triggering the development of ARMD. The monitoring of drusen characteristics and progression is crucial in identifying individuals at risk of vision loss and guiding timely intervention strategies.

Beyond just being unsightly, drusens have a negative impact on patients' health. The existence of soft drusens necessitates vigilant observation and control, even if minor drusens may not immediately cause apparent vision alterations. Patients with drusens, particularly those who are at risk of developing ARMD, may experience gradual central vision loss that affects their capacity to do daily activities.



Figure 2.3 Visualization of BRVO images from the RFMiD 2.0 dataset [3].

2.1.2.4 Diabetic Retinopathy (DR)

DR emerges as a significant ocular complication of diabetes, demanding close attention due to its prevalence, clinical significance, and profound effects on patients' vision and overall well-being. DR is a vascular disorder that develops as a consequence of a prolonged, constant rise in blood sugar, leading to damage in the blood vessels that supply the retina [15]. This condition holds particular relevance within the diabetic population and requires comprehensive understanding to mitigate its impact. The visualization of DR images is shown in Figure 2.4.

The frequency of DR is intricately connected to the escalating worldwide occurrence of diabetes, particularly in those with inadequately regulated blood sugar levels. As a prominent contributor to adult blindness, DR emphasizes the imperative of timely identification and treatment. The disorder encompasses varying stages, spanning from mild non-proliferative DR characterized by microaneurysms and minor hemorrhages, to advanced phases involving neovascularization and macular edema

[16]. Due to its widespread prevalence and propensity for advancement, DR possesses substantial clinical significance.

The clinical importance of DR stems from its capacity to induce permanent vision decline if not appropriately addressed. As DR progresses, aberrant blood vessels might emerge, giving rise to hemorrhages and fluid leakage into the retina. This can culminate in macular edema, where the central retina swells, negatively affecting central vision [17]. The more advanced proliferative stage of DR entails the development of fragile new blood vessels that can lead to bleeding into the vitreous, resulting in vision impairment. The impact of DR on patients' well-being extends broadly, encompassing both physical and emotional aspects. Those dealing with DR could experience not only visual impairment but also diminished self-sufficiency, lowered life quality, and emotional strain. The constraints posed by impaired vision can impede routine tasks, influencing patients' general functionality and self-assurance.

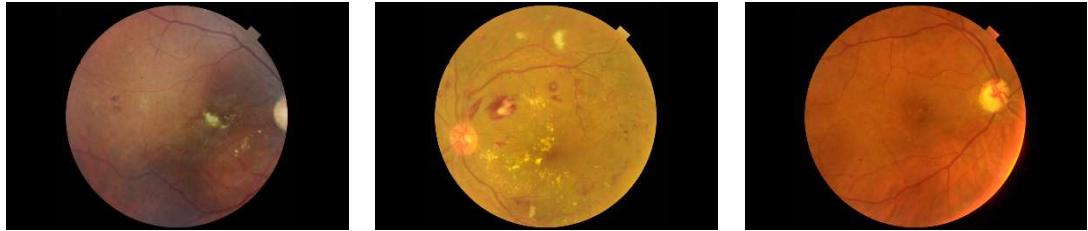


Figure 2.4 Visualization of DR images from the RFMiD 2.0 dataset [3].

2.1.2.5 Media Haze (MH)

Media Haze (MH) represents a noteworthy ocular phenomenon meriting examination due to its unique attributes and implications for visual sharpness. MH, colloquially known as "lens haze," involves the existence of opaqueness within the eye's ocular media, encompassing the cornea, lens, and vitreous humor [18]. This state can emerge from diverse factors like cataracts, corneal edema, or inflammation, impeding optimal light passage and subsequently inducing visual blurriness. The illustration of MH visuals is depicted in Figure 2.5.

The frequency of MH fluctuates in accordance with underlying reasons and patient characteristics. Cataracts, for instance, emerge as a widespread factor contributing to MH, particularly in the elderly population [19]. Furthermore, states like corneal edema,

characterized by fluid buildup in the cornea, can give rise to momentary media haze. MH has the potential to influence individuals across various age groups, presenting either as an independent occurrence or in conjunction with other ocular conditions.

The clinical importance of MH lies in its potential to significantly degrade visual acuity. The presence of opacities within the ocular media scatters and absorbs light, leading to reduced contrast sensitivity and blurriness. Individuals experiencing MH may report difficulty with tasks that demand clear vision, such as recognizing objects, or navigating unfamiliar environments. Timely assessment and accurate diagnosis are crucial in identifying the underlying causes of MH.

The effects of MH on patients' health extend beyond visual impairment to impact overall quality of life. The challenges posed by reduced visual clarity can result in diminished independence, frustration, and anxiety. Individuals with MH may struggle with maintaining their usual routines and engaging in activities they once enjoyed. Thus, the management of MH not only targets visual improvement but also aims to restore patients' functionality.

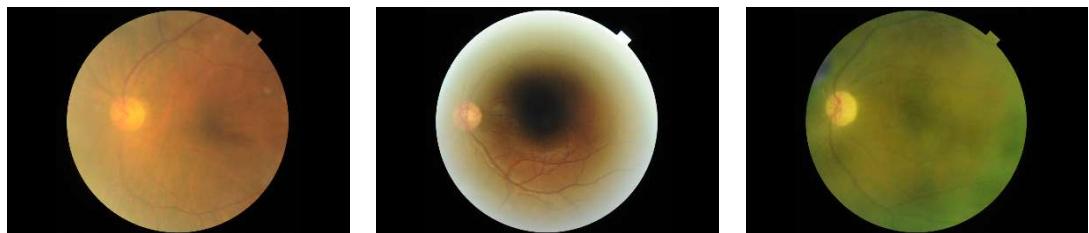


Figure 2.5 Visualization of MH images from the RFMiD 2.0 dataset [3].

2.1.2.6 Myopia (MYA)

Myopia (MYA), commonly referred to as nearsightedness, is a prevalent refractive error that holds significant importance in the context of visual health and its impact on individuals' daily lives [20]. Myopia occurs when the axial length of the eyeball is increased, causing the light to focus on the front of the retina instead of straight on it. This makes it harder to focus on distant objects while near vision remains relatively clear. The prevalence of myopia has been steadily increasing globally, particularly among young individuals in urbanized environments. The visualization of MYA images is shown in Figure 2.6.

Alarms are being raised by the rising prevalence of myopia because it affects a sizable portion of the population and has negative consequences on general health. The increase in myopia cases is related to factors like genetic predisposition, prolonged close-range activities like reading or utilizing screens, and environmental impacts [21]. Myopia typically develops in the early years and, if untreated, can gradually get worse, leading to more severe forms of myopia and other problems.

Myopia has clinical significance since it has the potential to impede people's functionality and cause visual discomfort. Myopia patients frequently struggle to distinguish distant things clearly, which makes it difficult for them to read traffic signs, watch presentations, or take in outdoor scenery [22]. If left untreated, severe myopia increases the chance of developing eye diseases such retinal detachment, glaucoma, and myopic maculopathy, which can damage vision or possibly cause blindness.

Vision impairment caused by myopia has implications beyond just physical limitations. Its impact on visual acuity can obstruct scholarly and vocational progress. Meanwhile, the emotional toll it exacts should not be dismissed. The self-confidence and self-esteem of children and teenagers with myopia could deteriorate due to potential social or academic impediments.



Figure 2.6 Visualization of MYA images from the RFMiD 2.0 dataset [3].

2.1.2.7 Optic Disc Cupping (ODC)

Optic Disc Cupping (ODC) is a distinctive ocular feature that merits investigation due to its clinical implications and relevance in the field of ophthalmology [23]. The optic disc, or the optic nerve head, is the location where the optic nerve exits from the eye and makes its connection to the brain [24]. Cupping refers to the hollowed-out or excavated appearance of the central portion of the optic disc, creating a cup-like depression. This occurrence can indicate underlying damage to the optic nerve and is

frequently connected with conditions like glaucoma. The illustration of ODC images is presented in Figure 2.7.

The occurrence of optic disc cupping fluctuates depending on the analyzed population and its root origins. It is predominantly noted among individuals afflicted with glaucoma, a collection of advancing eye disorders marked by optic nerve impairment and visual field decline. ODC can also manifest on its own, and its existence might encourage additional inquiry to ascertain the origin and possible connected vulnerabilities.

The clinical significance of optic disc cupping rests in its role as a pivotal marker for glaucoma, notably open-angle glaucoma, the predominant variant of the condition. In glaucoma, heightened intraocular pressure can cause gradual harm to the optic nerve and an ensuing reduction in the visual field. The degree of cupping can offer insights into the level of optic nerve impairment and the likelihood of disease advancement. Consistent tracking of optic disc cupping holds vital importance for the timely identification and control of glaucoma.

The implications of optic disc cupping for patients' well-being revolve around its connection to glaucoma and its potential influence on visual capabilities [25]. With the advancement of glaucoma, visual field abnormalities might surface, impacting both peripheral and central vision. Those with advanced optic disc cupping and glaucoma might encounter difficulties in activities necessitating peripheral awareness, like driving or navigating congested surroundings.



Figure 2.7 Visualization of ODC images from the RFMiD 2.0 dataset [3].

2.1.2.8 Optic Disc Edema (ODE)

Optic Disc Edema (ODE) stands as a noteworthy ocular presentation that calls for investigation owing to its clinical importance and repercussions for visual well-being. Edema pertains to tissue enlargement due to fluid buildup, and in the context of the optic disc, it leads to a distinctively enlarged appearance [26]. The depiction of ODE visuals is exhibited in Figure 2.8.

The occurrence of optic disc edema is influenced by its root origins and contributing elements. It can arise due to an array of conditions, including papilledema, often connected with heightened intracranial pressure stemming from intracranial tumors, cerebral hemorrhage, or pseudotumor cerebri. Inflammatory disorders, infections, and vascular irregularities all hold potential as culprits behind optic disc edema [27].

The clinical significance of optic disc edema rests in its capacity to indicate underlying health matters necessitating immediate assessment and care. Especially, papilledema holds significance as a noteworthy clinical indication that could point towards a severe brain-related condition, like heightened intracranial pressure [27]. Timely evaluation and identification of the root cause of optic disc edema are essential for suitable treatment and mitigating potential complications.

The implications of optic disc edema for patients' well-being revolve around its connection to underlying conditions and its influence on visual capabilities. In the case of papilledema, for instance, it can give rise to visual disruptions, such as blurred vision and temporary visual dimming, attributable to the compression of optic nerve fibers [27]. In the absence of treatment, it could lead to enduring vision decline. Different origins of optic disc edema might lead to diverse impacts on visual function, contingent on the extent and underlying pathology.



Figure 2.8 Visualization of ODE images from the RFMiD 2.0 dataset [3].

2.1.2.9 Optic Disc Pallor (ODP)

Optic Disc Pallor (ODP) represents an exceptional ocular manifestation meriting deeper investigation due to its clinical implications and significance within ophthalmology. Pallor refers to the fading or whitening of optic disc tissue resulting from diminished blood supply and neuronal damage [28]. The depiction of ODP visuals is depicted in Figure 2.9.

The frequency of optic disc pallor differs based on the root origins and contributing factors. Optic nerve atrophy, ischemic optic neuropathy, and neurodegenerative diseases are among the numerous conditions capable of inducing ODP [29]. Additional diagnostic assessments can be guided by the clinical display of ODP, which furnishes valuable details about the underlying pathology.

The clinical significance of optic disc pallor rests in its capability to signal substantial impairment of the optic nerve and vision decline. ODP can suggest conditions culminating in the deterioration of optic nerve fibers, such as multiple sclerosis or glaucoma. It can also be caused by ischemia events that impair the blood flow to the optic nerve, resulting in decreased oxygen and nutrient supply to the nerve tissue.

The implications of optic disc pallor for patients' well-being revolve around its connection to underlying conditions and its influence on visual capabilities. ODP frequently suggests irreversible harm to the optic nerve, potentially leading to enduring vision decline [30]. Those with ODP might encounter diminished visual acuity, alterations in visual field perception, and compromised color vision. The extent of vision loss hinges on the gravity of the underlying condition and the level of optic nerve impairment.

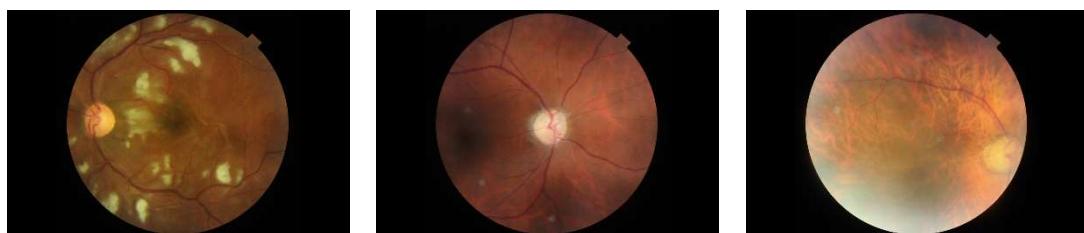


Figure 2.9 Visualization of ODP images from the RFMiD 2.0 dataset [3].

2.1.2.10 Tessellation (TSLN)

Tessellation (TSLN), sometimes referred to as fundus mottling, represents a notable ocular trait that merits investigation owing to its distinctive attributes and potential impacts on retinal well-being. Tessellation entails the existence of a geometric pigmentation or mottling pattern on the retinal surface, resulting in a mosaic-like visual aspect. This occurrence can be discerned through retinal imaging and assessments and might originate from diverse underlying factors. The depiction of TSLN visuals is portrayed in Figure 2.10.

The occurrence of tessellation fluctuates based on the population under scrutiny and the specific characteristics of retinal pigmentation [31]. Those with myopia, other retinal dystrophies, or related disorders may exhibit it. The dispersion of pigmentation and the interaction among retinal layers are a couple of instances that could impact the visual manifestation of tessellation.

The clinical significance of tessellation is rooted in its potential correlation with particular retinal disorders and its pertinence in diagnostic evaluations. Tessellation can function as a sign of retinal irregularities or degenerative alterations, particularly in conditions like retinitis pigmentosa or pattern dystrophies [32]. In specific instances, tessellation might even serve as a diagnostic indicator for certain genetic mutations or retinal pathologies.

The implications of tessellation for patients' well-being revolve around its possible connection to underlying retinal conditions and its role in diagnostic assessment. The manifestation of tessellation could trigger additional inquiries to ascertain the existence of any correlated retinal irregularities [33]. Grasping the correlation between tessellation and distinct retinal disorders can steer the implementation of suitable management approaches and interventions.



Figure 2.10 Visualization of TSLN images from the RFMiD 2.0 dataset [3].

2.1.2.11 Normal Retinal

Optical Coherence Tomography (OCT) is a key tool in eye medicine. It helps doctors see the retina in high detail. A normal OCT image of the retina gives insights into its structure and health, vital for spotting eye diseases and tracking eye conditions.

In a typical OCT retina image, you can see different layers and parts. The most noticeable is the retinal nerve fiber layer (RNFL). This is a thin layer of nerve fibers that carries what your eye sees to your brain [34]. It shows up as a light layer on the inside of the retina. Below the RNFL is the ganglion cell layer (GCL). It's darker and contains the cell bodies of nerve cells that transmit visual signals [35]. Further down, you find the inner plexiform layer (IPL) and the inner nuclear layer (INL). The IPL has connections where different retinal cells talk to each other [36]. The INL has the cell bodies of other eye neurons, like bipolar cells and horizontal cells. Deeper still, you'll see the outer plexiform layer (OPL) and the outer nuclear layer (ONL). The OPL is where photoreceptor cells (rods and cones) communicate with other retinal cells. The ONL has the cell bodies of photoreceptors. Its thickness tells us about retinal health [37]. At the bottom, you'll see the external limiting membrane (ELM) and the retinal pigment epithelium (RPE). The ELM separates photoreceptors from the RPE, which supports them. A normal OCT retinal image also shows the fovea, the central part of the retina that's responsible for clear vision.

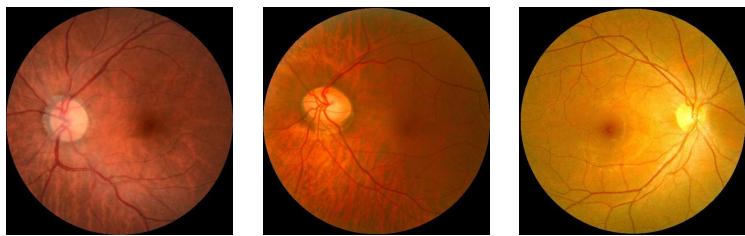


Figure 2.11 Visualization of normal retinal images from the Kaggle dataset [38].

2.1.3 Retinal Disease Treatment

The management of retinal illnesses encompasses a holistic effort involving a diverse array of approaches designed to uphold and restore vision, enhance patients' quality of life, and regulate underlying conditions. Treatment methods for retinal illnesses are contingent on the nature of the disease, its progression stage, and the

individual patient's requirements. This chapter extensively explores a variety of therapeutic choices, spanning from medicinal treatments to surgical interventions, which hold significant roles in the effective control of retinal disorders.

Medical treatments frequently form the foundation of retinal disease management, especially when inflammation, infection, or vascular irregularities are implicated. Pharmacological agents like anti-inflammatory medications, antibiotics, and anti-VEGF (vascular endothelial growth factor) drugs are employed to alleviate inflammation, counteract infection, or regulate aberrant blood vessel growth [39]. These interventions address the fundamental mechanisms underlying retinal diseases, consequently arresting disease advancement and enhancing visual results.

Surgical measures assume a pivotal role in tackling retinal diseases necessitating direct intervention within the ocular sphere. Vitrectomy, a surgical procedure encompassing the elimination of the vitreous humor, finds application in addressing conditions like retinal detachment, epiretinal membranes, and vitreous hemorrhage [40]. Laser therapy, exemplified by photocoagulation, is harnessed to seal leaky blood vessels in ailments like age-related macular degeneration (ARMD), thereby diminishing the likelihood of complications [41].

Technological progress has given rise to groundbreaking treatments for retinal diseases. An example is gene therapy, which strives to counter inherited retinal disorders by delivering operational genes to substitute for or complement mutated ones. Stem cell therapy shows potential for revitalizing compromised retinal tissues and potentially reinstating visual function in degenerative scenarios [42]. Moreover, novel approaches like retinal implants and prosthetics are under exploration to circumvent damaged retinal cells and reestablish vision via direct neural activation.

In some cases, retinal diseases may not have definitive cures, but management strategies focus on maintaining vision and preventing further deterioration. Regular monitoring, lifestyle adjustments, and supportive measures such as low vision aids and rehabilitation programs contribute to enhancing patients' visual capabilities and overall quality of life.

2.2 Deep Learning

DL algorithms are a subset of machine learning and artificial intelligence approaches that use artificial neural networks to understand complicated patterns and relationships in data [43]. Natural language processing (NLP), image or speech recognition, and even medical image analysis, including the earlier diagnosis of retinal degeneration, benefit greatly from these algorithms' propensity for handling complex patterns and enormous amounts of data.

A vast area of computer science called artificial intelligence (AI) is devoted to developing machines that mimic human thinking and perform actions and operations that normally require human understanding [44]. This includes a wide range of skills, such as the ability to comprehend natural language, spot patterns in data, make wise judgements, solve problems, and even continually develop oneself. Rule-based, expert-driven, or algorithmically driven, AI systems strive to imitate human abilities using a variety of techniques and methodologies.

Machine learning (ML) constitutes a subset of AI focused on enabling computers to learn and improve their performance from experience or data [45]. At its core, machine learning is about developing algorithms that allow systems to recognize patterns and relationships within data and then use this learned information to make predictions or decisions. Unlike traditional programming, where explicit rules are provided, machine learning involves training models on data to uncover underlying patterns and make informed predictions. Machine learning includes supervised learning (training with labelled data), unsupervised learning (identifying patterns in unlabeled data), and reinforcement learning (learning through interactions) [46].

Deep learning, a subset of machine learning as shown in Figure 2.11, has developed as an effective way to address complicated problems by utilizing artificial neural networks with numerous layers, sometimes known as "deep" networks. These networks are designed to automatically learn detailed features and representations from raw data, a process known as feature learning or representation learning. Deep neural networks' hierarchical nature enables them to extract and alter data progressively, capturing subtle correlations and patterns within the data.

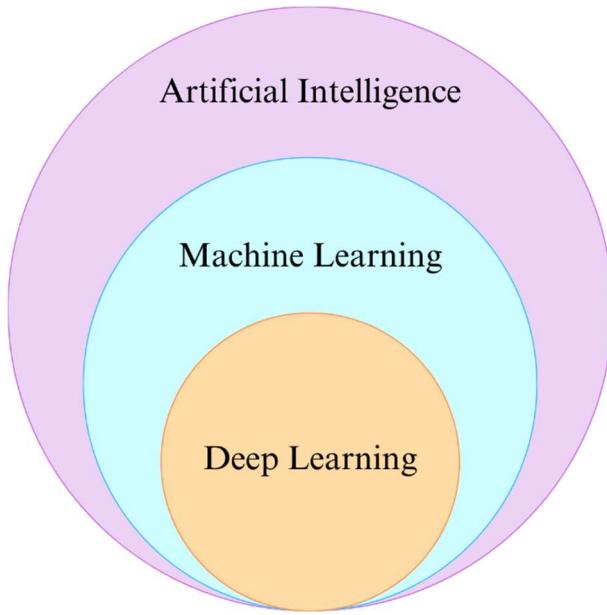


Figure 2.12 Deep learning, a subset of machine learning and artificial intelligence.

2.2.1 Convolutional Neural Network (CNN)

Front and center in the avant-garde advancements of artificial intelligence stand convolutional neural networks (CNNs), notably in image identification and interpretation [47]. These exceptional neural networks have given rise to a completely new understanding of how computers perceive visual data, deeming them crucial for a multitude of tasks, from recognizing objects to creating images. Image recognition, object classification, and photography all rely on the unique capabilities offered by CNNs.

At the heart of Convolutional Neural Networks (CNNs) lies the notion of convolution, which replicates the manner in which the human visual system processes information [48]. The design is influenced by the configuration of neurons in the visual cortex, enabling CNNs to autonomously grasp and derive characteristics from unprocessed pixel data. Within the network, convolutional layers execute filtering actions, detecting edges, textures, forms, and designs within the input images. This progressive strategy empowers the network to gradually acquire more intricate and conceptual attributes as it navigates deeper into the layers [49].

CNNs have demonstrated exceptional efficacy in tackling predicaments that conventional algorithms encountered difficulties with, like image recognition. For

example, they excel at classifying objects in images by mastering the discernment of features, thus enabling differentiation among diverse objects, animals, and even individuals. Additionally, CNNs have played a pivotal role in the realm of medical image analysis, as they diagnose illnesses from X-rays and MRIs. Furthermore, they have facilitated the development of perception systems for autonomous vehicles, enabling automobiles to identify pedestrians, traffic signs, and other vehicles on the road.

An impressive attribute of CNNs is their aptitude to extend acquired features to novel, unobserved data. This accomplishment is facilitated by procedures like pooling and normalization, which aid in diminishing sensitivity to minor discrepancies in data. Furthermore, CNNs commonly integrate fully connected layers in the network's concluding stages, permitting them to formulate forecasts grounded on the acquired features [50]. Figure 2.12 illustrates the architecture of the convolutional neural network.

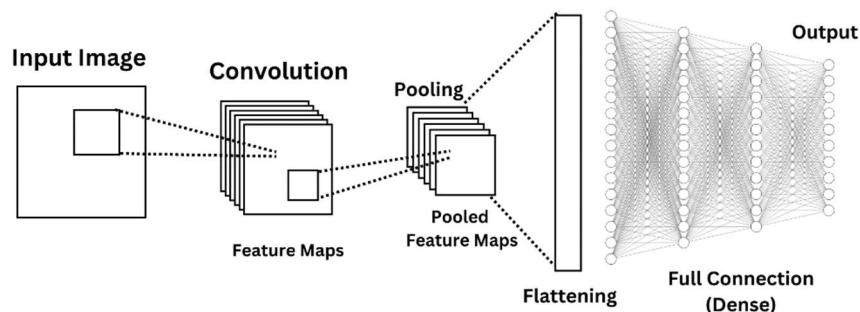


Figure 2.13 Convolutional Neural Network (CNN) Architecture.

Image from the following blog post - <https://www.pycodemates.com/2023/06/introduction-to-convolutional-neural-networks.html>

2.2.2 MobileNetV2 architecture

Convolutional neural networks (CNNs) must be used on devices with limited resources, such as mobile phones and embedded systems. The MobileNetV2 [51] architecture is a noteworthy achievement in the field of deep learning because it was created to address these issues. A development of MobileNetV1, MobileNetV2 aims to strike a compromise between model performance and computational efficiency, making

it the perfect solution for real-time applications with constrained computational resources.

At the core of the MobileNetV2 architecture lies a novel design philosophy that employs inverted residual blocks and linear bottleneck structures [51]. These innovations significantly reduce the computational load of the network while retaining its ability to capture intricate features from input data. Inverted residuals use a lightweight convolutional layer to first expand the number of channels, followed by a depthwise convolution to capture spatial relationships, and finally a pointwise convolution to compress the features back into fewer channels. This approach reduces computational requirements while maintaining feature richness.

MobileNetV2 caters to devices with diverse power capabilities, offering models aligned with their performance requirements. These span from the streamlined MobileNetV2-0.25 to the robust MobileNetV2-1.4. The array of options empowers developers to conveniently tailor the model intricacy to their distinct hardware, a flexibility that is instrumental in accommodating various setups.

MobileNetV2's versatility stands out as a key feature, enabling it to be effective in tackling a variety of computer vision problems, such as image classification, object detection, and semantic segmentation [52]. This versatility, coupled with its efficiency, has fueled its widespread adoption for implementing deep learning models on edge devices. Consequently, this progress has paved the way for the integration of applications like real-time object recognition, augmented reality, and medical diagnostics into mobile platforms.

MobileNetV2 also plays a role in democratizing artificial intelligence by narrowing the divide between robust server-based models and lightweight, on-device models. It facilitates resource-efficient utilization of device capabilities without sacrificing precision, thereby enhancing accessibility to AI-powered applications for a broader user base.

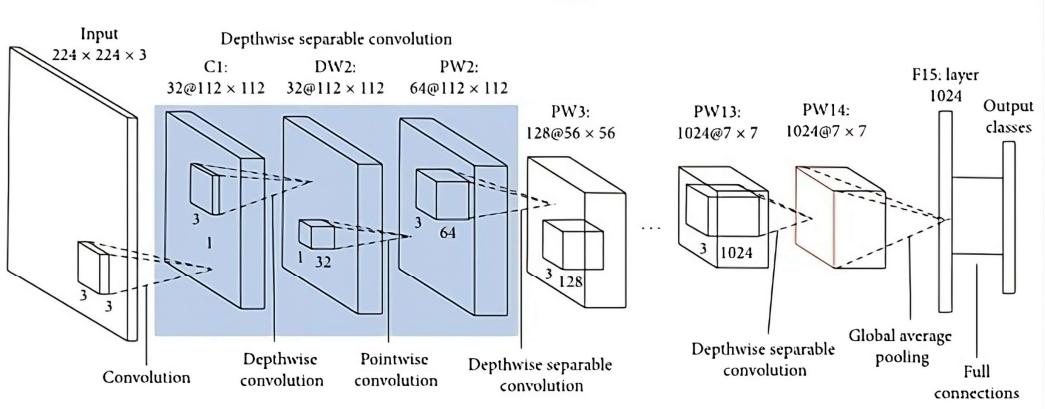


Figure 2.14 MobileNetV2 Architecture [53].

2.2.3 Data Augmentations for Biomedical Images

Data augmentation has emerged as a crucial method in the domain of biomedical image analysis, furnishing a flexible and potent means to elevate the efficacy of deep learning models [54]. Biomedical images, whether sourced from X-rays, MRIs, microscopic scans, or fundus images, frequently suffer from restricted availability due to data collection complexities and privacy considerations. Data augmentation tackles this restraint by creating enhanced renditions of extant images, thus effectively heightening the diversity and depth of the training dataset [55].

Within the domain of biomedical images, data augmentation involves implementing a sequence of modifications to the original images, all while upholding their innate biological attributes. These modifications might entail rotations, translations, flips, zooms, brightness alterations, and the introduction of noise. These diverse alterations emulate real-world situations, bolstering the model's learning process [56]. For instance, rotating an X-ray image simulates shifts in patient orientation, while the inclusion of noise replicates the inherent irregularities in medical imaging apparatus.

A foremost advantage of data augmentation in the context of biomedical images is its capacity to mitigate overfitting. Overfitting occurs when a deep learning model becomes excessively tailored to the training dataset, culminating in subpar generalization on novel, unobserved data. Through the integration of augmented images during training, the model acquires the skill to identify features in varied contexts,

thereby amplifying its aptitude to generalize across diverse patient groups and imaging situations.

Data augmentation also has a profound impact on model performance in the presence of limited training data. In the biomedical domain, where acquiring large datasets can be challenging, data augmentation effectively magnifies the training set without the need for extensive data collection [57]. This empowers deep learning models to capture a broader spectrum of features and nuances, improving their diagnostic accuracy and reliability.

Furthermore, data augmentation promotes the development of models that are more robust to variations and artifacts commonly encountered in biomedical images. It enables models to learn invariant features, unaffected by minor changes in lighting, orientation, or noise [58]. This is particularly crucial in medical diagnostics, where consistent and accurate assessments are paramount.

2.3 Transfer Learning

Pre-trained convolutional neural networks (CNNs) have demonstrated their ability to learn generic visual features from diverse image datasets, such as ImageNet. These features, extracted by lower-level layers of the CNN, capture fundamental shapes, textures, and edges that are relevant across different domains [59]. In retinal disease classification, these pre-trained features can be used as a starting point, sparing the need for models to learn these foundational features from scratch. By fine-tuning the higher-level layers of the CNN on retinal images, models can quickly learn to recognize disease-specific patterns and nuances [60].

Transfer learning significantly reduces the data requirements for training accurate retinal disease classifiers. Medical image datasets, especially those labeled for specific diseases, are often limited in size due to challenges in data collection and annotation [61]. With transfer learning, the pre-trained model's generalizable features are refined using the available medical images, making it possible to achieve impressive results even with smaller datasets.

Moreover, transfer learning enhances the interpretability of retinal disease classifiers. By starting with features learned from generic images, models gain a degree of robustness and adaptability to various imaging conditions, such as lighting

variations, orientations, and noise. This adaptability is essential for real-world applications, where medical images might exhibit diverse characteristics.

2.4 Non-Mydriatic Camera Images

Non-mydriatic camera images have emerged as a groundbreaking asset in the realm of ophthalmology, providing a patient-conducive and streamlined method for obtaining retinal images to facilitate diagnosis and screening. Historically, diagnosing retinal ailments and situations necessitated the application of mydriatic substances to dilate the pupils—an experience that can be discomfiting for patients and resource-intensive for healthcare practitioners. Non-mydriatic cameras, showcased in Figure 2.14, have revolutionized this procedure by enabling the acquisition of top-tier retinal images without the prerequisite of pupil dilation.

The importance of non-mydriatic camera images lies in their potential to optimize the diagnostic process and heighten patient well-being [62]. These cameras utilize sophisticated imaging technologies like infrared and multi-spectral imaging to obtain comprehensive retinal images without necessitating pupil dilation. This eradicates the unease linked with pupil dilation and substantially curtails the time needed for imaging, rendering retinal assessments more convenient and attainable for patients.

Non-mydriatic camera images assume a pivotal role in the timely detection of diseases and proactive healthcare [63]. These images enable ophthalmologists to oversee retinal well-being, discern irregularities, and diagnose ailments such as DR and ARMD. Through enabling routine screenings, non-mydriatic cameras promote early intervention and enhanced management of retinal conditions, thereby potentially averting irreversible vision impairment.

The integration of non-mydriatic camera images with artificial intelligence and deep learning technologies further amplifies their impact. These images serve as valuable training data for deep learning models designed to detect and classify retinal diseases automatically. By combining the efficiency of non-mydriatic imaging with the power of AI, clinicians can achieve rapid, accurate, and cost-effective diagnoses, particularly in resource-constrained settings where access to specialized care may be limited.



Figure 2.15 Non-mydriatic Fundus Camera [64].

2.5 Optical Coherence Tomography (OCT) Images

Optical Coherence Tomography (OCT) images stand as a notable innovation in medical imaging, especially within the field of ophthalmology. These images furnish intricate cross-sectional perspectives of diverse biological tissues, serving as indispensable resources for diagnosis and ongoing surveillance [65]. Among its primary uses, OCT excels at capturing images of the retina, making it a pivotal tool in the detection and monitoring of retinal ailments like macular degeneration, diabetic retinopathy, and glaucoma.

What sets OCT apart is its outstanding spatial resolution, which empowers it to unveil intricate microstructures within biological tissues. In the context of retinal OCT, it grants the ability to visualize discrete layers within the retina, encompassing the nerve fiber layer, ganglion cell layer, and photoreceptor layer, each holding a distinct function in visual processes. This heightened resolution equips clinicians with the capability to discern subtle alterations or irregularities within these layers, thus facilitating early disease diagnosis and ongoing monitoring [66].

Additionally, OCT distinguishes itself with its real-time imaging proficiency. It not only records stationary structural data but also offers insights into dynamic processes occurring within biological tissues. For example, it can depict the flow of blood within retinal vessels or track alterations in tissue thickness as they evolve over time. This real-

time feature of OCT holds immense value for the continuous monitoring of disease progression and the assessment of treatment efficacy [67].

Moreover, OCT offers a non-invasive imaging method, prioritizing patient safety and comfort. It obviates the need for surgical interventions or the utilization of contrast agents, thereby minimizing the inherent risks linked with invasive diagnostic procedures. This non-invasive nature positions it as an ideal option for regular check-ups and the extended monitoring of chronic conditions. Additionally, OCT yields quantitative data, allowing healthcare professionals to conduct meticulous measurements. Whether it pertains to determining retinal thickness, evaluating blood vessel diameters, or quantifying alterations in tissue structure, these quantitative measurements hold paramount importance in objective diagnosis and the monitoring of disease progression [68].

Furthermore, OCT images are often presented with color-coding and layer segmentation, enhancing their interpretability [69]. Color-coding can highlight specific features or regions of interest, such as blood vessels depicted in red, while layer segmentation automatically separates different tissue layers for easier analysis.

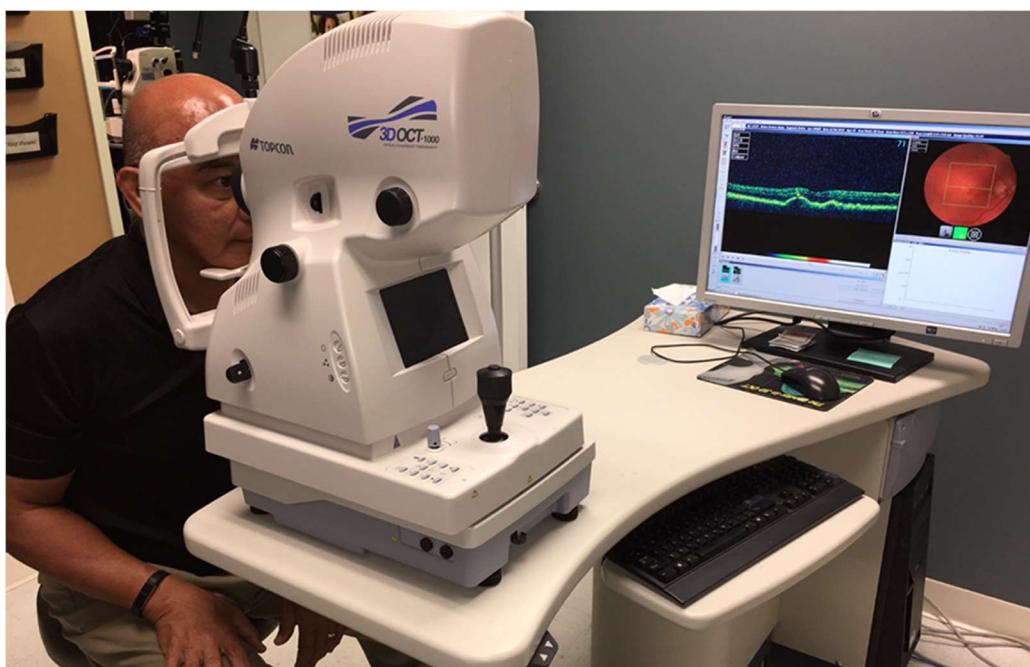


Figure 2.16 Optical Coherence Tomography (OCT) Workstation.

Source: <https://www.island-retina.com/diagnostics/spectral-domain-optical-coherence-tomography>

2.6 Data Cleaning

Addressing the issue of noisy labels has gained significant importance in the domain of deep learning, especially when working with datasets that encompass inaccuracies or errors in label assignments within the training data. The presence of noisy labels can exert considerable influence on model performance and dependability, resulting in diminished accuracy, inadequate generalization, and compromised decision-making capabilities. Dealing adeptly with noisy labels is a crucial requirement to uphold the resilience and credibility of deep learning models.

The challenge of noisy labels arises from various sources, including human annotation errors, ambiguities in labeling criteria, and even deliberate labeling inaccuracies [70]. These erroneous labels can mislead the learning process, causing the model to learn incorrect patterns and relationships from the data. Consequently, the model's performance on unseen data deteriorates, undermining its practical utility.

Numerous techniques have been developed to mitigate the influence of noisy labels and enhance the model's learning procedure. Data augmentation stands out as a widely adopted strategy where the training dataset is enriched with modified renditions of the original data. This augmentation expands the training pool and cultivates the model's capacity to withstand label noise. Another advantageous technique is the Local Outlier Factor (LOF) method, which assesses the uniqueness of data points within their localized neighborhoods [71]. Analogous to the Isolation Forest method, LOF can pinpoint instances harboring potentially erroneous labels and assign them outlier scores, effectively segregating them from the majority of accurately labeled data points. By harnessing LOF, the learning process gains resilience and a heightened focus on precisely labeled instances.

2.6.1 Isolation Forest

The Isolation Forest method has emerged as a potent and inventive solution for tackling the issue of noisy labels in machine learning datasets [72]. Initially developed for anomaly detection, this approach has been adapted to manage noisy labels by isolating instances with potential inaccuracies in their labels and minimizing their influence on the learning process. The Isolation Forest method capitalizes on the

inherent characteristics of anomalies or noisy instances to effectively differentiate them from the predominant, accurately labeled data points.

The foundational concept of the isolation forest method revolves around the creation of an isolation tree forest. Each individual tree is crafted by iteratively dividing the data space into subsets, with the objective of isolating instances bearing noisy labels. These instances necessitate fewer partitions to segregate them from the larger group of correctly labeled samples. In essence, instances marked by noisy labels are regarded as anomalies owing to their distinctiveness within the feature space.

Throughout the training phase, the Isolation Forest method attributes a score to each instance, denoting its degree of isolation within the forest. Instances boasting higher isolation scores are deemed potential anomalies or noisy labels [73]. By detecting these instances and handling them distinctively within the learning process, the method heightens the model's robustness against label noise.

An inherent strength of the Isolation Forest method lies in its capacity to manage noisy labels without necessitating extensive alterations to the model's architecture or loss functions [74]. Rather than concentrating on the noisy instances directly, the method centers its efforts on isolating these instances from the remainder of the data. This approach permits the learning process to proceed with diminished disruption stemming from label noise.

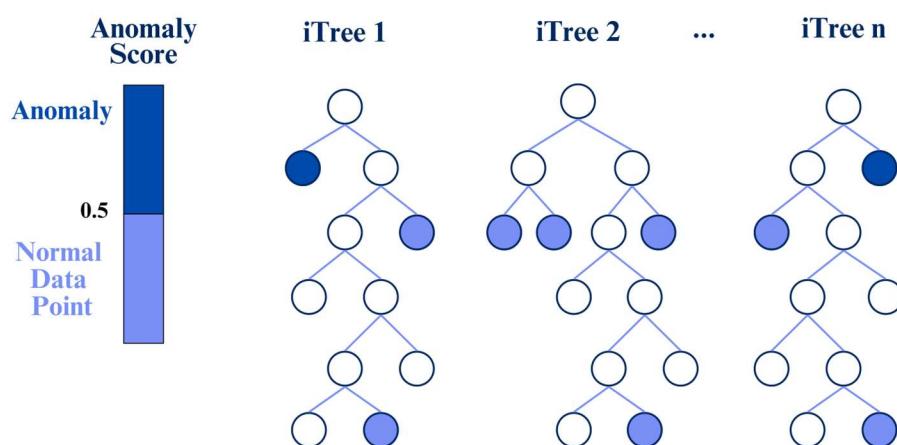


Figure 2.17 Isolation Forest Method. Source: <https://lekh-a-bhan88.medium.com/anomaly-detection-using-isolation-forest-and-local-outlier-factor-158985d6a45>

2.6.2 Local Outlier Factor (LOF)

The LOF method stands as a resilient and enlightening approach in the domain of outlier detection, presenting a data-centric method for pinpointing anomalies within datasets. Outliers, characterized by their substantial deviation from the bulk of data, hold pivotal importance in diverse contexts such as fraud detection, network security, and quality control. LOF offers a sophisticated viewpoint by evaluating not solely the distance of a data point from its neighboring points but also its density within its localized region [75].

LOF functions on the principle that outliers stand out due to their uniqueness within local neighborhoods. Conventional approaches frequently hinge on global statistics, which may not precisely encompass the nuances of data distribution. LOF surmounts this constraint by evaluating the level of isolation or extraordinariness exhibited by a data point concerning its immediate vicinity. This is achieved by gauging the ratio between the average density of a data point's neighbors and its own density.

The LOF algorithm [75] allocates a score to each data point, signifying its extent of deviation from the norm. Data points with elevated LOF scores are regarded as more probable outliers, whereas those with lower scores adhere more closely to the standard pattern. LOF's effectiveness stems from its aptitude to manage intricate data distributions and to detect outliers across diverse densities and shapes, areas where conventional distance-based methods could fall short.

While LOF stands as an effective technique, it does encounter challenges, especially when handling high-dimensional data and diverse densities. Nonetheless, it persists as a potent instrument for detecting outliers and supplements other methods, contributing to a more holistic understanding of data quality and reliability.

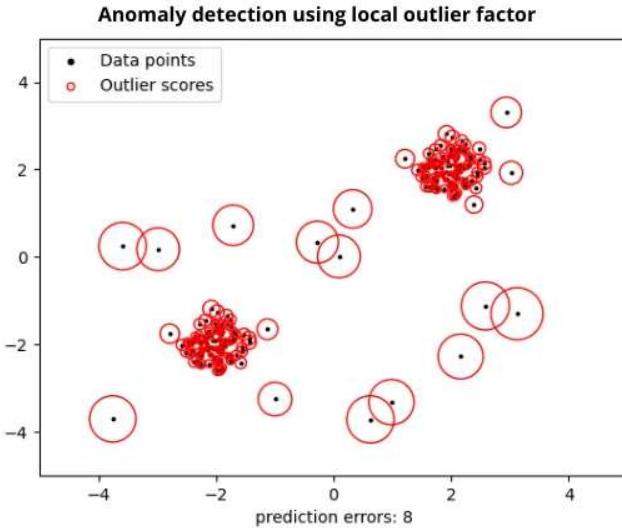


Figure 2.18 Local Outlier Factor (LOF) Method. Source: <https://lekha-bhan88.medium.com/anomaly-detection-using-isolation-forest-and-local-outlier-factor-158985d6a45>

2.7 Cross Validation

Cross-validation serves as a crucial method within deep learning and model assessment, addressing the task of evaluating a model's performance and its capability to generalize to unfamiliar data [76]. Given that models are trained and fine-tuned on a specific dataset, there's a potential danger of overfitting, wherein the model excels solely on the training data but struggles with generalizing to fresh data. Cross-validation counters this risk by methodically segmenting the dataset into multiple subsets, termed "folds," and repeatedly training and appraising the model on diverse fold combinations.

K-fold cross-validation is the most often used type of cross-validation. The dataset is separated into k equally sized folds using this technique. The model is trained on $k-1$ folds and then tested on the last fold [77]. This operation is done k times, with each fold acting as the validation set. Each iteration's performance measures are averaged to provide a full assessment of the model's performance.

Cross-validation assists in anticipating the model's performance on novel, previously unencountered data. It reveals both the model's bias and variance, and whether it's susceptible to overfitting or underfitting. Additionally, cross-validation contributes to the fine-tuning of hyperparameters since it delivers a more reliable estimation of the model's performance compared to evaluating it solely on a single validation set.

2.7.1 K-Fold Cross-Validation

One of its noteworthy advantages is its ability to decrease the fluctuation in performance assessment that could occur when using only a single validation set. This approach guarantees that the model is appraised using diverse samples by sequentially employing different data subsets as validation sets. This practice mitigates the influence of data variability on performance evaluations.

When dealing with a restricted dataset size, K-Fold Cross-Validation [78] proves valuable. It maximizes the utilization of the available data by involving each sample in both training and validation during multiple iterations. This becomes particularly significant when evaluating models that rely on a substantial validation set. The visualization of five-fold cross-validation is shown in Figure 2.17.

K-Fold Cross-Validation additionally aids in hyperparameter optimization. The technique offers a more accurate evaluation of a model's performance across diverse parameter configurations by doing numerous rounds of training and validation. Choosing the appropriate set of hyperparameters to provide the best generalization is made easier by doing this.

Although K-Fold Cross-Validation is a powerful method, it's worth noting that it comes with computational costs, as it requires training the model K times. However, its benefits in estimating model performance and preventing overfitting make it an essential tool for model evaluation and selection, ensuring that machine learning models are well-equipped to generalize to new and unseen data.

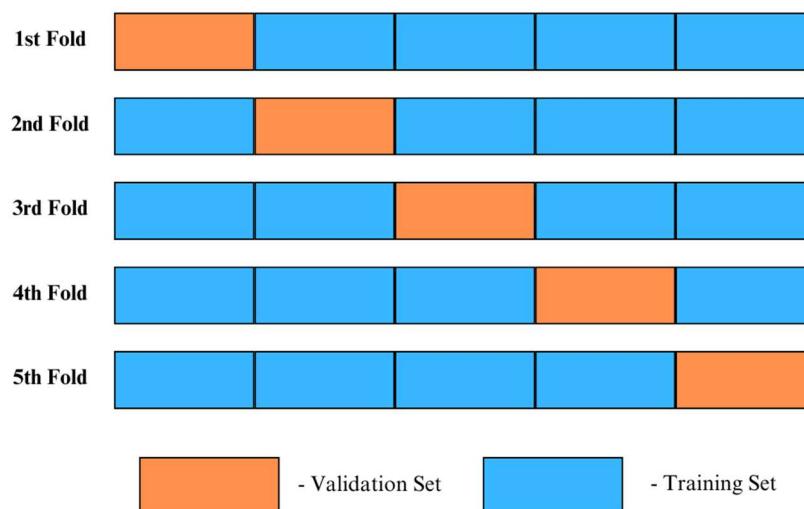


Figure 2.19 Visualization of Five-Fold Cross Validation.

2.8 Summary

In Chapter 1, the background and significance of the study, objectives, scope of the study, and report outline are proposed.

In this chapter, the state-of-the-art was categorized into retinal diseases (section 2.1), including the subsections of retinal disease screening (section 2.1.1), retinal diseases class breakdown (section 2.1.2), and retinal disease treatment (section 2.1.2). Observations on deep learning were interpreted (section 2.2). Transfer learning (section 2.3) is also explained, and the relevance of the state-of-the-art to non-mydriatic camera images (section 2.4), optical coherence tomography (OCT) images (section 2.5), data cleaning (section 2.6), cross-validation (section 2.7), and section 2.8 were summarized throughout the chapter.

The next chapter presents the original dataset, materials used for the experiment, methods, including evaluation criteria equations, k-fold cross-validation, and proposed methodology.

CHAPTER 3

MATERIALS AND METHODS

3.1 Introduction

This chapter describes the materials and methods. Firstly, the chapter describes the project's original dataset (section 3.2) and materials (section 3.3). The proposed methodology is then discussed (section 3.4). Creating a mobile application using Android Studio will be clarified (section 3.5). Lastly, Section 3.6 wraps up the entire chapter.

3.2 Dataset

Working with a Multilabel Dataset (MLD) can be tricky due to imbalances. This means that some symptoms or labels in the dataset are not evenly distributed. Tackling this issue in MLDs is a challenge, and the usual methods don't work well. Imbalanced datasets are a common problem in many real-world fields, like biomedical engineering. The RFMiD2.0 dataset complements our previous RFMiD dataset. RFMiD2.0 contains images of 49 different eye diseases, some common and some rare. These diseases have been identified by eye specialists and are part of the dataset. Figure 3.1 shows the steps involved in creating this database [3].

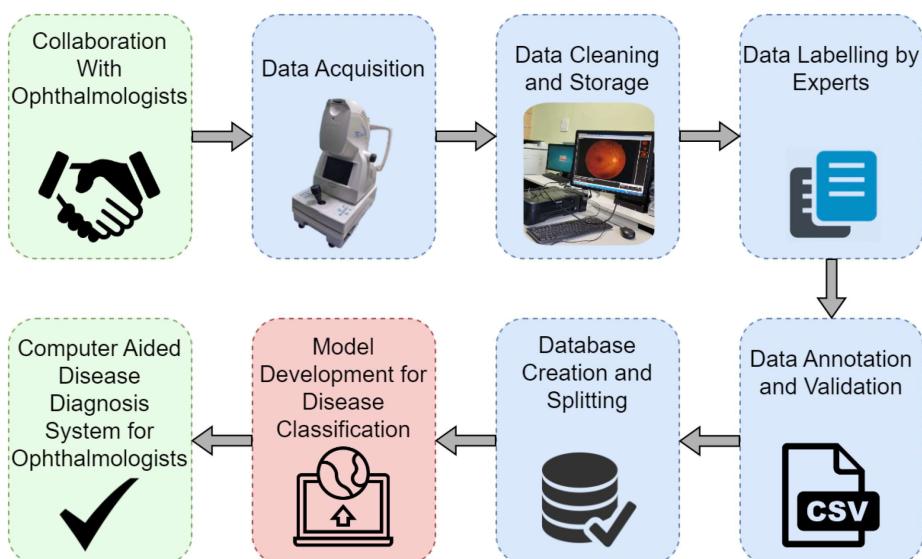


Figure 3.1 The procedure for establishing a dataset while obtaining consent from ophthalmologists for the development of a Computer-Aided Disease Diagnosis (CADD) System [3].

To create this dataset, the researchers collaborated with three eye specialists. These specialists collected retinal fundus images from the subjects, and the stored data was then handed over to the researchers. Afterward, the collected information underwent a cleaning process to remove any faulty images. The doctors labeled and annotated the remaining images. Ensuring accuracy was a top priority for the eye professionals. Once the labeled images were cross-verified, the dataset was split into training, validation, and testing. This dataset is a valuable resource for researchers worldwide aiming to develop disease classification models. Notably, the most significant applications of AI in research currently revolve around evaluation and diagnosis. Deep learning algorithms have made significant strides in automating medical image classification, achieving accuracy levels on par with human professionals. You can find the dataset's specifications in Table 3.1 [3]. The picture of the TRC-NW300 non-mydriatic retinal camera that is used for capturing retinal images is shown in Figure 3.2.

Table 3.1 Specifications of dataset [3].

Subject Area	The subject area pertains to Ophthalmology and Biomedical Data, specifically focusing on the classification of multiple diseases in retinal fundus images.
Type of Data	The data includes both image files in JPEG and PNG formats as well as CSV files for labeling and annotation.
Data Acquisition	The data was obtained using the TOPCON TRC-NW300 imaging system, with most patients undergoing mydriasis through the application of a 0.5% concentration of tropicamide. Some subjects, however, did not undergo mydriasis.
Experimental Features	The fundus images were captured with patients sitting upright, maintaining a working distance of 40.7 mm (TOPCON TRC-NW300) and 42 mm (CARL ZEISS FF450) between the lenses and the examined eye. This imaging was performed using a non-invasive fundus camera.
Data Source Location	The data was collected at the State of Art Eye Care Hospital, Shri Ganpati Netralaya, situated in Jalna, Maharashtra, India. Additionally, data processing and research were conducted at the Center of Excellence in Signal and Image Processing, SGGS Institute of Engineering and Technology, located in Nanded, Maharashtra, India.



Figure 3.2 TRC-NW300 Non-Mydriatic Retinal Camera.

Source: <https://www.beye.com/product/trc-nw300-non-mydriatic-retinal-camera>

3.3 Materials

In this study, the main emphasis was placed on coding, its execution, and the launch of a mobile app. A variety of tools and technologies were used, as mentioned below:

3.3.1 Integrated Development Environment (IDE)

To develop and execute code, Jupyter Notebook (version 7.0.2) was employed as our integrated development environment (IDE). Jupyter Notebook is widely favored by data scientists and researchers due to its interactive features and compatibility with multiple programming languages.

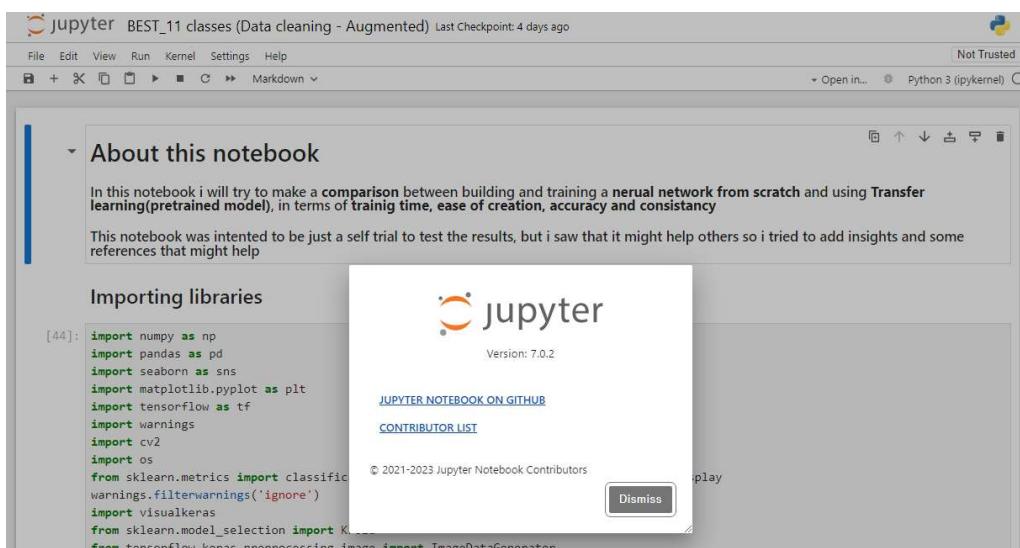


Figure 3.3 Jupyter Notebook Version 7.0.2 Interface.

3.3.2 Deep Learning Frameworks

Version 2.10.0 of the Keras library, along with TensorFlow (version 2.10.0), was employed for the implementation and training of deep learning models. Keras, recognized as a high-level neural network API that operates on top of TensorFlow, was used to facilitate the construction and experimentation with deep learning models. The logos of both Keras and TensorFlow are shown in Figure 3.3.

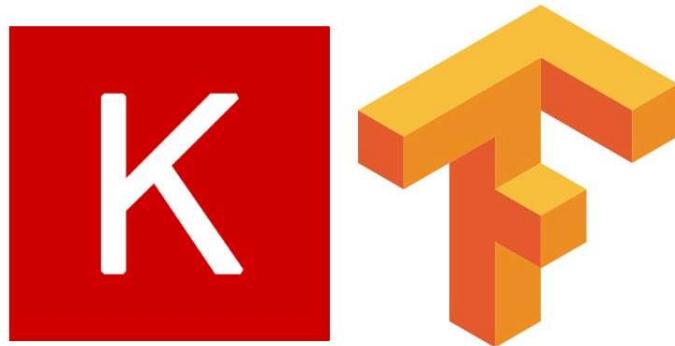


Figure 3.4 Keras (Left) and TensorFlow (Right) Logo. Source: <https://medium.com/ai%C2%B3-theory-practice-business/tensorflow-1-0-vs-2-0-part-3-tf-keras-ea403bd752c0>

3.3.3 Hardware Specifications

The experiments were conducted on a computer with the following hardware specifications:

- CPU: 12th Gen Intel(R) Core(TM) i3-12100F 3.30 GHz. The CPU plays a crucial role in model training and inference.
- RAM: 16GB. Sufficient RAM is essential for handling large datasets and complex models.
- GPU: NVIDIA GeForce GTX 1050Ti. A dedicated GPU accelerates deep learning tasks, especially training of neural networks.

3.3.4 Programming Language

The programming language utilized for coding the research project was Python 3.8.0. Python is favored within the machine learning and deep learning domains due to its rich libraries and strong community support.

3.3.5 Mobile Application Development

For the construction of the mobile application section of the research, the Huawei P20 mobile phone (Figure 3.5) was targeted, running Android version 10. This selection ensured compatibility and testing on a real-world mobile device.

Android Studio (version 2022.1.1 Patch 2) was used as the development environment for the creation of the Android application. Android Studio is the official integrated development environment for Android app development and offers tools for design, coding, testing, and deployment of Android apps. Android Studio interface is shown in Figure 3.6.



Figure 3.5 Huawei P20 Mobile Phone. Source: <https://www.notebookcheck.net/Huawei-P20.295626.0.html>

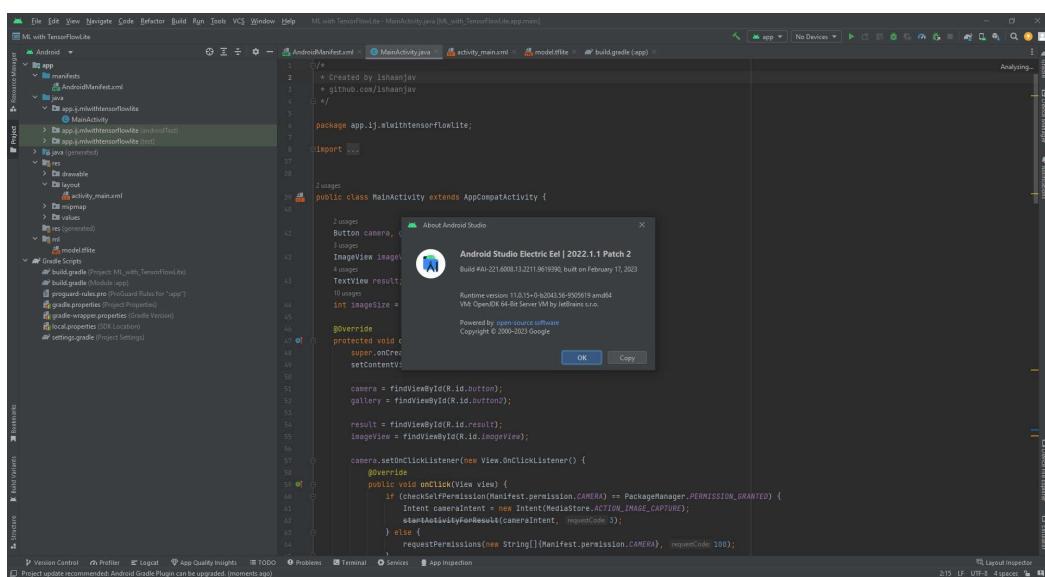


Figure 3.6 Android Studio Interface.

3.4 Proposed Methodology

For the proposed methodology, there are six different approaches that were tested experimentally to find the best testing accuracy. The below approaches were tested on the same dataset and the same hardware and software.

3.4.1 Parameter Settings

The MobileNetV2 architecture was employed in this work to train the image classification model. The input size for the training dataset was 224x224 pixel pictures. A batch size of 64 was employed to regulate how many samples were handled during each training cycle. The Adamax optimizer was used for weight optimization, and the sparse categorical cross-entropy loss function, which is commonly used for multi-class classification problems, was chosen. Softmax, a function suitable for multi-class classification, was chosen as the output layer activation function. The training technique, which was run across 100 epochs with a learning rate of 0.001, determined the step size for updating the model's weights during optimization. Table 3.2 lists these factors.

Table 3.2 Parameter Settings.

Parameters	Settings
Input size	224
Batch size	64
Activation	Softmax
Optimizer	Adamax
Loss	Sparse categorical cross-entropy
Epoch	100
Learning rate	0.001

3.4.2 Training with Unbalanced Dataset

The first method includes an examination of how a MobileNetV2 model is being trained using an imbalanced dataset. Pictures resized to a standard size of 224x224 pixels are employed as input data. The dataset was split into a training set of 80%, a validation set of 20%, and another 10% for the final testing set. A 10-fold cross-validation method is used to further increase the model's performance before training it

using those parameter settings in Table 3.2. This first method is shown in the visualization in Figure 3.7 below.

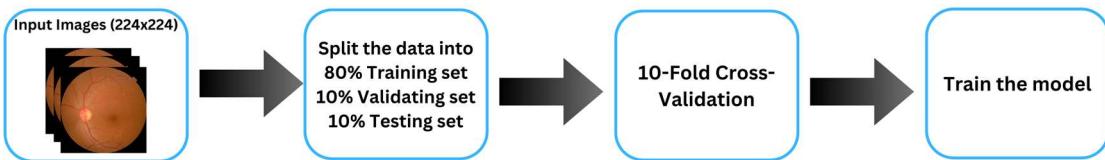


Figure 3.7 Training with Unbalanced Dataset.

3.4.3 Training with Unbalanced Dataset using Different Class Weight

The second method includes an examination of how a MobileNetV2 model is being trained using an imbalanced dataset. Pictures resized to a standard size of 224x224 pixels are employed as input data. The dataset was split into a training set of 80%, a validation set of 20%, and another 10% for the final testing set. A 10-fold cross-validation method is used to further increase the model's performance before training it using those parameter settings in Table 3.2, but this approach used a different class weight as an experiment. This second method is shown in the visualization in Figure 3.8 below.

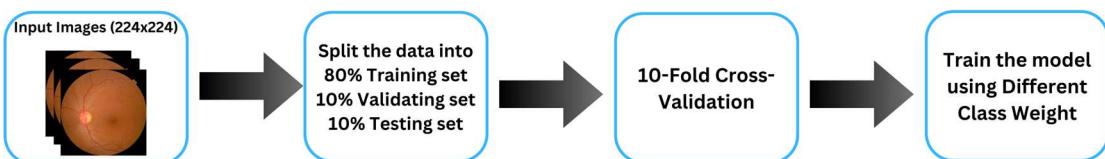


Figure 3.8 Training with Unbalanced Dataset using Different Class Weight.

3.4.4 Training with Unbalanced Dataset and using Data Cleaning

The third method includes an examination of how a MobileNetV2 model is being trained using an imbalanced dataset. Pictures resized to a standard size of 224x224 pixels are employed as input data. Before splitting the data, the data cleaning technique is applied to the Tessellation (TSLN) class first, and then the dataset is split into a training set of 80%, a validation set of 20%, and another 10% for the final testing set. A 10-fold cross-validation method is used to further increase the model's performance before training it using those parameter settings in Table 3.2. This third method is shown in the visualization in Figure 3.9 below.

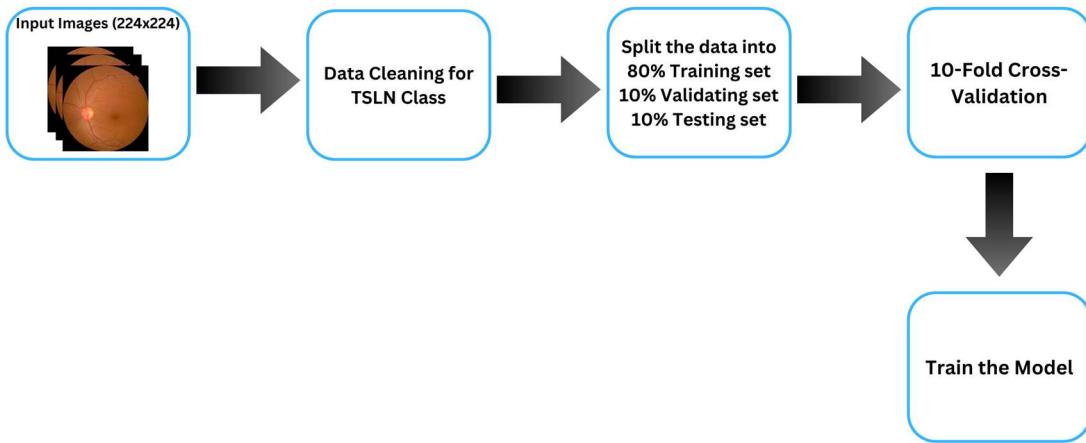


Figure 3.9 Training with Unbalanced and using Data Cleaning.

3.4.5 Training with Data Augmentation Only

The fourth method includes an examination of how a MobileNetV2 model is being trained using an imbalanced dataset. Pictures resized to a standard size of 224x224 pixels are employed as input data. First, the dataset is split into a training set of 80%, a validation set of 20%, and another 10% for the final testing set. Then data augmentation is applied to the training set. Next, a 10-fold cross-validation method is used to further increase the model's performance before training it using those parameter settings in Table 3.2. This fourth method is shown in the visualization in Figure 3.10 below.

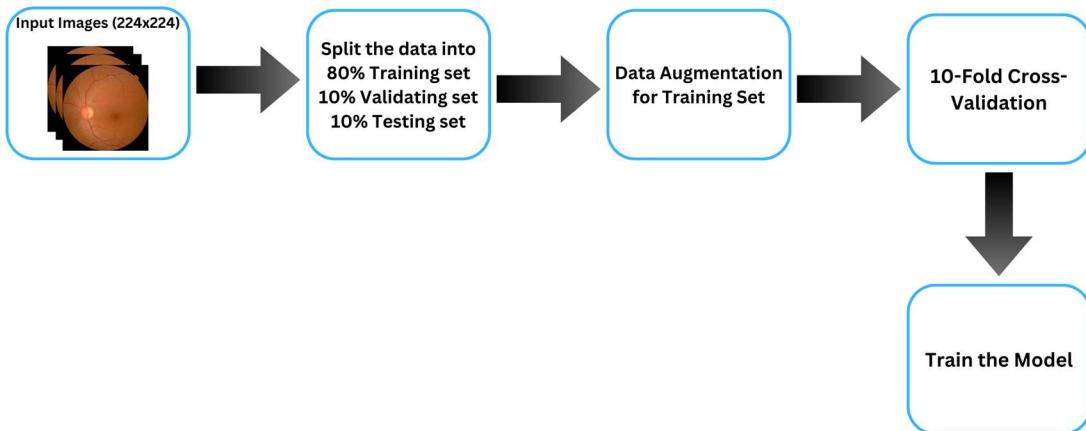


Figure 3.10 Training with Data Augmentation Only.

3.4.6 Training with Data Augmentation before using Data Cleaning

The fifth method includes an examination of how a MobileNetV2 model is being trained using an imbalanced dataset. Pictures resized to a standard size of 224x224 pixels are employed as input data. First, the dataset is split into a training set of 80%, a validation set of 20%, and another 10% for the final testing set. Then data augmentation is applied to the training set. Before the process of cross-validation, the data cleaning technique is used for the Tessellation (TSLN) class. Next, a 10-fold cross-validation method is used to further increase the model's performance before training it using those parameter settings in Table 3.2. This fifth method is shown in the visualization in Figure 3.11 below.

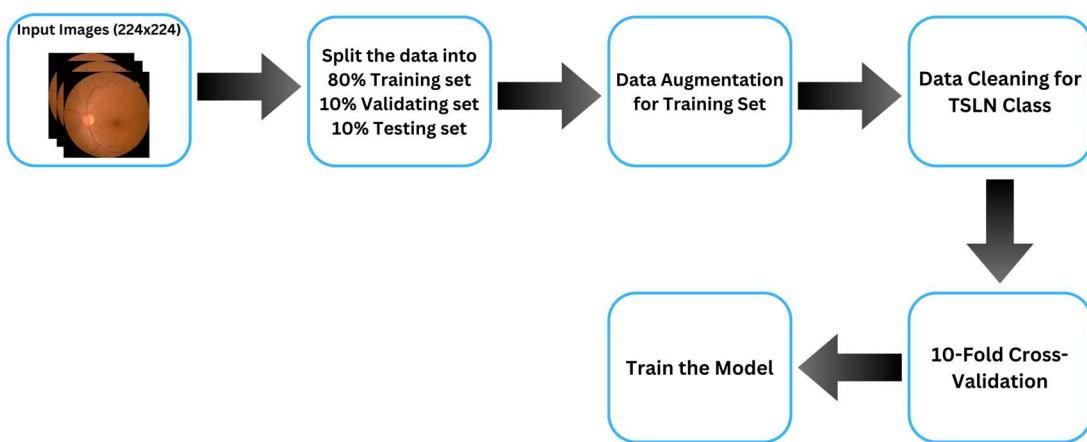


Figure 3.11 Training with Data Augmentation before using Data Cleaning.

3.4.7 Training with Data Cleaning before using Data Augmentation

The last method includes an examination of how a MobileNetV2 model is being trained using an imbalanced dataset. Pictures resized to a standard size of 224x224 pixels are employed as input data. First, the data cleaning technique is used for the Tessellation (TSLN) class and then the dataset is split into a training set of 80%, a validation set of 20%, and another 10% for the final testing set. Then data augmentation is applied to the training set. Next, a 10-fold cross-validation method is used to further increase the model's performance before training it using those parameter settings in Table 3.2. This last method is shown in the visualization in Figure 3.12 below.

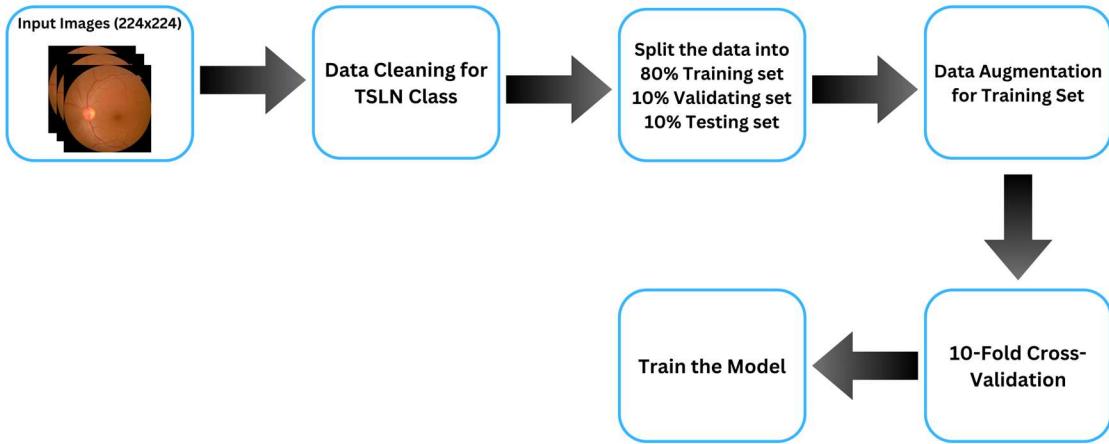


Figure 3.12 Training with Data Cleaning before using Data Augmentation.

3.5 Creating a Mobile Application using Android Studio

Users may load test images into the program, which will then classify them using the trained model. The model's accurate prediction of the correct class for the test image resulted in a high prediction score. This successful integration of the trained model into the Android application demonstrates the model's capacity for real-world deployment in a practical environment, as well as its practical applicability and usability for real-time picture classification tasks on mobile devices.

3.5.1 Design Application Interface

A mobile application has been created using Android Studio, and within the app's interface, two buttons have been included: "Take Picture" and "Launch Gallery." After the user inputs an image, it is processed, and the predicted class name, along with the prediction score, is displayed, with the user-input image showcased at the top of the screen. This application offers users a straightforward means of obtaining predictions for their input images, thereby enhancing their overall interaction and engagement with the app. This design is shown in Figure 3.13 below.

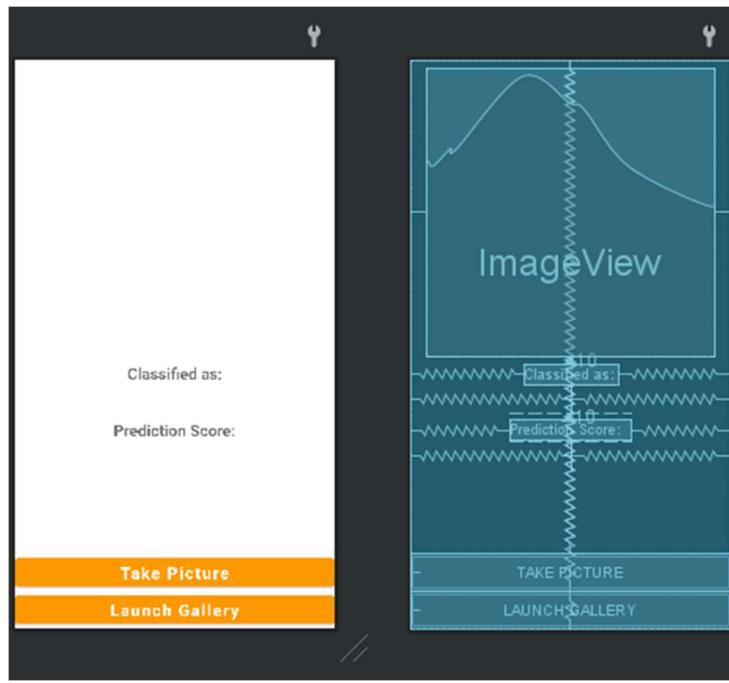


Figure 3.13 Mobile Application Design on Android Studio.

3.6 Summary

This chapter describes the high-level requirements and design of the system that was used to do this research. The chapter started by describing how the dataset was gathered. The material was then discussed in section 3.3, followed by the proposed methodology in section 3.4 and creating a mobile application using Android Studio (section 3.5).

Experimental results are covered in further detail in Chapter 4 which describes the accuracy of training, validating, and testing. Additionally, training loss and validation loss will be investigated in the next chapter.

CHAPTER 4

EXPERIMENTAL RESULT

4.1 Introduction

The fourth chapter describes the experimental results from the last chapter. It will show how the experiments were tested and show the outcomes, including evaluation metrics of accuracy, precision, recall, and F1-score to evaluate. First in Section 4.2, show the evaluation criteria equations, and then in Section 4.3, show the results of the proposed methodologies. After that, Section 4.4 compares all the different methods that were tested to figure out which one is best. Lastly, the results of the experiments are summarized in Section 4.5.

4.2 Evaluation Criteria Equations

In this study, the model's predictions were evaluated using a method that took into account the ratio of accuracy on the training data to the probability gained from the test data. Standard measures such as accuracy, precision, recall, and F1-score were used to assess the performance of the classification job. These measures provide useful information about the model's overall accuracy, ability to minimise false positives, ability to capture real positives, and the trade-off between precision and recall. These commonly used deep learning evaluation metrics give a quantifiable assessment of the model's predicted accuracy and test its performance against expected outcomes or ground truth. Equations (1-4) assess accuracy, precision, recall, and F1-score [79]:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (2)$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (3)$$

$$F1 - score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

where TP stands for true positive, TN for true negative, FP for false positive (Type I error), and FN for false negative (Type II error).

4.3 Results of Proposed Methodology

This chapter visualizes the results of the proposed methodology. In sections 4.3.1 to 4.3.6, a series of experiments were described that demonstrated the best outcome.

4.3.1 Result of Training with Unbalanced Dataset

Retinal images applied in this study were collected from the Retinal Fundus Multi-Disease Image Dataset (RFMiD) 2.0 [3], which mostly comes from Indian patients. The image dataset was divided into three categories, with 80% of the images used for model training, 10% for model validation, and the remaining 10% for model testing. The images were categorized into eleven classes: ARMD with 136 images (4.2%), BRVO with 96 images (2.9%), DN with 181 images (5.6%), DR with 478 images (14.7%), MH with 401 images (12.3%), MYA with 133 images (4.1%), ODC with 347 images (10.7%), ODE with 78 images (2.4%), ODP with 90 images (2.8%), TSLN with 243 images (7.5%), and normal with 1,074 images (33%). As seen in Figure 4.1, the dataset is unbalanced, and the majority class is normal.

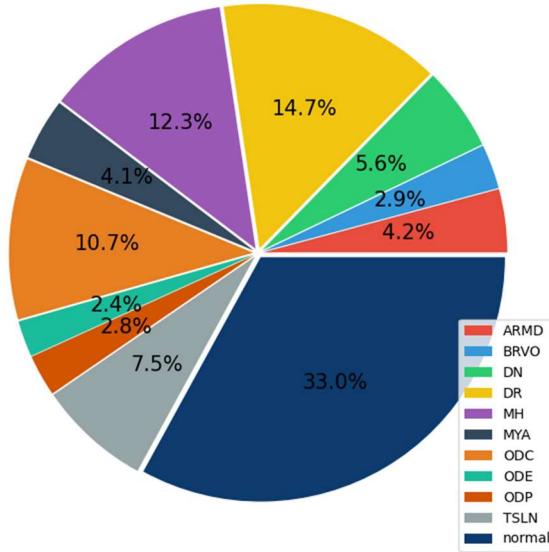


Figure 4.1 The percentage of the images in each class for Training with Unbalanced Dataset Method.

After training with the unbalanced dataset method and using the parameter settings mentioned in Table 3.2, the best epoch was determined to be epoch 51 after 71

epochs of training a model with early stopping enabled. 83.57% training accuracy, 85.94% validation accuracy, and 82.81% testing accuracy were attained by the model. The training loss was at 0.4968 and the validating loss was at 0.4957. Training and Validation Accuracy vs. Epoch and Training and Validation Loss vs. Epoch of MobileNetV2 for this approach are shown in Figure 4.2 and Figure 4.3 respectively. The confusion matrix of this approach is shown in Table 4.1, and the classification report is shown in Table 4.2.

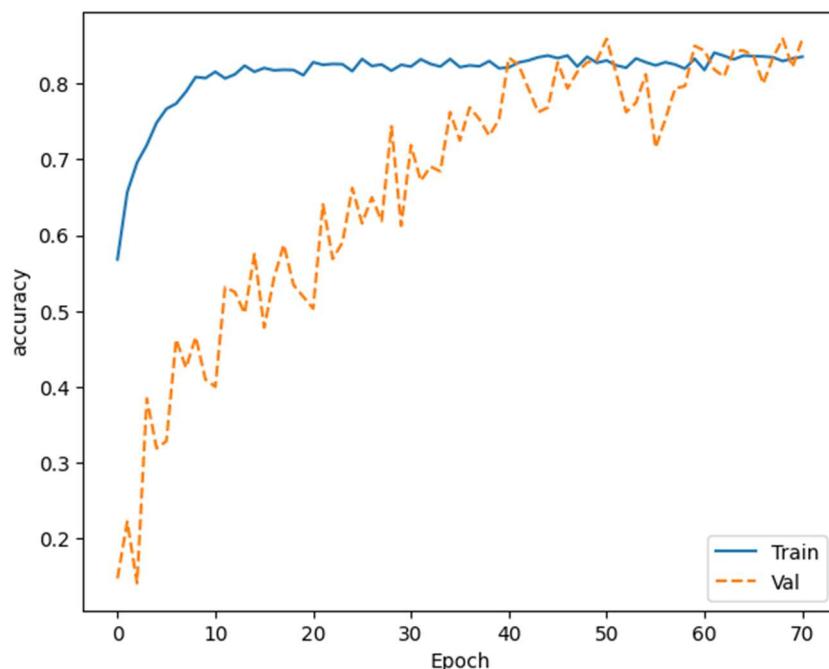


Figure 4.2 Training and Validation Accuracy vs. Epoch using Training with Unbalanced Dataset Method.

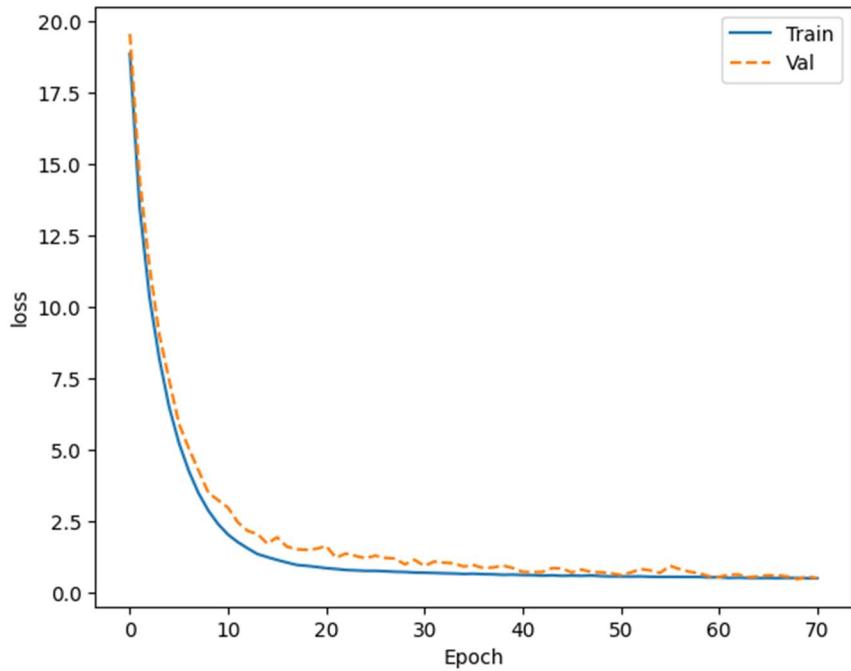


Figure 4.3 Training and Validation Loss vs. Epoch using Training with Unbalanced Dataset Method.

Table 4.1 Confusion Matrix of MobileNetV2 using Training with Unbalanced Dataset Method.

Class	<i>ARMD</i>	<i>BRVO</i>	<i>DN</i>	<i>DR</i>	<i>MH</i>	<i>MYA</i>	<i>ODC</i>	<i>ODE</i>	<i>ODP</i>	<i>TSLN</i>	<i>Normal</i>
<i>ARMD</i>	3	0	0	0	0	2	3	0	0	2	0
<i>BRVO</i>	0	6	0	0	1	0	1	0	0	1	0
<i>DN</i>	0	0	7	0	0	0	5	0	0	2	0
<i>DR</i>	0	0	0	25	3	0	9	0	0	3	0
<i>MH</i>	0	0	0	0	33	0	3	0	0	2	0
<i>MYA</i>	0	0	0	0	2	7	1	0	0	0	0
<i>ODC</i>	0	0	0	0	0	0	41	0	0	5	0
<i>ODE</i>	0	0	0	0	0	0	0	9	0	0	0
<i>ODP</i>	0	0	0	0	0	0	0	0	5	1	0
<i>TSLN</i>	0	0	1	0	4	0	4	0	0	15	0
<i>Normal</i>	0	0	0	0	0	0	0	0	0	0	114

Table 4.2 Classification Report on The Testing Set using Training with Unbalanced Dataset Method.

Class	Precision	Recall	F1 - score
ARMD	1.00	0.30	0.46
BRVO	1.00	0.67	0.80
DN	0.88	0.50	0.64
DR	1.00	0.62	0.77
MH	0.77	0.87	0.81
MYA	0.78	0.70	0.74
ODC	0.61	0.89	0.73
ODE	1.00	1.00	1.00
ODP	1.00	0.83	0.91
TSLN	0.48	0.62	0.55
Normal	1.00	1.00	1.00
<i>macro avg</i>	<i>0.87</i>	<i>0.73</i>	<i>0.76</i>
<i>weighted avg</i>	<i>0.87</i>	<i>0.83</i>	<i>0.83</i>

4.3.2 Result of Training with Unbalanced Dataset using Different Class Weight

Just like the previous method, the dataset used in this method is the same size. The image dataset was divided into three categories, with 80% of the images used for model training, 10% for model validation, and the remaining 10% for model testing. The images were categorized into eleven classes: ARMD with 136 images (4.2%), BRVO with 96 images (2.9%), DN with 181 images (5.6%), DR with 478 images (14.7%), MH with 401 images (12.3%), MYA with 133 images (4.1%), ODC with 347 images (10.7%), ODE with 78 images (2.4%), ODP with 90 images (2.8%), TSLN with 243 images (7.5%), and normal with 1,074 images (33%). As seen in Figure 4.4, the dataset is unbalanced, and the majority class is normal.

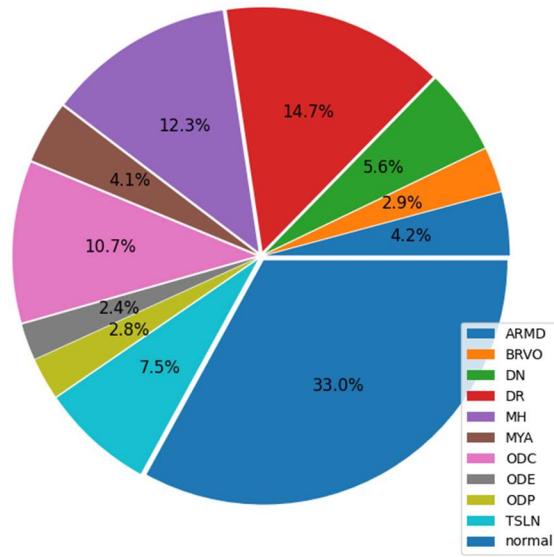


Figure 4.4 The percentage of the images in each class for Training with Unbalanced Dataset using Different Class Weight Method.

After training with the unbalanced dataset using different class weight method while simultaneously using the parameter settings mentioned in Table 3.2, the best epoch was determined to be epoch 41 after 61 epochs of training a model with early stopping enabled. 83.11% training accuracy, 83.13% validation accuracy, and 80% testing accuracy were attained by the model. The training loss was at 0.6074, and the validating loss was at 0.5904. However, the accuracy is somehow lower than not using different class weights, so it can be concluded that this method is not the best approach. Training and Validation Accuracy vs. Epoch and Training and Validation Loss vs. Epoch of MobileNetV2 for this approach are shown in Figures 4.5 and 4.6, respectively. The confusion matrix of this approach is shown in Table 4.3, and the classification report is shown in Table 4.4.

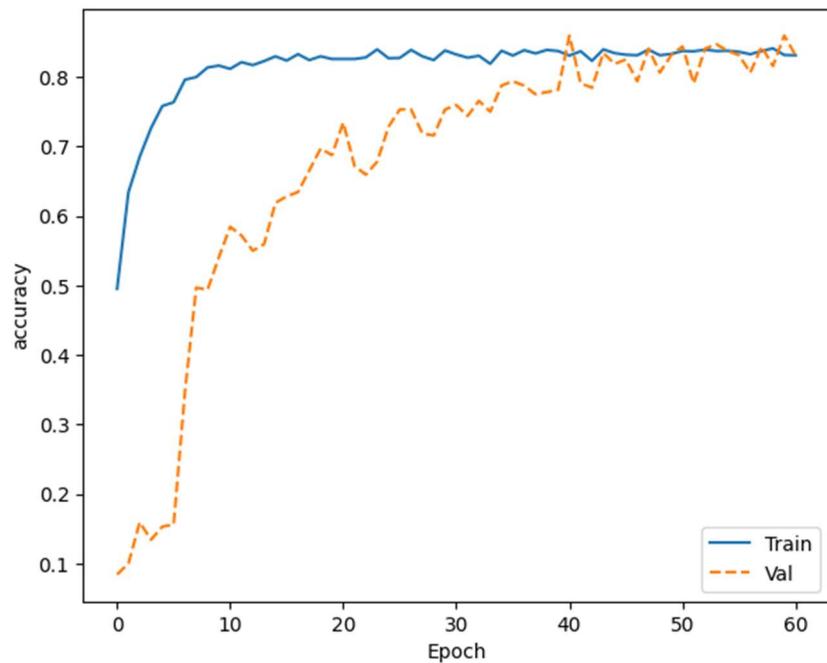


Figure 4.5 Training and Validation Accuracy vs. Epoch using Training with Unbalanced Dataset using Different Class Weight Method.

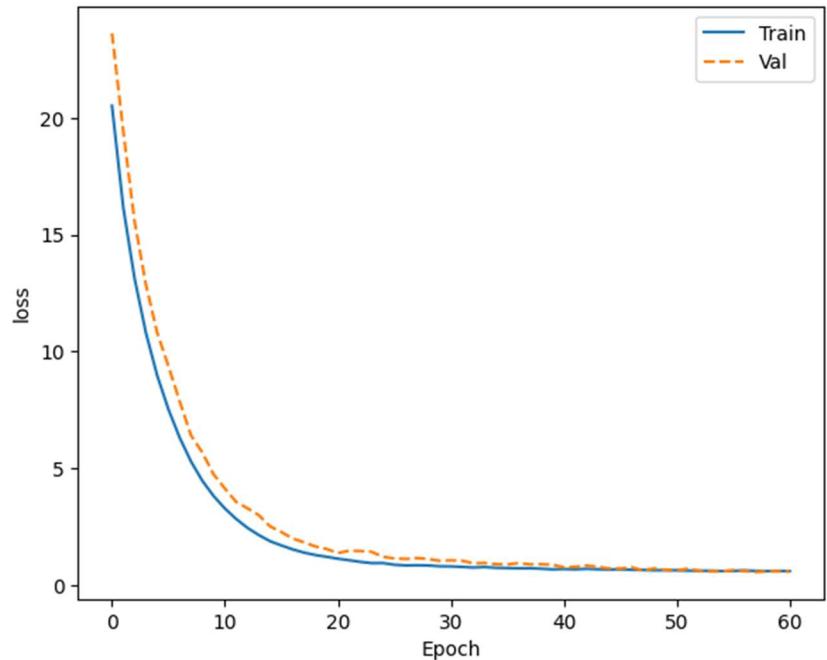


Figure 4.6 Training and Validation Loss vs. Epoch using Training with Unbalanced Dataset using Different Class Weight Method.

Table 4.3 Confusion Matrix of MobileNetV2 using Training with Unbalanced Dataset using Different Class Weight Method.

Class	<i>ARMD</i>	<i>BRVO</i>	<i>DN</i>	<i>DR</i>	<i>MH</i>	<i>MYA</i>	<i>ODC</i>	<i>ODE</i>	<i>ODP</i>	<i>TSLN</i>	<i>Normal</i>
<i>ARMD</i>	3	0	0	0	0	4	0	0	0	3	0
<i>BRVO</i>	0	2	0	0	1	1	4	0	0	1	0
<i>DN</i>	0	0	7	0	1	0	3	0	0	3	0
<i>DR</i>	0	0	0	23	1	0	7	0	1	8	0
<i>MH</i>	1	0	0	0	30	0	3	0	0	4	0
<i>MYA</i>	0	0	0	0	0	10	0	0	0	0	0
<i>ODC</i>	0	0	1	3	2	3	31	0	0	6	0
<i>ODE</i>	0	0	0	0	0	0	0	9	0	0	0
<i>ODP</i>	0	0	0	0	0	0	0	0	4	2	0
<i>TSLN</i>	1	0	1	0	4	0	0	0	0	23	0
<i>Normal</i>	0	0	0	0	0	0	0	0	0	0	114

Table 4.4 Classification Report on The Testing Set using Training with Unbalanced Dataset using Different Class Weight Method.

Class	<i>Precision</i>	<i>Recall</i>	<i>F1 - score</i>
ARMD	0.60	0.30	0.40
BRVO	1.00	0.22	0.36
DN	0.88	0.50	0.64
DR	0.88	0.57	0.70
MH	0.86	0.79	0.82
MYA	0.56	1.00	0.71
ODC	0.65	0.67	0.66
ODE	1.00	1.00	1.00
ODP	0.80	0.67	0.73
TSLN	0.46	0.96	0.62
Normal	1.00	1.00	1.00
<i>macro avg</i>	0.79	0.70	0.69
<i>weighted avg</i>	0.84	0.80	0.80

4.3.3 Result of Training with Unbalanced Dataset and using Data Cleaning

In this method, the data cleaning technique is then applied to only the TSLN class of the unbalanced dataset. The image dataset was divided into three categories, with 80% of the images used for model training, 10% for model validation, and the remaining 10% for model testing. The images were categorized into eleven classes: ARMD with 136 images (4.2%), BRVO with 96 images (3%), DN with 181 images (5.6%), DR with 478 images (14.8%), MH with 401 images (12.5%), MYA with 133 images (4.1%), ODC with 347 images (10.8%), ODE with 78 images (2.4%), ODP with 90 images (2.8%), TSLN with 206 images (6.4%), and normal with 1,074 images (33.4%). As seen in Figure 4.7, the dataset is unbalanced, and the majority class is normal.

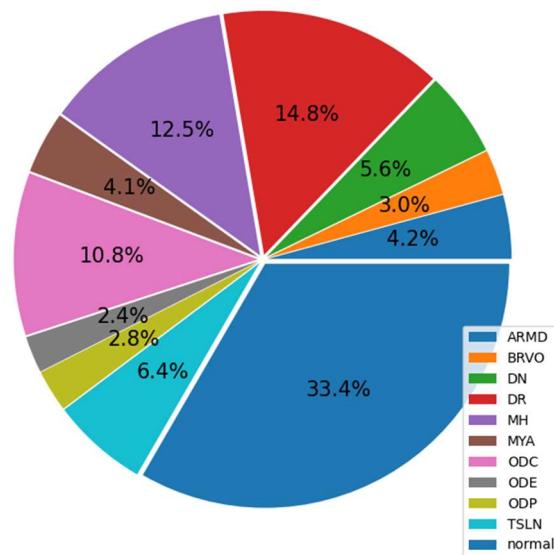


Figure 4.7 The percentage of the images in each class for Training with Unbalanced Dataset and using Data Cleaning Method.

After training with the unbalanced dataset and using the data cleaning method while simultaneously using the parameter settings mentioned in Table 3.2, the best epoch was determined to be epoch 78 after 98 epochs of training a model with early stopping enabled. 90.62% training accuracy, 90.94% validation accuracy, and 92.5% testing accuracy were attained by the model. The training loss was at 0.3188, and the validating loss was at 0.2755. Training and Validation Accuracy vs. Epoch and Training and Validation Loss vs. Epoch of MobileNetV2 for this approach are shown in Figures

4.8 and 4.9, respectively. The confusion matrix of this approach is shown in Table 4.5, and the classification report is shown in Table 4.6.

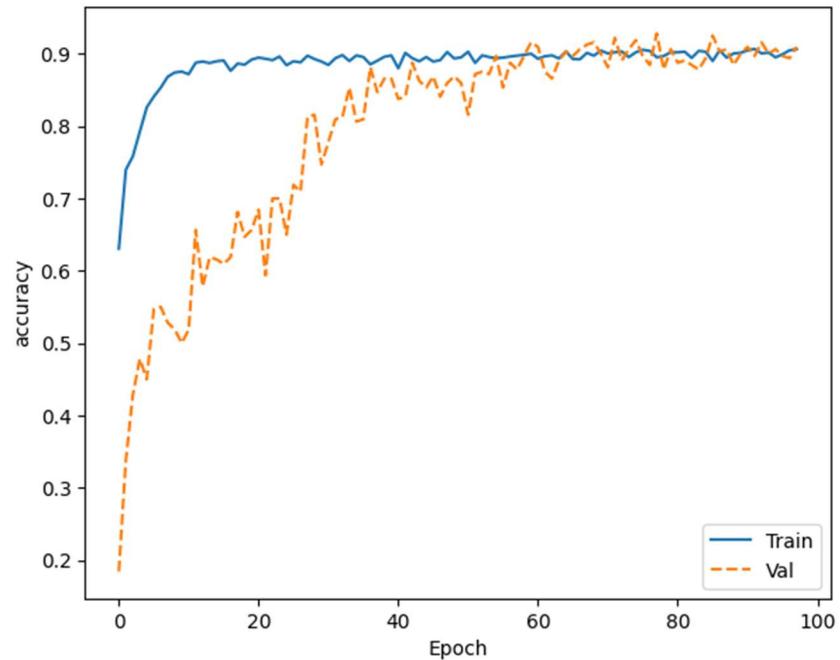


Figure 4.8 Training and Validation Accuracy vs. Epoch using Training with Unbalanced Dataset using Data Cleaning Method.

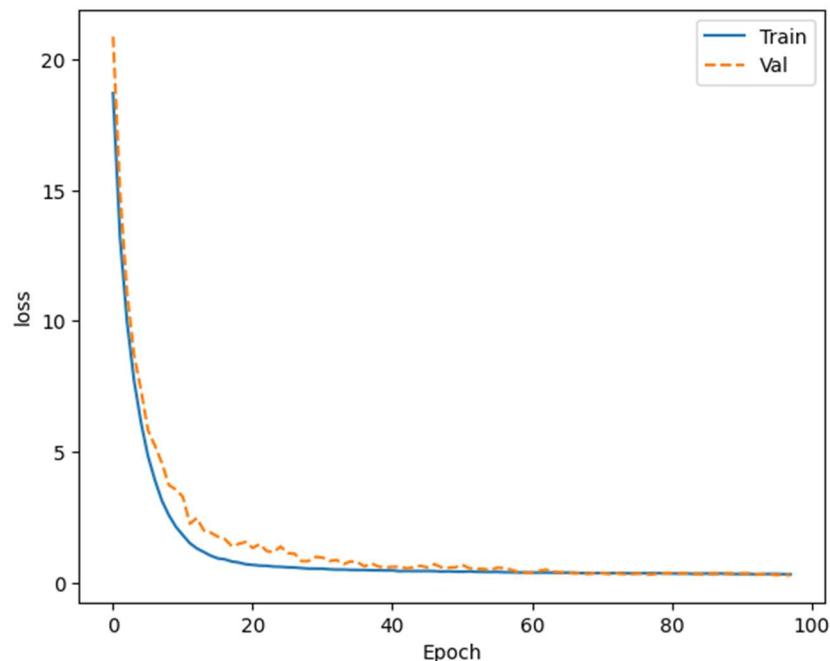


Figure 4.9 Training and Validation Loss vs. Epoch using Training with Unbalanced Dataset using Data Cleaning Method.

Table 4.5 Confusion Matrix of MobileNetV2 using Training with Unbalanced Dataset using Data Cleaning Method.

Class	ARMD	BRVO	DN	DR	MH	MYA	ODC	ODE	ODP	TSLN	Normal
ARMD	12	0	0	0	0	0	0	0	0	0	0
BRVO	0	7	0	0	0	0	0	0	0	0	0
DN	2	0	15	0	1	0	1	0	0	0	0
DR	0	0	0	44	2	0	1	0	3	0	0
MH	0	0	0	0	37	0	1	0	1	0	0
MYA	3	0	0	1	0	7	0	0	0	0	0
ODC	1	1	1	0	5	1	24	0	0	0	0
ODE	0	0	0	0	0	0	0	8	0	0	0
ODP	0	0	0	0	0	0	0	0	9	0	0
TSLN	1	0	0	0	0	0	0	0	0	21	0
Normal	0	0	0	0	0	0	0	0	0	0	112

Table 4.6 Classification Report on The Testing Set using Training with Unbalanced Dataset using Data Cleaning Method.

Class	Precision	Recall	F1 - score
ARMD	0.67	1.00	0.80
BRVO	0.88	1.00	0.93
DN	0.94	0.83	0.88
DR	0.98	0.88	0.93
MH	0.82	0.95	0.88
MYA	0.88	0.64	0.74
ODC	0.92	0.73	0.81
ODE	1.00	1.00	1.00
ODP	0.69	1.00	0.82
TSLN	1.00	1.00	1.00
Normal	1.00	1.00	1.00
<i>macro avg</i>	0.89	0.91	0.89
<i>weighted avg</i>	0.94	0.93	0.93

4.3.4 Result of Training with Data Augmentation Only

For this method, the dataset was then augmented to images in each class to make it balance. The image dataset was divided into three categories, with 80% of the images used for model training, 10% for model validation, and the remaining 10% for model testing. The images were categorized into eleven classes: ARMD with 1,088 images (9%), BRVO with 1,056 images (8.8%), DN with 1,086 images (9%), DR with 1,064 images (8.8%), MH with 1,203 images (10%), MYA with 1,064 images (8.8%), ODC with 1,041 images (8.6%), ODE with 1,092 images (9.1%), ODP with 1080 images (9%), TSLN with 1,215 images (10.1%), and normal with 1,074 images (8.9%). As seen in Figure 4.10, the dataset is balanced.

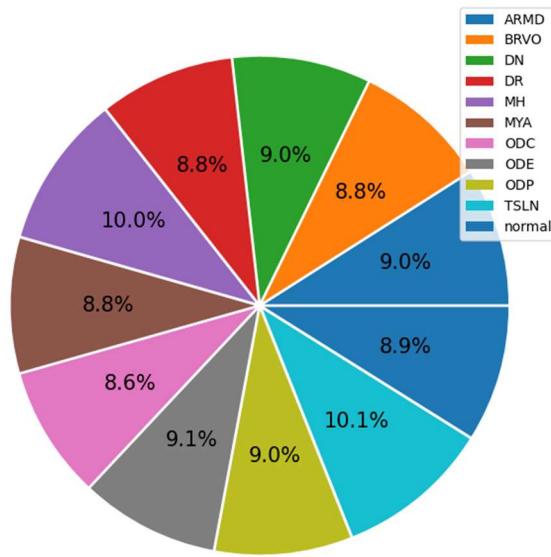


Figure 4.10 The percentage of the images in each class for Training with Data Augmentation Only Method.

After training with the data augmentation only method and using the parameter settings mentioned in Table 3.2, the best epoch was determined to be epoch 89 after 109 epochs of training a model with early stopping enabled. 95.38% training accuracy, 95.14% validation accuracy, and 95.64% testing accuracy were attained by the model. The training loss was at 0.1790, and the validating loss was at 0.1778. Training and Validation Accuracy vs. Epoch and Training and Validation Loss vs. Epoch of MobileNetV2 for this approach are shown in Figures 4.11 and 4.12, respectively. The

confusion matrix of this approach is shown in Table 4.7, and the classification report is shown in Table 4.8.

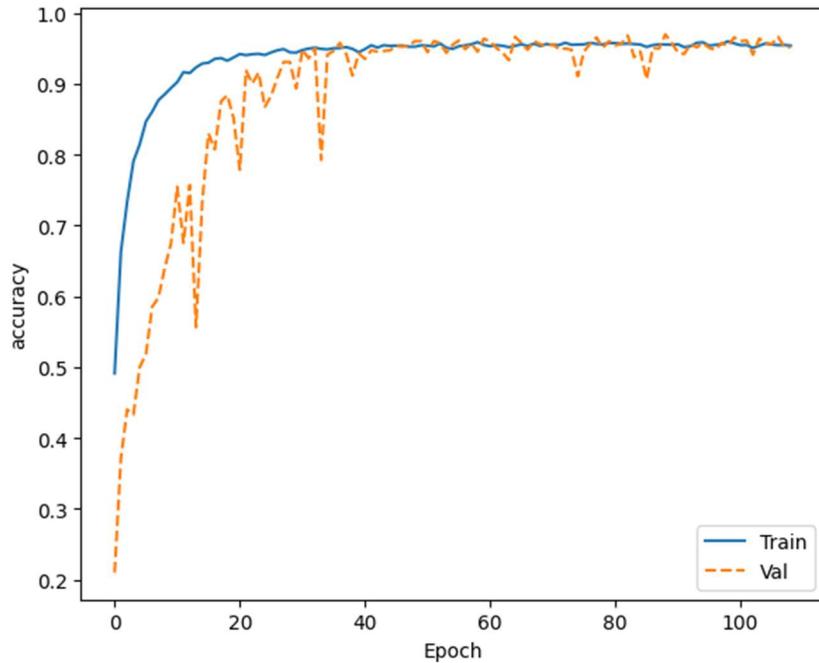


Figure 4.11 Training and Validation Accuracy vs. Epoch using Training with Data Augmentation Only Method.

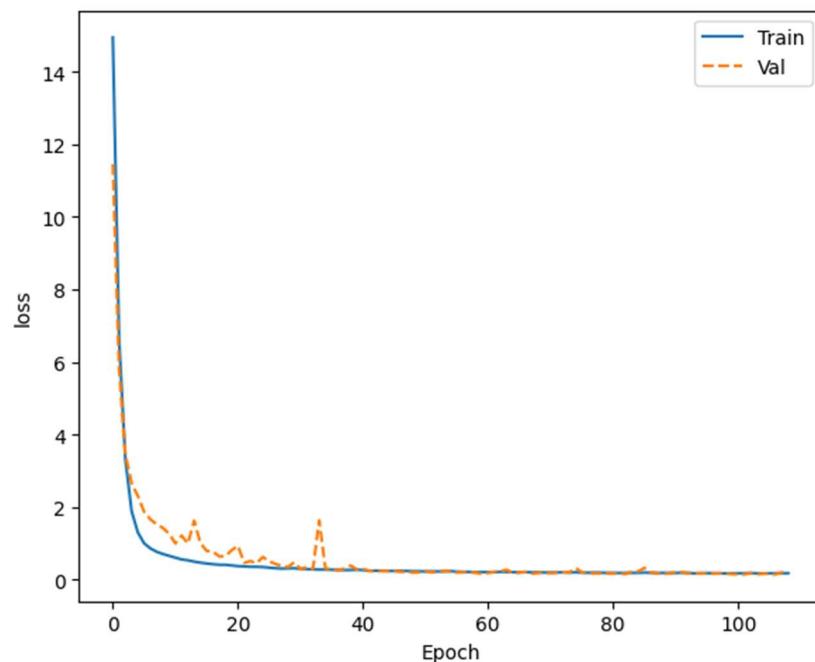


Figure 4.12 Training and Validation Loss vs. Epoch using Training with Data Augmentation Only Method.

Table 4.7 Confusion Matrix of MobileNetV2 using Training with Data Augmentation Only Method.

Class	ARMD	BRVO	DN	DR	MH	MYA	ODC	ODE	ODP	TSLN	Normal
ARMD	104	0	0	0	0	0	0	0	0	2	0
BRVO	0	94	1	0	1	0	0	0	0	1	0
DN	0	0	109	0	2	0	0	0	0	1	0
DR	0	0	0	116	1	0	0	0	0	0	0
MH	0	0	1	0	122	0	0	0	0	7	0
MYA	5	0	0	2	3	101	2	0	1	0	0
ODC	1	1	1	1	4	0	73	0	0	2	0
ODE	0	0	0	0	0	0	0	101	0	0	0
ODP	1	0	1	0	0	0	0	0	117	0	0
TSLN	1	0	3	5	2	0	0	0	0	109	0
Normal	0	0	0	0	0	0	0	0	0	0	117

Table 4.8 Classification Report on The Testing Set using Training with Data Augmentation Only Method.

Class	Precision	Recall	F1 - score
ARMD	0.93	0.98	0.95
BRVO	0.99	0.97	0.98
DN	0.94	0.97	0.96
DR	0.94	0.99	0.96
MH	0.90	0.94	0.92
MYA	1.00	0.89	0.94
ODC	0.97	0.88	0.92
ODE	1.00	1.00	1.00
ODP	0.99	0.98	0.99
TSLN	0.89	0.91	0.90
Normal	1.00	1.00	1.00
<i>macro avg</i>	0.96	0.96	0.96
<i>weighted avg</i>	0.96	0.96	0.96

4.3.5 Result of Training with Data Augmentation before using Data Cleaning

For this method, the dataset was then augmented to images in each class to make it balanced, and then the data cleaning technique for only the TSLN class, mentioned in Section 2.6 was applied. The image dataset was divided into three categories, with 80% of the images used for model training, 10% for model validation, and the remaining 10% for model testing. The images were categorized into eleven classes: ARMD with 1,088 images (9.1%), BRVO with 1,056 images (8.8%), DN with 1,086 images (9%), DR with 1,064 images (8.9%), MH with 1,203 images (10.1%), MYA with 1,064 images (8.9%), ODC with 1,041 images (8.7%), ODE with 1,092 images (9.1%), ODP with 1080 images (9%), TSLN with 1,105 images (9.2%), and normal with 1,074 images (9%). As seen in Figure 4.13, the dataset is balanced.

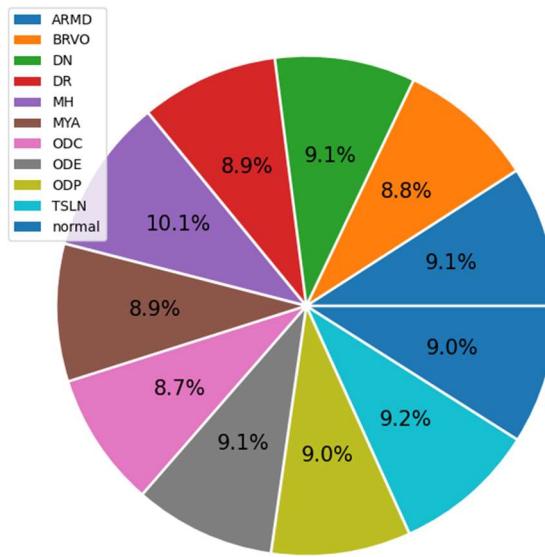


Figure 4.13 The percentage of the images in each class for Training with Data Augmentation before using Data Cleaning Method.

After training with the data augmentation before using data cleaning for only TSLN class method while simultaneously using the parameter settings mentioned in Table 3.2, the best epoch was determined to be epoch 56 after 76 epochs of training a model with early stopping enabled. 95.36% training accuracy, 94.53% validation accuracy, and 95.23% testing accuracy were attained by the model. The training loss was at 0.2152, and the validating loss was at 0.2166. Training and Validation Accuracy

vs. Epoch and Training and Validation Loss vs. Epoch of MobileNetV2 for this approach are shown in Figures 4.14 and 4.15, respectively. The confusion matrix of this approach is shown in Table 4.9, and the classification report is shown in Table 4.10.

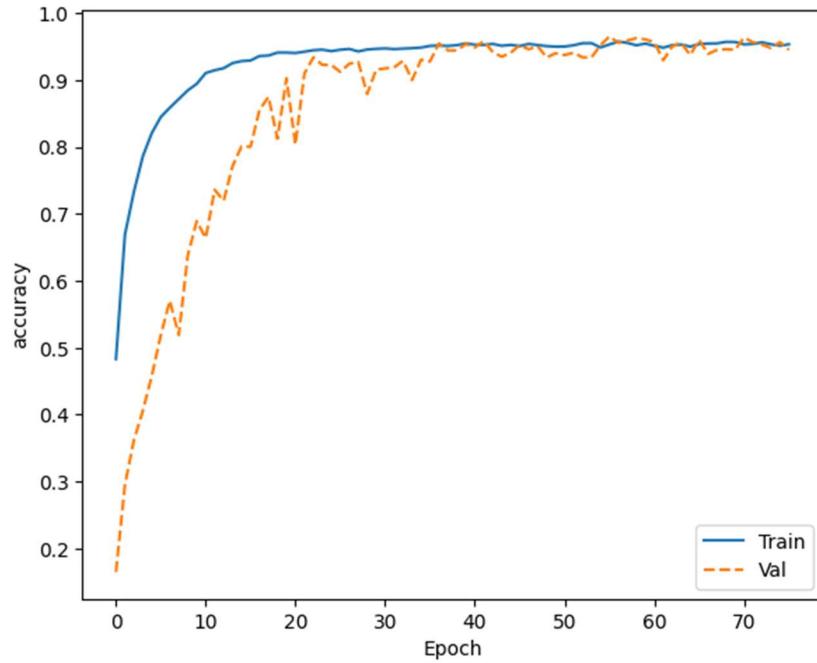


Figure 4.14 Training and Validation Accuracy vs. Epoch using Training with Data Augmentation before using Data Cleaning Method.

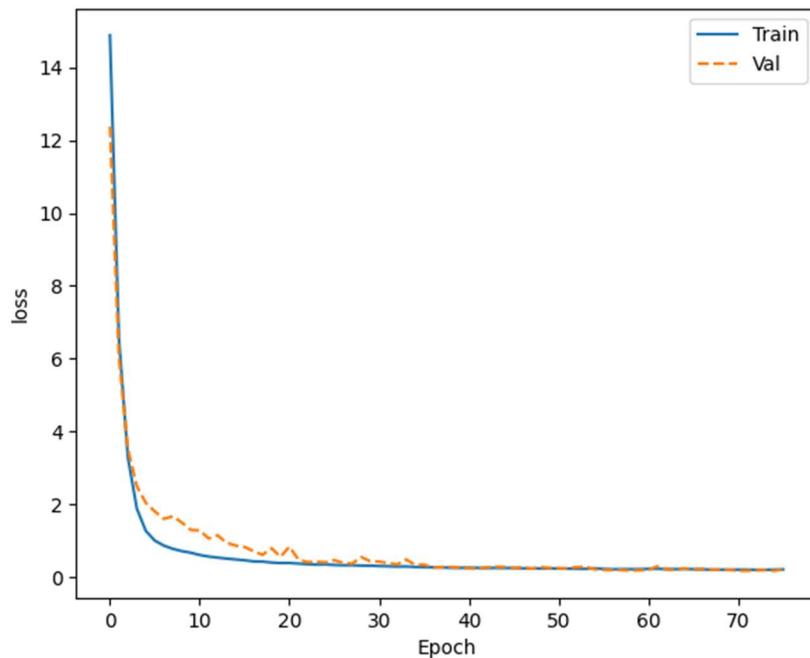


Figure 4.15 Training and Validation Loss vs. Epoch using Training with Data Augmentation before using Data Cleaning Method.

Table 4.9 Confusion Matrix of MobileNetV2 using Training with Data Augmentation before using Data Cleaning Method.

Class	ARMD	BRVO	DN	DR	MH	MYA	ODC	ODE	ODP	TSLN	Normal
ARMD	103	0	0	0	1	1	1	0	0	1	0
BRVO	0	103	0	0	0	0	0	0	0	0	0
DN	2	0	113	0	1	0	1	0	1	1	0
DR	0	0	0	97	4	0	0	0	0	0	0
MH	0	0	0	0	124	0	0	0	0	0	0
MYA	2	0	0	1	2	95	5	0	0	0	0
ODC	0	1	0	6	7	0	94	0	0	0	0
ODE	0	0	0	0	0	0	0	113	0	0	0
ODP	0	0	0	0	0	0	0	0	111	0	0
TSLN	4	1	0	4	6	0	4	0	1	94	0
Normal	0	0	0	0	0	0	0	0	0	0	111

Table 4.10 Classification Report on The Testing Set using Training with Data Augmentation before using Data Cleaning Method.

Class	Precision	Recall	F1 - score
ARMD	0.93	0.96	0.94
BRVO	0.98	1.00	0.99
DN	1.00	0.95	0.97
DR	0.90	0.96	0.93
MH	0.86	1.00	0.92
MYA	0.99	0.90	0.95
ODC	0.90	0.87	0.88
ODE	1.00	1.00	1.00
ODP	0.98	1.00	0.99
TSLN	0.98	0.82	0.90
Normal	1.00	1.00	1.00
<i>macro avg</i>	0.96	0.95	0.95
<i>weighted avg</i>	0.96	0.95	0.95

4.3.6 Result of Training with Data Cleaning before using Data Augmentation

For the last method, the dataset was cleaned for only the TSLN class before data augmentation to make the class balanced. The image dataset was divided into three categories, with 80% of the images used for model training, 10% for model validation, and the remaining 10% for model testing. The images were categorized into eleven classes: ARMD with 1,088 images (9%), BRVO with 1,056 images (8.8%), DN with 1,086 images (9%), DR with 1,064 images (8.8%), MH with 1,203 images (10%), MYA with 1,064 images (8.8%), ODC with 1,041 images (8.6%), ODE with 1,092 images (9.1%), ODP with 1080 images (9%), TSLN with 1,215 images (10.1%), and normal with 1,074 images (8.9%). As seen in Figure 4.16, the dataset is balanced.

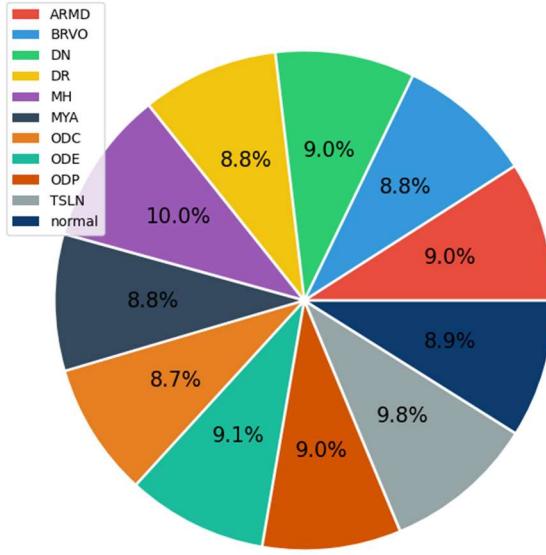


Figure 4.16 The percentage of the images in each class for Training with Data Cleaning before using Data Augmentation Method.

After training with the data cleaning for only the TSLN class before using data augmentation method while simultaneously using the parameter settings mentioned in Table 3.2, the best epoch was determined to be epoch 45 after 65 epochs of training a model with early stopping enabled. 95.75% training accuracy, 94.70% validation accuracy, and 96.30% testing accuracy were attained by the model. The training loss was at 0.2020, and the validating loss was at 0.2226. Training and Validation Accuracy vs. Epoch and Training and Validation Loss vs. Epoch of MobileNetV2 for this

approach are shown in Figures 4.17 and 4.18, respectively. The confusion matrix of this approach is shown in Table 4.11, and the classification report is shown in Table 4.12.

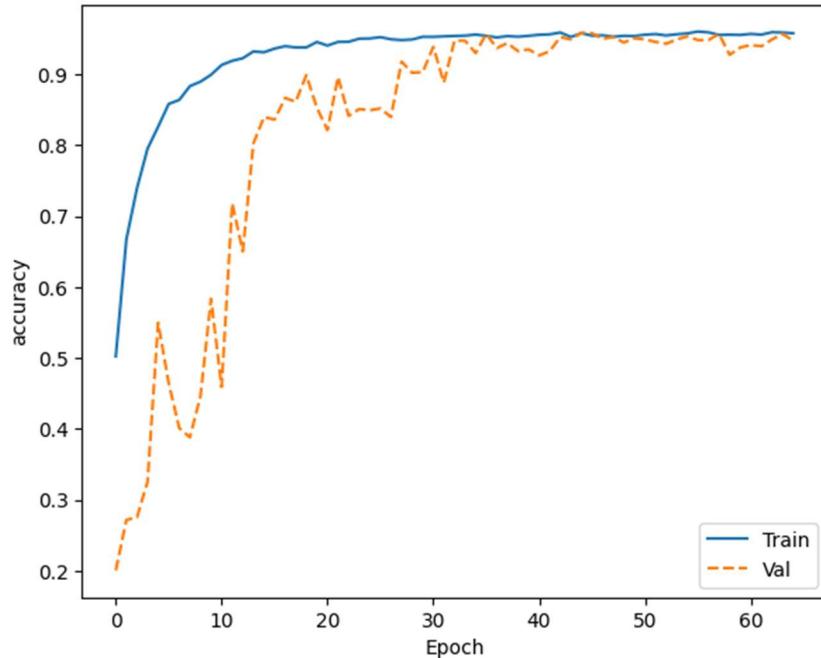


Figure 4.17 Training and Validation Accuracy vs. Epoch using Training with Data Cleaning before using Data Augmentation Method.

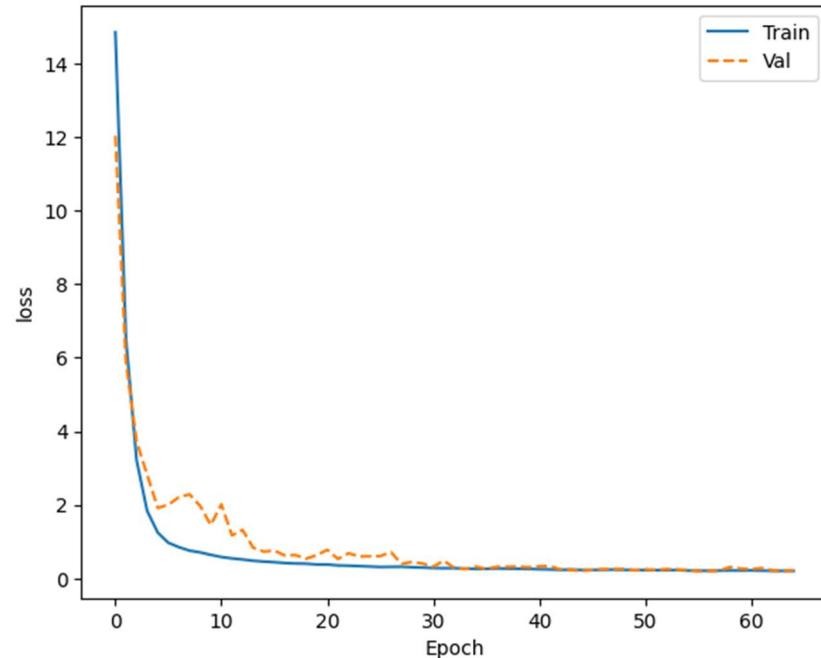


Figure 4.18 Training and Validation Loss vs. Epoch using Training with Data Cleaning before using Data Augmentation Method.

Table 4.11 Confusion Matrix of MobileNetV2 using Training with Data Cleaning before using Data Augmentation Method.

Class	<i>ARMD</i>	<i>BRVO</i>	<i>DN</i>	<i>DR</i>	<i>MH</i>	<i>MYA</i>	<i>ODC</i>	<i>ODE</i>	<i>ODP</i>	<i>TSLN</i>	<i>Normal</i>
<i>ARMD</i>	100	0	0	0	1	0	0	0	0	0	0
<i>BRVO</i>	0	107	0	0	0	0	0	0	0	0	0
<i>DN</i>	0	0	104	0	2	0	2	0	0	2	0
<i>DR</i>	0	0	0	104	4	0	0	0	0	0	0
<i>MH</i>	0	0	0	0	118	0	7	0	0	0	0
<i>MYA</i>	3	0	0	0	1	93	0	0	0	0	0
<i>ODC</i>	1	4	1	4	0	0	117	0	0	4	0
<i>ODE</i>	0	0	0	0	0	0	0	109	0	0	0
<i>ODP</i>	0	1	0	1	0	1	1	0	97	0	0
<i>TSLN</i>	2	1	0	1	1	0	1	0	0	99	0
<i>Normal</i>	0	0	0	0	0	0	0	0	0	0	123

Table 4.12 Classification Report on The Testing Set using Training with Data Cleaning before using Data Augmentation Method.

Class	<i>Precision</i>	<i>Recall</i>	<i>F1 - score</i>
ARMD	0.94	1.00	0.97
BRVO	0.95	1.00	0.97
DN	0.99	0.95	0.97
DR	0.95	0.96	0.95
MH	0.94	0.94	0.94
MYA	0.99	0.96	0.97
ODC	0.91	0.89	0.90
ODE	1.00	1.00	1.00
ODP	1.00	0.96	0.98
TSLN	0.94	0.94	0.94
Normal	1.00	1.00	1.00
<i>macro avg</i>	<i>0.96</i>	<i>0.96</i>	<i>0.96</i>
<i>weighted avg</i>	<i>0.96</i>	<i>0.96</i>	<i>0.96</i>

4.4 Comparison of All Proposed Methodologies

Table 4.13 Performance Metrics for Different Classification Approaches.

Methods	Training Accuracy	Validation Accuracy	Testing Accuracy	Training Loss	Validation Loss	Precision	Recall	F1-Score
Unbalanced Dataset	83.57%	85.94%	82.81%	0.4968	0.4957	0.87	0.83	0.83
Unbalanced Dataset using Different Class Weight	83.11%	83.13%	80%	0.6074	0.5904	0.84	0.80	0.80
Unbalanced Dataset & using Data Cleaning	90.62%	90.94%	92.50%	0.3188	0.2755	0.94	0.93	0.93
Data Augmentation Only	95.38%	95.14%	95.64%	0.1790	0.1778	0.96	0.96	0.96
Data Augmentation before using Data Cleaning	95.36%	94.53%	95.23%	0.2152	0.2166	0.96	0.95	0.95
Data Cleaning before using Data Augmentation*	95.75%	94.70%	96.30%	0.2020	0.2226	0.96	0.96	0.96

*Best Approach

The best method was determined to be the last method, or ‘Data Cleaning before using Data Augmentation’. 95.75% training accuracy, 94.70% validation accuracy, and 96.30% testing accuracy were attained by the model. The training loss was at 0.2020, and the validating loss was at 0.2226. According to the results, the model worked well, with high accuracy on both the training and validation sets. The best epoch was determined based on validation accuracy, which may have signaled the start of a performance plateau for the model. Because the model’s accuracy and validation accuracy were continuously high, it was able to successfully generalize to previously unreported data. The model’s excellent performance was aided by the use of early stopping, data augmentation, data cleaning, and 5-fold cross-validation, which also helped prevent overfitting.

Mobile Application

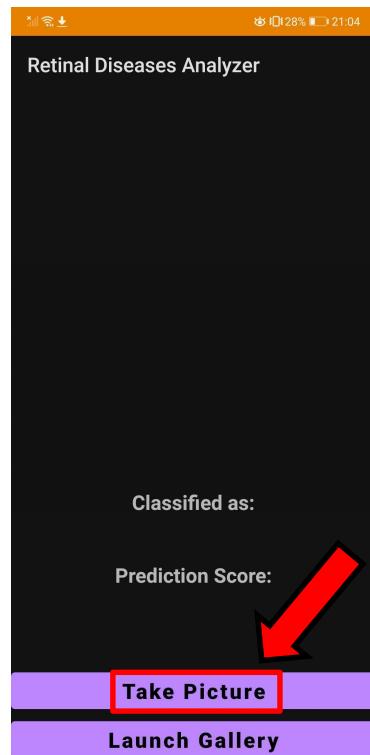
Finally, an Android application was developed with Android Studio and deployed to a mobile device. Users may load test photos into the app, which then classifies them using the learned model. The model's accurate prediction of the correct class for the test image resulted in a high prediction score. This successful integration of the trained model into the Android application demonstrates the model's capacity for real-world deployment in a practical environment, as well as its practical applicability and usability for real-time picture classification tasks on mobile devices.



Figure 4.19 Retinal Diseases Analyzer Mobile Application Icon.

4.5.1 Use a Testing Image by Taking a Picture

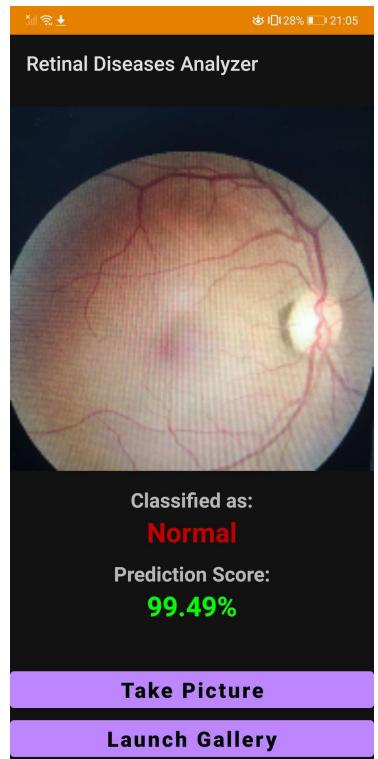
1.) When an application is opened, the interface will look like this: After that, tap on the button 'Take Picture' at the bottom of the screen to use your mobile phone camera to take a picture.



2.) The camera screen will pop up and be ready to take a picture of retinal image.

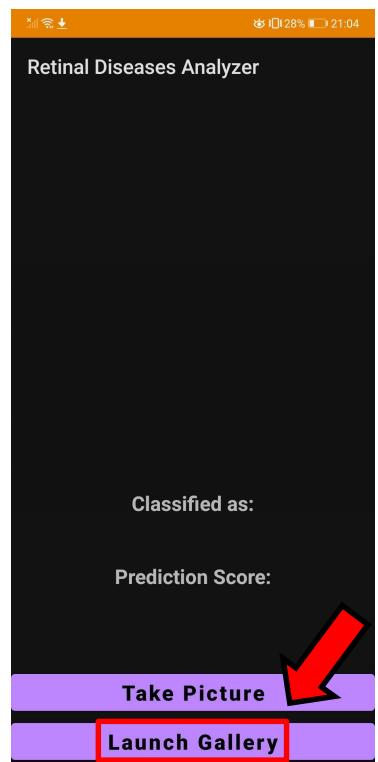


3.) After taking the picture, the app will automatically classify the retinal disease class and show a prediction score; in this case, it is classified as an 'Normal' class with a prediction score of 99.49%.

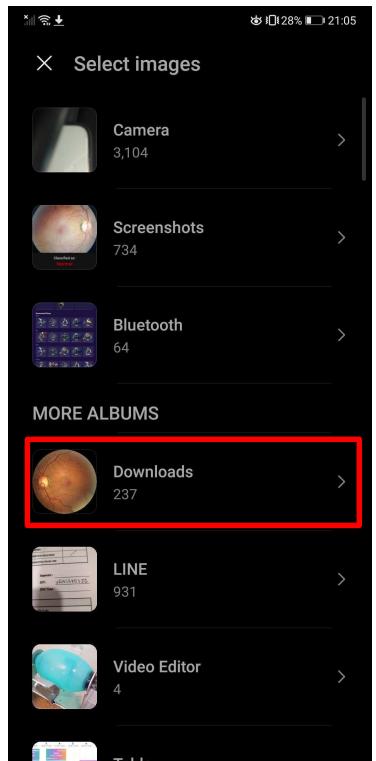


4.5.2 Use a Testing Image by Browsing the Gallery

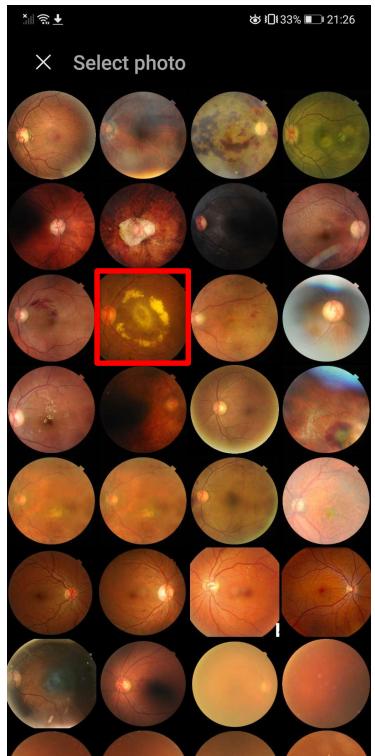
1.) When an application is opened, the interface will look like this: After that, tap on the button 'Launch Gallery' at the bottom of the screen to use your mobile phone camera to choose the testing image from your mobile's storage.

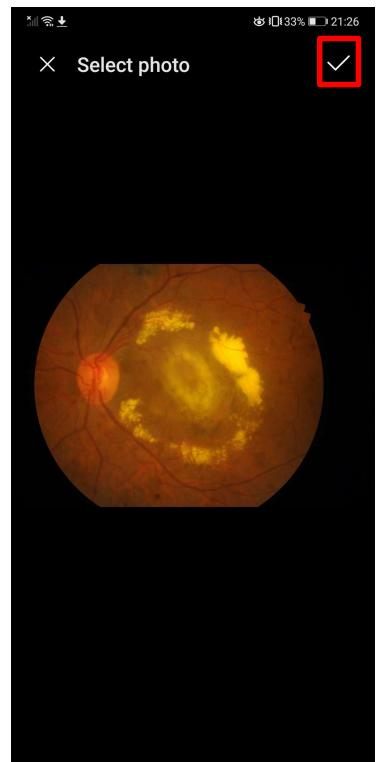


2.) Select the folder where you save your testing image.

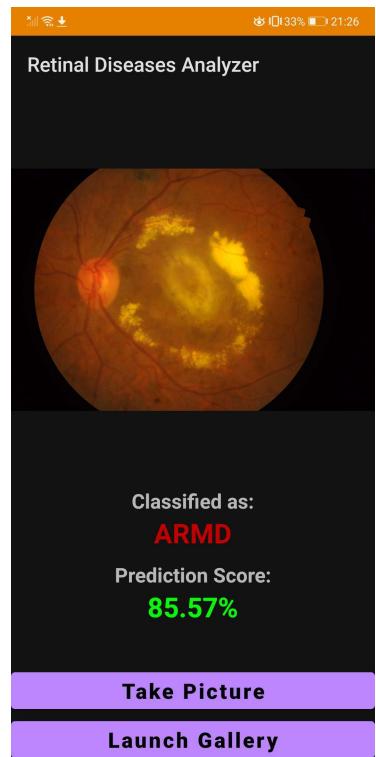


3.) Choose the retinal image that you want to predict.





4.) Tap on the check mark to confirm.



4.5.2.1 Prediction for the Rest of the Class using ‘Launch Gallery Method’

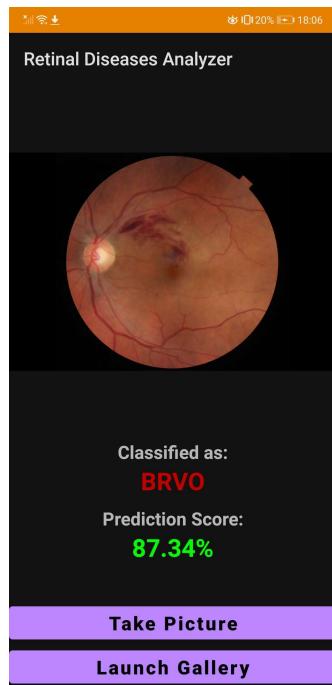


Figure 4.20 Classification of BRVO with a Prediction Score of 87.34%.

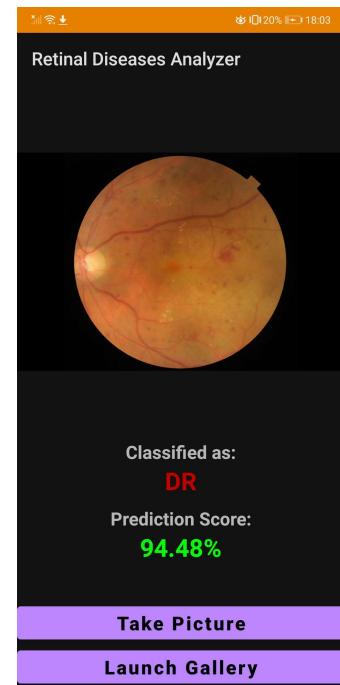


Figure 4.21 Classification of DR with a Prediction Score of 94.48%.

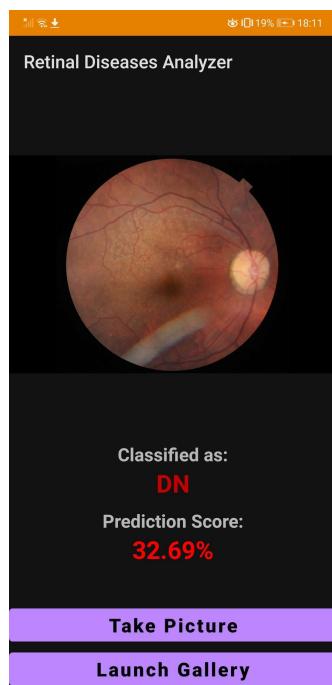


Figure 4.22 Classification of DN with a Prediction Score of 32.69%.

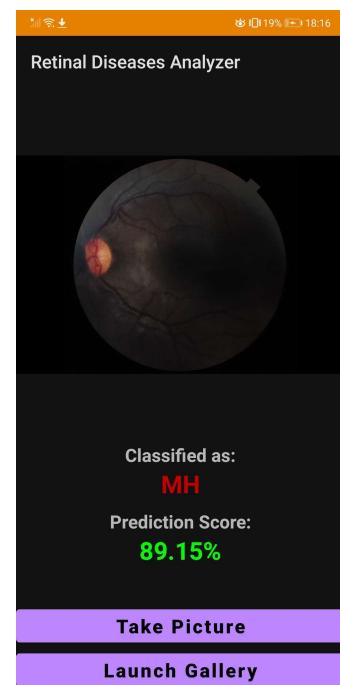


Figure 4.23 Classification of MH with a Prediction Score of 89.15%.

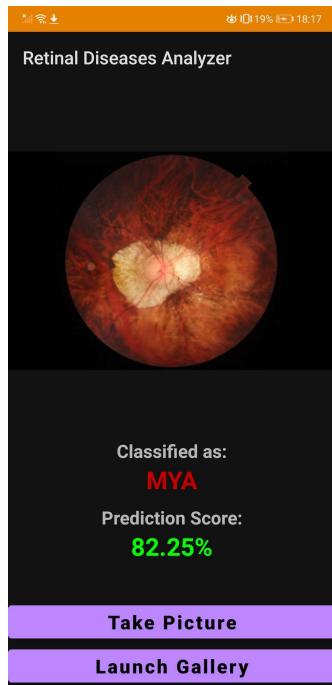


Figure 4.24 Classification of MYA with a Prediction Score of 82.25%.

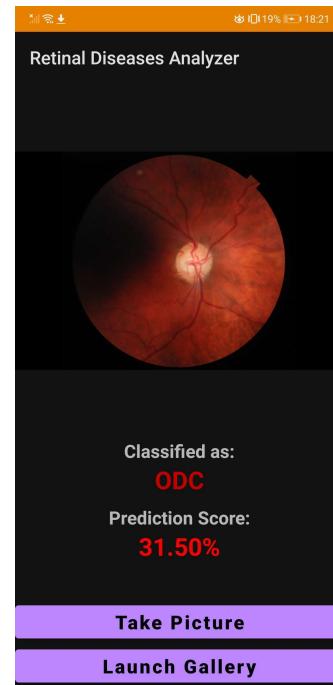


Figure 4.25 Classification of ODC with a Prediction Score of 31.50%.

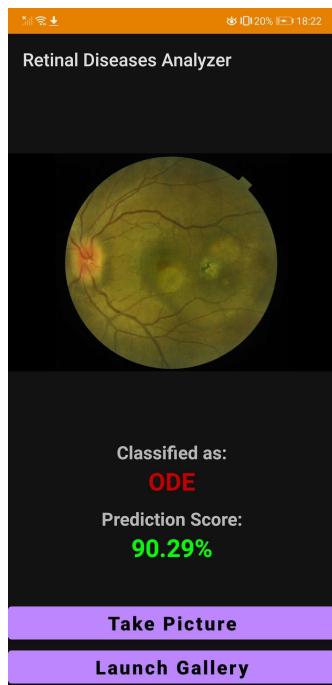


Figure 4.26 Classification of ODE with a Prediction Score of 90.29%.

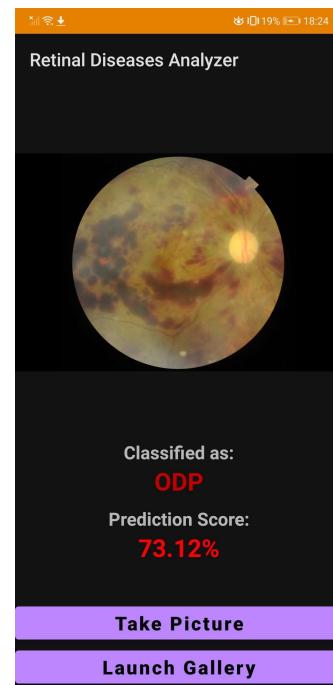


Figure 4.27 Classification of ODP with a Prediction Score of 73.12%.

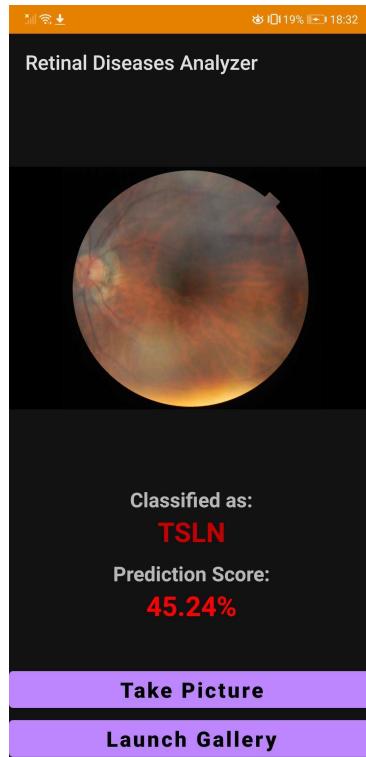


Figure 4.28 Classification of TSLN with a Prediction Score of 45.24%.

4.5 Summary

This chapter shows all six proposed method results (Section 4.3) in detail using evaluation criteria equations including accuracy, precision, recall, and F1-score. Moreover, the best approach or method was determined to be ‘Data Cleaning before using Data Augmentation’ in Section 4.4. Furthermore, the implementation of the prediction results in a mobile application is then visualized in Section 4.5.

The next chapter will talk about the discussion, which will compare these results with the literature reviews, and also conclude the experiment in the conclusion section.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Introduction

In the last chapter, the discussion will be explained in Section 5.2. A number of conclusions are drawn about key parts of the work undertaken in Section 5.3, and finally, Section 5.4 will discuss future work to improve the methods and experiments leading to better patient outcomes in the way of classifying retinal diseases.

5.2 Discussion

Table 5.1 Comparison Between the Literature Review and Our Research.

Ref (Year)	Dataset	Number of classes:	Based Deep- learning Model	Performance (Validation Accuracy)
Taufiqurrah man et al. [80] (2020)	APPOS 2019	5	MobileNetV2- SVM	85%
Patel and Chaware [81] (2020)	APPOS 2019	5	MobileNetV2	81%
Yildirim et al. [82] (2022)	Firat University Hospital	5	MobileNetV2	87.40%
Our Method (2023)	RFMiD 2.0	11	MobileNetV2	94.70%

We compared our proposed method with some related research in the literature [80], [81], and [82], and the comparative results are described in Table 4.13. Superior performance can be discerned from this table, with a validation accuracy of 94.70% being achieved on the RFMiD 2.0 dataset with eleven classes by our method. Even

though the same model architecture, MobileNetV2, was utilized by all the methods in Table 4.13, better outcomes were achieved for three main reasons: (i) MobileNetV2 was employed in a transfer learning style with the utilization of ImageNet pretrained weights; (ii) Our dataset was adjusted to achieve better balance and reduce bias, thereby enhancing the model's generalization ability; and (iii) Ten-fold cross-validation was applied to ensure increased robustness and the reduction of overfitting. However, it should be acknowledged that this comparison comes with limitations, as the methods mentioned operated on different datasets, which precludes calling it a fair comparison. The unavailability of their dataset prevented the testing of our method on it.

5.3 Conclusion

In conclusion, using the MobileNetV2 model, this research proposes a promising technique for identifying retinal disorders using deep learning and retinal pictures, with a training accuracy of 95.75%, a validation accuracy of 94.70%, and a testing accuracy of 96.30%. Our approach's excellent accuracy in diagnosing various retinal disorders has the potential to improve patient outcomes through early and accurate detection.

5.4 Future Scope

Nonetheless, more research and testing in clinical settings are required to validate its efficacy. The findings given in this work are thought to contribute to the growing body of knowledge in the field of research on deep learning in the detection of retinal disorders and related studies. The performance of the classifier suggested in this study can be enhanced in future work by adding additional photos or classes and experimenting with other models.

REFERENCES

- [1] Asiri, N., Hussain, M., Al Adel, F., & Alzaidi, N. (2019). Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey. *Artificial Intelligence in Medicine*, 99, 101701. <https://doi.org/10.1016/j.artmed.2019.07.009>
- [2] Yagi, M., Yamanouchi, K., Fujita, N., Funao, H., & Ebata, S. (2023). Revolutionizing Spinal Care: Current Applications and Future Directions of Artificial Intelligence and Machine Learning. *Journal of Clinical Medicine*, 12(13), 4188. <https://doi.org/10.339>
- [3] Panchal, S., Naik, A., Kokare, M., Pachade, S., Naigaonkar, R., Phadnis, P., & Bhange, A. (2023). Retinal Fundus Multi-Disease Image Dataset (RFMiD) 2.0: A Dataset of Frequently and Rarely Identified Diseases. *Data*, 8(2), 29. <https://doi.org/10.3390/data80>
- [4] Nowak, J. Z. (2013). Oxidative stress, polyunsaturated fatty acidsderived oxidation products and bisretinoids as potential inducers of CNS diseases: Focus on age-related macular degeneration. *Pharmacological Reports*, 65(2), 288–304. <https://doi.org/10.1016>
- [5] Holz, F. G., Strauss, E. C., Schmitz-Valckenberg, S., & Van Lookeren Campagne, M. (2014). Geographic Atrophy. *Ophthalmology*, 121(5), 1079–1091. <https://doi.org/10.1016/j.ophtha.2013.11.023>
- [6] MacGillivray, T. J., Trucco, E., Cameron, J. R., Dhillon, B., Houston, J. G., & Van Beek, E. J. R. (2014). Retinal imaging as a source of biomarkers for diagnosis, characterization and prognosis of chronic illness or long-term conditions. *The British Journ*
- [7] Wong, T. Y., & Sabanayagam, C. (2020). Strategies to Tackle the Global Burden of Diabetic Retinopathy: From Epidemiology to Artificial Intelligence. *Ophthalmologica*, 243(1), 9–20. <https://doi.org/10.1159/000502387>
- [8] Kong, Y., Liu, P., Li, Y., Nolan, N. D., Quinn, P. M. J., Hsu, C., Jenny, L. A., Zhao, J., Cui, X., Chang, Y., Wert, K. J., Sparrow, J. R., Wang, N.,

- & Tsang, S. H. (2023). HIF2A activation and mitochondrial deficit due to iron chelation cause retinal atrophy. *EMBO Molecular Medicine*, 15(2), e16525. <https://doi.org/10.15252/emmm.202216525>
- [9] Girmens. (2012). Dry age-related macular degeneration: A currently unmet clinical need. *Intractable & Rare Diseases Research*. <https://doi.org/10.5582/irdr.2012.v1.3.103>
- [10] Binder, S., Stanzel, B. V., Krebs, I., & Glittenberg, C. (2007). Transplantation of the RPE in AMD. *Progress in Retinal and Eye Research*, 26(5), 516–554. <https://doi.org/10.1016/j.preteyeres.2007.02.002>
- [11] Nguyen, Q. D., Das, A., Do, D. V., Dugel, P. U., Gomes, A., Holz, F. G., Koh, A., Pan, C. K., Sepah, Y. J., Patel, N., MacLeod, H., & Maurer, P. (2020). Brolucizumab: Evolution through Preclinical and Clinical Studies and the Implications for the Management of Neovascular Age-Related Macular Degeneration. *Ophthalmology*, 127(7), 963–976. <https://doi.org/10.1016/j.ophtha.2019.12.031>
- [12] Song, P., Xu, Y., Zha, M., Zhang, Y., & Rudan, I. (2019). Global epidemiology of retinal vein occlusion: A systematic review and meta-analysis of prevalence, incidence, and risk factors. *Journal of Global Health*, 9(1), 010427. <https://doi.org/10.7189/jogh.09.010427>
- [13] Mettu, P. S., Wielgus, A. R., Ong, S. S., & Cousins, S. W. (2012). Retinal pigment epithelium response to oxidant injury in the pathogenesis of early age-related macular degeneration. *Molecular Aspects of Medicine*, 33(4), 376–398. <https://doi.org/10.1016/j.mam.2012.04.006>
- [14] Sarks, J. P., Sarks, S. H., & Killingsworth, M. C. (1994). Evolution of soft drusen in age-related macular degeneration. *Eye*, 8(3), 269–283. <https://doi.org/10.1038/eye.1994.57>
- [15] Fenwick, E. K., Pesudovs, K., Khadka, J., Dirani, M., Rees, G., Wong, T. Y., & Lamoureux, E. L. (2012). The impact of diabetic retinopathy on quality of life: Qualitative findings from an item bank development project. *Quality of Life Research*, 21(10), 1771–1782. <https://doi.org/10.1007/s11136-012-0110-1>

- [16] Usman, T. M., Saheed, Y. K., Nsang, A., Ajibesin, A., & Rakshit, S. (2023). A systematic literature review of machine learning based risk prediction models for diabetic retinopathy progression. *Artificial Intelligence in Medicine*, 143, 102617. <https://doi.org/10.1016/j.artmed.2023.102617>
- [17] Rübsam, A., Parikh, S., & Fort, P. (2018). Role of Inflammation in Diabetic Retinopathy. *International Journal of Molecular Sciences*, 19(4), 942. <https://doi.org/10.3390/ijms19040942>
- [18] Owsley, C., & McGwin, G. (1999). Vision Impairment and Driving. *Survey of Ophthalmology*, 43(6), 535–550. [https://doi.org/10.1016/S0039-6257\(99\)00035-1](https://doi.org/10.1016/S0039-6257(99)00035-1)
- [19] Monteiro, J. P., Santos, F. M., Rocha, A. S., Castro-de-Sousa, J. P., Queiroz, J. A., Passarinha, L. A., & Tomaz, C. T. (2015). Vitreous humor in the pathologic scope: Insights from proteomic approaches. *PROTEOMICS - Clinical Applications*, 9(1–2), 187–202. <https://doi.org/10.1002/prca.201400133>
- [20] Gordon-Shaag, A., Shneor, E., Doron, R., Levine, J., & Ostrin, L. A. (2021). Environmental and Behavioral Factors with Refractive Error in Israeli Boys. *Optometry and Vision Science*, 98(8), 959–970. <https://doi.org/10.1097/OPX.0000000000001755>
- [21] Tirayoh, N. C., & Dja'far, V. H. (2021). INCREASED MYOPIA PREVALENCE IN YOUNGSTERS AS A RESULT OF ONLINE LEARNING. *TRANSFORMATIONAL LANGUAGE, LITERATURE, AND TECHNOLOGY OVERVIEW IN LEARNING (TRANSTOOL)*, 1(1), 39–46. <https://doi.org/10.55047/transtool.v1i1.106>
- [22] Szpiro, S., Zhao, Y., & Azenkot, S. (2016). Finding a store, searching for a product: A study of daily challenges of low vision people. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 61–72. <https://doi.org/10.1145/2971648.2971723>
- [23] Andrade, T. D. S., Araújo, R. B. D., Rocha, A. A. D. N., Mello, L. G. M., Cunha, L. P., & Monteiro, M. L. R. (2022). Bruch Membrane Opening

- Minimum Rim Width and Retinal Nerve Fiber Layer Helps Differentiate Compressive Optic Neuropathy From Glaucoma. American Journal of Ophthalmology, 234, 156–165. <https://doi.org/10.1016/j.ajo.2021.08.008>
- [24] Hayreh, S. S. (1969). Blood supply of the optic nerve head and its role in optic atrophy, glaucoma, and oedema of the optic disc. British Journal of Ophthalmology, 53(11), 721–748. <https://doi.org/10.1136/bjo.53.11.721>
- [25] Chauhan, B. C. (2001). Optic Disc and Visual Field Changes in a Prospective Longitudinal Study of Patients With Glaucoma: Comparison of Scanning Laser Tomography With Conventional Perimetry and Optic Disc Photography. Archives of Ophthalmology, 119(10), 1492. <https://doi.org/10.1001/archopht.119.10.1492>
- [26] Spaide, R. F., Gemmy Cheung, C. M., Matsumoto, H., Kishi, S., Boon, C. J. F., Van Dijk, E. H. C., Mauget-Faysse, M., Behar-Cohen, F., Hartnett, M. E., Sivaprasad, S., Iida, T., Brown, D. M., Chhablani, J., & Maloca, P. M. (2022). Venous overload choroidopathy: A hypothetical framework for central serous chorioretinopathy and allied disorders. Progress in Retinal and Eye Research, 86, 100973. <https://doi.org/10.1016/j.preteyeres.2021.100973>
- [27] Xie, J. S., Donaldson, L., & Margolin, E. (2022). Papilledema: A review of etiology, pathophysiology, diagnosis, and management. Survey of Ophthalmology, 67(4), 1135–1159. <https://doi.org/10.1016/j.survophthal.2021.11.007>
- [28] Selhorst, J., & Chen, Y. (2009). The Optic Nerve. Seminars in Neurology, 29(01), 029–035. <https://doi.org/10.1055/s-0028-1124020>
- [29] Barbosa, B. G., Cassimiro, A. V. N. A., Rocha, B. G. S., Alvarenga, B. M., Birbrair, A., & Ribas, V. T. (2022). Pericyte Biology in the Optic Nerve and Retina. Current Tissue Microenvironment Reports, 3(3), 37–50. <https://doi.org/10.1007/s43152-022-00036-5>
- [30] Theodossiadis, G. P., Grigoropoulos, V. G., Liarakos, V. S., Rouvas, A., Emfietzoglou, I., & Theodossiadis, P. G. (2012). Restoration of the

- photoreceptor layer and improvement of visual acuity in successfully treated optic disc pit maculopathy: A long follow-up study by optical coherence tomography. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 250(7), 971–979. <https://doi.org/10.1007/s00417-011-1918-z>
- [31] Lyu, H., Chen, Q., Hu, G., Shi, Y., Ye, L., Yin, Y., Fan, Y., Zou, H., He, J., Zhu, J., & Xu, X. (2021). Characteristics of Fundal Changes in Fundus Tessellation in Young Adults. *Frontiers in Medicine*, 8, 616249. <https://doi.org/10.3389/fmed.2021.616249>
- [32] Chou, H.-D., Wu, A.-L., Cheng, Y.-C., & Wang, N.-K. (2020). Retinitis Pigmentosa. In G. Cheung (Ed.), *Heredity Chorioretinal Disorders* (pp. 1–43). Springer Singapore. https://doi.org/10.1007/978-981-15-0414-3_1
- [33] Jonas, J. B., Wei, W. B., Xu, L., Rietschel, M., Streit, F., & Wang, Y. X. (2018). Self-rated depression and eye diseases: The Beijing Eye Study. *PLOS ONE*, 13(8), e0202132. <https://doi.org/10.1371/journal.pone.0202132>
- [34] Kim, J. S., Ishikawa, H., Sung, K. R., Xu, J., Wollstein, G., Bilonick, R. A., Gabriele, M. L., Kagemann, L., Duker, J. S., Fujimoto, J. G., & Schuman, J. S. (2009). Retinal nerve fibre layer thickness measurement reproducibility improved with spectral domain optical coherence tomography. *British Journal of Ophthalmology*, 93(8), 1057–1063. <https://doi.org/10.1136/bjo.2009.157875>
- [35] Van Dijk, H. W., Verbraak, F. D., Kok, P. H. B., Garvin, M. K., Sonka, M., Lee, K., DeVries, J. H., Michels, R. P. J., Van Velthoven, M. E. J., Schlingemann, R. O., & Abràmoff, M. D. (2010). Decreased Retinal Ganglion Cell Layer Thickness in Patients with Type 1 Diabetes. *Investigative Ophthalmology & Visual Science*, 51(7), 3660. <https://doi.org/10.1167/iovs.09-5041>
- [36] Kolb, H., & Dekorver, L. (1991). Midget ganglion cells of the parafovea of the human retina: A Study by electron microscopy and serial section reconstructions. *The Journal of Comparative Neurology*, 303(4), 617–

636. <https://doi.org/10.1002/cne.903030408>
- [37] Beier, C., Hovhannyan, A., Weiser, S., Kung, J., Lee, S., Lee, D. Y., Huie, P., Dalal, R., Palanker, D., & Sher, A. (2017). Deafferented Adult Rod Bipolar Cells Create New Synapses with Photoreceptors to Restore Vision. *The Journal of Neuroscience*, 37(17), 4635–4644. <https://doi.org/10.1523/JNEUROSCI.2570-16.2017>
- [38] Doddi, G. V. (2022). Eye_diseases_classification [dataset]. <https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification>
- [39] Bolinger, M., & Antonetti, D. (2016). Moving Past Anti-VEGF: Novel Therapies for Treating Diabetic Retinopathy. *International Journal of Molecular Sciences*, 17(9), 1498. <https://doi.org/10.3390/ijms17091498>
- [40] Parke, D. W., & Lum, F. (2018). Return to the Operating Room after Macular Surgery. *Ophthalmology*, 125(8), 1273–1278. <https://doi.org/10.1016/j.ophtha.2018.01.009>
- [41] Peña, J. S., & Vazquez, M. (2022). Harnessing the Neuroprotective Behaviors of Müller Glia for Retinal Repair. *Frontiers in Bioscience-Landmark*, 27(6), 169. <https://doi.org/10.31083/j.fbl2706169>
- [42] Kharbikar, B. N., Mohindra, P., & Desai, T. A. (2022). Biomaterials to enhance stem cell transplantation. *Cell Stem Cell*, 29(5), 692–721. <https://doi.org/10.1016/j.stem.2022.04.002>
- [43] Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: Machine intelligence approach for drug discovery. *Molecular Diversity*, 25(3), 1315–1360. <https://doi.org/10.1007/s11030-021-10217>
- [44] Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence*, 2(1), 4. <https://doi.org/10.1007/s44163-022-00022-8>
- [45] Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, 1142, 012012. <https://doi.org/10.1088/1742-6596/1142/1/012012>

- [46] R. Saravanan and P. Sujatha, "A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 945-949, doi: 10.1109/ICCONS.2018.8663155.
- [47] Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering*, 19(1), 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- [48] Y. Tang, Q. Teng, L. Zhang, F. Min and J. He, "Layer-Wise Training Convolutional Neural Networks With Smaller Filters for Human Activity Recognition Using Wearable Sensors," in *IEEE Sensors Journal*, vol. 21, no. 1, pp. 581-592, 1 Jan.1, 2021, doi: 10.1109/JSEN.2020.3015521.
- [49] Forno, E., Fra, V., Pignari, R., Macii, E., & Urgese, G. (2022). Spike encoding techniques for IoT time-varying signals benchmarked on a neuromorphic classification task. *Frontiers in Neuroscience*, 16, 999029. <https://doi.org/10.3389/fnins.2022.999029>
- [50] I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari and N. Navab, "Deeper Depth Prediction with Fully Convolutional Residual Networks," 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 2016, pp. 239-248, doi: 10.1109/3DV.2016.32.
- [51] Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
- [52] Cazzato, D., Cimarelli, C., Sanchez-Lopez, J. L., Voos, H., & Leo, M. (2020). A Survey of Computer Vision Methods for 2D Object Detection from Unmanned Aerial Vehicles. *Journal of Imaging*, 6(8), 78. <https://doi.org/10.3390/jimaging6080078>
- [53] Aji, Sani & Kumam, Poom & Siricharoen, Punnarai & Bukar, Ali & Adamu, Mohammed Sani. (2021). Deep Transfer Learning for Automated Artillery Crater Classification. *Thai Journal of Mathematics*.

19. 1068-1081.
- [54] F. Xing, Y. Xie, H. Su, F. Liu and L. Yang, "Deep Learning in Microscopy Image Analysis: A Survey," in IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 10, pp. 4550-4568, Oct. 2018, doi: 10.1109/TNNLS.2017.2766168.
- [55] Alzu'bi, A., Albalas, F., AL-Hadhrami, T., Younis, L. B., & Bashayreh, A. (2021). Masked Face Recognition Using Deep Learning: A Review. *Electronics*, 10(21), 2666. <https://doi.org/10.3390/electronics10212666>
- [56] Yeh, F.-C. (2023). Brain MRI Segmentation using Template-Based Training and Visual Perception Augmentation. <https://doi.org/10.48550/ARXIV.2308.02363>
- [57] D. Karimi, G. Nir, L. Fazli, P. C. Black, L. Goldenberg and S. E. Salcudean, "Deep Learning-Based Gleason Grading of Prostate Cancer From Histopathology Images—Role of Multiscale Decision Aggregation and Data Augmentation," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 5, pp. 1413-1426, May 2020, doi: 10.1109/JBHI.2019.2944643.
- [58] Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258. <https://doi.org/10.1016/j.array.2022.100258>
- [59] Hu, F., Xia, G.-S., Hu, J., & Zhang, L. (2015). Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sensing*, 7(11), 14680–14707. <https://doi.org/10.3390/rs71114680>
- [60] Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B. S., Waldstein, S. M., & Bogunović, H. (2018). Artificial intelligence in retina. *Progress in Retinal and Eye Research*, 67, 1–29. <https://doi.org/10.1016/j.preteyeres.2018.07.004>
- [61] Y. Zhou, B. Wang, L. Huang, S. Cui and L. Shao, "A Benchmark for Studying Diabetic Retinopathy: Segmentation, Grading, and Transferability," in IEEE Transactions on Medical Imaging, vol. 40, no. 3, pp. 818-828, March 2021, doi: 10.1109/TMI.2020.3037771.

- [62] Ting, D. S. W., Cheung, G. C. M., & Wong, T. Y. (2016). Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: a review: Global burden of diabetic eye diseases. *Clinical & Experimental Ophthalmology*, 44(4), 260–277. <https://doi.org/10.1111/ceo.12696>
- [63] Newton, M. J. (2014). The promise of telemedicine. *Survey of Ophthalmology*, 59(5), 559–567. <https://doi.org/10.1016/j.survophthal.2014.02.003>
- [64] Mackay, D. D., & Bruce, B. B. (2016). Non-mydriatic fundus photography: A practical review for the neurologist. *Practical Neurology*, 16(5), 343–351. <https://doi.org/10.1136/practneurol-2016-001443>
- [65] Monroy, G. L., & Won, J. (2017). Clinical translation of handheld optical coherence tomography: Practical considerations and recent advancements. *Journal of Biomedical Optics*, 22(12), 1. <https://doi.org/10.1117/1.JBO.22.12.121715>
- [66] Schneider, E., Zimmermann, H., Oberwahrenbrock, T., Kaufhold, F., Kadas, E. M., Petzold, A., Bilger, F., Borisow, N., Jarius, S., Wildemann, B., Ruprecht, K., Brandt, A. U., & Paul, F. (2013). Optical Coherence Tomography Reveals Distinct Patterns of Retinal Damage in Neuromyelitis Optica and Multiple Sclerosis. *PLoS ONE*, 8(6), e66151. <https://doi.org/10.1371/journal.pone.0066151>
- [67] Ang, M., Baskaran, M., Werkmeister, R. M., Chua, J., Schmidl, D., Aranha Dos Santos, V., Garhöfer, G., Mehta, J. S., & Schmetterer, L. (2018). Anterior segment optical coherence tomography. *Progress in Retinal and Eye Research*, 66, 132–156. <https://doi.org/10.1016/j.preteyeres.2018.04.002>
- [68] M. D. Abràmoff, M. K. Garvin and M. Sonka, "Retinal Imaging and Image Analysis," in IEEE Reviews in Biomedical Engineering, vol. 3, pp. 169-208, 2010, doi: 10.1109/RBME.2010.2084567.
- [69] Považay, B., Hofer, B., Hermann, B., Unterhuber, A., Morgan, J. E., Glittenberg, C., Binder, S., & Drexler, W. (2007). Minimum distance mapping using three-dimensional optical coherence tomography for

- glaucoma diagnosis. *Journal of Biomedical Optics*, 12(4), 041204. <https://doi.org/10.1117/1.2773736>
- [70] N. M. Müller and K. Markert, "Identifying Mislabelled Instances in Classification Datasets," 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 2019, pp. 1-8, doi: 10.1109/IJCNN.2019.8851920.
- [71] Paulauskas, N., & Bagdonas, A. F. (2015). Local outlier factor use for the network flow anomaly detection. *Security and Communication Networks*, 8(18), 4203–4212. <https://doi.org/10.1002/sec.1335>
- [72] Ilyas, I. F., & Rekatsinas, T. (2022). Machine Learning and Data Cleaning: Which Serves the Other? *Journal of Data and Information Quality*, 14(3), 1–11. <https://doi.org/10.1145/3506712>
- [73] Carletti, M., Terzi, M., & Susto, G. A. (2023). Interpretable Anomaly Detection with DIFFI: Depth-based feature importance of Isolation Forest. *Engineering Applications of Artificial Intelligence*, 119, 105730. <https://doi.org/10.1016/j.engappai.2022.105730>
- [74] Karimi, D., Dou, H., Warfield, S. K., & Gholipour, A. (2020). Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis. *Medical Image Analysis*, 65, 101759. <https://doi.org/10.1016/j.media.2020.101759>
- [75] Cheng, Z., Zou, C., & Dong, J. (2019). Outlier detection using isolation forest and local outlier factor. *Proceedings of the Conference on Research in Adaptive and Convergent Systems*, 161–168. <https://doi.org/10.1145/3338840.3355641>
- [76] Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>
- [77] S. Yadav and S. Shukla, "Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification," 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 2016, pp. 78-83, doi: 10.1109/IACC.2016.25.

- [78] S. Yadav and S. Shukla, "Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification," 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 2016, pp. 78-83, doi: 10.1109/IACC.2016.25.
- [79] Sonarra, W., Vongmanee, N., Wanluk, N., Pintavirooj, C., & Visitsattapongse, S. (2022). Detection and Classification of COVID-19 Chest X-rays by the Deep Learning Technique. In 2022 14th Biomedical Engineering International Conference (BMEiCON).IEEE. <https://doi.org/10.1109/bmeicon56653.2022.10012094>
- [80] CA: University Science, 1989 Taufiqurrahman, S., Handayani, A., Hermanto, B. R., & Mengko, T. L. E. R. (2020). Diabetic Retinopathy Classification Using A Hybrid and Efficient MobileNetV2-SVM Model. 2020 IEEE REGION 10 CONFERENCE (TENCON), 235–240. <https://doi.org/10.1109/TENCON50793.2020.9293739>
- [81] Patel, R., & Chaware, A. (2020). Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy. 2020 International Conference for Emerging Technology (INCET), 1–4. <https://doi.org/10.1109/INCET49848.2020.9154014>
- [82] Yıldırım, H., ÇeliKer, Ü., Güngör Kobat, S., Dogan, S., Baygin, M., Yaman, O., Tuncer, T., & Erdağ, M. (2022). An automated diabetic retinopathy disorders detection model based on pretrained MobileNetv2 and nested patch division using fundus images. Journal of Health Sciences and Medicine, 5(6), 1741–1746. <https://doi.org/10.32322/jhsm.1184981>