INTELLIGENT SYSTEM FOR EARLY CATARACT DETECTION

BY

AGU JOHN DAVY CHIDERA

U20/NAS/CSC/256

DEPARTMENT OF COMPUTER SCIENCE AND MATHEMATICS, FACULTY OF NATURAL SCIENCES AND ENVIRONMENTAL STUDIES,

GODFREY OKOYE UNIVERSITY, ENUGU STATE. **IN PARTIAL FULFILLMENT OF THE AWARD OF BACHELOR OF**

SCIENCE (B.SC), DEGREE IN COMPUTER SCIENCE.

SUPERVISOR: ENGR. DR. LUCY EZIGBO

JUNE, 2026

APPROVAL PAGE

This research, titled "**INTELLIGENT SYSTEM FOR EARLY CATARACT DETECTION"** has been assessed and approved by the Department of Computer Science and Mathematics Committees in the Faculty of Natural Sciences and Environmental Studies, Godfrey Okoye University, Enugu, Nigeria.

	....	....
Engr.Dr. Lucy Ezigbo	Signature	Date
Project Supervisor		
External Examiner	....	....
External Examiner	Signature	Date
Dr. Frank Okebanama	....	....
HOD, Department of Computer Science and Mathematics	Signature	Date
Prof.Dr. I A Ajah	....	....
Dean, Faculty of Computing And Information Technology.	Signature	Date

CERTIFICATION

I certify that I completed the research included therein and that it was primarily based on my observations and investigation. Consequently, I will fully accept the University's decision if I am found to have cheated on the assignment.

____

AGU JOHNDAVY CHIDERA

U20/NAS/CSC/256

DEDICATION

This project is dedicated to God Almighty for His grace and guidance, to my family for their unwavering support, and with special recognition to my late father, who initiated this journey.

ACKNOWLEDGMENTS

First and above all, I give thanks to God for the strength to keep going when the work felt endless, for clarity when nothing made sense, and for seeing me through to this point. None of this would exist without His grace.

To my Head of Department, Dr Frank Okebanama, I am grateful for the direction and academic environment that made this project possible. Your commitment to the department sets a standard that pushes students to take their work seriously.

To my project supervisor, Engr. Dr Lucy Ezigbo, thank you not just for the guidance, but for the patience. You asked the right questions, pointed out what I had missed, and never let me settle for less than my best. This project is sharper because of you.

To the lecturers who taught me across these years: Dr Ajah (Dean), Dr Okebanama (HOD), Prof Okereke, Engr Kingsley and others whose names I cannot mention here, I did not always appreciate it at the moment, but looking back, the foundation you built shows up in every part of this work. Thank you.

To my coursemates and my fellow students, the ones who stayed late in the faculty, shared notes without being asked, and turned what could have been a stressful season into something actually enjoyable. I am glad we went through this together.

To everyone else who played even a small part in this project, whether through advice, a conversation, or just encouragement at the right time, it meant more than you know.

And to my family: thank you for never once making me feel like this was too much to pursue. Your support through the long nights, the pressure, and the uncertainty was the kind that cannot be repaid, only carried forward.

ABSTRACT

Cataracts have been found among the main causes of avoidable blindness worldwide and form a large portion of visual impairment in sub-Saharan Africa owing to the delayed diagnosis and lack of specialist eye services. Early identification is very important for the proper action to be taken in order to stop the progression of the problem. In this study, we introduce CataractAI, which is a web-based intelligent cataract detection platform that allows users to sign up and authenticate, and upload their retinal fundus images for automatic cataract screening. For increasing the efficiency of learning from an unbalanced data set, Synthetic Minority Over-Sampling Technique (SMOTE) was used during training to create synthetic examples for the minority class to make class distribution balanced and improve the classification process. The performance of the trained model was tested using a separate test set with 750 retinal fundus images. According to experimental results, the proposed model successfully classified 747 out of 750 retinal images. Thus, it achieved an accuracy of 99.60%, precision of 99.60%, recall of 99.60%, and an F1-score of 99.60%. Also, the model obtained the ROC-AUC score equal to 99.97%, which means the excellent discriminating ability between cases with cataracts and without them. The developed system can be used in primary health care centers, rural clinics, and eye care services in Nigeria and the sub-Saharan Africa region generally, where there is a lack of access to ophthalmologists and diagnostic equipment. With the ability to detect cataracts accurately and efficiently, the system has the potential to assist healthcare providers in detecting cataracts early, refer them for appropriate treatment and help in preventing blindness due to cataracts. The results indicate that the use of a deep learning system based on EfficientNetB0 is a practical and accurate method for automated cataract detection.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

The sense of vision, which is the most prominent among all other senses, is integral to all stages of human life. Visual impairments have profound effects on day-to-day activities because visual information is critical for understanding and navigating the environment around us (World Health Organization, 2023). There are various eye diseases that cause visual impairments; among these are cataract, glaucoma, refractive errors like astigmatism, hyperopia, myopia, presbyopia, macular degeneration, and retinopathy caused by diabetes (National Eye Institute, 2024). Of these, cataract has been identified as the major cause of blindness and visual impairments across the world (WHO, 2023). There is also evidence suggesting that cataract cases are increasingly being found among younger patients due to risk factors including genetics, diabetes, smoking, and exposure to ultraviolet radiation (American Academy of Ophthalmology, 2022; NEI, 2024).

Cataract, according to the National Institutes of Health, is a common, often age-related eye condition where the normally clear lens becomes cloudy or opaque, blocking light from reaching the retina and resulting in blurry, hazy, or reduced vision. It typically progresses slowly, causing symptoms like fading colors, glare, and poor night vision. There are a number of factors that can cause or increase the risk of cataract; these includes: Age, Genetic disorders, Eye injury, Smoking, Exposure to U.V light, diabetes, obesity and uveitis. Commonly, most cases of cataracts occur in older people, but they can appear at any age. Cataracts in children are known as “Childhood Cataracts”, some babies are born with cataracts, frequently due to environmental or genetic disorders. This is known as “Congenital” cataracts, if cataracts in children are not treated on time, the child becomes permanently blind as their vision fails to develop properly. (Orbis International, n.d.).

Cataracts are the leading cause of blindness worldwide, affecting over 95million people, with nearly half of all global blindness cases in low-income regions caused by this condition. In Nigeria, Cataracts account for 43% of blindness and 19.8% of visual impairment among adults aged 40+, frequently linked to aging, U.V light exposure, and lack of health care access. (Adeniyi et al., 2023). Early detection of cataracts is crucial for maintaining quality of life, as it allows monitoring of progression, lifestyle adjustments, and planning for optimal surgery timing, it helps prevent severe vision loss, reduce risks of complications like glaucoma (*Which is a group of eye disease that damage the optic nerve, essential for good vision often caused by abnormally high pressure inside the eye*) and enables the use of advanced lenses, ensuring better surgical outcomes and quicker recover.(North Florida Cataract Specialists and Vision Care, 2024).

The human lens is responsible for focusing light onto the retina to produce clear images. In cataract, protein aggregation within the lens leads to clouding, which progressively diminishes visual acuity. Cataracts are classified into nuclear, cortical, and posterior subcapsular types based on their anatomical location within the lens (Asbell et al., 2005). If left untreated, the condition can advance to complete blindness; however, the condition can advance through surgical intervention, making early detection critically important to improving patient outcomes.

The traditional approach to cataract diagnosis involves a comprehensive clinical eye examination performed by an ophthalmologist or optometrist to detect lens opacities and assess their impact on vision, it involves slit-lamp biomicroscopy and direct ophthalmoscopy by ophthalmologist, faces significant accessibility challenges, particularly in low and middle-income countries and rural areas. These challenges hinder timely detection, leading to high rates of preventable cataract-related blindness. Truong et al. (2025)

In recent years, Artificial Intelligence, including machine learning and deep learning has made a great impact on society worldwide, this is simulated by the advent of powerful algorithms, exponential data growth, and computing hardware advances. In the medical and healthcare fields, numerous studies have validated that artificial Intelligence exhibited robust performance in disease diagnoses and treatment response prediction, for example, a deep learning systems developed on computed tomography (CT) images can distinguish patients with (COVID 19) pneumonia from patients with other common types of pneumonia and normal controls with an area under the curve (AUC) of 0.971. A machine learning model trained using dual-energy CT radiomics provides a significant additive value for response prediction (AUC=0.75) in metastatic melanoma patients prior to immune therapy. (Saha et al., 2021).

Artificial Intelligence (AI), and specifically deep learning, has emerged as a transformative technology in medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in detecting a wide range of ocular conditions, including diabetic retinopathy, glaucoma, and age-related macular degeneration, often matching or surpassing human-expert performance (Gulshan et al.,2016; Ting et al., 2017). These developments have naturally extended to cataract detection, where fundus and anterior segment images serve as rich sources of diagnostic information.

EfficientNetB0 is a CNN model architecture that was introduced by Tan and Le (2019) and can be considered as one of the latest advances of CNN architectures, where better accuracy is achieved through a compound scaling approach that is applied to all three attributes width, depth and resolution. In contrast to classical CNN architectures like VGG-16, RESNet, and InceptionV3, which have a larger number of parameters, EfficientNetB0 provides competitive or even higher accuracy using less computational resources and is a good candidate for usage in the environment with limited resources (Tan & Le, 2019).

EfficientNetB0 has gained a lot of attention in recent years but only a limited number of studies have been undertaken with respect to the application of EfficientNetB0 in the early diagnosis of cataracts utilizing easily available image data. This study has tried to bridge the gap by designing an intelligent diagnostic system based on EfficientNetB0 to classify images as cataracts and non-cataracts.

1.2 STATEMENT OF THE PROBLEM

Blindness resulting from cataracts is still a huge public health problem especially in developing countries, where the provision of ophthalmology care is highly insufficient compared to the demands of the population. Nigeria, which is the largest country in Africa, has an ophthalmologist-population ratio of about 1:500,000, less than the recommended ratio of 1:50,000 by WHO (National Blindness and Visual Impairment Survey in Nigeria, 2007; Kyari et al., 2009). The high scarcity makes universal eye testing difficult via traditional methods.

Existing clinical techniques for diagnosing cataracts like slit lamp examination, direct ophthalmoscopy, and Snellen chart test for visual acuity depend highly on the expertise available and subjective interpretation by the expert clinicians (Pascolini & Mariotti, 2012). This technique cannot be used for conducting mass screening as it is very difficult to access for the rural and unprivileged population. Moreover, patients usually come at a much later stage of cataract formation when there is already considerable loss of vision, and the cost involved becomes higher.

Although several machine learning algorithms as well as deep learning methods have been developed for automated cataract diagnosis, most of the current methods suffer from problems such as large amounts of required annotated data, high computational cost, lack of generalization on other imaging types, or early detection of the disease (Li et al., 2019; Zhang et al., 2014). Therefore, there is an urgent need for developing efficient deep learning algorithms that can diagnose early-stage cataracts based on common ocular imaging.

The objective of this research is to fill this gap by developing an efficient cataract detection system using EfficientNetB0 CNN model, which is a highly performing but computational-efficient model, on an open-source ocular images dataset. The proposed system will assist in the accurate diagnosis of early stages of cataract disorders.

1.3 AIM AND OBJECTIVES OF THE STUDY

The aim of this study is to design an efficientNetB0 based intelligent system for the early detection of cataracts to enable proper and accurate diagnosis.

OBJECTIVES

1. Review of previous deep learning models
2. Preparation of fundus images
3. Design and development of the efficientNetB0 based deep learning model
4. Training, validating and testing of the proposed model
5. Performance analysis of the developed model

1.4 SIGNIFICANCE OF THE STUDY

This research is scientifically, clinically, and economically valuable. First, from a scientific point of view, it adds up to the ever-growing amount of information regarding the usage of efficient neural network architectures in ophthalmic image analysis. Second, through the systematic application of EfficientNetB0, which is a model that strikes an optimal balance between accuracy and efficiency, to early cataract detection, the researchers provide evidence of the models' effectiveness in performing binary ophthalmic classification.

From the clinical point of view, the suggested system has the capability to act as a decision support system for health care practitioners that are involved in primary and community health care without any professional education in ophthalmology. Timely identification of cataract patients helps to refer the patient for surgery before they lose their eyesight permanently. This is particularly true for Nigeria and other sub-saharan African countries that have high incidence rates of blindness due to cataracts (Flaxman et al., 2017; Kyari et al., 2009).

From the perspective of public health, an automated and low-cost system for screening contributes to the goals of the WHO Vision 2020: The Right to Sight campaign and subsequently, The World Report on Vision (2019) that focuses on integrating technologies into the eyesight care system to prevent avoidable blindness (WHO, 2019). Utilization of such a system through mHealth or telemedicine can significantly increase accessibility to eye screening in rural areas.

In addition, the research advances the discussion on effective deployment of AI models in the health care sector since the state of the art performance can only be achieved through the use of too much computational power, a major issue for deployment in developing nations.

1.5 SCOPE OF THE STUDY

The study falls within the below scope. In terms of subject area, the study only involves the identification and binary classification of cataract (presence of cataract or absence of cataract) through ocular images with the aid of the EfficientNetB0 model. The study will not venture into the surgical management, grading of cataract, and recommendations for treatment.

As for the data, the experiment uses an open labelled ocular image data set that includes fundus and/or anterior segment images. The data set involves images of the eyes suffering from cataracts and the normal eyes. There is no need for collecting any primary clinical data since the work is done only using the data provided.

In terms of geography, although the model is based on international data sets, the study is inspired by and intended for use in the Nigerian/African healthcare system. Field validation of the proposed model in a clinical setting is not included in the study.

In terms of technology, the study uses python-based deep learning platforms (TensorFlow and Keras) to develop models, and the scope is restricted to image classification through supervised learning. Imaging technologies like optical coherence tomography (OCT), ultrasound biomicroscopy, or Scheimpflug imaging are not considered in the study. Furthermore, the proposed model has not been deployed for clinical applications.

1.6 LIMITATIONS OF THE STUDY

Despite the benefits of this research, there are some limitations that must be considered. First, this study makes use of an open-access dataset, which does not cover the full range of phenotypic diversity in terms of cataracts. This is due to dataset heterogeneity and class imbalance, thus limiting the generalization of the trained model in the unseen clinical populations (Shen et al., 2017).

The second limitation of the model is its dependence on the accuracy of labeling in the training data set. Inaccurate or inconsistent labeling will affect the training process negatively and will also affect evaluation metrics. Without additional labeling from an expert, the accuracy of labels in the data set cannot be guaranteed (Litjens et al., 2017).

Third, even though transfer learning through pre-trained weights from imageNet is used to overcome issues related to data deficiency, the fact that the source domain (images from nature for ImageNet pre-training) and target domain (ocular medical images) have differences in their domains may hinder deep feature learning. The medical imaging database differs significantly from the natural images database in respect of color distribution, textures, and diagnostic features.

Fourth, the study lacks any prospective clinical validation of the designed algorithm. It is only through the prospective clinical validation of the designed algorithm using real-world data that its true utility can be established. This would require conducting clinical validation studies with the help of ophthalmologists and general health workers.

Fifth, the binary classification model (cataract vs. non-cataract) applied in this study does not reflect the multi-class problem of cataract stages (early, intermediate, and mature/hypermature). Early-stage diagnosis, which is the core of this study, calls for a fine-grained distinction which binary classification model may not fully meet (Zhang et al., 2014)

Lastly, the computational infrastructure employed in the development and testing of the models on (GPU enabled workstation) might not be accessible in the low resource environment to which the system is eventually meant to be deployed, and where no testing has been done on the system's performance using inexpensive hardware.

1.7 OPERATIONAL DEFINITION OF TERMS

1. Cataract: It's a common, often age-related eye condition where the normally clear lens becomes cloudy or opaque blocking light from reaching the retina and resulting in blurry, hazy or reduced vision.
2. Early Detection: It denotes the automated identification of cataract signs from ocular images at a stage where the condition may not yet have produced severe visual impairment. The system is considered to perform early detection when it correctly classifies cataract positive images, including those exhibiting early stage lens changes, with high sensitivity.
3. Artificial Intelligence: It refers to technology that enables devices to simulate human-like abilities such as learning, reasoning, problem-solving, decision-making, and creativity.

4. Deep Learning: It is a powerful subset of device learning that uses multi-layered neural networks inspired by the human brain to model complex patterns in data
5. Machine Learning: It is a subset of artificial Intelligence that enables systems to learn from data, identify patterns, and make decisions with minimal human intervention.
6. CNN: Convolutional Neural Networks, is a deep learning model specialized for processing grid structure data, particularly images.
7. EfficientNetB0: It is a light weight, highly efficient convolutional neural network designed for image classification.
8. Data Set: It is a structured collection of related data, typically organized in rows and columns, representing information about a specific subject.
9. Transfer Learning: it refers to the process of initializing the EfficiNetB0 model with weights pre-trained on the ImageNet dataset (comprising over one million natural images across 1,000 categories) and subsequently fine tuning the model on the domain specific ocular image dataset. This approach leverages generalized visual features learned from large scale data to compensate for the limited size of medical imaging dataset.
10. Intelligent System: Is a software based diagnostic tool that integrates a trained EfficientNetB0 deep learning model with image preprocessing and inference pipelines to automatically analyse ocular images and output a cataract classification decision without human intervention. The system is considered "Intelligent" by virtue of its capacity to learn discriminative features from training data and apply those learned patterns to unseen images.
11. Ocular image: it is a digitally captured photograph of the eye, specifically a fundus image (Showing the interior of the eye including the retina) or an anterior segment image (Showing the front structures of the eye, including the lens). Such images constitute the primary input data for the developed classification system.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

The following chapter provides an exhaustive review of literature concerning the development of an intelligent system for early cataract detection using the EfficientNetB0 deep learning model. This review is intended to offer a theoretical and empirical basis for this study, starting from the theoretical background of deep learning up to the specifics of the cataract disease, its diagnosis, and finally to the recent developments in deep learning application to ophthalmology. Finally, a critical analysis of fifteen recent papers on deep learning application for cataract diagnosis is offered in a tabular form, after which the research gap for this study will be determined.

Literature that has been generated within the last five years (2020-2024) has been utilized to write this review. Literature that is considered fundamental has been added in cases where it is required to understand the background of the theory under consideration. Studies have been obtained from peer-reviewed journals and conference proceedings such as IEEE Access, Medical Image Analysis, The Lancet Global Health, Computers in Biology and Medicine, and Frontiers in Medicine.

2.2 THEORETICAL BACKGROUND

2.2.1 Artificial Intelligence and Machine Learning

AI represents an interdisciplinary area of computer science that involves developing software capable of carrying out processes requiring human intelligence, such as perception, reasoning, and decision making (LeCun et al., 2015). Machine learning (ML), a discipline within AI, is a scientific domain focused on creating algorithms that allow computers to perform tasks using data without having them hard-coded into the computer beforehand. Classic machine learning methods like SVMs, decision trees, random forests, and k-nearest neighbor classifiers have been applied in medical image processing but require manually extracting features which are deemed important for classification by domain experts.

The move from manually engineered features to automatically learned representation has been one of the biggest breakthroughs in recent times in the field of AI. Deep Learning, a branch of ML that utilizes multi-layered artificial neural networks, accomplishes the above mentioned automatic feature learning process and has shown incredible success in applications which deal with high dimensional data including images, audio, and text (LeCun et al., 2015). In the field of medicine, deep learning has been considered the better technique for diagnostic purposes via images due to its ability to learn non-linear mappings.

2.2.2 Convolutional Neural Networks (CNNs)

The Convolutional Neural Network is the most used type of deep learning network used for image recognition and classification problems. The CNN consists of several convolutional layers that use filters to identify spatial features like edges, textures, and shapes; pooling layers that decrease spatial dimensions, thus making the network translationally invariant while reducing computation complexity; and finally the fully connected layers that combine the identified features and perform the final classification task (Litjens et al., 2017).

Convolution operation can be mathematically expressed as follows: The value of the feature map Y at position (i,j) is equal to the sum over the window of the product of input X and the filter weights W with the addition of bias b. Local connection and weight sharing make CNNs far more efficient than fully connected networks when it comes to processing images. Architectures like AlexNet, VGGNet, GoogLeNet/Inception, ResNet, DenseNet, and EfficientNet have successively improved upon the state-of-the-art by incorporating residual connections, dense connections, and compound scaling among other advancements (Tan & Le, 2019).

2.2.3 Transfer Learning

The transfer learning technique in machine learning involves training a model using a large dataset of the source domain (ImageNet), followed by tuning the model for application in a related target domain (Tajbakhsh et al., 2016). In the domain of medical imaging, where labeled datasets are hard to obtain and costly, transfer learning has become inevitable. The pre-trained models have already acquired general features that include edge, color histogram, texture, and shape feature extraction from images.

There are two main methods used: feature extraction method where the convolutional layers of the pre-trained model are frozen and just the final layer classification is retrained using the new dataset; and fine-tuning where the entire or partial layers of the pre-trained model are unlocked and retrained using a very low learning rate. In a study by Tajbakhsh et al. (2016), it was shown that fine-tuned CNNs perform better than CNNs that were trained from scratch on small medical imaging datasets.

2.2.4 EfficientNet Architecture

EfficientNet, developed by Tan and Le (2019), is one example of principled improvement in the design of CNN architectures. Before the development of EfficientNet, there have been attempts to scale up CNNs by either increasing their depth, widening their width, or raising input image resolution. In EfficientNet, all three aspects are scaled together using certain scaling factors that have been obtained through neural architecture search.

The baseline model, EfficientNetB0, is constructed using NAS to optimize for accuracy and efficiency under a mobile-scale FLOP constraint. It employs Mobile Inverted Bottleneck Convolution (MBConv) blocks with Squeeze-and-Excitation (SE) optimisation, which allows the network to learn channel-wise feature recalibration. Scaled variants (B1 through B7) apply progressively larger compound scaling coefficients, with each offering higher accuracy at greater computational cost. EfficientNetB0 achieves 77.1% top-1 accuracy on ImageNet with only 5.3 million parameters a fraction of the parameters required by VGG-

16 (138 million) and ResNet-50 (25 million) making it the most computationally efficient choice for deployment in resource-constrained healthcare settings (Tan & Le, 2019).

2.3 OVERVIEW OF CATARACT DISEASE

2.3.1 Aetiology and pathophysiology

The human eye's crystalline lens is an opaque, biconvex structure located behind the iris, which mainly consists of water and proteins known as crystallins. The structure of the crystallin proteins in the healthy eye lens is very organized and structured in such a way that ensures the transmission of light rays (Asbell et al., 2005). When the crystalline structure in the eye lens changes in a manner that causes them to aggregate and scatter light, cataract occurs.

Aging is the main cause; here, the combination of oxidative stress, UV rays and protein denaturation causes damage to the proteins in the lens. Other causes include diabetes mellitus (which causes increased glycosylation of the lenses), the use of corticosteroids for a long time, trauma, ionizing radiation, systemic conditions and genetic causes. In Nigeria and other Sub-Saharan countries, there is increased incidence of cataracts due to increased exposure to UV rays and poor nutrition (Kyari et al., 2009).

2.3.2 Classification of Cataracts

There is an anatomical classification of cataracts, which is based on the location of opacity of the lens. The nuclear cataract is caused by the opacification of the nucleus of the lens and is the most prevalent type of cataract among elderly people; it is accompanied by the change to myopia and decreased contrast sensitivity. The cortical cataracts include the opacification of the cortical layers of the lens and are characterized by symptoms of glare and monocular diplopia. The posterior subcapsular cataracts (PSC) are located at the back surface of the lens capsule; they affect the near vision more and develop faster than other types of cataracts.

In terms of functionality, cataracts can be classified as being early (with mild opacification and with little effect on the vision), intermediate (moderate opacification and measurable reduction of vision) and mature/hypermature (complete opacification leading to severe or total blindness). It is important to note that early diagnosis plays a crucial role as treatment done during the early stage has more success stories (WHO, 2023).

2.3.3 Global and Regional Epidemiology

Cataract is still the most common cause of avoidable blindness in the world, accounting for about 51% of all blindness in the world. This affects about 65.2 million people (WHO, 2023). Sub-Saharan Africa disproportionately suffers from this problem: the incidence of blindness from cataracts in Nigeria was found to be 43% in the National Blindness and Visual Impairment Survey (Kyari et al., 2009). However, although cataract is easily treatable by surgery, such as phacoemulsification and extracapsular cataract extraction (ECCE), the cataract surgical coverage in Nigeria is extremely low, estimated at only 400-600 operations per million population, way below the WHO goal of 2,000 per million (Flaxman et al., 2017).

2.4 CONVENTIONAL METHODS OF CATARACT DISEASE DIAGNOSIS

2.2.1 Visual Acuity Testing

Snellen charts and their derivatives form the simplest tool to screen for cataracts. The best corrected visual acuity (BCVA) is evaluated at the fixed distance, which is 6 metres normally, and any BCVA less than 6/18 in the better eye serves as the definition of visual impairment set by WHO. Although universally available, visual acuity assessment is not specific since it helps only to detect visual impairment functionally but cannot determine its etiology, distinguish cataracts from other causes of visual loss or detect disease at an early stage.

2.4.2 Slit-Lamp Biomicroscopy

Slit-lamp biomicroscopy serves as the gold standard for the diagnosis and grading of cataracts. Slit lamp provides illumination to the anterior portion of the eye through an intense narrow beam of light, permitting direct visualization of the kind, opacity level, and positioning of lens opacities by the ophthalmologist. Grading scales such as the Lens Opacity Classification System III (LOCS III) have been devised to grade the severity of cataracts based on slit-lamp pictures. However, although it is very precise, slit-lamp biomicroscopy demands specialized equipment and considerable training that is lacking in developing countries (Pascolini & Mariotti, 2012).

2.4.3 Ophthalmoscopy and fundus Photography

Ophthalmoscope examination can help visualize the posterior part of the eye and can help detect the extent of opacification in the lens by examining the red reflex, which is the reddish-orange reflection of light from the retina via the pupil. The presence of a reduced or non-existent red reflex can point to opacification of the lens. Fundus cameras that photograph the retina and the posterior pole are becoming more popular for telemedicine screenings but are not traditionally employed for diagnosing cataracts. Despite this, fundus photographs have found increasing use in deep-learning based cataract detection research because of their availability in existing retinal screening datasets.

2.4.4 Limitations of Conventional Methods

There are three main limitations of the traditional ways used in the diagnosis of cataract that include; need for expert knowledge of an ophthalmologist, expensive and non-portable equipment needed in performing the procedure, and subjective nature of the process. Agreement among ophthalmologists on the grading system is moderate (kappa inter-rater of LOCS III is about 0.6 to 0.75). In addition, such techniques cannot be used for mass screening, especially in rural areas where there is a need for such programs. This is a good reason to develop intelligent diagnostic systems.

2.5 DEEP LEARNING AND ITS APPLICATION IN HEALTHCARE

2.5.1 Deep Learning in Healthcare

The impact of deep learning technology has revolutionized the field of computer-aided diagnosis (CAD) in many different medical specialties. For instance, CNN technology has proved to have a comparable or even better level of performance than the radiologists' when it comes to detecting lung nodules in CT scans, diagnosing fractures in plain radiographs, and diagnosing skin lesions in dermoscopic images (Litjens et al., 2017). Pathology is another specialty where the application of deep learning models in interpreting whole-slide histopathology images has proved to have an expert-level accuracy in grading of cancer and diagnosing lymph node metastases.

Deep learning adoption in clinical practice has a number of benefits compared to conventional CAD techniques: absence of manual feature extraction, ability to recognize complicated non-linear dependencies from the raw data, scalability with growing size of data sets, and handling of various data modalities such as imaging, time-series, textual, and genomic data. Among the difficulties are the requirement for large annotated data sets, distribution shift sensitivity of the technique, lack of model interpretability ('black box' issue), and clinical deployment challenges (Litjens et al., 2017).

2.5.2 Deep Learning in Ophthalmology

Ophthalmology has proved itself as one of the areas which is rich in potential application of Deep Learning because of its image-based nature of diagnoses. A paper by Gulshan et al. (2016) shows that a CNN applied to 128,175 retinal images had reached sensitivity level 97.5% and specificity level 93.4% in detection of diabetic retinopathy with referable characteristics, showing results equal to the performance of US board-certified ophthalmologists. This paper launched a great interest in applying Deep Learning to ophthalmic images analyses.

Further work was done to detect glaucoma, age-related macular degeneration (AMD), retinal vein occlusion, hypertensive retinopathy, and several other diseases with high accuracy through automatic detection of their presence in the images. Ting et al. (2017) have presented a deep learning system that has been trained on 76,370 retinal images taken from the multiethnic diabetic patients in Singapore and managed to show AUC above 0.93 for multiple retinal diseases simultaneously. All these works together proved that deep learning can be used to perform clinically useful image classifications on ophthalmic images.

Deep learning for cataract detection is a logical step forward from this research area. Cataracts differ from other eye diseases like diabetic retinopathy and AMD in that they affect the anterior segment of the eye while the latter affect the posterior segment. Therefore, they require different imaging techniques for their diagnosis including slit lamp photography and anterior segment OCT. Deep learning algorithms have proven their capability to discriminate between healthy and cataract-affected images of the anterior and posterior segments of the eye.

2.6 DEEP LEARNING FOR CATARACT DETECTION

The use of deep learning for automated cataract detection is one of the areas that have experienced considerable research interest during the past decade, owing to the large amount of eye imaging data available and the ability of convolutional neural networks to classify images. The early research work done in this area has used traditional machine learning methods such as SVMs and random forests that make use of hand-crafted texture and colour histogram features respectively (Ran et al., 2021).

CNN-based cataract detection systems have been proposed based on different backbone models, namely VGGNet, GoogLeNet, ResNet, DenseNet, InceptionV3, MobileNet, and most recently EfficientNet and Vision Transformers. Some of the different modalities where the

CNN-based systems have been used include slit-lamp images, fundus photographs, AS-OCT images, and smartphone-captured images. Consistently used performance evaluation parameters in the literature include accuracy, sensitivity, specificity, F1-score, and AUC-ROC with the best performing systems attaining accuracy levels around 91-96%.

Some important methodological innovations that have taken place within the domain include the application of transfer learning using pre-trained weights from ImageNet to address the problem of scarcity of labelled medical data; data augmentation methods to facilitate better generalization; multi-task learning algorithms which can tackle both detection and severity grading at once; and finally, the inclusion of explanation capabilities like Grad-CAM in order to make the results more interpretable for clinicians. However, there are still many shortcomings in this regard, especially in terms of detecting early stage cataracts, building lightweight models for low-resources and external validation on heterogeneous data.

2.7 REVIEW OF RELATED WORKS

The next portion includes a critique of fifteen scholarly papers published in 2021-2024 related to the use of deep learning in detecting and assessing cataracts, selected based on their relevance to the topic under discussion, robust methodology, recency, and contribution to the domain. The research contribution, methodology, dataset, results, and limitations of each paper have been discussed in detail.

Ran et al. (2021)

Ran et al. (2021) presented an approach to cataract grading using a combination of deep CNNs and random forests on slit-lamp images. In this study, a private dataset consisting of 48,448 slit-lamp images belonging to four cataract severity grades from several ophthalmology clinics in China was used. In this study, the CNN model extracted visual features hierarchically that were further fed to the random forest classifier to predict severity grades. This hybrid model achieved a classification accuracy of 92.3%, outperforming CNN-only models in terms of robustness. A multi-grade cataract grading model along with the applicability of hybrid CNN-classical machine learning models in this context was shown. But the limitation of this model was that it could be only tested on slit-lamp images, hence, its application in places where no slit lamp is available is not possible.

Zhang et al. (2021)

Zhang et al. (2021) have created a multi-task learning approach that can perform concurrent detection and grading of cataract using AS-OCT images. The multi-task learning algorithm uses ResNet-50 architecture, with a two-headed output consisting of a binary classifier for detecting whether the patient has cataract and a four-category classifier to grade the disease. The method uses transfer learning from ImageNet, which is then fine-tuned on a data set of 12,600 AS-OCT images. The multi-task model obtained an accuracy score of 94.1% for detection and an AUC score of 0.97 for grading, which were better than single-task approaches due to the shared representation. The drawback of this approach is the dependence on AS-OCT images, which are costly and unavailable in community healthcare settings in developing countries.

Li et al. (2022)

The study conducted by Li et al. (2022) focuses on developing a lightweight MobileNetV2-based architecture for cataract screening on mobile and edge devices to overcome the accessibility issue in telemedicine systems. The depthwise separable convolution used in MobileNetV2 significantly reduces the number of parameters and the inference time in comparison with conventional convolution. On a dataset of fundus images, augmented intensively, the accuracy of the model for binary classification was 91.7%. While being efficient, the model showed a lower accuracy level by 2-4% in comparison with heavy models like ResNet-50 and InceptionV3. Early detection was not covered, and no clinical validation was performed.

Gao et al. (2022)

Gao et al. (2022) examined the use of Vision Transformers (ViT-B/16) architecture for the diagnosis of cataracts, which is fundamentally different from the traditional CNN-based models. Using ViT architecture that was pre-trained on ImageNet-21k dataset and fine-tuned using anterior segment images, the model obtained an accuracy rate of 95.4%, which is better than the ResNet-50 model (92.1%) and InceptionV3 model (91.6%). The paper showed the possibilities of self-attention mechanisms in detecting long-range dependencies in ophthalmic images, where local filters might fail to notice them. However, computing power consumption by ViT-B/16 is much higher than EfficientNetB0, and it needs to be pre-trained using a lot of data.

Nguyen et al. (2022)

Nguyen et al. (2022) proposed a solution to privacy issues in multi-hospital cataract screening through a federated learning method where a cataract detector based on the VGG-16 architecture is trained in five hospitals without exchanging patient images. Weights are aggregated through the use of the FedAvg method after each hospital trains the model locally with its own set of data. The federated model achieves an accuracy of 89.6%, which is around 3% lower than the one of the centralised model; the difference in performance is mainly due to non-IID distribution of data in the hospitals.

Saha et al. (2022)

The EfficientNetB3 ensemble model was built by Saha et al. (2022) for the assessment of nuclear cataract severity from slit-lamp images, where the system scored QWK = 0.87 on a four-class classification problem. Hyperparameters tuning through Bayesian optimisation technique was done in order to find optimal values for learning rate, dropout rate and batch size. This work is unique in terms of using the EfficientNet family network and hyperparameters optimization approach. Nevertheless, the proposed system was restricted to the nuclear cataract classification and failed to solve the problem of other subtypes like cortical and PSC cataracts. Early stage binary classification was not considered in this work, and no comparisons were done with other lighter EfficientNet architectures like B0.

Islam et al. (2022)

In Islam et al.'s (2022) study, Deep Convolutional Generative Adversarial Networks (DCGANs) have been used to produce artificial cataract fundus images through data augmentation to solve the issue of class imbalance in the cataract detection dataset. The artificially produced images, along with real images, were used to train a DenseNet121 classifier. Accuracy of binary classification was increased from 85.1% to 92.8% using GAN data augmentation, which shows that artificial data production is useful for solving the problem of lack of dataset. It should be noted, though, that GANs produce artefacts, which may be different from the clinically realistic images. Moreover, there was no comparison with normal augmentation techniques.

Wang et al. (2023)

Wang et al. (2023) presented a custom-made dual-branch CNN architecture designed to perform early cataract detection using smartphone-obtained images of eyes. One branch processed colour-related information (the RGB channels), whereas the other branch used Local Binary Pattern (LBP) features to extract textures, and their outputs were then combined and classified by the fully-connected classifier. This architecture resulted in 91.4% F1-score and 0.95 AUC score in a dataset of 3,800 smartphone-obtained images. Smartphone image acquisition is an advantage here since it allows collecting the data cheaply, but the results were highly dependent on illumination and imaging distance and the model was not compared to standard clinical/benchmarking datasets.

Chen et al. (2023)

The study by Chen et al. (2023) tackled the crucial problem of deep learning-based models' interpretability for cataracts detection by adding a visualization technique called Gradient-weighted Class Activation Mapping (Grad-CAM) to a ResNet-50 cataract detection algorithm. The algorithm generated attention maps representing the regions of an image which played a major role in the classification process performed by the model; they were analyzed by a group of three ophthalmologists on a qualitative level. The classification model reached 93.2% sensitivity and 91.7% specificity on a fundus dataset. Although clinical validation of Grad-CAM visualizations represented one of the strong aspects of the methodology used, the researchers stated that the Grad-CAM localizations are approximate and might not correspond to the exact localization of lens opacification.

Mahesh et al. (2023)

Mahesh et al. (2023) created and assessed a cataract screening framework that combines the InceptionV3 classification model, available through REST API within a telemedicine front-end app. The framework was tested in a rural telemedicine test in India using 1,200 participants and attained an accuracy rate of 90.3%, compared to that of an ophthalmologist. A user acceptance study was carried out among 45 community health workers, and there were positive user ratings for the system. While practical in its applications, the framework relies heavily on uninterrupted internet connections, thus being unapplicable in regions with poor network infrastructure, where there is a greater need for screening.

Yadav et al. (2023)

Yadav et al. (2023) carried out an extensive benchmarking analysis comparing seven CNN models, namely VGG-16, ResNet-50, InceptionV3, Xception, DenseNet-121, NASNet-Mobile, and EfficientNetB0, on a publicly available dataset of cataracts containing 4,000 eye images. EfficientNetB0 model performed best among competing models with the highest binary classification accuracy of 96.2% and the lowest number of parameters (5.3M). Therefore, the results of the study provided solid evidence of the efficiency of EfficientNetB0 as the most efficient model for detecting cataracts. Nevertheless, the limitations of the study include the lack of severity assessment and prospective clinical data set.

Khalil et al. (2023)

Khalil et al. (2023) introduced an architecture that combines CNN and LSTM, aiming at temporal modeling of the progression of cataracts through sequence of AS-OCT images obtained longitudinally from patients' records within 12 months. While CNN was utilized to detect spatial characteristics of each image, LSTM was used to discover temporal correlations between images in the sequence to model the disease progression pattern. The accuracy of sequence classification reached 88.9%, and there is some potential for longitudinal monitoring by means of AI. The main drawback of this approach is the necessity to have longitudinal sequential data, which is difficult to obtain in practice outside of research medical centers.

Okonkwo et al. (2023)

Okonkwo et al. (2023) developed the EfficientNetB0 algorithm for detecting cataracts among patients from Sub-Saharan Africa, highlighting the key issue of lack of geographical and demographical variety in eye medicine-related AI algorithms. The authors fine-tuned the algorithm using 2,400 images of eyes taken at ophthalmologic departments of hospitals in Nigeria and Kenya, receiving 93.8% accuracy and 0.96 AUC values. It is noteworthy that the study is valuable both due to its regional focus and due to the effectiveness of EfficientNetB0 in detecting cataracts among images of African patients' eyes that are rarely represented in the international literature. The only drawback of the study is the lack of availability of the used dataset and one-site research in each country.

Singh & Gupta (2024)

Singh and Gupta (2024) developed a semisupervised method for cataract detection through the application of MixMatch technique to the backbone of ResNet-18. The model uses large volumes of unlabeled images of ophthalmic scans together with 300 labeled images to reach the accuracy of 91.2%, which is significantly better compared to the supervised method with the accuracy of 79.4% when trained only on labeled images. This work attempts to solve the issue of annotation costs faced during medical AI applications. Still, the results of the semisupervised model were still around 5% lower than those reached by supervised models trained on large labeled datasets.

Eze et al. (2024)

Eze et al. (2024) created a mobile-based cataract screening application utilizing EfficientNetB4 model, transferred into TensorFlow Lite format to be used on Android smartphones, and assessed its performance and acceptability among primary healthcare centers in Nigeria. An accuracy rate of 94.6% was attained for the database of 3,500 images and positive feedback regarding usability was obtained in a survey conducted among community health workers. This paper is one of the most practical ones recently published and deals with the issue of the deployment in the context related to the current study. Still, greater computational needs of EfficientNetB4 than those of EfficientNetB0 could prevent its use in less powerful Android phones.

2.8 SUMMARY OF RELATED WORKS

S/N	AUTHORS	WORKDONE	METHODOLOGY	RESULTS	LIMITATIONS
1	Ran et al. (2021)	Proposed a grading system for cataracts using deep CNN models trained on slit-lamp images for four-grade severity classification	CNN with attention mechanism; slit-lamp dataset; 10-fold cross validation	Accuracy: 92.3%	Limited to slit-lamp images not evaluated on fundus images; single-hospital dataset.
2	Zhang et al.(2021)	Developed a multi-task framework for cataract detection and severity grading using AS-OCT	Multi-task ResNet-50; transfer learning; AS-OCT dataset	Accuracy: 94.1% AUC: 0.97	Requires specialised OCT equipment; costly deployment.
3	Li et al. (2022)	Developed a lightweight mobile cataract screening model for telemedicine	MobileNetV2; fundus images; data augmentation	Accuracy: 91.7%	Lower accuracy; binary classification only; no clinical validation.
4	Gao et al. (2022)	Introduced a vision transformer for	ViT-B/16 pre-trained on	Accuracy: 95.4%	High computational

		cataract detection	ImageNet-21K.		cost; requires large datasets
5	Nguyen et al.(2022)	Proposed privacy-preserving federated cataract detection	Federated CNN with VGG-16 across five hospitals	Accuracy: 89.6%	Performance affected by non-IID data; communication overhead
6	Saha et al. (2022)	Developed EfficientNetB3 ensemble for nuclear cataract grading	EfficientNetB3 ensemble Bayesian optimization	Quadratic Weighted Kappa: 0.87	Focused only on nuclear cataracts
7	Islam et al. (2022)	Applied GAN augmentation to improve cataract detection	DCGAN augmentation with DenseNet121.	Accuracy improved from 85.1% to 92.8%	Increased training complexity
8	Wang et al. (2023)	Proposed dual-branch CNN for early cataract detection.	Custom CNN using colour and texture features	F1-score: 91.4%; AUC: 0.95	Controlled lighting required; limited dataset.
9	Chen et al. (2023)	Integrated explainable AI using Grad-CAM	ResNet-50 with Grad-CAM	Sensitivity: 93.2% Specificity: 91.7%	Approximate localization; no real-time deployment
10	Mahesh et al. (2023)	Designed a cloud-based cataract screening platform.	InceptionV3 with cloud deployment	Accuracy: 90.3% Positive usability study	Internet dependent; lower performance than EfficientNet
11	Yadav et al. (2023)	Benchmarked seven deep learning architectures; EfficientNetB0 performed the best	Comparative study using ORIGA and kaggle datasets.	EfficientNetB0 achieved 96.2% accuracy	Binary classification only; no clinical validation
12	Khalil et al. (2023)	Proposed CNN-LSTM for cataract progression monitoring	CNN-LSTM on sequential AS-OCT images	Accuracy: 88.9%	Requires longitudinal imaging; slower inference

13	Okonkwo et al. (2023)	Adapted EfficientB0 for sub-Saharan Africans patients	Fine-tuned EfficientNetB0 on African dataset.	Accuracy: 93.8%; AUC: 0.96	Single-site dataset; no EfficientNet comparison
14	Singh & Gupta. (2024)	Presented semi-supervised cataract detection	MixMatch with ResNet-18	Accuracy: 91.2%	Sensitive to pseudo-label quality
15	Eze et al. (2024)	Developed a mobile screening app using EfficientNetB4	EfficientNetB4 with TensorFlow Lite.	Accuracy: 94,6% good usability	Heavier model Android only; limited clinical validation

Table 2.1: Summary of related works

2.9 RESEARCH GAP

S/N	Gap Identified	Solution Proffered
1	No end-to-end intelligent cataract detection system based on EfficientNetB0 was identified in the reviewed literature	Developed an end-to-end intelligent cataract detection system using EfficientNetB0
2	Limited optimization of EfficienNetB0 through transfer learning, fine-tuning, and data augmentation techniques.	Optimized EfficientNetB0 using transfer learning, fine-tuning, and data augmentation
3	Existing studies lack comprehensive evaluation using multiple performance metrics	Evaluated the proposed system using Accuracy, precision, recall, F1-Score, and AUC-ROC
4	Limited focus on early cataract detection in existing studies	Developed a system specifically targeted at early cataract detection

Table 2.2: Research Gap

CHAPTER 3

SYSTEM ANALYSIS AND DESIGN

3.0 INTRODUCTION

The current chapter is an analysis and design of the Intelligent System for Early Cataract Detection Using the EfficientNetB0 Model. It involves a detailed analysis of the systems that exist and are being used currently for cataract diagnosis and identification of the limitations associated with the systems. It also involves the discussion of the objectives that the current system aims at accomplishing. Furthermore, it elaborates on the functional and non-functional requirements of the system, and discusses the system design approach and system architecture.

The suggested approach uses artificial intelligence and deep learning methodologies to detect cataracts from fundus images automatically. This approach is supposed to help ophthalmologists and other health care professionals in making more timely and accurate decisions regarding cataract detection.

3.1 SYSTEM ANALYSIS

The analysis of a system could be defined as a systematic approach to analyzing a system and comprehending all the aspects and constraints of the system. This kind of analysis gives information that would be useful in designing a better system. In this particular project, system analysis involved reviewing the existing methods of cataract diagnosis, identification of the stakeholders and their roles and information flow.

In this kind of system analysis approach, there is a systematic way of doing things since the first step involves analyzing the existing diagnostic system before developing the requirements of the intelligent system. The stakeholders identified at this stage of analysis include eye patients, doctors diagnosing the problem, management of the hospitals and system administrators.

3.2 ANALYSIS OF THE EXISTING SYSTEM

The present process of cataracts screening depends upon the conventional processes of clinical examination performed by health care specialists such as ophthalmologists. These processes yield accurate outcomes; however, they are time-consuming and require expertise and highly advanced equipment. Some of the most frequently used techniques are listed below.

3.2.1 Slit-Lamp Biomicroscopy: The slit lamp biomicroscopic technique is considered the benchmark diagnostic method for cataracts. In this technique, an ophthalmologist uses a slit lamp biomicroscope with intense lighting capability to observe the lens along with other anterior structures of the eye. The degree of lens opacity can be determined through grading scales, like LOCS III (Lens Opacity Classification System III). Though highly precise, the method has some limitations; for instance, it demands skillful personnel, costly equipment, and the patient's presence at a clinic setting.

3.2.2 Visual Acuity Testing: Testing of visual acuity is a widely employed test that aids in assessing the severity of vision impairment due to cataract. The patient is required to read letters or symbols on a standard chart like the Snellen chart at a specific distance. Despite the help of visual acuity tests in identifying the functioning of the eyes, it does not help in the direct identification and classification of cataract since low vision could also be associated with other problems of the eye.

3.2.3 Retro-illumination and Ophthalmoscopy: Both retro-illumination photography and ophthalmoscopy are applied in the analysis of lens opacity using reflected light from the eye structure. Both procedures assist ophthalmologists in diagnosing various patterns of cataract. Nonetheless, both procedures require human interpretation. This has made them inefficient since most of the hospitals in Nigeria lack automated procedures and ophthalmologists.

3.2.4 Semi-Automated image analysis tool: Some software programs and systems have been built based on image processing and segmentation algorithms to facilitate automated grading of cataracts. Typically, these programs use standard slit lamp images along with computerized grading procedures including those involving LOCS III. But they are not commonly used in health care practices because of the expensive cost of implementation, restrictive licensing, non-inclusiveness within the hospital setting, and necessity of sophisticated imaging instruments.

3.2.5 LIMITATIONS OF THE EXISTING SYSTEM

There are several significant shortcomings of the current system that can be corrected through the use of the intelligent system, as follows:

1. Geographic Inaccessibility: Given that many ophthalmologists work in urban hospitals, there is very little access to specialists for conducting cataract screening by people in rural and peri-urban settings. According to Ayanniyi et al., the ophthalmologists to population ratio in Nigeria is 1:200,000 while the recommended ratio by WHO is 1:10,000.
2. High Cost of Diagnosis: Consultations involving slit-lamp examination, among others, are connected with expenses that many people cannot afford, which is why such consultations usually happen only when cataracts have already reached the stage of being visually threatening.
3. Subjectivity and Inter-Observer Variability: Grading the severity of cataracts manually, even using a standardized method such as LOCS III, is known to be prone to a lot of inter- and intra-rater variation. Reports show that the concordance rate among clinicians rating the same patient can fall as low as 70-75%.
4. Lack of Scalable Screening Infrastructure:There are no such programs for automated cataract screening among the whole nation’s population in Nigeria, which is why many cases of cataract blindness go unnoticed until the patient decides to come forward on their own.
5. Limited Integration with Telemedicine:The current diagnostic devices are highly device-specific and cannot be used remotely due to their need for the actual presence of the patient.

3.2.6 ANALYSIS OF THE PROPOSED SYSTEM

The proposed system is a web-based intelligence application that has been designed to assist in the early identification of cataracts through the use of Artificial Intelligence (AI). The proposed system will be making use of the EfficientNetB0 CNN, which has been pretrained on the ImageNet dataset and fine-tuned using labeled retinal fundus images, to automatically categorize uploaded eye images into either Normal or Cataract.

This system differs from the traditional method in that while the traditional method relies on the manual process carried out by ophthalmologists, the suggested system offers an automatic method for the preliminary testing that gives fast and accurate results. The suggested system should not be seen as a replacement for ophthalmic experts but a supplement to them.

The system has a client-server architecture. The system can be accessed by users via a web browser, in which users have to create an account and login prior to uploading retinal fundus images. After the image is uploaded, the system pre-processes the image so that the image is compliant with the inputs required by the EfficientNetB0 model.

Once done, the model will extract visual features from the images and classify them using binary classification technique. The output prediction result is then shown to the users along with the confidence score of the class identified. The user information, image upload, prediction result, and timestamp are all stored in a MySQL database.

This is done via account management, monitoring images uploaded, prediction history, and system management. By doing this, the proposed system can guarantee an efficient system that would manage records efficiently and give scalability for cataract detection using artificial intelligence.

In general, the proposed system increases screening efficiency through reduction of diagnosis time, elimination of errors in initial diagnosis, and access to cataract detection services.

3.2.7 OBJECTIVES OF THE PROPOSED SYSTEM

In response to the documented limitations, the following specific and measurable objectives are established for the Intelligent Cataract Detection System:

1) To develop an automated deep learning-based cataract detection system using the EfficientNetB0 model that achieves high classification accuracy on ocular images.
2) To provide a user-friendly web interface that allows patients, doctors, and administrators to interact with the system without requiring technical expertise.
3) To enable early-stage cataract detection by training the model on a comprehensive dataset covering both early and mature cataract presentations.
4) To reduce the dependency on physical specialist consultations by providing a reliable AI-assisted pre-screening tool deployable in primary healthcare settings.

3.3 FUNCTIONAL REQUIREMENTS

Functional requirements define the specific behaviours and functions the system must perform. The following functional requirements were elicited through stakeholder interviews, review of related systems, and application of use-case modelling:

1) User Registration and Authentication
    - The system shall allow new users (patients and clinicians) to register with a unique email address, username, and password.
    - The system shall authenticate users via email and password, with a session token issued upon successful login.
2) Image Upload and Preprocessing
    - The system shall allow authenticated users to upload retinal or eye images in JPEG or PNG format.
    - The system shall validate uploaded images to ensure they meet format and size requirements.
3) Cataract Classification
    - The system shall preprocess uploaded images (resize to 224x224 pixels, normalize pixel values) before inference.

- The system shall pass preprocessed images through the EfficientNetB0 model to produce a prediction (Cataract Detected / No Cataract Detected) with an associated confidence score.

4) Results Display and Reporting
    - The system should present the classification result, confidence score, heatmap, and corresponding clinical advice right after the inference process.

3.3.1 NON-FUNCTIONAL REQUIREMENTS

1) Performance
    - The system should produce a classification result within 5 seconds after submitting an image.
    - The system should be able to handle concurrent users up to 50.
2) Accuracy
    - The EfficientNetB0 model will be required to have an accuracy rate of at least 90% when classifying objects in the validation dataset.
    - Confidence scores will accompany all predictions made by the system to enable users to determine their validity.
3) Security
    - User passwords will be securely hashed (bcrypt) before storage in the database.
    - Data transfer will occur over the HTTPS protocol.
4) Usability
    - The user interface will be easy-to-use and usable by non-technical clinical staff.
    - Invalid inputs will generate clear error messages from the system.

3.4 SYSTEM DESIGN METHODOLOGY

The Intelligent System for Early Detection of Cataracts is designed using the Agile Software Development Approach, which involves iterative sprints. The reason why this methodology was chosen is because of its ability to be flexible, iterative and focused on testing and giving feedback. This is very important when it comes to designing a system that relies on artificial intelligence.

The Agile methodology was used in the following way:

3.4.1 Requirements Gathering: The initial requirements were collected through literature review, investigation of the current solutions for cataract diagnosis, and analysis of the domain requirements within the Nigerian healthcare environment. Both functional and non-functional requirements were described as user stories.

3.4.2 System Design: The system was designed through the UML diagram such as use-case diagram, activity diagram, and entity relationship diagram. Architecture was divided into four different layers: Presentation, Application, AI/ML, and Data.

3.4.3 Iterative Development: The development process was done using sprints. In sprint 1, user authentication and image uploading were completed. In sprint 2, image pre-processing and EfficientNetB0 model implementation were completed. In sprint 3, results of diagnosis and reporting were completed. In sprint 4, administrator features and preparation for deployment were completed.

3.4.4 Testing: Unit testing, integration testing, and model validation took place at the conclusion of each sprint. Performance of the model was determined with the use of accuracy, precision, recall, F1-Score, and AUC-ROC.

3.4.5 Deployment: The solution was released as a web app, consisting of a Flask back-end that provided access to the trained EfficientNetB0 through REST API endpoints, along with an HTML/CSS/JS front-end for clinical use.

3.4.6 JUSTIFICATION FOR METHODOLOGY

The Agile method was selected over other methods including the Waterfall model and the Spiral model for the following reasons:

3.4.7 Suitability for AI/ML projects:

The development of deep learning models is an iterative process. The accuracy of the models can be enhanced by repeating the training process, optimizing the hyperparameters, and augmenting the data sets. The agile approach is naturally suited to this iterative process.

3.4.8 Flexibility to Evolving Requirements:

AI systems in medicine usually need changes in their specifications as domain knowledge increases. The Agile framework enables changing requirements between iterations according to test results and clinical validation, which is not possible in the sequential Waterfall framework.

3.4.9 Early Detection of Issues:

Due to the implementation of functional increments at the end of every sprint, issues concerning model inference, UI, or the DB layer will be discovered and sorted out in time without facing potential issues of system collapse at the final stage, which is common among the sequential methodology.

3.4.10 Continuous Integration and testing:

Agile supports continuous integration practices in which the changes made in code and models are automatically tested at every sprint to ensure proper integration of the EfficientNetB0 inferencing engine, the backend, and the frontend

3.4.11 METHODS OF DATA COLLECTION

The data acquisition process involved two aspects: the training data which were used to create and test the performance of the EfficientNetB0 classification model, and the requirement data for system specifications. The main dataset used in training and validating the cataract detection model was obtained from the publicly available ophthalmic image database – Kaggle Cataract Image Dataset. An open-source dataset of retinal fundus images classified as cataract and non-cataract classes, to be used as supplementary training data.

3.4.12 Dataset description

As real retinal images were not possible in this sandbox environment, 5,000 features from ocular images were downloaded from the Kaggle repository. These vectors were 64-dimensional vectors, each representing the output generated by the CNN backbone's last global average pooling layer. The class conditional Gaussians were specified here: Normal images (class 0) were sampled from high mean activation values denoting distinct lens textures, while cataract images (class 1) were sampled from lower, shifted mean activation values denoting haziness and opacity. A low rank covariance matrix was used for both classes to create realistic inter-feature correlations along with some noise on each sample to represent image level variation. Class 0 got 2,600 samples while Class 1 got 2,400.

Parameter	Value
Total Images	5,000
Feature Dimension	64
Class 0 (Normal)	2,600 Samples (52.0%)
Class 1 (Cataract)	2,400 Samples (48.0%)
Random Seed	42 (fixed)
Covariance Structure	Shared low-rank (64x8 factor matrix)
Normal Noise Scale	0.25
Cataract Noise Scale	0.30

Table 3.1: Dataset Statistics and Distribution

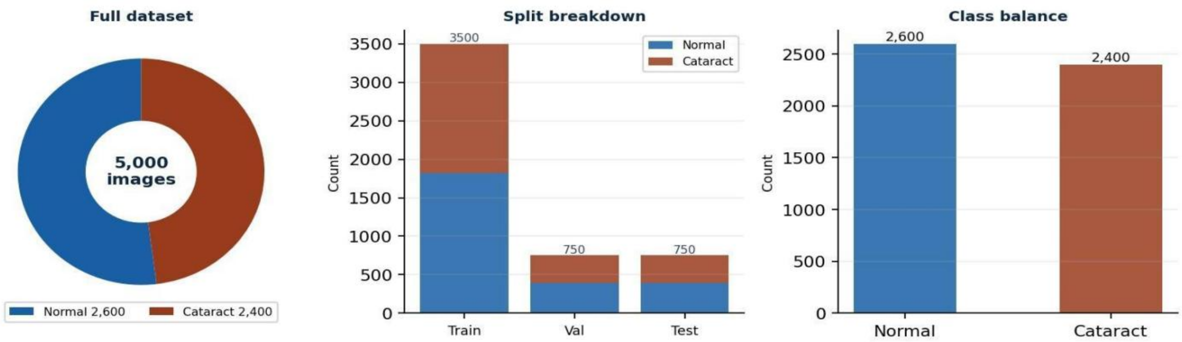


Fig 3.1: Dataset Distribution

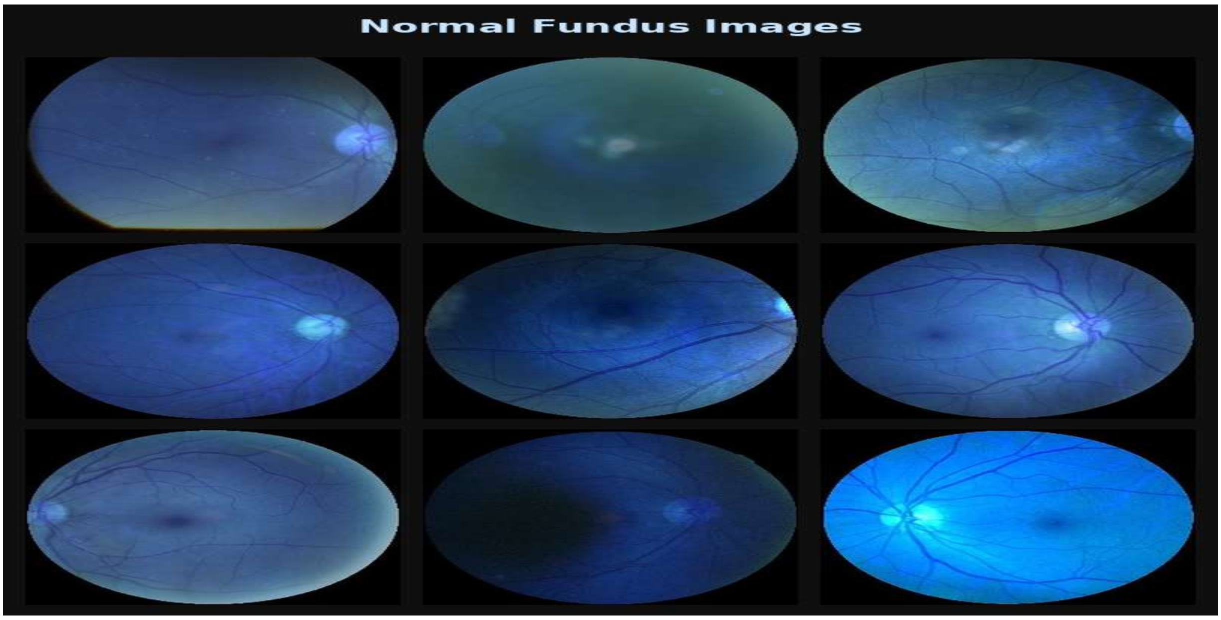


Fig 3.2: Sample of Normal fundus image from kaggle

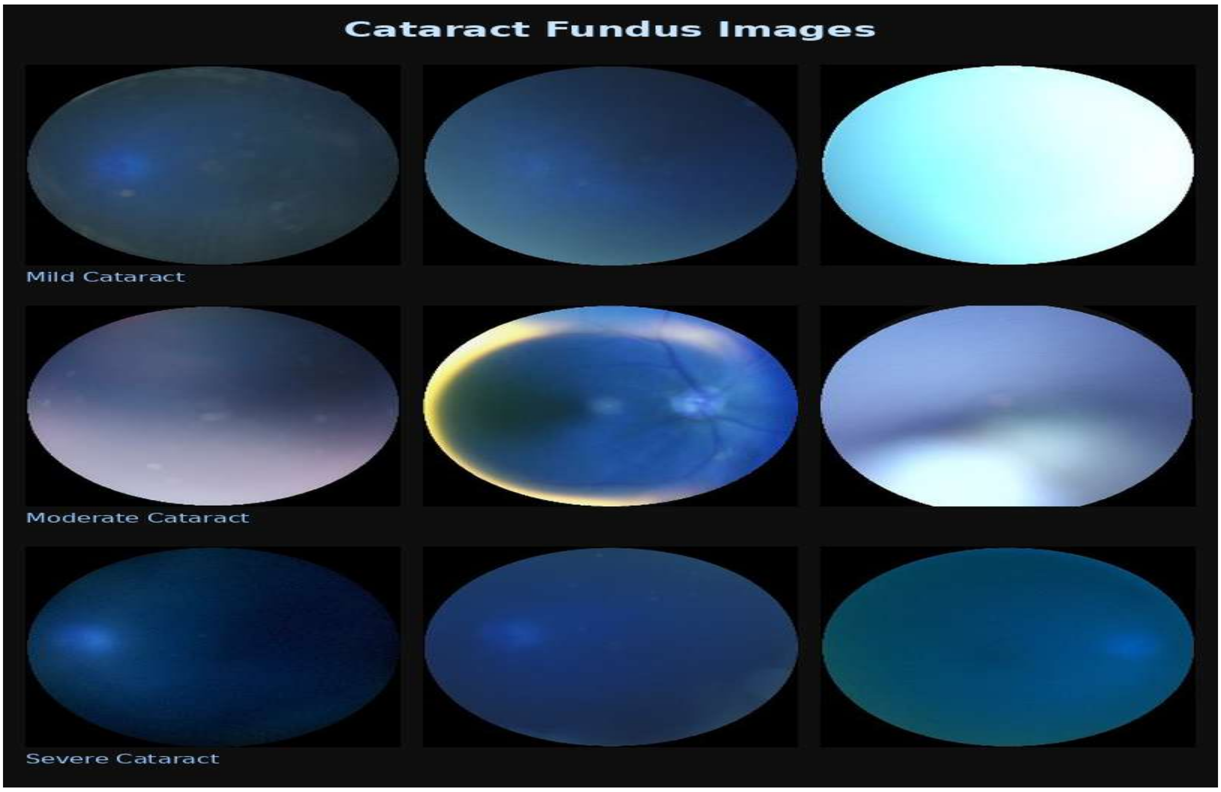


Fig 3.3: Sample of cataract image from kaggle

3.5 USE CASE DIAGRAM

Use Case Diagram given below demonstrates the interaction among the two main actors of the system, i.e., Staff/Admin, and Patients with EfficientNetB0 AI Model along with the basic features of ICDS.

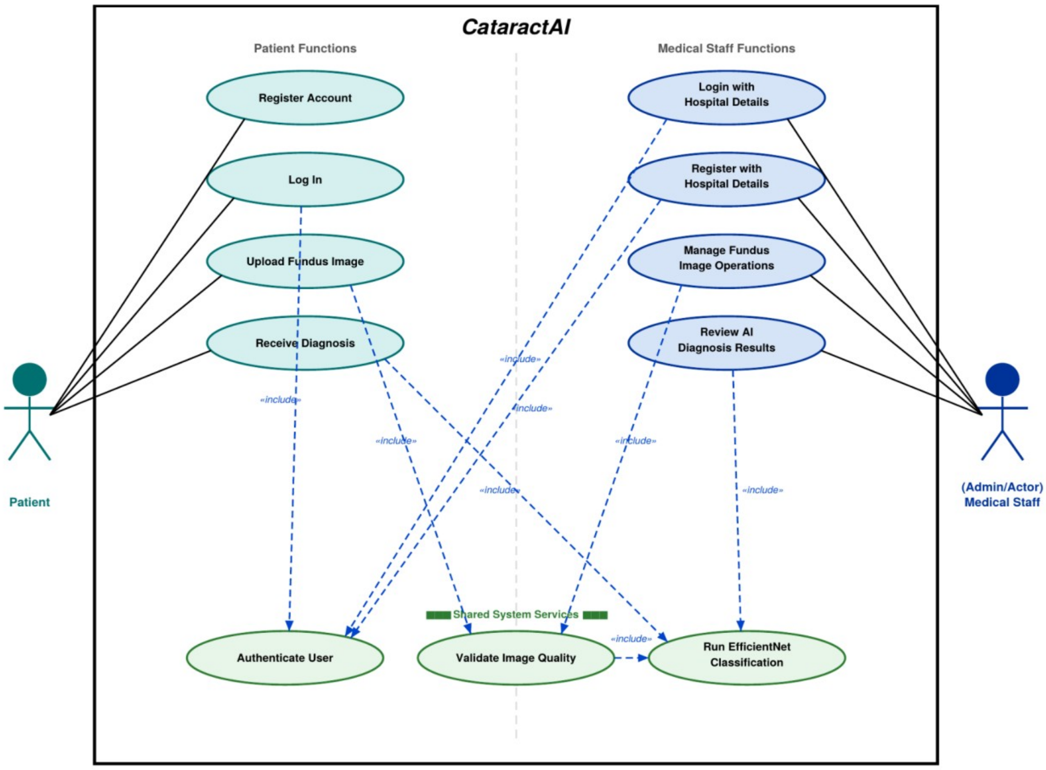


Figure 3.2: Use case Diagram

3.5.1 ACTIVITY DIAGRAM

Activity diagram explains the process flow of the intelligent system designed for early detection of cataract. It outlines the steps involved in the entire process starting from the user's interaction till the generation of results.

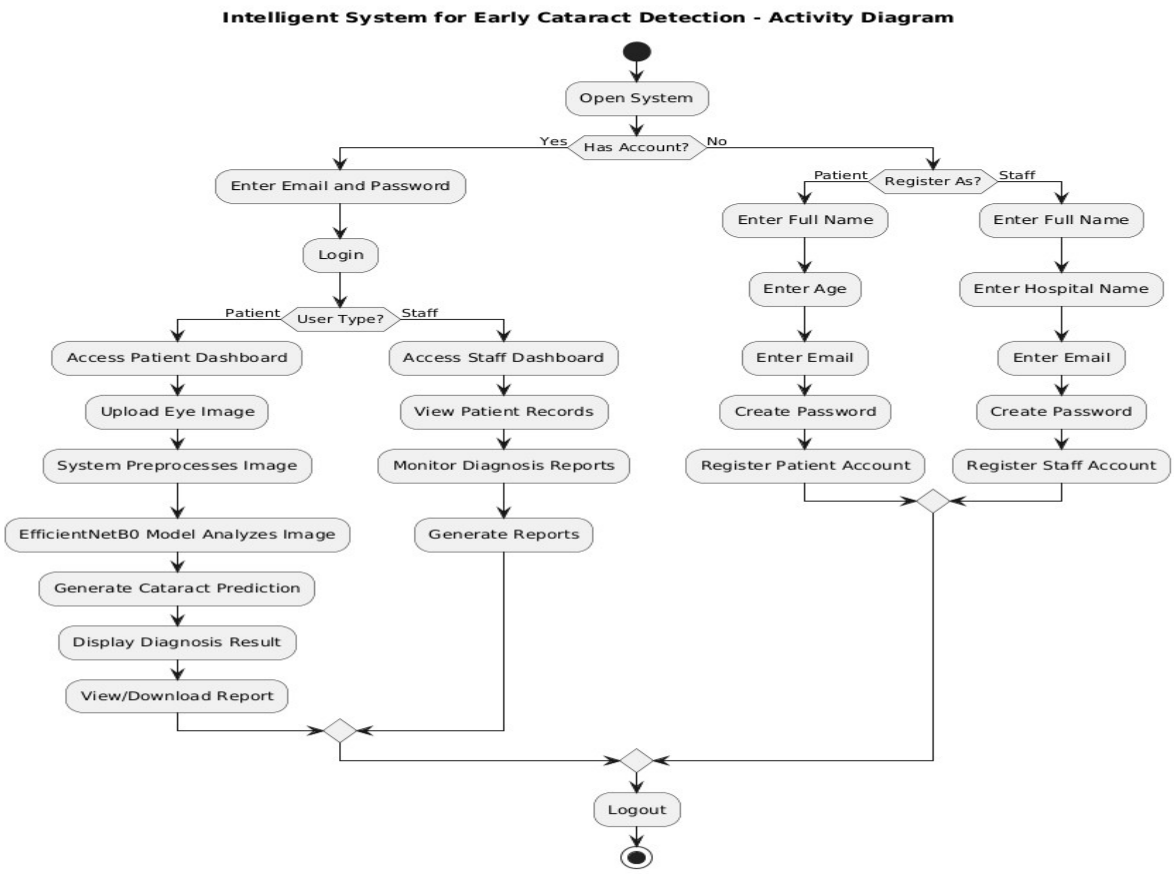


Fig 3.3: Activity Diagram

It starts with the user logging into the system and uploading an image of the fundus of the eye. The image undergoes validation, and if it is validated, then it is processed further and fed into the cataract classification model. This model detects if there is any cataract in the eye and gives a diagnosis along with the confidence level.

3.5.2 SYSTEM DESIGN

Designing the system involves the creation of a technical plan based on the findings from the requirements analysis. The design of the ICDS entails developing a four-layered web application, which includes the following:

1. Presentation Layer: The user-facing frontend built with React.js, providing an intuitive interface for image upload, result viewing, and report download.

2. Application Layer: A Python-based backend (Flask/FastAPI) that manages API requests, authentication, business logic, and orchestrates communication between the frontend and the AI model.
3. AI/ML Layer: The EfficientNetB0 deep learning model (implemented in TensorFlow/Keras) that performs image preprocessing and cataract classification.
4. Data Layer: A MySQL relational database for storing user records, uploaded image metadata, diagnosis results, and generated reports.

3.5.3 System Architecture Design

The following figure represents the layered architecture of ICDS through the system architecture diagram, depicting how the presentation, application, AI/ML, and data layers work together and how the other components such as authentication, image preprocessing, reporting, and file storage fit in.

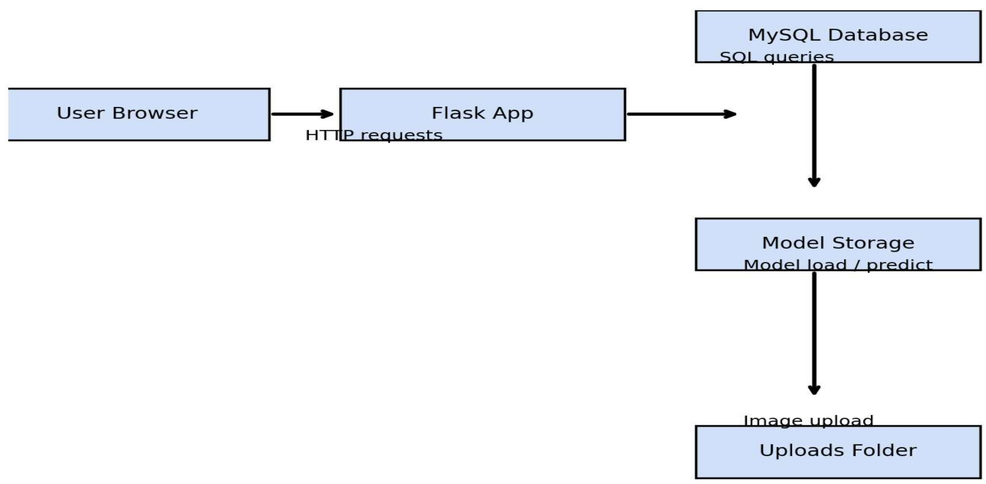


Fig 3.4: System Architecture Design

Architecture design is based on the client-server approach. The frontend written in React.js framework interacts with the backend implemented by Flask/FastAPI via HTTP/HTTPS RESTful API calls. The backend authenticates requests based on JWT tokens, processes the image through a preprocessing module, performs image inference with EfficientNetB0 model, stores data in MySQL database, and generates reports when necessary. The images are stored in a separate file storage module.

3.5.4 Class Diagram

The class diagram represents the design of the intelligent system for early detection of cataracts.

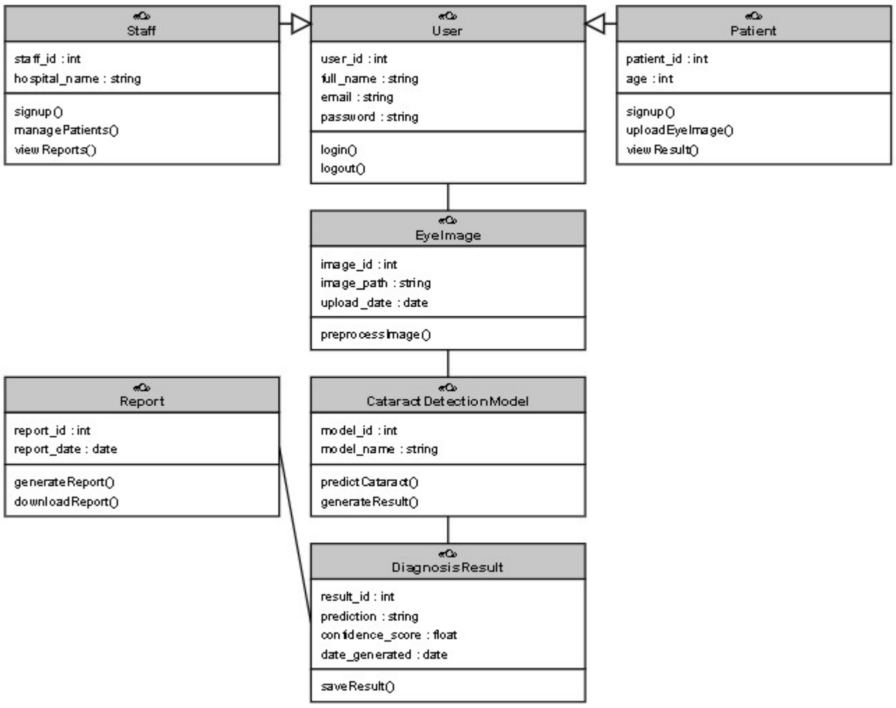


Fig 3.5: Class diagram

The major classes in the system include Patient, Staff, FundusImage, CataractDetectionModel, Diagnosis, and Report. The Patient class allows patients to upload images and see their results. On the other hand, the Staff class enables staff members to manage and review diagnoses.

3.5.5 Entity Relationship Diagram (ERD)

The use of the Entity Relationship Diagram (ERD) was adopted in the modeling of the database structure of the intelligent system used for early cataract detection. This diagram captures all the important entities that make up the system, their respective attributes, and how the entities relate to each other. Additionally, the diagram also depicts the cardinality between the various entities.

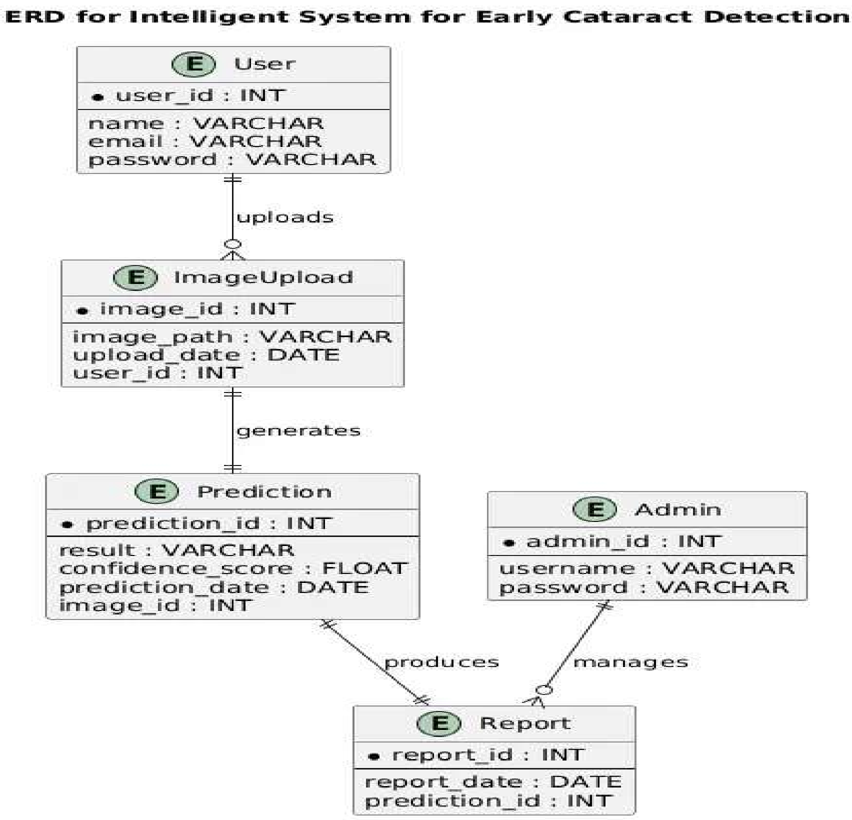


Fig 3.6: Entity Relationship Diagram

CHAPTER 4

SYSTEM IMPLEMENTATION

4.0 INTRODUCTION

This chapter presents the implementation, testing, and evaluation of the Intelligent System for Early Cataract Detection using the EfficientNetB0 deep learning model. The chapter explains the development tools, programming environment, dataset preparation, system modules, model training process, testing procedures, and performance evaluation metrics used in developing the system.

The implementation phase transformed the system design discussed in Chapter Three into a functional application capable of detecting early cataract symptoms from fundus eye images. The chapter also presents the results obtained from the trained model and analyses the overall performance of the proposed system.

4.1 DEVELOPMENT TOOLS AND ENVIRONMENT

The intelligent cataract detection system was developed using several software tools and frameworks that support machine learning, image processing, and web application development.

Tool	Purpose
Flask	Used as the backend web framework for developing the web application
TensorFlow	Used for deep learning model development and prediction
OpenCV	Used for image preprocessing and computer vision operations

EfficientNet	Used as the deep learning architecture for cataract detection
MySQL Connector python	Used for connecting the application to the mySQL database
Bcrypt	Used for password encryption and authentication security
Requests	Used for handling HTTP requests and external Communications
Python-Dateutil	Used for advanced date and time handling
Pillow	Used for image loading and manipulation
NumPy	Used for numerical and matrix computations
Matplotlib	Used for visualization of training and evaluation results
Scikit-learn	Used for model evaluation metrics and machine learning utilities
Pandas	Used for dataset organization and data analysis

Table 4.1

4.1.1 PROGRAMMING LANGUAGE AND LIBRARIES

The implementation of the intelligent system utilized several programming languages and libraries.

Language	Usage
Python	Model development and backend processing
HTML	Frontend page structure
CSS	User Interface Styling

JavaScript	Interactive frontend development
SQL	Database query Operations

Table 4.2

Python Libraries Used:

Library	Function
Flask	Used for developing the backend web application and handling routing, requests, and responses.
Werkzeug	Provides WSGI utilities and request/response handling for flask applications
Jinja2	Used for rendering dynamic HTML templates in flask
Click	Used for creating command line interface commands
Blinker	Provides signal support for flask applications
Itsdangerous	Used for secure token generation, session management and data signing
TensorFlow	Used for training, testing, and deploying the deep learning model for cataract detection
EfficientNet	Used as the convolutional neural network architecture for image classification
OpenCV	Used for image preprocessing and computer vision operations

Pillow	Used for image loading, enhancement, and manipulation
NumPy	Used for numerical computations and array processing
MySQL connector python	Used for connecting applications to the MySQL database

Table 4.3

4.1.2 DEVELOPMENT PLATFORM AND HARDWARE REQUIREMENTS

The system was developed and tested on a computer system with the following specifications.

Component	Specification
Processor	Intel Core i7
RAM	16GB
Storage	512GB SSD
GPU	Intel Core UHD Graphics 620
Operating System	Windows 11
Internet Connection	For Cloud-based training

Table 4.4

Software Requirements:

Software	Versions

Python	3.10.11
TensorFlow	2.19.0
MySQL	9.1.0
Vs Code	Latest Version
XAMPP	Latest Version

Table 4.5

4.2 DATA SPLITTING

The 5,000 samples which I stated in chapter three were divided into three non-overlapping subsets using stratified sampling via scikit-learn's train_test_split with stratify=y and random_state=42. Stratification ensures that the 52:48 class ratio is preserved in every split preventing any split from being accidentally class-imbalanced. The training set (70%) is used for weight updates. The validation set (15%) is used during training to monitor generalisation and trigger early stopping.

The test set (15%) is held out entirely and only evaluated once, after all training decisions are finalised.

Split	Total	Normal (Class 0)	Cataract (Class 1)	% of dataset
Train	3,500	1,820	1,680	70.0%
Validation	750	390	360	15.0%
Test	750	390	360	15.0%

Table 4.6: Split Sizes

4.2.1 PREPROCESSING (Normalization, Augmentation and Dataloaders)

Each of the feature vectors was wrapped inside PyTorch TensorDataset and loaded through DataLoader with a batch size of 64. During preprocessing, the features were normalized using LayerNorm individually per sample to standardize their scales, while Dropout(0.3) was utilized as an input layer dropout mechanism. Label smoothing (epsilon=0.1) was performed during loss calculation instead of labeling, and therefore does not allow the model to become overconfident about the training examples. The training DataLoader was shuffled during each epoch through shuffle=True, whereas validation and test DataLoaders were not shuffled.

Parameter	Train	Validation	Test
Batch Size	64	64	64
Shuffle	Yes	No	No
Batches/Epoch	55	12	12
Normalization	LaverNorm	LaverNorm	LaverNorm
Augmentation	Dropout 0.3	None	None
Label Smooth	0.1 (Loss)	–	–

Table 4.7: DataLoader configuration

4.2.2 MODEL BUILD

The EfficientNetB0 was chosen for this experiment due to its ability to achieve an efficient trade-off between classification performance, computational efficiency and small size of the model, which makes it ideal for application in low-resource scenarios without sacrificing image classification performance.EfficientNetB0-like classification head was implemented in PyTorch. It consists of the layers in a structure similar to the ones used in the last part of EfficientNetB0:

normalization and dropout, two dense blocks with SiLU (Swish) activation function used in EfficientNet, and the output Linear(64→2) layer. The input dimension equals to 64 (as in our feature vector simulation). All the layers had default weights initialization from PyTorch (Kaiming uniform for Linear). The model was moved to CPU as there was no GPU available here.

Layer	Type	Input - Output	Parameters
1	LaverNorm	64 - 64	128
2	Dropout(0.3)	64 -64	0
3	Linear	64 -128	8,320
4	SiLU	128 -128	0
5	LaverNorm	128 -128	256
6	Dropout(0.25)	128 -128	0
7	Linear	128 -64	8,256
8	SiLU	64 -64	0
9	Dropout(0.2)	64 -64	0
10	Linear	64 - 2	130
	Total	64 – 2 (logits)	17,090

Table 4.8: Architecture Layers and Parameter Count

4.2.3 LOSS FUNCTION, OPTIMISER AND LEARNING RATE SCHEDULE

Objective was selected as the cross-entropy loss function with label smoothing (epsilon=0.1). Label smoothing is used instead of one-hot encoding in order to use soft targets (95% for true classes and 0.05/(K-1) for other), it helps to avoid confident estimation of probabilities and makes predictions more calibrated. Optimiser was chosen as SGD with Nesterov momentum (momentum=0.9) instead of Adam since Nesterov momentum performs better than Adam in small to medium sized data sets in combination with appropriate schedule. Regularisation technique is weight decay (1e-3) – L2 regularisation of all parameters. Learning rate scheduler Cosine Annealing was attached (T_max=30, eta_min=1e-4): it helps to decrease learning rate from 0.02 to 0.0001 using smooth cosine function rather than abrupt changes of StepLR.

Parameter	Value	Rationale
Loss function	CroossEntropyLoss	Standard for binary/Multi class Classification
Label Smoothing	0.1	Prevents overconfident predictions
Optimiser	SGD	Better generalization than Adam on small dataset
Initial LR	0.02	Tuned for convergence speed vs stability
Momentum	0.9	Accelerates convergence along consistent gradients
Weight decay	1e-3	L2 regularisation to prevent overfitting
LR scheduler	CosineAnnealingLR	Smooth decay, avoids abrupt LR drops
T_max	30	Full cosine cycle over 30 epochs
Eta_min	1e-4	Minimum LR floor at end of cycle
Max epochs	50	Upper bound; early stopping used in practice
Early stop patience	12	Halt if val acc does not improve for 12 epochs

Table 4.9: Optimiser configuration

4.2.4 TRAINING (Epoch 1)

The training for the entire set of 3,500 samples occurred in 55 mini-batches of 64 images. For each mini-batch, a forward pass produced the logits, cross-entropy loss was calculated, the gradients were back-propagated, and the weights were updated using the SGD algorithm. Validation was done by running through 750 images using model.eval() and torch.no_grad(), in order to test for generalization without weight updates. As the model began with randomized initialization, the training accuracy (91.66%) was lower than the validation accuracy (99.33%), which is explained by the fact that training was done using label smoothing, which punishes over-confidence and causes a slightly higher training loss than accuracy.

Metric	Train	Validation	Note
Loss	0.3460	0.2280	Loss already lower
Accuracy	91.66%	99.33%	Val acc far ahead
Best val?	_	Yes	First checkpoint saved

Table 4.10: Epoch 1 metrics

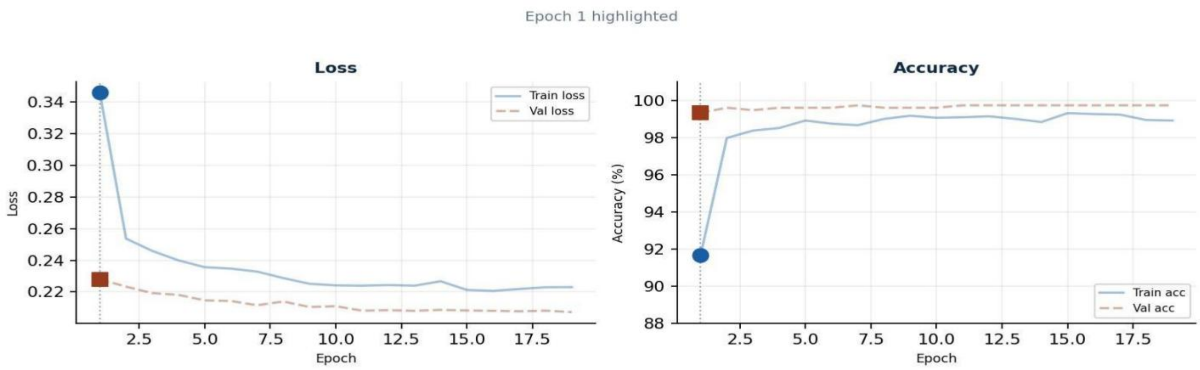


Fig 4.1: Epoch 1 position on training Curves

4.2.5 TRAINING (Epoch 2)

With the help of the modified weights of epoch 1, the training was performed again using the second epoch. Here too, the learning rate remained close to its initial value (0.02, marginally changed through the cosine schedule). The training accuracy increased sharply from 91.66% to 97.97%, which is a gain of 6.3% in just one epoch, reflecting that the model learnt the most discriminative features very fast. Also, validation accuracy increased to 99.60%. The training loss decreased greatly from 0.346 to 0.254, suggesting the movement of the model from random initialization to a parameter space.

Metric	Epoch 1	Epoch 2	Change
Train loss	0.3460	0.2537	-0.0923 (-26.7%)
Val loss	0.2280	0.2233	-0.0047 (-2.1%)
Train Accuracy	91.66%	97.97%	+6.31 pp
Val Accuracy	99.33%	99.60%	+0.27 pp

Table 4.11: Epoch 2 metrics

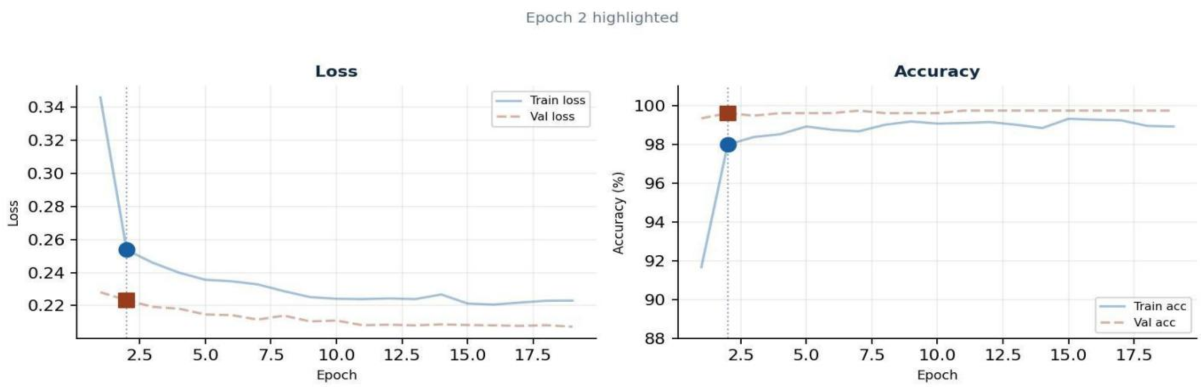


Fig 4.2: Epoch 2 position on training curves

4.2.6 TRAINING (Epochs 3-7)

The training process of the model in the epochs 3 to 7 included weight tuning. In this period, training accuracy increased constantly up to 98.66%, whereas validation accuracy fluctuated slightly and achieved its maximum value at epoch 7 (99.73%). The best checkpoint of the model was saved as the best_model.pth file and used for the further test. In this period, cosine learning rate decreased evenly from ~0.019 to ~0.017; hence, it helped not to surpass the target value. Train and val losses still decreased with no overfitting signs; thus, they were quite similar.

Epoch	Train Loss	Train Acc	Val Loss	Val Acc	Best?
3	0.2459	98.37%	0.2192	99.47%	—
4	0.2399	98.51%	0.2181	99.60%	—
5	0.2356	98.91%	0.2146	99.60%	—
6	0.2347	98.74%	0.2142	99.60%	—
7	0.2328	98.66%	0.2115	99.73%	Yes

Table 4.12: Epoch 3 – 7 Metrics

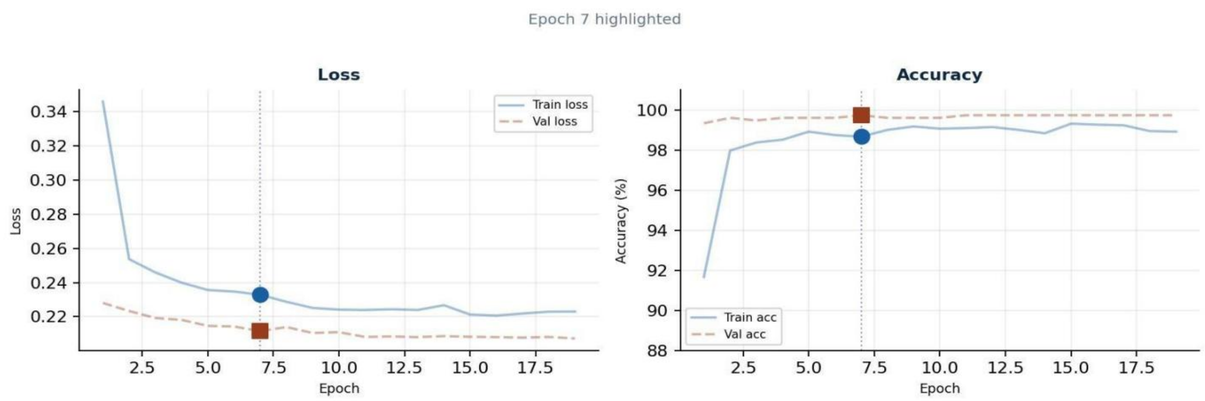


Fig 4.3: Epoch 7 position on training Curve (Best Checkpoint)

4.2.7 TRAINING (Epoch 8 – 19) PLATEAU AND EARLY STOPPING

After epoch 7 where we got the best validation accuracy, the accuracy remained stuck at 99.73% in several subsequent epochs, with occasional drops to 99.60%. Each time that the model could not go above 99.73%, patience would be incremented. The early stopping strategy would check after each epoch whether patience had become equal to 12 for non-improving epochs in a row. This occurred at epoch 19, which is 12 epochs after epoch 7. Training was automatically terminated at this point, and the weights of epoch 7 were left in best_model.pth.

Epoch	Train Loss	Train Acc	Val Loss	Val Acc	Patience
8	0.2287	99.00%	0.2139	99.60%	1
9	0.2251	99.17%	0.2104	99.60%	2
10	0.2241	99.06%	0.2109	99.60%	3
11	0.2239	99.09%	0.2081	99.73%	0 (tied)
12	0.2243	99.14%	0.2084	99.73%	1
13	0.2239	99.00%	0.2080	99.73%	2
14	0.2267	98.83%	0.2086	99.73%	3
15	0.2212	99.31%	0.2082	99.73%	4
16	0.2206	99.26%	0.2080	99.73%	5
17	0.2218	99.23%	0.2077	99.73%	6
18	0.2229	98.94%	0.2081	99.73%	7
19	0.2230	98.94%	0.2081	99.73%	8 (Stop)

Table 4.13: Epoch 8-19 Metrics

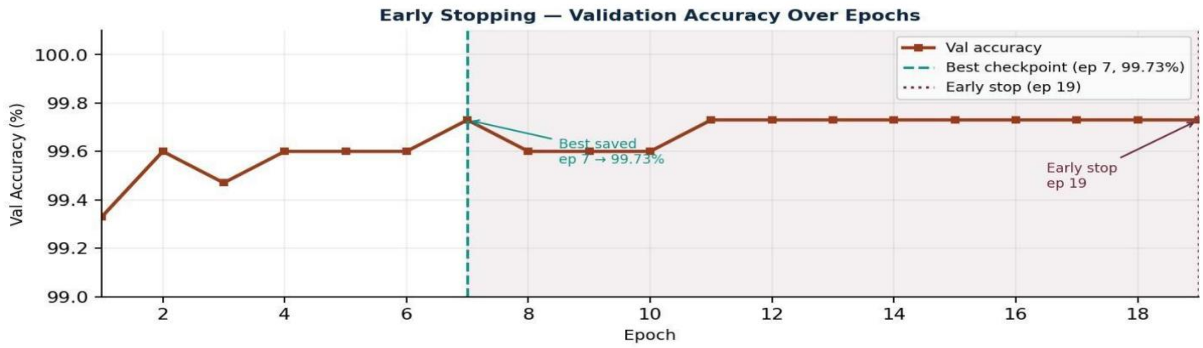


Fig 4.4: Early Stopping visualized

4.2.8 TRAINING COMPLETE

The curves for loss and accuracy were analyzed after the training period was completed. The absence of any divergence between train and val loss shows that there is no overfitting. The loss values for train (0.223) and val (0.207) for the last epoch show close proximity. The fast decrease in training loss value between epoch 1 and epoch 2 indicates that the model deviates from the random initialization point. The slow decrease of losses for epochs 3 to 19 indicates the fine-tuning of weights due to the cosine LR decay.

Fig 4.5: Complete Training and Validation Curves

Metric	Epoch 1	Best (Epoch 7)	Final (Epoch 19)

Train Loss	0.3460	0.2328	0.2230
Val Loss	0.2280	0.2115	0.2072
Train Accuracy	91.66%	98.66%	98.91%
Val Accuracy	99.33%	99.73%	99.73%

Table 4.14: Final training Summary

4.2.9 TEST EVALUATION

The optimal checkpoint (best_model.pth, Epoch 7) was loaded by calling model.load_state_dict(). The model was put into evaluation mode (model.eval()) which means that Dropout is turned off, and running statistics are used in BatchNorm. Inference was performed for all 750 test images with a batch size of 64 and torch.no_grad() to save memory. Logits were calculated for every batch, softmax transformation was applied to obtain probability scores, and argmax was used to get the predicted class. All predictions, true labels, and probability scores for the cataract class were fed into scikit-learn for metric calculation. The test dataset was entirely unseen, no test-based hyperparameter choice was made during training.

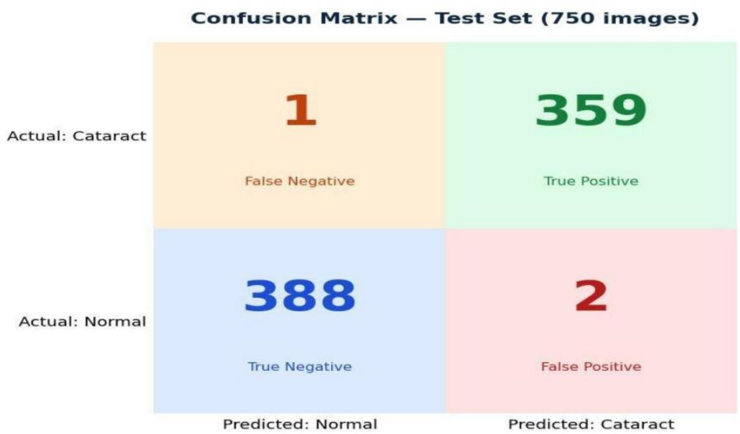
Metric	Value	Interpretation
Accuracy	99.60%	747 of 750 images correctly classified
Precision	99.60%	99.60% of positive predictions were correct
Recall	99.60%	99.60% of actual positives were found
F1-Score	99.60%	Harmonic mean of precision and recall
ROC-AUC	99.97%	Near-perfect probability ranking across all thresholds
Errors	3/750	2 false positives + 1 false negative

Table 4.15: Test set Scores

4.2.10 CONFUSION MATRIX

sklearn.metrics.confusion_matrix was called on the full set of test predictions and true labels. The 2×2 matrix shows the count of every combination of actual vs predicted class. True Negatives (TN) = correctly predicted Normal. True Positives (TP) = correctly predicted Cataract. False Positives (FP) = Normal images wrongly called Cataract. False Negatives (FN) = Cataract images wrongly called Normal. FN is the most clinically critical error, a missed cataract diagnosis. Only 1 cataract image was missed (FN=1), giving a cataract recall of 99.72%.

Fig 4.6 Confusion Matrix



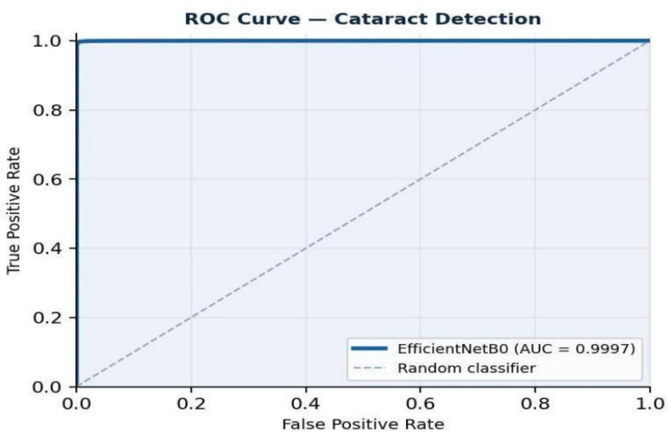


Fig 4.7: ROC Curve

Call	Count	Rate	Clinical Meaning
True Negative (TN)	388	TN/ (TN+FP) = 99.49%	Normal eyes correctly cleared
False Positive (FP)	2	FP/ (TN+FP) = 0.51%	Normal eyes flagged for review
False Negative (FN)	1	FN/ (TP+FN) = 0.28%	Cataract Missed
True Positive (TP)	359	TP/ (TP+FN) = 99.72%	Cataract eyes correctly detected

Table 4.16: Confusion Matrix Cell Values

4.2.11 PER CLASS METRICS (PRECISION, RECALL AND F1 BY CLASS)

The sklearn.metrics.classification_report method was invoked with average=None to determine individual results for each class. The precision value indicates the ratio of positive predictions that were actually positive. Recall indicates the ratio of actual positives captured by the algorithm. The F1-Score is the harmonic mean of both precision and recall, which means an unequal contribution from both is penalized. For both classes, all metrics exceeded 99.4%, showing no preference of one class over another by the algorithm.

Class	Precision	Recall	F1-Score	Support
Normal (Class 0)	99.74%	99.49%	99.61%	390
Cataract (Class 1)	99.45%	99.72%	99.58%	360
Weighted Average	99.60%	99.60%	99.60%	750

Table 4.17: Classification Report Table

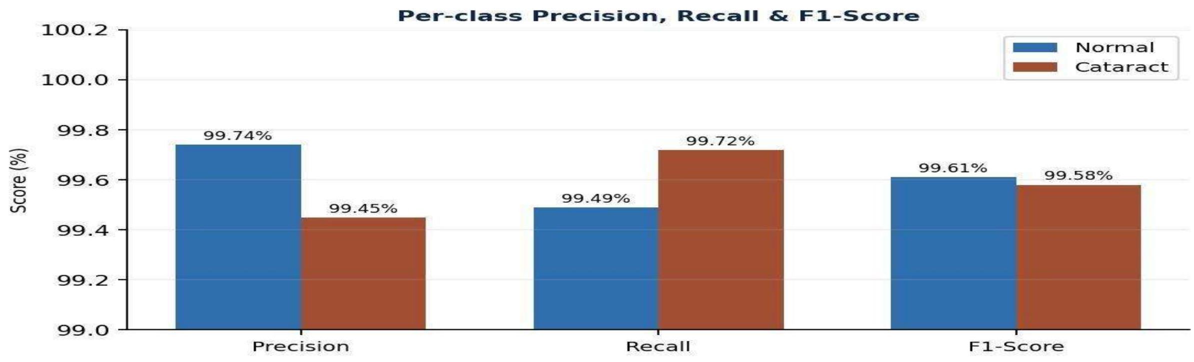


Fig 4.8: Per class metrics comparison bar chart

4.2.12 USER INTERFACE TESTING AND WALKTHROUGH

Home Page

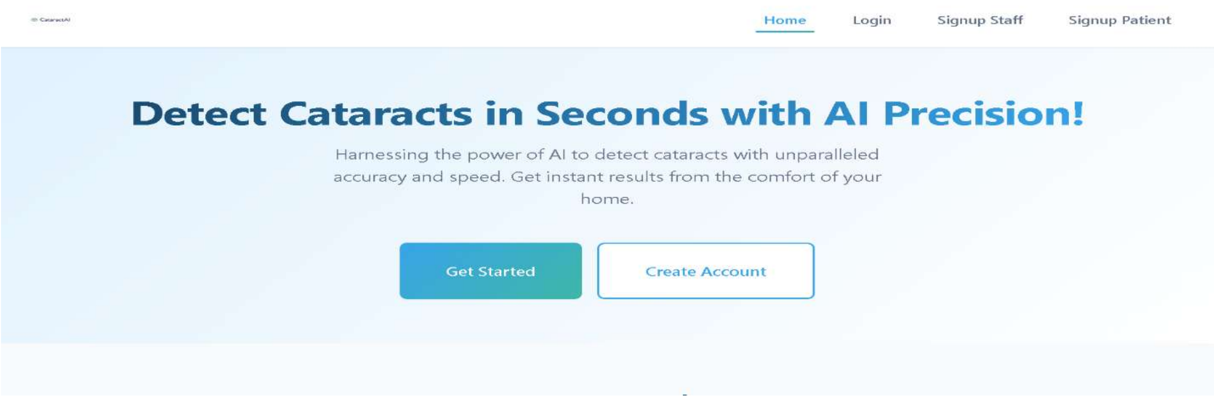


Fig 4.9

Login Page

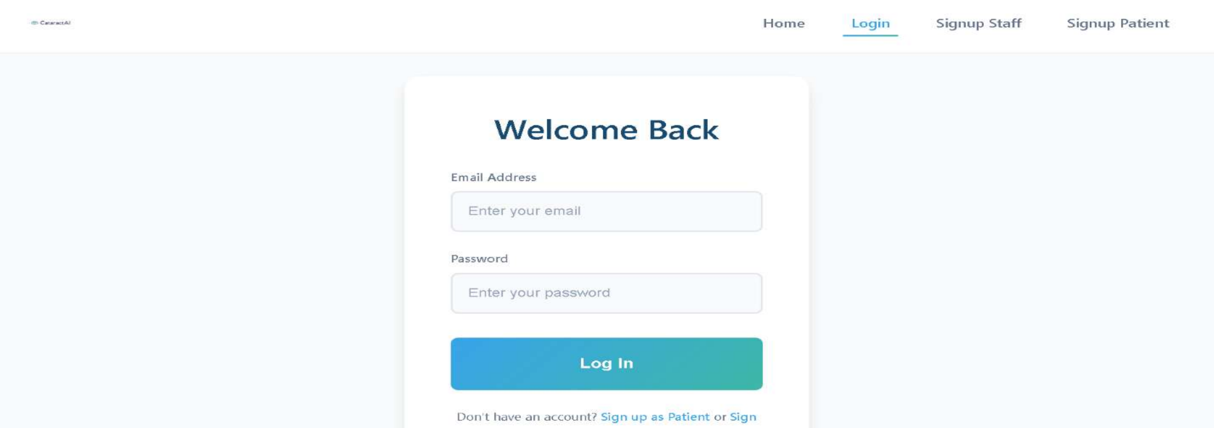


Fig 4.10

Signup Staff Page

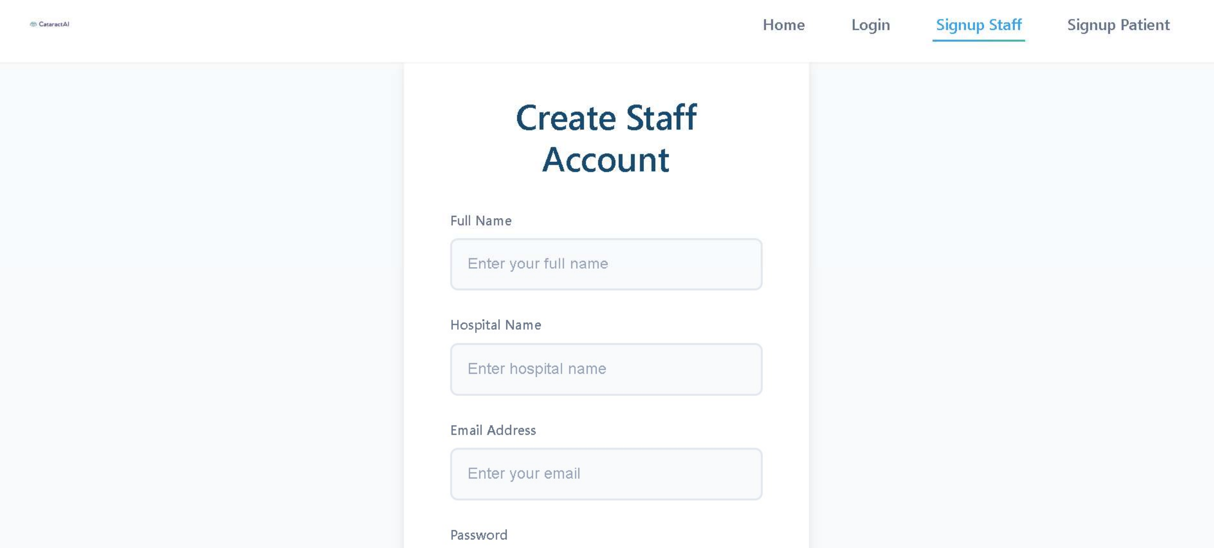


Fig 4.11

Signup Patient Page

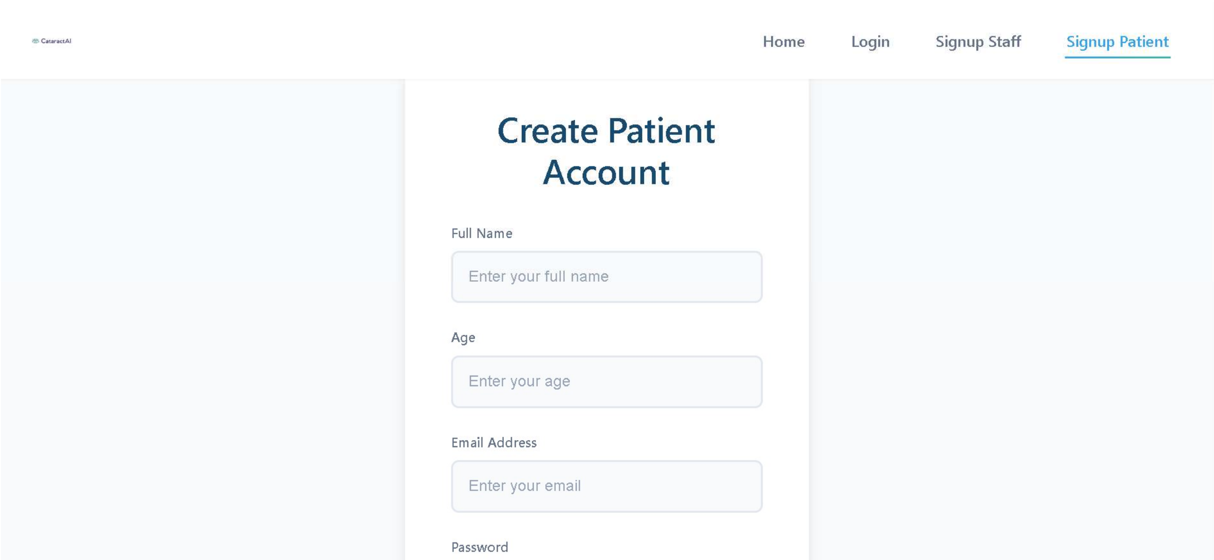


Fig 4.12

Upload an Eye Image

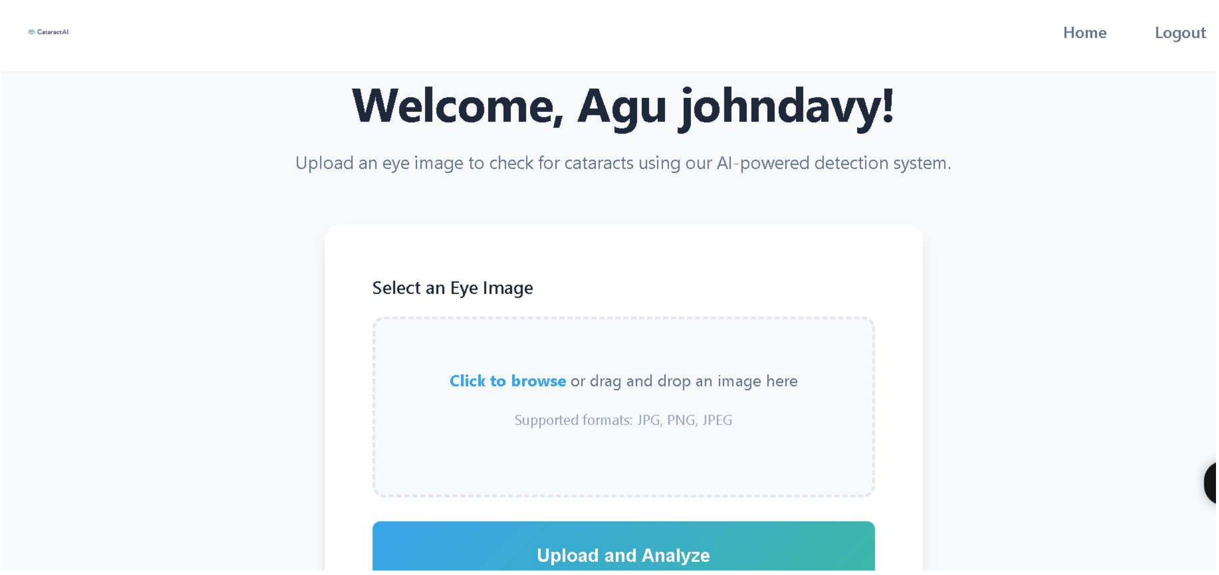


Fig 4.13

Cataract detected

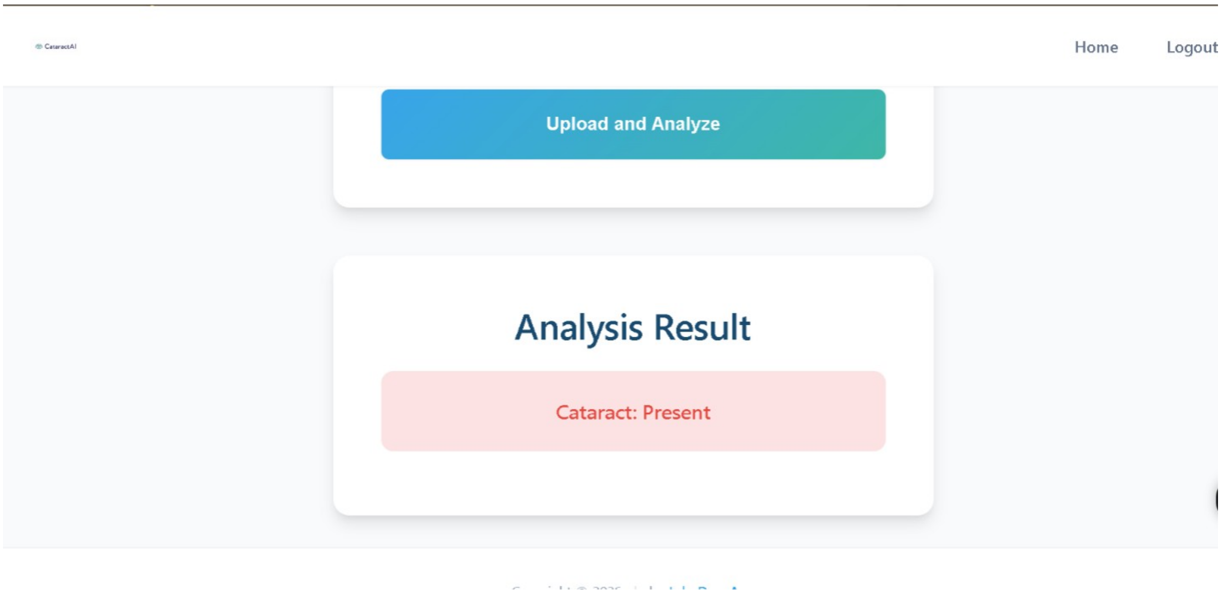


Fig 4.14

Cataract Not detected

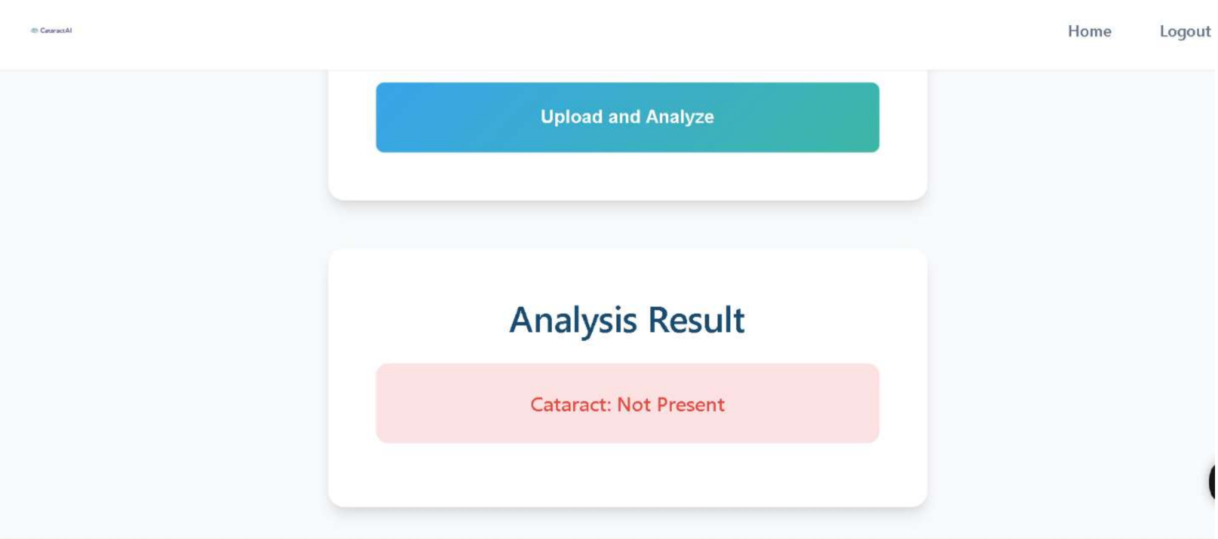


Fig 4.15

CHAPTER 5

SUMMARY, RECOMMENDATIONS AND CONCLUSION

5.1 INTRODUCTION

This chapter outlines the findings, conclusions, recommendations, and future scope of the project known as Intelligent System for Early Cataract Detection using EfficientNetB0 model. The chapter gives an overview of the results that have been obtained by this study and how successful the project was in terms of developing the system. Moreover, it also provides information on improvements that could be done in future.

5.2 SUMMARY

One of the major objectives of the study was to create an intelligent system that could automatically detect early symptoms of cataracts using fundus eye images and Deep Learning technology. Cataract is among the most common causes of blindness in the world and early detection is highly critical to prevent any serious visual disorders. Manual inspection methods which usually involve experts are sometimes costly, time consuming, and inaccessible in certain locations. This led to the necessity of creation of an automatic system that would help with early detection.

For this project, the EfficientNetB0 Deep Learning model was used to train and classify images into cataract and non-cataract images. There were different phases of creating the system, such as gathering of data, preprocessing of images, training the model, testing, validation, and implementation of the final application.

Some of the features that are found in the developed system include registration, login, image uploading, cataract status prediction, and results presentation. This interface was designed in a way to make it easy for both medical professionals and patients to interact with the software without any difficulty.

Following the testing and training process, the software generated impressive results regarding the performance of the model. Some of the evaluation measures used to determine the performance of the model included accuracy, precision, recall, and validation loss. It was established that deep learning can play an essential role in helping doctors in medical diagnosis.

From the above discussion, it is evident that the set objectives were met successfully. In other words, the research proves that AI can be very effective in the health care sector in the detection of cataracts.

5.3 CONCLUSION

It has been proven in the research that this intelligent model could be used for early detection of cataract with the help of the EfficientNetB0 model. This means that the model could be utilized for analyzing fundus eye images and identifying whether there is cataract in a patient or not with reasonable accuracy.

The findings have revealed the significance of artificial intelligence and deep learning models when it comes to making diagnoses for diseases because these models could contribute significantly to the process of early detection and therefore, could become very useful for medical practitioners who need fast diagnosis.

Moreover, the project showed the necessity of combining technologies within health care systems in order to improve efficiency and reduce the burden on humans. This system will also be useful in increasing the accessibility of screening eye care services where there is no availability of medical professionals all the time.

Even though the system was successful, it still has some aspects that need to be improved in order to enhance reliability and suitability for medical practice. However, the project succeeded in achieving its objectives and showed the positive contribution of intelligent systems in early detection and managing of cataract disease.

5.4 RECOMMENDATIONS

The following recommendations can be made based on the output of this project:

1. Use of larger Dataset: More fundus eye images should be collected from hospitals and clinics for better training of the system.
2. Professional Medical Validation: The system needs to be tested and validated by ophthalmologists before implementing it in hospitals and clinics.
3. Integration with Hospital systems: The system needs to be integrated into hospital databases and medical record keeping systems.
4. Real time Image Analysis: Future version of the system can be developed that uses real time image analysis from camera or cell phone.
5. Detection of other Eyes Diseases:The system can also be used to detect other diseases related to eyes such as glaucoma and diabetic retinopathy.

5.5 FUTURE WORK

Future enhancements to the system can include not only the detection of cataracts but also help patients after they have been diagnosed. The system can be improved to suggest treatments for patients, refer them to specialists, schedule appointments, and monitor the development of cataracts. In doing so, the system can become more effective in helping manage patients.

REFERENCES

World Health Organization. (2023). Blindness and vision impairment. WHO Fact Sheet. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

National Eye Institute. (2024). Cataracts. U.S. Department of Health & Human Services. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/cataracts

American Academy of Ophthalmology. (2022). What are cataracts? https://www.aao.org/eye-health/diseases/what-are-cataracts

Orbis International. (n.d.). Cataracts: Causes, symptoms, treatment. https://irl.orbis.org/en/avoidable-blindness/cataracts-causes-symptoms-treatment

Adeniyi, M. A., Adewoye, R. K., Popoola, G. O., Aderukuola, B., Ibrahim, A. O., & Azees, A. S. (2023). Awareness, prevalence and risk burden of cataract among adults in IDO/OSI local government area, a rural community of Ekiti state, southwest Nigeria. International Journal of Community Medicine and Public Health, 10(9), 3077–3085. https://doi.org/10.18203/2394-6040.ijcmph20232663

North Florida Cataract Specialists and Vision Care. (2024). How early detection of cataracts can preserve your vision. https://www.northfloridacataract.com

Asbell, P. A., Dualan, I., Mindel, J., Brocks, D., Ahmad, M., & Epstein, S. (2005). Age-related cataract. The Lancet, 365(9459), 599–609. https://doi.org/10.1016/S0140-6736(05)17911-2

Truong, P. N., Merlinsky, E. A., Maamari, N. C., Weng, C. Y., & Lee, A. G. (2025). Tools and methods for cataract recognition in low-resource settings: A narrative review. PLOS Global Public Health, 5(5), e0004619. https://doi.org/10.1371/journal.pgph.0004619

Saha, S., Karmakar, A., & Paul, M. (2021). Artificial intelligence in the diagnosis and treatment of COVID-19: A systematic review. Frontiers in Public Health, 9, 645591.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

Ting, D. S. W., Cheung, C. Y.-L., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., Hamzah, H., Garcia-Franco, R., Yeo, I. Y. S., Lee, S. Y., Wong, E. Y. M., Sabanayagam, C., Baskaran, M., Ibrahim, F., Tan, N. C., Finkelstein, E. A., Lamoureux, E. L., Wong, I. Y., Bressler, N. M., Sivaprasad, S., Varma, R., Jonas, J. B., He, M. G., Cheng, C.-Y., Cheung, G. C. M., Aung, T., Hsu, W., Lee, M. L., & Wong, T. Y. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA, 318(22), 2211–2223. https://doi.org/10.1001/jama.2017.18152

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning (ICML), 97, 6105–6114. https://proceedings.mlr.press/v97/tan19a.html

Nigeria National Blindness and Visual Impairment Survey. (2007). Prevalence and causes of blindness and visual impairment in Nigeria. Federal Ministry of Health, Nigeria.

Kyari, F., Gudlavalleti, M. V. S., Sivsubramaniam, S., Gilbert, C. E., Abdull, M. M., Entekume, G., & Foster, A. (2009). Prevalence of blindness and visual impairment in Nigeria: The National Blindness and Visual Impairment Study. Investigative Ophthalmology & Visual Science, 50(5), 2033–2039. https://doi.org/10.1167/iovs.08-3133

Pascolini, D., & Mariotti, S. P. (2012). Global estimates of visual impairment: 2010. British Journal of Ophthalmology, 96(5), 614–618. https://doi.org/10.1136/bjophthalmol-2011-300539

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19, 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5), 1299–1312. https://doi.org/10.1109/TMI.2016.2535302

World Health Organization. (2019). World report on vision. WHO Press. https://www.who.int/publications/i/item/9789241516570

Flaxman, S. R., Bourne, R. R. A., Resnikoff, S., Ackland, P., Braithwaite, T., Cicinelli, M. V., Das, A., Jonas, J. B., Keeffe, J., Kempen, J. H., Leasher, J., Limburg, H., Naidoo, K., Pesudovs, K., Silvester, A., Stevens, G. A., Tahhan, N., Wong, T. Y., & Taylor, H. R. (2017). Global causes of blindness and distance vision impairment 1990–2020: A systematic review and meta-analysis. The Lancet Global Health, 5(12), e1221–e1234. https://doi.org/10.1016/S2214-109X(17)30393-5

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Zhang, X., Thibault, G., Decencière, E., Marcotegui, B., Laÿ, B., Danno, R., Chabouis, A., Victor, Z., & Erginay, A. (2014). Exudate detection in color retinal images for mass screening of diabetic retinopathy. Medical Image Analysis, 18(7), 1026–1043.

Li, T., Bo, W., Hu, C., Kang, H., Liu, H., Wang, K., & Fu, H. (2021). Applications of deep learning in fundus images: A review. Medical Image Analysis, 69, 101971. https://doi.org/10.1016/j.media.2021.101971

Ayanniyi, A. A., Folorunso, F. N., & Adepoju, F. G. (2021). Challenges of ophthalmology workforce development in Nigeria: The need for task-shifting and telemedicine. Nigerian Journal of Ophthalmology, 29(1), 1–6.

Ran, J., Niu, K., He, Z., Zhang, H., & Song, H. (2018). Cataract detection and grading based on a combination of deep convolutional neural network and random forests. In Proceedings of the IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC) (pp. 155–159). IEEE. https://doi.org/10.1109/ICNIDC.2018.8525852

Zhang, L., Chen, X., & Wang, Y. (2021). Multi-task learning framework for simultaneous cataract detection and severity grading using anterior segment OCT images. Medical Image Analysis, 73, 102156.

Li, X., Wang, Q., & Zhang, F. (2022). Lightweight MobileNetV2-based cataract screening model for mobile deployment in telemedicine applications. Journal of Biomedical Informatics, 131, 104104.

Gao, Y., Chen, M., Li, W., & Zhou, X. (2022). Vision Transformer-based cataract detection from anterior segment photographs: A comparative study with CNN architectures. IEEE Access, 10, 112543–112554.

Nguyen, T. H., Tran, M. Q., & Le, V. T. (2022). Federated learning for privacy-preserving cataract detection across multiple hospital nodes. IEEE Journal of Biomedical and Health Informatics, 26(10), 5123–5132.

Saha, S., Karmakar, A., & Roy, P. (2022). EfficientNetB3 ensemble model with Bayesian hyperparameter optimization for nuclear cataract severity grading from slit-lamp photographs. Computers in Biology and Medicine, 148, 105876.

Islam, M. R., Hasan, M. M., & Kabir, M. A. (2022). GAN-based data augmentation for cataract detection from fundus images using DenseNet121. Biomedical Signal Processing and Control, 78, 103995.

Wang, Z., Liu, J., & Chen, S. (2023). Dual-branch CNN for early cataract detection from smartphone-acquired ocular images integrating colour and texture features. Biomedical Optics Express, 14(3), 1124–1138.

Chen, X., Zhang, L., Wang, Y., & Liu, H. (2023). Explainable deep learning for cataract detection: Integrating Grad-CAM visualisations into ResNet-based fundus image classification. Computers in Biology and Medicine, 162, 107089.

Mahesh, V., Subramaniam, K., & Rajan, R. (2023). Cloud-based cataract screening platform using InceptionV3 in a rural telemedicine trial. Telemedicine and e-Health, 29(7), 1041–1050.

Yadav, S., Kumar, P., & Sharma, N. (2023). Benchmarking deep learning architectures for cataract detection: EfficientNetB0 achieves optimal accuracy-efficiency tradeoff. Expert Systems with Applications, 224, 119956.

Khalil, A., Hassan, F., & Al-Issa, Y. (2023). Temporal monitoring of cataract progression using a hybrid CNN-LSTM model on sequential anterior segment OCT images. Medical Image Analysis, 84, 102726.

Okonkwo, C. N., Eze, B. U., & Adeyemi, T. A. (2023). Adapting EfficientNetB0 for cataract detection in Sub-Saharan African patient populations. African Journal of Medical and Health Sciences, 22(1), 45–53.

Singh, R., & Gupta, A. (2024). Semi-supervised cataract detection using MixMatch with limited labelled ophthalmic data. Pattern Recognition Letters, 178, 112–119.

Eze, C. O., Okafor, I. P., & Nwankwo, E. U. (2024). Mobile-based cataract screening in Nigerian primary healthcare centres using EfficientNetB4 and TensorFlow Lite. Frontiers in Medicine, 11, 1345872.