IKRAME BELCADI

I-SCANNER: DETECTION AND CLASSIFICATION OF OCULAR DISEASES USING ARTIFICIAL INTELLIGENCE AND INTERNET OF THINGS

2024

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF BIOMEDICAL ENGINEERING

I-SCANNER: DETECTION AND CLASSIFICATION OF OCULAR DISEASES USING ARTIFICIAL INTELLIGENCE AND INTERNET OF THINGS

M.Sc. THESIS

Ikrame BELCADI

Nicosia

November, 2024

NEAR EAST UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF BIOMEDICAL ENGINEERING

I-SCANNER: DETECTION AND CLASSIFICATION OF OCULAR DISEASES USING ARTIFICIAL INTELLIGENCE AND INTERNET OF THINGS

M.Sc. THESIS

Ikrame BELCADI

Supervisor
Assoc. Prof. Dr. Abdullahi UMAR IBRAHIM

Nicosia
November, 2024

APPROVAL

We certify that we have read the thesis submitted by Ikrame Belcadi titled "**I-SCANNER: Detection and Classification of Ocular Diseases Using Artificial Intelligence and Internet of Things**" and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Biomedical Engineering Sciences.

Examining Committee	Name-Surname	Signature
Head of the Committee:	Asst. Prof. Dr. OMID MIRZAE	
Committee Member*:	Asst. Prof. Dr. HUZAIFA UMAR	
Supervisor:	Assoc. Prof. Dr. ABDULLAHI UMAR IBRAHIM	

Approved by the Head of the Department

30/01/2025

.....

Assoc. Prof. Dr. SÜLEYMAN AŞIR

Head of the Department

Approved by the Institute of Graduate Studies

...../...../ 202

Prof. Dr. Kemal Hüsnü Can Başer

Head of the Institute of Graduate Studies

DECLARATION OF ETHICAL PRINCIPLES

I hereby declare that all information, documents, analysis and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of Institute of Graduate Studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

Ikrame Belcadi

…../…../2024

ACKNOWLEDGMENTS

In the name of God, the Most Gracious, the Most Merciful.

Above all, Alhamdulillah, the Almighty, for His endless guidance, blessings, and strength throughout this journey. Every achievement in this work and others is by His will and mercy. Alhamdulillah for every moment of inspiration, every spark of understanding, and every challenge that shaped me into who I am today.

To my dearest parents, father, Abdelkader Belcadi and mother, Meriem Benmessaoud, thank you for your unconditional love, support, and encouragement. Dad, your faith in me has always given me strength, and Mom, your wisdom and care have guided me every step of the way. I am deeply grateful to you both for being my rock and my inspiration.

To my supervisor, Assoc. Prof. Dr. Abdullahi Umar Ibrahim, thank you for your invaluable guidance, patience, and motivation. Your mentorship has been instrumental in shaping this work, and I am deeply appreciative of your belief in my abilities.

A huge thank you to my brother Oussama Belcadi for being such an incredible help and assistance. He really came through for me, jumping in when I needed it most. I couldn't have done it without him, and I'm so grateful for his support, advice, and encouragement along the way. To my dear brothers, Yahya and Yasser, thank you for always being there for me. You have been my pillars of strength and sources of endless joy. Having all three of you in my corner has been a blessing beyond measure.

To my closest friend, Manal Azibou, thank you dear for your constant support and encouragement. Your presence has made this journey lighter, and your faith in me has kept me going during the toughest times.

Finally, I thank everyone who has supported and encouraged me throughout this journey, my extended family, friends, and colleagues. Your prayers and kind words have meant so much to me. This work is a result of all the love, support, and encouragement I've been blessed with. Thank you all, and may Allah bless you abundantly.

Ikrame Belcadi

ABSTRACT

The human retina plays a fundamental part in the vision process, it receives light using special retinal receptors and transforms it to electrical impulses that are sent to the brain which makes it a fundamental organ for detecting various ocular pathologies. Retinal imaging modalities, and notably colour fundus photography, are widely employed for the non-invasive screening of eye diseases that can be very burdensome in healthcare. Some of these eye conditions include diabetic retinopathy, cataract, and glaucoma. The integration of advanced computer-aided diagnosis frameworks that enables the usage of artificial intelligence-based technologies like convolutional neural networks had a great effect on impacting the analysis of medical scans considering that it provides an enhanced accuracy and efficiency in diagnosing retinal conditions. This study proposes an automated diagnostic tool combining artificial intelligence and internet of medical things innovations to classify retinal images into four categories of normal images, diabetic retinopathy images, cataracts images, and glaucoma images. Using transfer learning by implementing previously trained convolutional neural network architectures, such as ResNet-50, ResNet-101, VGG16, and DenseNet121 transfer learning, this diagnosis tool aims to support ophthalmologists in clinical settings by enabling a fast, accurate, and scalable diagnosis modality to assist and support early detection of ocular diseases. The aim of this proposed work is to contribute to advancing ocular health globally by integrating these technologies into healthcare systems and using them with the aim for facilitating early diagnosis and therefore allowing for more effective treatment strategies.

Key words: Artificial intelligence, colour fundus photography, ocular pathologies, diagnosis, internet of medical things.

TABLE OF CONTENT

CHAPTER IV: RESULTS AND DISCUSSION

CHAPTER V: LIMITATIONS AND FUTURE WORK

CHAPTER VI: CONCLUSION

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Adam: Adaptive Moment Estimation

AI: Artificial Intelligence

ANN: Artificial Neural Networks

CNNs: Convolutional Neural Networks

DME: Diabetic Macular Edema

GAP: Global Average Pooling

HRF: High-Resolution Fundus

IDRiD: Indian Diabetic Retinopathy Image Dataset

IoMT: Internet of medical things

IOP: Intraocular Pressure

IoT: Internet of Things

NPDR: Non-Proliferate Diabetic Retinopathy

OCT: Optical Coherence Tomography

OCTA: Optical Coherence Tomography angiography

ODIR: Organizational for Development Innovation & Research

PDR: Proliferate Diabetic Retinopathy

ReLU: Rectified Linear Unit

RGB: Red, Green, and Blue

ROC: Receiver Operating Characteristic curve

SCM: Spatial Correlation Module

SGD: Stochastic Gradient Descent

API: Application Programming Interface

CLI: Command Line Interface

GPU: Graphics Processing Unit

WSGI: Web Server Gateway Interface

URL: Uniform Resource Locator

ResNet: Residual Network

VGG: Visual Geometry Group

SQL: Structured Query Language

HTML: HyperText Markup Language

CSS: Cascading Style Sheets

CHAPTER I
INTRODUCTION

1.1 Introduction

The human eye is an intricate organ, with the retina located at the back part of the eye which acts as a vital component for visual perception, it converts the incoming light into neural impulses and then transmits them to the brain's visual cortex. This thin layer of tissue plays a crucial role for recognizing objects and scenes visually (Abràmoff et al., 2010). The deep understanding and analysis of the retinal tissue is fundamental for revealing important details and information related to the presence of various ocular pathologies and maintaining overall health (Gour & Khanna, 2021). In order to analyse retinal images, the processing of different imaging modalities is required; these include colour fundus photography, fundus fluorescein angiography (Burlina et al., 2017), and Optical Coherence Tomography (OCT) (Alsaih et al., 2017). Fundus imaging stands out among these modalities for its non-invasive nature as well as its cost-effectiveness, making it an extremely useful tool when it comes to the detection and screening of eye conditions. This imaging modality is heavily relied on by ophthalmologists when it comes to diagnosing a wide range of eye conditions such as diabetic retinopathy, cataract, hypertension, glaucoma, and other abnormalities (*ODIR-2019 - Grand Challenge*, 2020).

In recent years, the area of retinal imaging has known significant advancements encouraging the development of Computer-aided Diagnosis (CAD) based tools specifically developed to detect various eye conditions. These tailored systems' main goal is to assist medical professionals by facilitating the diagnosis process while enhancing its accuracy, thereby reducing time and effort. Fundus images are able to depict the 3-dimension retinal structure into a 2-dimension coloured Red, Green, and Blue (RGB) format, therefore providing essential visual indicators that are necessary to detect diseases (Yannuzzi et al., 2004). In order to prevent vision loss and blindness, early detection of ocular disease is a fundamental concept. The progression of ocular pathologies is mainly tracked through abnormal changes occurring in anatomical structures within the retina where we can find the fovea, optic disc, the macula, and the retinal blood vessels. These abnormalities can appear in the three coloured channels of fundus images, holding invaluable diagnostic information (Gour & Khanna, 2021). Implementing deep learning-based algorithms has become

significantly common in the field of medical image diagnosis considering the resulting remarkable performance across multiple tasks such as features recognition, medical scans classification (J. He et al., 2021b), and disease detection (Alam & Khan, 2021). Deep learning has emerged in the past few years as a popular technology of computer vision in several fields including eye disease diagnosis, attracting considerable interest from scholars considering the promising outcomes in the evolution of advanced technologies for medical image processing and analysis (Z. Li et al., 2018; Ouda et al., 2022; Richards et al., 2019; Z. Wang et al., 2022).

As opposed to machine learning techniques, the segmenting lesions and extracting features manually is not needed in deep learning, and this can prove to be inefficient when it comes to examining retinal fundus images (Ouda et al., 2022). But using deep learning also enables the creation of CAD algorithms in a highly efficient manner. Convolutional Neural Networks (CNNs) played a fundamental role in this shift considering how it enabled the extraction of disease characterizing features directly from image data (Sultan et al., 2020). These emerging networks are able master the identification of subtle patterns and image structures from basic edges to complex features indicating multiple ophthalmic pathologies (Diaz-Pinto et al., 2019; Simonyan & Zisserman, 2015).

Internet of Medical Things (IoMT), which is known as the combination of Internet of Things (IoT) and healthcare (Maitra & Chatterjee, 2006), provides unprecedented opportunities to detect and classify diseases in multiple fields including ocular health. The seamless integration of IoT devices with cloud computing resources, IoMT gives access to remote patient monitoring and facilitates real-time collaboration between doctors and medical professionals (Sriram et al., 2015). The combination of IoMT with the cloud virtualizing capabilities can prove to be very useful in handling practical constraints such as storage, power, and the efficient processing of data, which eventually enhances the scalability and efficiency of diagnosing and detecting diseases. This combination opens new doors and opportunities for the creation and development of novel healthcare services and applications specifically designed for particular needs and medical purposes. The healthcare industry is currently witnessing transformative changes in medical imaging analysis that is being reinforced and facilitated by the IoMT advancements along with their cloud computing infrastructures which provide reliable big data analytics and Artificial Intelligence (AI) assistance (Rajasekaran &Indirani, 2021).

The aim and objectives of this presented thesis is to propose an automated diagnostic modality with the aid of AI and IoT, that can classify retinal fundus images and identify some

of the most significant ocular diseases such as diabetic retinopathy, cataracts, and glaucoma that are burdensome and greatly affect global vision health. This proposed thesis study aims to develop a four-class ocular disease classification tool of: normal images, diabetic retinopathy, glaucoma, and cataracts. The used data is retrieved from the Kaggle database, which provides a rich repository of labelled fundus images publicly available for research usage. By using IoT and pretrained deep learning architectures such as ResNet-50 and ResNet-101, VGG-16 and VGG-19, and DenseNet121, the ultimate goal of the work proposed in this thesis extends beyond research and development, aiming to implement the result of this work as a diagnostic modality in real-world medical settings to assist ophthalmologists in detecting and classifying ocular pathologies in a fast and more accurate and effective way.

1.2 Overview of Ocular Diseases

1.2.1 Fundus Anatomy

The fundus is a fundamental area of the eye when it comes to understanding eye diseases like diabetic retinopathy, glaucoma, cataract, and others. It mainly consists of the retina, the macula, the optic disc, and retinal vascular structures. The retina is the photoreceptive component that is located at the posterior segment of the eye, using photoreceptor cells (called rods and cones), it detects the light and transforms it into electrical and neural signals initiating the process of vision. Its anatomy allows for the tracking of microcirculation is highly essential for vision. The retina contains several layers including a layer of photoreceptors which are responsible for detecting light and colour. The health of the retina is critical and any damage to it can cause serious vision complications and even vision loss (Shabbir et al., 2013).

The optic disc is another component of the fundus and it's the point where all the retinal veins and arteries originate from and where ganglion cell axons exit the eye. It has a circular or oval shape appearance in fundus images. Detecting the optic disc is vital when it comes to diagnosing ocular diseases; it serves as a reference point to locate other structures of the fundus. The anatomical relationship between the optic disc and the other fundus components is also significant for an accurate disease diagnosis (Shabbir et al., 2013).

The macula is the small area on retina that mainly enables central vision and colour perception and is very crucial for tasks such as reading or features recognition which require detailed vision. Anatomical and structural changes in the macula can be a sign of multiple eye

diseases which further emphasises the importance of examining it for disease analysis. Eye blood vessels including arteries and veins act as a supply medium to provide the essential nutrients and oxygen to the retina and discard any waste products. Ensuring the health of retinal blood vessels is critical considering that any blockage or damage to them can lead to serious vision problems due to conditions such as retinal artery or vein occlusion. The analysis of blood vessels is a key component when it comes to diagnosing and investigating retinal diseases. Understanding fundus components including the retina, optic disc, macula, and blood vessels and their functions serves has a vital part in the early diagnosis of ocular disease including glaucoma, cataract, diabetic retinopathy, and other eye diseases which can be very burdensome for vision healthcare and can lead to severe impairment if not addressed properly (Shabbir et al., 2013).

1.2.2 Diabetic Retinopathy

Diabetic retinopathy can be identified as the most prevalent complication among patients suffering from diabetes as well as the primary microvascular threat to people diagnosed with diabetes (Madsen-Bouterse & Kowluru, 2008; W. Wang & Lo, 2018). If not diagnosed early and treated properly diabetic retinopathy can lead to various eye disorders including an impaired vision and partial blindness. Therefore, a good understanding of the mechanisms behind diabetic retinopathy is significantly important for an accurate diagnosis and assessment of the disease as well as its management (Porta & Bandello, 2002).

The progression of diabetic retinopathy starts with mild abnormalities which are usually characterized by leaking, advancing to moderate and then severe non-proliferate diabetic retinopathy that is marked by an increasing fluid leakage or loss/damage of retinal blood vessels, which eventually leads to parts of the retina not getting enough blood and therefore resulting in retinal ischemia. This progression may eventually lead to proliferate diabetic retinopathy which is a condition that is mainly distinguished by the emergence of new vascular structures on the retina. These newly emerged vessels usually trigger the formation of fibrous tissue leading to vitreous haemorrhage, which refers to the bleeding into the gel inside the eye, and the deviation of the retina from its normal position which is also called tractional retinal detachment. Regardless of whether proliferate diabetic retinopathy is treated or left untreated, it will eventually reach an inactive stage. The resulting level on clarity of vision, also referred to as visual acuity, is highly dependent on the degree of damage occurring at that particular region of the retina (Patz et al., 1978).

Regarding symptoms, patients suffering from non-proliferate diabetic retinopathy are usually asymptomatic. The patient may experience a sudden loss of vision due to vitreous hemorrhage if the condition develops further. Patients may also notice a more prolonged loss of vision if there is fluid buildup in the retina (Fung et al., 2022).

Diabetic retinopathy is known to be primarily divided into non-proliferate diabetic retinopathy (NPDR) which involves changes occurring in the capillaries inside the retina, and proliferate diabetic retinopathy (PDR) that is primarily distinguished by the emergence of new retinal veins and arteries on different parts of the retina (Wilkinson et al., 2003). The swelling that occurs in the central area of the retina is called diabetic maculae edema (DME), and can occur in both NPDR and PDR (Amoaku et al., 2020). Classification of diabetic retinopathy can be listed as follows (Wilkinson et al., 2003).

- No retinopathy: No abnormal indicators found on dilated ophthalmoscopy
- Mild NPDR: The presence of only small abnormal bulges in blood vessels called microaneurysms.
- Moderate DR: the presence of more abnormalities other than microaneurysms that are overall less severe than NPDR
- Severe NPDR: no new blood vessel growth but shows one of the following:
    - Over 20 area of intraretinal hemorrhage in each of four sections.
    - Visible abnormalities in the retinal veins in two or more sections
    - Prominent abnormal blood vessels inside the retina in one or more quadrants.
- PDR: Shows new vascular growth (angiogenesis) outside the retina, including hemorrhage in the vitreous gel or in the front of the retina.
- Mild DME: The retina thickens, or deposits are present out of the central part of the macula.
- Moderate DME: The retina thickens or deposits near the macula centre are present.
- Severe DME: Swelling or deposits reaching the macula's central part.

Diagnosis of Diabetic Retinopathy

A thorough eye examination for a person with diabetic retinopathy usually involves testing vision by checking visual acuity, and intraocular pressure, evaluating the front part of the eye by performing slit-lamp biomicroscopy, as well as performing dilated funduscopic test to check the back of the eye after dilating the pupil. Using an ophthalmoscope, even non-specialists can check and examine the fundus. Doctors are then able to diagnose and classify

diabetic retinopathy mostly by searching for specific abnormalities during the test (Fung et al., 2022). For NPDR, indicating signs include microaneurysms, small bleeding spots, and residues referred to as exudates. In some cases, cotton wool spots indicating patches of damaged nerve fibres, as well as swollen veins and an abnormal capillary growth can be present. Meanwhile in PDR, new blood vessels may grow on the optic nerve or elsewhere in the retina. Growth of fibrous tissue can occur in severe cases causing the retina to deviate from its normal position and thus distort vision (Fung et al., 2022).

1.2.3 Glaucoma

Glaucoma is a serious eye condition which can contribute to significant visual impairment and even vision loss if not diagnosed and managed early and properly. It is mainly characterised by progressive degeneration to the optic nerve which eventually also affects the transmission of visual information from the retina to the brain. This damage results most of the times from the increased eye pressure, also referred to as intraocular pressure (IOP), although other factors can also be responsible for the disease progression (H. S, 2017; Jayaram et al., 2023; Sahu, 2024).

The symptoms of glaucoma can be quite subtle and often developing gradually as time progresses. Many patients may not notice any vision changes until later stages of the disease. The most common symptoms of glaucoma include peripheral vision loss, which is considered as one of the first signs, where patients may notice that they're losing their side vision gradually. Another symptom of glaucoma is blurred vision and halos around lights, where patients may experience blurriness in low light conditions and see halos especially at night. Patients with glaucoma can also experience eye pain or discomfort in the eye, particularly in acute forms of glaucoma. (Cohen & Pasquale, 2014; Supuran, 2019) Because of the insidious nature of the disease regular eye examinations are essential for early diagnosis and appropriate control of this condition. (Lee & Higginbotham, 2005)

The underlying mechanisms behind glaucoma involve multiple cellular complex interactions withing the eye. The hypothesis that is generally accepted states that an elevated IOP exerts force and tension on the head of the optic nerve that leads eventually to the degeneration of the retinal ganglion cells through a process called apoptosis (Sahu, 2024). In addition to that, other factors contributing to this disease include:

- Oxidative stress: refers to the damage to retinal cells caused by free radicals.

- Inflammation: chronic inflammation in the eye can also cause significant harm to the optic nerve.
- Neurotrophic Factor Deficiency: has potential to hinder the survival of retinal ganglion cells due to a lack of essential growth factors.
- Mitochondrial Dysfunction: impaired production of energy in retinal cells can lead to cell death and cause further glaucoma progression (Sahu, 2024).

Diagnosis of Glaucoma

Diagnosing glaucoma requires a thorough eye examination process. Intraocular pressure measurement is a critical test considering that intraocular pressure can increase the likelihood of development of eye conditions like glaucoma. Visual field testing is another diagnosis approach used to assess peripheral vision and identify any loss that may be a sign of glaucoma (Lim, 2022). Optic nerve evaluation can also be conducted using imaging techniques such as ophthalmoscopy, also known as fundus photography, fluorescein angiography, or optical coherence tomography which allow for a detailed examination of the optic nerve and detect any structural changes or damage (Borrás, 2012).

1.2.4 Cataract

Cataracts are a major cause of reversible blindness and visual impairment worldwide. These conditions are mainly characterized by blurring in the ocular lens, which causes a decreased transparency and clarity of vision. Cataract can be identified by an abnormality in the eye's lens that is essentially marked by a reduced transparency and increased cloudiness (Lam et al., 2015). This condition can manifest in different forms; mature cataracts can usually be classified into two classes: brunescent cataract and white cataract. This classification is mainly determined by whether it is the nucleus or the cortex of the lens that was affected and became opaque (Chakrabarti et al., 2000; Song et al., 2014; Vasavada et al., 1998).

The development of cataract involves complex mechanisms that cause structural changes in the crystal lens. Proteins inside of the lens tend to denature over time leading to the formation of opacities that prevent the light from reaching retina (Chakrabarti et al., 2000; Vasavada et al., 1998). If not treated properly cataract can worsen progressively and eventually compromise visual function leading to secondary diseases such as glaucoma or

uveitis. Common signs of this disease include a hazy vision, trouble seeing in low-light conditions, an increased sensitivity to bright lights, and the perception of light halos. Some patients may also experience changes in colour perception with colours appearing faded or yellowed. At later stages of the disease, these symptoms can impair daily activities and quality of life significantly (Lam et al., 2015).

Diagnosis of Cataract

The diagnosis of cataracts generally includes a detailed eye examination conducted by specialized physicians (ophthalmologists). This may involve tests to measure visual acuity, slit-lamp exams, and assessments of the lens's clearness (Chuck et al., 2021). Additionally, some advanced imaging modalities like optical coherence tomography, fundus photography, or fluorescein angiography can be utilized as well to evaluate the extent of cataract formation and its impact on the eye structure (Gus et al., 2000; Rocha et al., 2007).

1.3 AI and Deep Learning in the Detection of Ocular Diseases

1.3.1 AI in Ocular Disease Detection

AI has known several advancements during last decades increasing the interest in the computer science and medical fields. AI involves the design and creation of systems that are able to imitate human-like cognitive abilities and solving problems. During last years, the area of retinal imaging has known significant advancements encouraging the development of CAD based tools specifically designed for the detection of various eye conditions. These tailored systems' main goal is to assist medical professionals by facilitating the diagnosis process while enhancing its accuracy and therefore reducing time and effort. Fundus images are able to depict the 3-dimension retinal structure into a 2-dimension coloured RGB format, therefore providing essential visual indicators that are necessary to detect diseases (Yannuzzi et al., 2004).

In relation to the prevention of vision loss and blindness, we can say that early diagnosis of ocular disease is a fundamental concept. The progression of ocular pathologies is mainly tracked and monitored through abnormal changes occurring in the anatomical structures within the retina including the macula, fovea, optic disc, and retinal vascular structures. These abnormalities hold invaluable information and details about potential eye diseases and can be visualised in the three coloured channels of fundus images (Gour & Khanna, 2021). The process of implementing deep learning-based systems has become significantly common in

research related to the analysis and processing of medical images considering their ability in demonstrating remarkable performance across different tasks such as features detection, recognition of objects, medical scans classification (J. He et al., 2021b), and disease detection (Alam & Khan, 2021). Deep learning has emerged in the past few years as a popular technology of computer vision in several fields including eye disease diagnosis, capturing considerable interest from scholars inn research considering its potential and promising outcomes in the development of advanced algorithms that can be able to analyse medical images process them (Z. Li et al., 2018; Ouda et al., 2022; Richards et al., 2019; Z. Wang et al., 2022).

As opposed to the usual machine learning methods, deep learning discards the need for segmenting lesions and extracting features manually which can prove to be inefficient when it comes to examining retinal fundus images (Ouda et al., 2022). Using deep learning enables the innovation and evolution of CAD algorithms in a highly efficient manner. CNNs played a fundamental role in this shift considering how it enabled the extraction of disease characterizing features and patterns directly from image data (Sultan et al., 2020). These emerging networks are able to master the identification of subtle patterns and image structures from basic edges to complex anatomical features indicating multiple ophthalmic pathologies (Diaz-Pinto et al., 2019; Simonyan & Zisserman, 2015).

One of the applications of AI in ophthalmology is the diagnosis and detection of diabetic retinopathy. While previous attempts have been made to integrate and make use of computerized systems into diabetic retinopathy detection process, recent advancements in deep learning have encouraged many countries to renew their approach in diagnosis of using AI for diabetic retinopathy diagnosis. The advantages and promising results of these research works speak for themselves seeing how the outcome often exceed standard screening guideline recommendations in terms of sensitivity and specificity (Wong & Bressler, 2016).

1.3.2 Deep Learning

A rapidly growing area within machine learning models are deep learning models. They use artificial neural networks (ANN) to extract subtle features and patterns from image data with architectures of multiple layers (Dutta et al., 2018). These models have proven highly effective across various tasks in computer vision and biomedical imaging analysis like it was demonstrated by various studies (Guo et al., n.d.; Zhang et al., 2019). One of the most

popular choices for researchers was deep CNNs as they've emerged as a favourite, especially for the classification of natural images and success in medical image classification. For instance, CNN models have shown promising results by successfully classifying fundus images into NPDR achieving high performance metrics percentages. Diabetic retinopathy grading systems have been made more efficient, accessible, and cost-effective through some added improvements that have been validated across large datasets of high-quality images and various settings which surpassed traditional tailored methods based on features (Abbas et al., 2018; Galveia et al., 2018).

Deep learning is significantly important for enhancing intelligence by automating various processes such as environmental control and simplifying disease detection within the medical industry. Recently, several automated methods were developed with the aim of detecting diabetic retinopathy (Islam et al., 2020). When specialists diagnose fundus images manually, they search for blood vessels abnormalities, deposits, and leaking of substances such as blood and other fluids. As a consequence, much of research studies have been directed towards the automated identification and classification of these lesions without human intervention (Ahmad et al., 2014).

CHAPTER II
RELATED WORK

2.1 Literature Review

Multiple studies have been conducted in the research area related to ocular disease classification to detect and predict several eye conditions with the usage of deep learning and machine learning algorithms. These were conducted using different approaches and datasets. Table 2.1 summarizes the work done on multiple papers.

Table 2.1 Summary of related work

References	Type of Images	Classes	Approach	Performance
Tayal et al. (2022)	OCT	4	Five-layered CNN Seven-layered CNN Nine-layered CNN	Accuracy: 96.5% sensitivity: 94.47% specificity 98.16% F1 Score: 95.80%
Gour & Khanna (2021)	Fundus images	8	ResNet-50, InceptionV3, MobileNet, VGG16	Accuracy: 85.34%, precision: 84.5%, recall: 83.7%, F1-score: 84.1%, AUC 84.93%
J. He et al. (2021)	Fundus images	8	ResNet-18, ResNet-34, ResNet-50, ResNet-101, SCM	AUC: 93% F1-score: 91.3%.

J. Wang et al. (2020)	Fundus images	8	EfficientNetB3	Accuracy: 89%, precision: 63%, recall: 58%, AUC: 73%, F1-score: 89%
Cheng et al. (2020)	Fundus images	8 (diabetic retinopathy lesions)	ResNet-101, GCN	Average overall F1-score: 80.8%
Dipu et al. (2021)	Fundus images	8	ResNet-34, MobileNetV2, EfficientNet	Accuracy: 97.23%
F. Li et al. (2022)	Fundus images	2	Inception-V4	AUC: 97.2%, sensitivity: 92.3% for referable diabetic retinopathy.
Reguant et al. (2021)	Fundus images	5	Inception-v3, ResNet50, InceptionResNet50, Xception	Accuracy: 95%, AUC: 98%, specificity: 0.96%, sensitivity: 0.86%
Pektaş, M. (2023).	Fundus images	8	MobileNet, EfficientNet, SqueezeNet	Accuracy: 96.64%, F1-score: 0.6870

Ryu et al. (2021)	OCTA	4	ResNet101, Machine learning-based classifier	Accuracy: 95.4%, AUC: 96.7% sensitivity: 98.1%, specificity: 98.1%

Tayal et al. proposed an approach to develop a diagnostic tool using deep learning-based models to automatically recognise and sort ocular diseases into four categories: Normal, DME, choroidal neovascularization, and drusen by analysing OCT scan images. The data set of OCT retinal scans used in this study were retrieved from a public source (Mendeley database) which was published in Kermany et al. ‘s work. Three different CNN models were employed with a varying number of layers (five, seven, and nine layers). These models were trained on the pre-processed OCT retinal scans to recognize patterns and detect different features characterizing each ocular pathology. The pre-processing steps involve resizing images to 150×150 pixels, then center-cropping them to 128×128 pixels. The images were divided into training (90.16%), validation (1.84%), and testing (8%) sets. All models used Rectified Linear Unit (ReLU) functions to improve gradient flow, they also reduced special dimensions using max pooling and at the they used dense layers to combine the learned features and make final predictions. A training rate of 0.001 was also set using the Adam (Adaptive Moment Estimation) optimizer (Kingma, 2014) . The trained CNN models were then evaluated on separate sets of OCT scans to evaluate their capability in categorizing the four ocular diseases. Comparison of the presented approach with manual ophthalmological diagnosis was also conducted demonstrating a high classification accuracy of 96.5%, an F1 score of 95.80%, 94.47% in sensitivity, and 98.16% in specificity. The limitations of this work include:

- **Data Diversity**: The dataset used in this research was retrieved from a single population and this limits the diversity of the ocular structures that were presented. This lack of diversity can also affect the model's generalizability to different populations and races and therefore lead to biased results.
- **Image Type Specificity**: This study focused exclusively on OCT scans, where other types of imaging like fundus photographs or angiographic images were not included.

This indicates that the model may need retraining to effectively classify diseases using these different imaging modalities.

- **Uniform Scanning Techniques**: All scans in the dataset were taken using the same scanning settings and techniques. This may not realistically represent the variability encountered in real-world clinical settings that make usage of different equipment and techniques.
- **Model Efficacy**: The current model's efficacy across different systems and conditions is not yet proven. The study suggests that more research is still needed to try and test multiple options for dimension reduction and enhance the models' robustness.
- **Limited Disease Analysis**: The presented approached focuses on a limited set of ocular pathologies. While it successfully identifies DME, drusen, and choroidal neovascularization, it still doesn't include other significant diseases including diabetic retinopathy, glaucoma, or age-related macular degeneration. Expanding this approach to include the above mentioned conditions can further improve its clinical utility.
- **Image Size Reduction**: This study reduced the input image size to (128,128) dimension to optimize input variables. While this approach can help with reducing computational load it can also lead to a loss of critical and essential information that could be important for an accurate diagnosis.

Gour & Khanna presented a study with an automated CNN-based approach to detect eye pathologies from retinal fundus images which is a common modality employed in clinical ophthalmology. This study uses an 8-class classification modality by automatically detecting cataract, diabetic retinopathy, age-related macular degeneration, myopia, glaucoma, and normal images. The proposed method involves the usage of four pre-trained CNNs on a dataset of fundus images acquired from the Organizational for Development Innovation & Research (ODIR) database using transfer learning. Notably, this approach mainly focuses on multi-classification and multi-labelling of fundus image pairs of both eyes of each patient. A double approach was presented in this work: Model-1 processes individually the right and left eyes of each patient in the dataset using pre-trained CNN frameworks while Model-2 puts the left and right eyes concatenated as an input. The extraction and classification of the final features was performed using global average pooling (GAP) along with the sigmoid activation and loss functions. The preprocessing for this approach involves resizing all images to 224x224 pixels, applying normalization, and augmenting the dataset by applying rotation,

flipping, and zooming operations to reduce overfitting. The images were divided into training, validation, and testing using a (70: 15: 15) ratio to evaluate the model's performance, ensuring the reproducibility by setting a random seed of 42.The final results of this study have demonstrated that VGG16 with the Stochastic Gradient Descent (SGD) optimizer performs best across all labels and categories, with a promising outcome for a clinical practice integration as a CAD tool to detect and classify ocular pathologies. The results demonstrated an accuracy of 85.34%, a precision of 84.5%, a recall of 83.7%, and an F1-score of 84.1%. Those results emphasize the effectiveness of the model in classifying ocular diseases, demonstrating its potential for reliable application in clinical environments. The limitations of this study can be listed as follows:

- **Generalization Capability**: The MobileNet model does not generalize well for the ODIR database despite being lightweight. On the other hand, the VGG16 architecture despite being heavier, provides a better outcome in terms of performance for this specific dataset.
- **Class Imbalance**: The ODIR database shows a class imbalance problem considering that certain disease classes such as glaucoma and cataract, have significantly less images compared to others like diabetic retinopathy. This imbalance can lead to a biased model performance in which the model may generalize well on classes with more data but poorly on under-represented classes.
- **Overfitting**: This study mentioned that the concatenated input approach leads to overfitting, which is a common problem when using deep learning. This can also be a sign that the model may not generalize effectively to unseen data which can limit its practical application in clinical settings.
- **Limited Testing Set**: The test data set for evaluation does not contain ground truth labels and this issue can affect the reliability of evaluation parameters like the F1-score and the Area Under the Curve (AUC). This limitation puts in question the validity of the results reported.
- **Performance Evaluation**: The class-based performance analysis is conducted on a limited validation set, which might not accurately reflect the model's true capability across all classes. This can also lead to misleading conclusions about the model's effectiveness.

- **Complexity of Multi-label Classification**: The paper transforms a multi-labelling task of identifying multiple diseases in a single image into a multi-class task where only one disease can be identified from a single image. And this might not actually reflect the complexities of the image features. This simplification can also limit the effectiveness of the model when it comes to classifying images having multiple diseases.

J. He et al. proposed a CAD approach where they presented a dense correlated network (DCNet) to categorize coloured retinal fundus images. Their work used the ODIR-2019 database which contains seven classes of ocular pathologies. They employed a DCNet architecture consisting of a backbone CNN that is primarily assigned to extract features, in addition to a Spatial Correlation Module (SCM) which computes pixel-wise correlations between left fundus and right fundus features sets to refine and update them, along with classifier that uses fully connected layers to produce an eight-category output of ocular diseases. The pre-processing of fundus images in this work includes resizing them to 224x224 pixels, and normalizing them, applying rotation, shifting, zooming, and flipping to add more the diversity to training set and enhance the model's robustness. The images were also split into a (70: 15: 15) ratio for training, testing and validation respectively. Their approach mainly relies on trying various backbones of pre-trained architectures consisting of multiple versions of the ResNet model with varying depths (18 layers, 34 layers, 50 layers, and 101 layers). The model with the ResNet101 backbone showed the best results with an AUC of 93%and an F1 Score of 91.3%. This study faced a major data imbalance issue that can potentially affect the resulting performance of the used models. The advantage of this study is its applicability to multi-modal image analysis; however, their use of a patient-based methodology made the comparison with other related studies challenging and not possible. The limitations of this study can be listed as follows:

- **Limited Training Samples**: The study acknowledges that the performance of the models may be hindered by the limited set of available training images. This can lead to inadequate training of the network, affecting its ability to generalize effectively when introduced to unseen images.
- **Network Depth and Performance**: The research indicates that increasing the depth of the backbone CNN does not always lead to improved performance. Specifically, the

transition from ResNet-50 to ResNet-101 shows only minor enhancements, suggesting that deeper networks may not always be beneficial due to issues like gradients becoming too small to effectively learn and features not being used efficiently.

- **Computational Complexity**: This paper highlights the fact that deeper networks, even if potentially they're more powerful, they also come with an increased computational complexity. For instance, using a ResNet-101 backbone may not be necessary if computational resources are limited, which can potentially limit the model's applicability in low resources environments.
- **Focus on Specific Ocular Diseases**: The model in this study primarily addresses certain ocular diseases and this can limit its applicability to a wider range of conditions. This study also highlights that existing works often focus on a single or a few disease categories which doesn't necessarily reflect the challenges of real-life scenarios where patients are likely to carry multiple eye diseases.
- **Image Quality**: The performance of the proposed model is mainly dependent on the quality of colour fundus photographs. Poor image quality can significantly impact the model's performance by making it less reliable in practical settings where conditions of image acquisition may vary.

J. Wang et al. presented a transfer learning modality to extract features from coloured fundoscopy images from the ODIR5K dataset that are labelled into eight categories. The images were divided into 90% as a training set, and 10% as a validation set, they additionally underwent normalization, cropping to a 1:1 ratio, and resizing to 448x448 pixels. They additionally employed data augmentation to enhance the dataset by applying rotations of 45 and 90 degrees as well as random translations and histogram equalization to improve contrast. This study used the EfficientNetB3 pre-trained CNN architecture and was integrated into a transfer learning framework with adjusted weights to better capture the specific features for multi-label classification. They additionally integrated two weak classifiers independently trained as an approach to improve generalization and improve the overall performance. The model showed promising results when tested on 40 fundus images from the ODIR images with an accuracy of 89%, a precision of 63%, 58% recall, an AUC of 73%, and an F1-score of 89%. The challenge of data imbalance was also addressed in this paper; however, its major disadvantage is the issues with low network performance due to the "other diseases" category

and the uncommon optic discs in the images of their used dataset. The limitations of this work include:

- **Dataset Limitations**: The ODIR-2019 dataset is used in this work. this dataset has a limited amount of data for certain eye diseases. One of the labels, 'O' which represents (other diseases) class, includes a variety of uncommon fundus diseases. This also makes it challenging to improve the model's performance for these cases.
- **Overfitting**: It was indicated that the model might be overfitting during training. The results of the validation set were significantly higher than those on the testing set which suggests that the model learned too well from the training data and may not generalize effectively to new data.
- **Black Box Nature of Deep Learning**: A fundamental limitation of deep learning networks, including the ones used in this study, is their "black box" nature. Even if the network can automatically learn features from images, the specific features it learns remain unknown. This unpredictability issue can limit the clinical understanding of the model's decision-making procedure.

Cheng et al. employed a graphical convolution network for to classify eight diabetic retinopathy lesions from colour fundoscopy images. The data used in this study was collected from multiple hospitals with a total of 7459 fundus images. Each image was labelled by two ophthalmologists into their respective diabetic retinopathy lesion. Using the ResNet-101 CNN architecture to extract features from the retinal images, the data were split into a ration of (70: 15: 15) for training, testing, and validation respectively, and were resized to 1024x1024 pixel sizes. The SGD optimizer was employed with the loss function. The model achieved a superior performance with a highest F1-score of 80.8%, and a highest AUC score of 98.6% for the laser scars lesion. This study succeeded in achieving better accuracy for specific lesions but struggled in detecting others due to their appearance characteristics. The limitations of this study can be listed as follows:

- **Detection Challenges for Specific Lesions**: The model in this paper demonstrated better detection results for certain lesions compared to others. This indicates that the model's ability to detect all types of lesions in an accurate way is inconsistent which can potentially lead to misdiagnosis in clinical settings.

- **Visual Similarity of Microaneurysms**: Microaneurysms are particularly hard for the model to detect because they appear as small red spots in the blood vessels of the retina. This visual similarity to the background can make it confusing for the model to point them out especially when the original images are resized for input.
- **Complexity of Coexisting Lesions**: The presence of multiple lesions in a single fundus image complicates the model's capability to extract features in an efficient way. As an example, soft and hard exudates are often present along with other lesions and can obscure their detection and make it difficult.
- **Dependence on Image Quality**: The quality of fundus photographs used in this work can significantly affect the model's performance. If the images are blurry, noisy, or poorly lit, the model may struggle to correctly identify and classify lesions. Similar issues have been noted in previous research, where image imperfections reduced the accuracy of detecting specific features because of the added difficulty of distinguishing between noise and meaningful patterns.

Dipu et al. showed a transfer learning approach using several deep learning models' architectures to detect and classify eight ocular pathologies from retinal fundus images acquired from the ODIR2019 database. In their work, they have reported the accuracy obtained from each model, however their evaluation was limited to the accuracy metric alone.

F. Li et al. proposed a work where they improved the Inception-V4 structure to make it compatible with the task of classifying two ocular pathologies: Diabetic retinopathy and DME. They used retinal fundus images from 2,966 patients and applied some data augmentation including flipping and rotations after cropping them to remove borders. The images were then divided into training, validation, and testing splitting with no overlap between patients across the splits. They succeeded in achieving a high performance with an AUC of 97.2% and a sensitivity of 92.3% for referable diabetic retinopathy. The limitations of this study can be listed as follows:

- **Grader Bias**: In this study, six experienced ophthalmologists graded the retinal images, and their majority decision was used as the reference standard for training and validating the model. While this ensures expert input, it might also introduce biases based on the graders' individual perspectives or tendencies, which could affect the accuracy of the true labels the model learns from.

- **Pre-screening by Human Graders**: Before adding images to the dataset, human graders checked them for quality and signs of other diseases. While this ensured a clean dataset, it might have excluded images that could have been helpful for training, reducing the variety of the training data.
- **Insufficient Representation of Clinical Practice**: Even though the dataset is high quality, it might not fully represent the challenges and conditions seen in real-world clinical settings. This could limit how well the model performs in actual clinical use.
- **Dataset Size and Performance**: The study suggests the model could perform better with a larger dataset. The current dataset might not be big enough to fully evaluate the model's potential, highlighting the importance of further validation using bigger and more varied datasets.

In Reguant et al.'s work, they visualized the decision process of neural network and they assessed features of retinal images in order to classify diabetic retinopathy into five stages. The images from the datasets had varying quality, dimensions, and aspects ratios, so the pre-processing step involved removing blurred, overly dark and bright images. The fundus images were then resized to 512x512 pixel values after experimentation with different dimensions, and the non-symmetric images were cropped to square shape. A ratio of 80:10:10 was used to divide the dataset into training, testing, and validation respectively. Data augmentation was also implemented to address class imbalance by applying oversampling and undersampling techniques to even out the number of images in each class, adjusting their number to 500 images per class. Other techniques for augmentation like rotation, height shift, width shift, scaling, and flipping were also applied to increase the diversity in the dataset. This study experimented with four CNN architectures: Inception-v3, ResNet50, InceptionResNet50, and Xception. They employed transfer learning for models' initialization by loading pre-trained ImageNet weights. Weights were also initialized using Glorot uniform initializer with the top layers of the base CNN architectures. Evaluation metrics showed that the Inception-v3, ResNet-50, InceptionresNet50, and Xception models achieved accuracies between 89% to 95% and AUCs ranging from 95% to 98%, with Xception being the model with the winning performance as it achieved 95% in accuracy, and an AUC of 98%. The limitations of this study can be listed as follows:

- **Limited Datasets**: The training and validation datasets in this study were quite small. While transfer learning helped improve the model's performance, the small dataset

size meant the models relied heavily on pre-trained weights. This suggests that using larger datasets could lead to significantly better results.

- **Model Architecture Constraints**: The study faced some limitations with the design of the model. In order to make the model easier to understand, the CNN layers were placed near the output of the models which made it impossible to combine multiple CNN models. This issue may have restricted the model's overall performance for the sake of simplifying its structure.
- **Image Quality Issues**: This study worked with high-quality images, but these don't always match the kind of images seen in real-world situations. Since the tool doesn't check for image quality, low-quality images could then result in wrong predictions. This shows the importance of having tools that can assess image quality before making predictions.
- **Predictive Variability Across Classes**: The models performed differently depending on the diabetic retinopathy grade. For example, the models struggled with grade zero, where the features are very subtle. This uneven performance means the models might not be as reliable for lower grades, which could impact clinical decisions.

In Pektaş's study, they experimented with multiple versions of three CNN architectures: MobileNet, EfficientNet, and SqueezeNet. This study used retinal fundus images from the ODIR-5K database to classify them into eight categories. The images were pre-processed and resized to 224x224 pixels and normalized to a 1:1 aspect ratio, to split them later into 70% for training and 30% for testing. The CNNs were trained using Adam optimizer along with early stopping with the goal of preventing any potential overfitting. Different data augmentation methods were tested; however, the best results were achieved without data augmentation. The best performing model was EfficientNetB3 with a training/testing split ratio of 90:10 and was selected based on the accuracy and F1-score metrics with values of 96.64% and 0.6870 respectively, outperforming both the MobileNet and SqueezeNet models. The limitations of this study can be listed as follows:

- **Dataset Limitations**: The study highlights that many researchers face challenges in achieving good performance for diabetic retinopathy classification because of limited training data and inconsistent annotations. This issue can make it harder for the model to perform well on different datasets.

- **Focus on Specific Models**: This research mainly focuses on the EfficientNetB3 model, which showed high accuracy and good performance. However, it doesn't explore other models or architectures that might work better in different situations or with other datasets.
- **Hyperparameter Tuning**: While this study mentions plans to use automatic hyperparameter tuning in the future, the current approach may not fully optimize the model. This means the results might not show the model's best possible performance.
- **Generalizability of Results**: The findings are based on the ODIR-5K dataset, which may limit how applicable they are to other datasets or real-world settings. The study doesn't explore how the model would perform on different populations or imaging conditions, which is important for its clinical use.
- **Potential Overfitting**: The model's high reported accuracy (96.94%) raises concerns about overfitting, especially since the dataset used for training is relatively small. Without thorough validation on new, unseen data, the model's reliability isn't fully clear.

Ryu et al. proposed a diabetic retinopathy diagnostic approach using a CNN model by detecting features from optical coherence tomography angiography (OCTA) scans, and they succeeded in achieving a relatively high performance. Their approach included preprocessing the images to remove scans of low quality and using different OCTA image resolutions (3×3 mm² and 6×6 mm²). A four-fold cross-validation was employed to train and assess the two models they experimented with: a ResNet101-based classifier with OCTA images as an input, and a machine learning classifier using extracted local features from the OCTA scans. The model which achieved the best performance was the ResNet101-based model with accuracies and AUCs of 95.4% and 96.7% respectively for diabetic retinopathy, and 97.5% and 97.6% for referable diabetic retinopathy. The sensitivity and specificity of the model were notable as well, reaching 98.1% sensitivity and 98.1% specificity for referable diabetic retinopathy detection. The limitations of this study can be listed as follows:

- **Small Sample Size**: A relatively small group of patients was included, which might limit how well the results apply to other populations. While the sample size is similar to other studies using OCTA, it's still a factor that could affect the reliability of the findings.

- **Potential for Misclassification**: The study included a relatively small group of patients, which might limit how well the results apply to other populations. While the sample size is similar to other studies using OCTA, it's still a factor that could affect the reliability of the findings.
- **Subjectivity in Traditional Methods**: Traditional methods for diagnosing diabetic retinopathy can be subjective, leading to inconsistent results. This shows the need for automated tools but also highlights the weaknesses of current manual approaches.
- **Generalization Issues**: This work mentions that models based on handcrafted features may not perform well on different datasets due to overfitting. While the CNN model shows potential, its ability to work in a variety of clinical settings may still be limited.

2.2 Limitations of Existing Studies

The existing studies on ocular disease detection and classification that were mentioned above face various limitations. First, several studies rely on limited datasets in terms of them not representing the full range of ocular pathologies that a patient may have. This can be challenging for AI systems to accurately detect and correctly classify eye conditions. Secondly, even though some of these studies focus on specific conditions such as glaucoma, diabetic retinopathy, or AMD, they do not consider the possibility that some patients may be carrying multiple eye disease at once, which indicates that the developed AI systems in these studies are incapable of handling cases where multiple eye pathologies are present. In addition to that, some of the existing studies in this area use deep learning-based approaches without necessarily validating their effectiveness on diverse datasets, and this can make it challenging for other researchers to build upon their findings.

Another limitation is class imbalance; several studies used datasets that show significant imbalances with certain diseases (e.g., diabetic retinopathy in the ODIR-5K database) more represented than others. This issue can lead to models performing better on the more represented diseases while struggling with the underrepresented diseases, which can significantly affect the overall performance. Overall, despite the fact that deep learning classification-based systems prove to be very promising in detecting and diagnosing ocular pathologies, more search is needed to address these challenges and make sure that these emerging technologies are effectively used in clinical environment.

CHAPTER III
METHODOLOGY

3.1 Overall Methodology

This section covers the overall methodology/experimental set-up which is divided into 5 stages. Stage 1 revolves around data collection from publicly accessible domains such as Kaggle, stage 2 will involve data preparation and pre-processing. Stage 3 will involve models' construction, evaluation, training and visualization. Pre-trained deep learning architectures that are used in this work are ResNet-50, ResNet-101, VGG-16, VGG-19, and DenseNet121. Stage 4 will focus on evaluation of the generalizability of the proposed approaches using performance metrics and final stage will involve the development of IoT-based framework. The general methodology is summarized in figure 3.1.

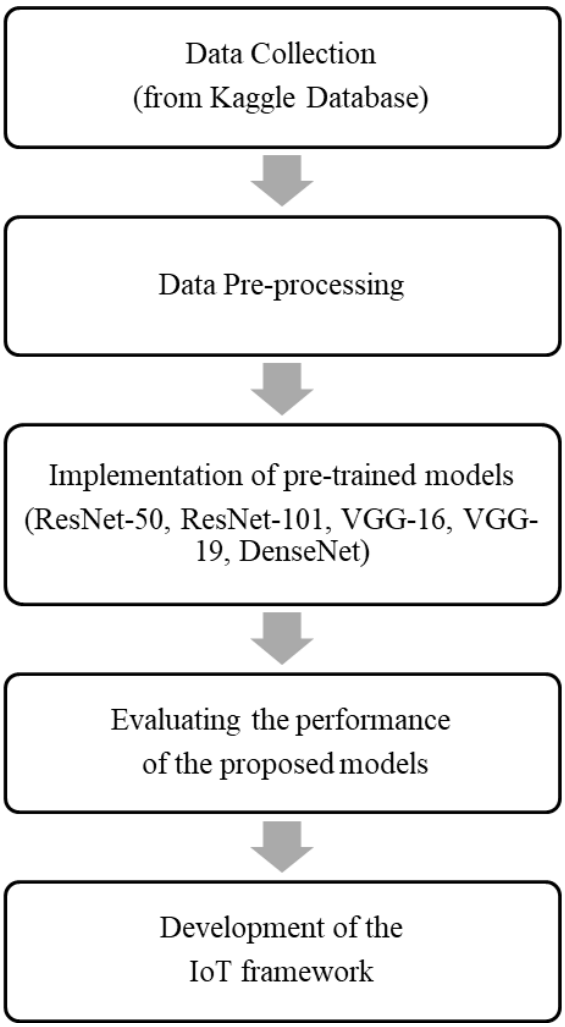


Figure 3.1 Overall methodology

3.2 Data Description

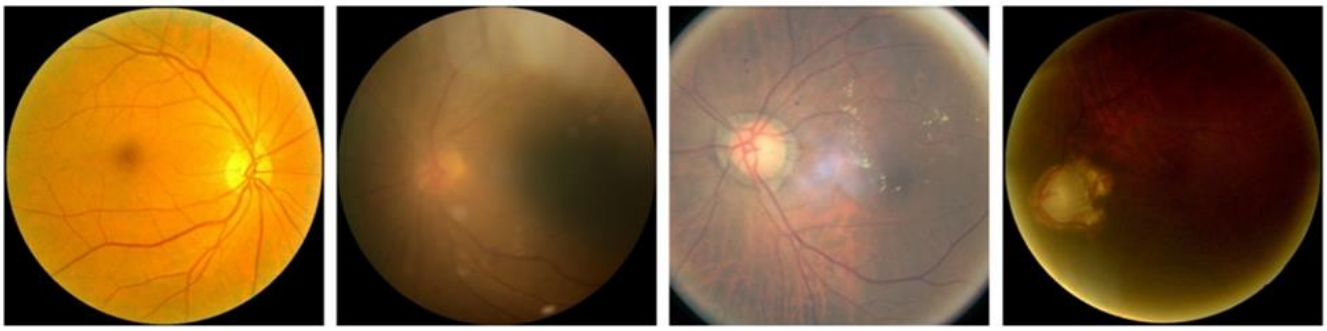
The dataset is curated from Kaggle repository title “eye_diseases_classification”. The dataset comprises 4,217 retinal colour funduscopy retinal images retrieved from diverse databases like the Indian Diabetic Retinopathy Image Dataset (IDRiD) (Porwal et al., 2018), High-Resolution Fundus (HRF) image database (Budai et al., 2013), and other sources. The retinal fundus images in this dataset are organized in 4 folders: normal, diabetic retinopathy, glaucoma, and cataract. Table 2 provides the website source where the data was collected from and gives a brief description of the characteristics of the collected data.

Table 3.1 Description of dataset

Repository	Characteristics	Website
“eye_diseases_classification”	Colour fundus images grouped into 4 classes: normal, diabetic retinopathy, cataract, and glaucoma	https://www.kaggle.com/datasets/gunavenkatdoddi/eye-disease

3.3 Data Visualisation

A sample image from each class was displayed as shown in figure 3.2 to give a quick visual confirmation of data structure. The dataset contains a total number of 4,217 images of the retina, with 1038 images belonging to cataract, 1098 images to diabetic retinopathy, 1007 images belonging to glaucoma, and 1074 to normal fundus. To better understand the dataset distribution, the number of images in each disease folder is calculated and visualized using a bar plot as shown in figure 3.3. This step is important to verify that the classes are balanced and gives insight into potential class imbalances that could affect training. Additionally, images in this dataset have varying size dimensions; a bar plot is also used to display the number of fundus images in each size dimension as shown in figure 3.4. It shows how this dataset contains images with pixel sizes of (512×512), (256×256), (2592×1728), (2464×1632), and (1848×1224).



(a) (b) (c) (d)

Figure 3.2 Samples of colour fundus retinal images from the dataset where (a) represents normal; (b) represents cataract; (c) represents diabetic retinopathy; and (d) represents glaucoma

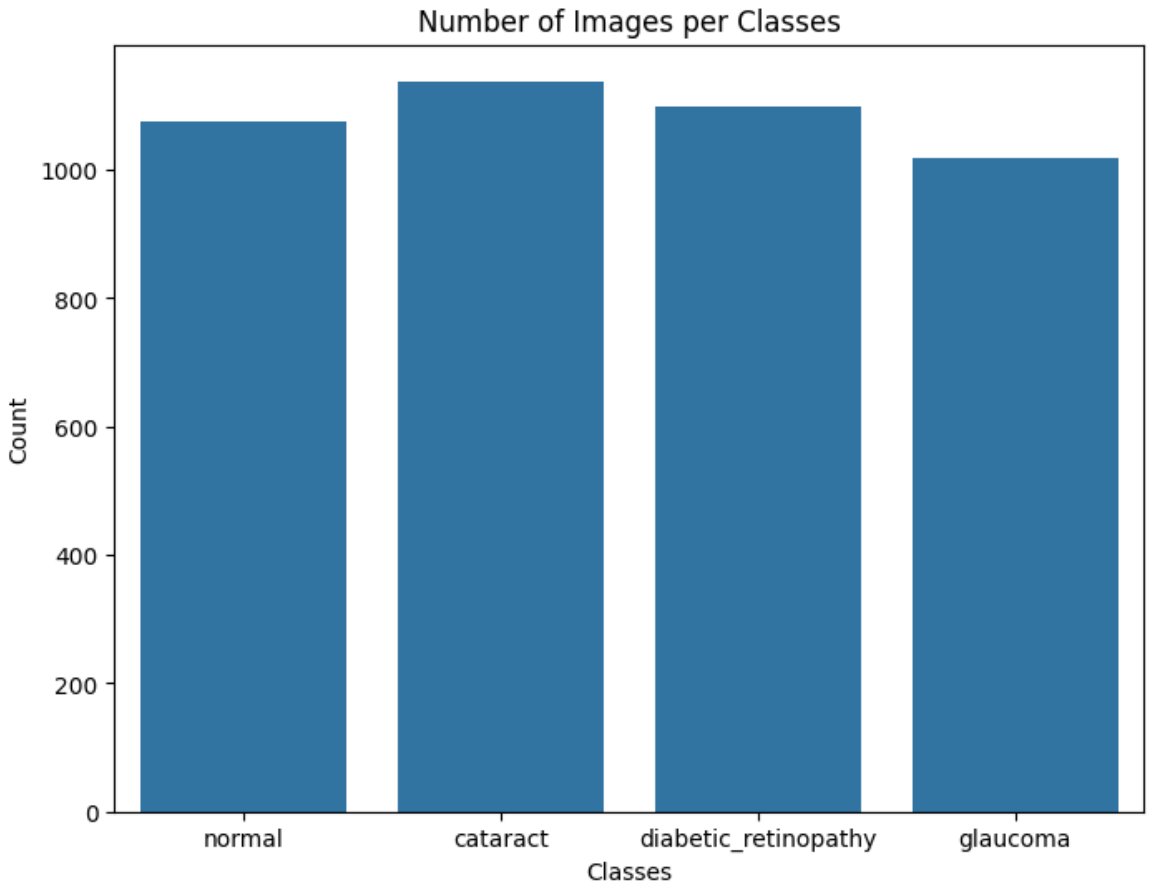


Figure 3.3 Shows the number of fundus images in each class

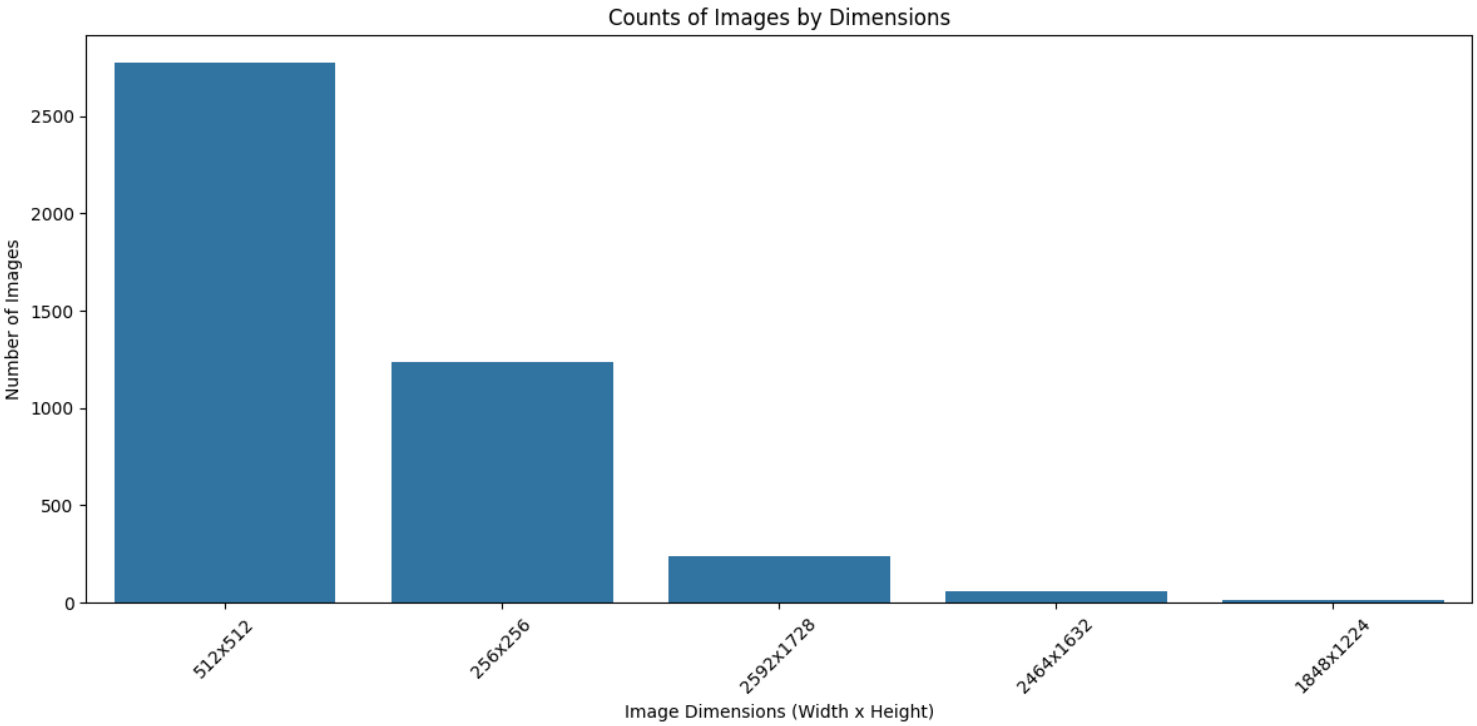


Figure 3.4 Shows the different image sizes in the dataset and number of images with each size dimension

3.4 Data Pre-processing

The repository containing the dataset was stored in a drive file which was then mounted to access the dataset from the dataset folders "dataset4cat" which contains images of four classes: Normal, Diabetic Retinopathy, Cataract, and Glaucoma. The dataset directory is defined as a starting point for the loading and following processing steps. The code iterates through each class folder reading the images. Images in this work are resized to 224×224 pixels to ensure that all images conform to the expected input size for the deep learning models intended to use in this work like the ResNet, VGGNet, and DenseNet models that require size inputs of this dimension. The next step involves the conversion of images into a NumPy arrays where each resized image is appended to an array 'x', while its corresponding class label is added to an array 'y'. This step ensures that images are in a uniform format suitable for the models' input layers.

The preprocessing procedure are listed and explained in the following steps:

1. Loading images; the code iterates through each class folder and reads images with OpenCV.
2. Reading files from the directory
3. Resizing images into (224×224) pixel size using OpenCV to adjust their size to the required input size for the models.
4. Conversion of images into Arrays for use in the models

3.5 Data Splitting and Label Encoding

The data was initially divided into a conventional splitting of 80% images for training, and 10% for testing and validation each, however, the results we got from this ratio were unsatisfactory, so a splitting of 67% for training and 33% for testing was opting to try a different approach. A 10% split was then taken from the training data for validation. This resulted in splitting the whole data into 60% for training, 7% for validation, and 33% for testing. To train the models, the images from the training data were used, the validation set was used to evaluate the model during training, while the test set is reserved for the final evaluation after training. Label encoding or binarization is then performed to prepare for the classification where labels for each image representing a disease class are encoded into one-hot format using Label Encoder. The class labels in 'y' are encoded as integers (from 0 to 3, one for each class) and are then one-hotcoded.

3.6 Proposed CNN Models using Transfer Learning

The idea of CNNs which was initially introduced by Lecun et al., uses convolution operations rather than simple matrix multiplications. CNNs have become essential for achieving high performance in medical image classification, and other procedures like enhancement and segmentation. CNN architectures involve the usage of convolutional layers, batch normalization (BN), activation functions such as ReLU, and other convolutional operations such as pooling operations for deep feature extraction and complex data analysis. To achieve high and better performance, CNNs often require large datasets like ImageNet for pre-training especially when dealing with complex tasks like medical image classification (Gour & Khanna, 2021). Transfer learning enables the collection and usage the acquired knowledge and of learning om these previously trained models regardless of the area or field they were trained on intially (Gour & Khanna, 2021).

In this study, the pre-trained CNN architectures that we aim to implement are the ResNet models (50 and 101 layers) (K. He et al., 2016), VGGNet models (16 and 19 layers) (Simonyan & Zisserman, 2015), and the DenseNet (121 layers) (Huang et al., 2017) in order to classify coloured images of the retina into four classes: normal, cataract, diabetic retinopathy, and glaucoma. These CNNs are usually set to train using big databases such as ImageNet in order to achieve a high classification performance (Gour & Khanna, 2021). The proposed CNN architectures are briefly explained in this section.

3.6.1 ResNet-50 and ResNet-101

ResNets, also known as Residual Networks are a modality of convolutional neural network frameworks consisting of multiple layers each containing a residual unit. This introduced the concept of residual connections that allow the training of complex multi-layered networks resulting in better learning and improving the quality of image recognition (H. He et al., 2016).

ResNet-50 is another layered version of the ResNet architecture with 50 convolutional layers including shortcut connections and batch normalization that enables more effective training and a higher performance on recognition tasks. On the other hand, ResNet-101 extends ResNet-50 with 101 convolutional layers in total, which allows for the extraction of more complex patterns and features, and therefore resulting in a better performance with more difficult tasks (H. He et al., 2016).

3.6.2 VGG-16 and VGG-19

This is another CNN algorithm, also known as Visual Geometry Group (VGG) architectures that are mostly renowned for their simple and uniform framework, which consists of several convolutional and pooling layers (Simonyan & Zisserman, 2015).

VGG-16 is also a variant of VGGNet consisting of 16 convolutional layers as well as some max-pooling layers. The convolutional units of VGG-16 are able to reduce image dimensions due to the fact that they contain several layers with small filter sizes of 3 versus 3, some ReLU functions, and max-pooling layers. VGG-19 on the other hand contains three more layers than the VGG-16 with 19 convolutional layers in total. This extension results in a deeper architecture which facilitate the extraction of more complex patterns and a better performance of challenging tasks (Simonyan & Zisserman, 2015).

3.6.3 DensNet121

DenseNet is also a CNN architecture that is mainly characterized by its dense connections between its layers, where each layer gets feature maps from the preceding layers, and passes down its own generated maps to all the following layers within a dense unit. This structure facilitates reusing features and strengthens their propagation and flow throughout the system resulting in high level performance with fewer parameters (Huang et al., 2017). In this work, the 121-layer version of DenseNet is used.

3.7 Models Implementation and Training

In this work Tesla T4 GPU was used, and models were implemented and trained using TensorFlow and Keras with data processing steps including resizing, array conversion, and test-train-validation splitting at specified ratios. The training employed the Adam optimizer for its efficient convergence properties. The metrics that the models were analysed and evaluated based on are accuracy, precision, recall, F1 score, and AUC in order to provide a comprehensively evaluation of the classification performance across different eye disease categories. The ResNet50 model used in this work has 50 layers with 3.67M trainable parameters, the ResNet101 model with 101 layers and 3.67M trainable parameters, respectively, VGG16 (16 layers and 2.1M trainable parameters), VGG19 (with 19 layers, and 2.1M trainable parameters), and DenseNet121 (121 densely connected layers and 2.6M trainable parameters). These architectures provide a range of parameters and network depths and enables the assessment of model's effectiveness according to its layer complexity.

The pre-trained CNNs were loaded with pre-trained weights to identify and extract key patterns and recognize characterizing features in the images. The initial pre-trained layers of the CNN models are frozen in order to prevent their weight from being updated during training. The model is the customized by adding several layers on top of the models' base. The input shape (224, 224, 3) is also specified to ensure compatibility with the images being used after resizing them and including the 3 colour channels (RBG). GAP layer is added to decrease the dimensions of the feature map and condensing it into a single vector by averaging the feature values across the spatial dimensions. Three fully connected dense layers are also added with 1024, 1024, and 512 neurons for each layer accordingly, and each using a ReLu function to allow further learning on more complex patterns. A fully connected (dense) layer with 4 neurons is included along with the Softmax activation function in the output layer

to represent the four possible predictions and present the probability distribution across all four classes. A learning rate of 0.0001 was set to compile the models using the Adam optimizer along with loss function, making it suitable for tuning (Meng et al., 2018). Training accuracy and loss are mainly tracked during the training process through every epoch.

Verbose output was included to show training progress for each epoch, and validation accuracy and validation loss were also monitored through the 30 epochs that the training process went through.

The models' implementation and training process are listed and explained in the following steps:

1. Loading the pre-trained CNNs
2. Freezing the top layers of the models
3. Adding Custom Layers: GAP layers, fully connected layers, Softmax Activation layers.
4. Model compilation: Adam optimizer/ Cross-Entropy Loss
5. Model training

3.8 Performance Metrics

The evaluation of the proposed models was conducted using some performance evaluation parameters and metrics. These include accuracy, F1 score, precision, and recall. Plots of the confusion matrices and the Receiver Operating Characteristic (ROC) curves were also generated for each model to further evaluate their performances across all classes. The evaluation metrics are briefly explained and discussed in this section.

Accuracy: An evaluation metric that assesses the rate of correctly classified images among all the classifications in the datasets (Sokolova & Lapalme, 2009).

Area Under the Curve (AUC): Determines to which extent models are able to set apart and distinguish between various classes. The higher the value, the better the assessment ability (Fawcett, 2006)

F1-Score: Offers a comprehensive evaluation of false positives and negatives by combining both precision and recall into one single value. It basically determines how well the model is doing in identifying positives and at the same time avoiding false positives (Powers, 2020).

Precision: Gives an evaluation of the correct positive predictions among all the classes predicted as positive (Sokolova & Lapalme, 2009).

Recall: Also known as sensitivity, it's used for calculating the rate of how many actual positive predictions the model succeeded in identifying among all the true positives. (Sokolova & Lapalme, 2009).

As a summary of the methodology process, a flow chart diagram is presented in Figure3.5 to describe the steps conducted from step one of inputting data to the step of evaluating the results.

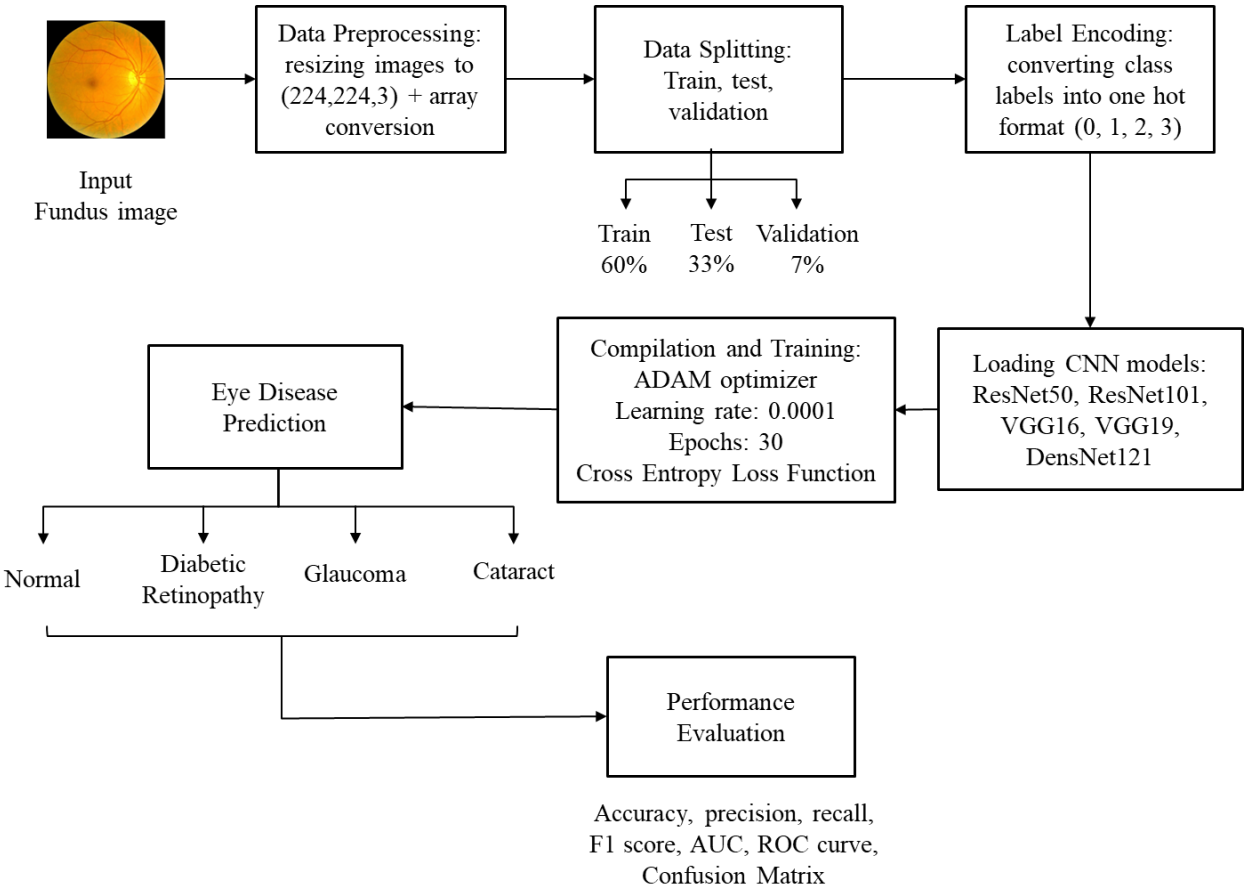


Figure 3.5 Flow Chart Diagram summarizing the Methodology Process

CHAPTER IV
RESULTS AND DISCUSSION

4.1 Evaluation Metrics and ROC Curve

Five CNN models were used in this study to classify coloured fundoscopy images into four categories: normal, diabetic retinopathy, cataract, and glaucoma. The results shown in this section are of the performance metrics including the accuracies, F1 score, precision, recall, AUC, ROC curves and confusion matrices plots of each model. The results are compared and discussed in this chapter.

The evaluation metrics and performances of each model (ResNet50, ResNet101, VGG16, VGG19, and DenseNet121) are compared in Tables 4.1, 4.2, and 4.3; where Table 4.1 displays the evaluation metric values of each CNN architecture. The performance of each CNN model and their ability to classify fundus images into four categories: normal, cataracts, diabetic retinopathy, and glaucoma has been analysed through the above mentioned evaluation metrics as shown in Table 4.1.

Overall, the ResNet50 and ResNet101 models delivered the top results in accuracy, precision, recall, and F1 score with ResNet50 slightly outperforming ResNet101. ResNet50 achieved a test accuracy of 91.25%, while ResNet101 reached a similar percentage of 91.04%. Both models showed very high AUC scores (ResNet50 at 0.9858 and ResNet101 at 0.9863), which indicates strong classification ability and model robustness. VGG16 also performed well and achieved a test accuracy of 90.34% and an AUC value of 0.9825, demonstrating a competitive performance despite having fewer layers than the ResNet models. However, VGG19 which is an extended version of VGG16 with more layers, achieved slightly lower accuracy and AUC values, possibly due to overfitting with additional layers that did not improve the model's generalization ability. This demonstrates that layer depth doesn't necessarily guaranty a big gap in terms of good performance. As for DenseNet121, while achieving good accuracy for certain classes, it demonstrated the lowest overall accuracy and AUC values, and this highlights potential limitations when using this model architecture for the provided dataset.

Table 4.1 Evaluation metrics

Model	Accuracy	Precision	Recall	F1 Score	AUC
ResNet50	**0.9125**	**0.9133**	**0.9125**	**0.9126**	0.9858
ResNet101	0.9104	0.9114	0.9104	0.9105	**0.9863**
VGG16	0.9034	0.9057	0.9034	0.9022	0.9825
VGG19	0.8880	0.8890	0.8880	0.8884	0.9820
DensNet121	0.8684	0.8709	0.8684	0.8686	0.9678

Table 4.2 highlights the training, validation, and testing accuracies for each CNN architecture. It can be observed that in terms of training accuracy, all models achieved very high accuracies (approximately around 99%), which demonstrates that they learned the training data effectively. However, there is a noticeable gap between the training and validation/test accuracies, particularly with VGG19 and DenseNet121, which suggests that these models may likely suffer from slight overfitting. For example, VGG19 have shown a training accuracy of 99.96% but a test accuracy of only 88.80%, while DenseNet121 had the lowest test accuracy of 86.84%, reflecting a significant difference between training and validation/test accuracy. On the other hand, the ResNet models demonstrated more stable performance across training, validation, and test sets with minimal gaps, which demonstrates that these models generalize better on unseen data compared to the rest of the models.

Table 4.2 Training, validation, and test accuracies

Model	Training Accuracy	Validation Accuracy	Test Accuracy
ResNet50	**0.9996**	0.9172	**0.9125**
ResNet101	**0.9996**	**0.9207**	0.9104
VGG16	0.9992	0.8897	0.9034
VGG19	**0.9996**	0.8793	0.8880
DensNet121	0.9854	0.8483	0.8684

Table 4.3 provides the accuracy of each model across the four specific disease categories. For the normal class, VGG16 has clearly achieved the highest accuracy in the normal category, reaching a value of 95.65%, and indicating a high ability to accurately classify healthy fundus images. ResNet101 followed closely with an accuracy of 94.63%, and

ResNet50 also performed well at 93.61%. DenseNet121 had performed the least in this category with an accuracy of 85.93%. For cataract, all models performed remarkably well in detecting cataracts with DenseNet121 achieving the highest accuracy of 99.17%. ResNet50 and VGG16 also showed high performance and achieved high accuracies of 98.90% and 98.34%, respectively. This high performance suggests that cataracts is easier to detect and distinguish in fundus images considering that all models, regardless of their architecture's depth, showed high accuracy in this category.

Table 4.3 Accuracy values of each class across all five models

Model	Normal accuracy	Cataracts accuracy	Diabetic retinopathy accuracy	Glaucoma accuracy
ResNet50	0.9361	0.9890	0.8889	**0.8263**
ResNet101	0.9463	0.9779	0.8860	0.8204
VGG16	**0.9565**	0.9834	**0.9123**	0.7455
VGG19	0.9207	0.9669	0.8450	0.8084
DensNet121	0.8593	**0.9917**	0.8596	0.7545

In diabetic retinopathy the models had an overall moderate performance, with VGG16 achieving the highest accuracy of 91.23%. ResNet50 and ResNet101 performed similarly at 88.89% and 88.60%, respectively, and DenseNet121 had the lowest performance at an accuracy of 85.96%. These results indicate that the models faced challenges in accurately identifying diabetic retinopathy. As for classifying glaucoma, it has proved to be challenging for all models considering the lower accuracy values compared to the other categories. ResNet50 and ResNet101 achieved accuracies of 82.63% and 82.04%, respectively, showing relatively better performance in this category. VGG16 achieved an accuracy of 74.55%, as DenseNet121 struggled with an accuracy of only 75.45%.

This performance difference could be due to the discretion and variability of glaucoma features in fundus images, and this may require more specific features for a more accurate classification. To better visualize the model's ability to tell apart different classes and distinguish between them, a graphical representation of the diagnostic ability is illustrated in Figure4.1 which shows the ROC curve plots of each CNN architecture. The ROC curve of the

normal class is represented in yellow, the curve for diabetic retinopathy is represented in green, the cataract curve is portrayed in red, and the glaucoma curve is portrayed in blue.

4.2 Confusion Matrix

As demonstrated in Figure 4.2, the confusion matrices for each of the five CNNs used in this study (ResNet50, ResNet101, DenseNet121, VGG16, and VGG19) indicate their ability to classify the four eye disease categories: Normal, Cataract, Diabetic Retinopathy, and Glaucoma. ResNet50 demonstrates excellent performance with the highest accuracy in classifying the "Normal" class with 366 true positives, and the "Cataract" class with 356 true positives. However, there are still some misclassifications observed in “Glaucoma” with 13 false negatives as well as "Diabetic Retinopathy" with 6 false negatives. Nonetheless, these errors remain minimal compared to the overall classification performance.

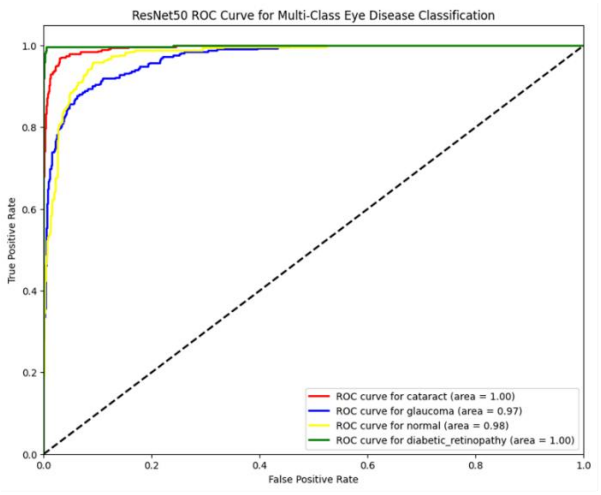
ResNet101 shows a slightly higher performance than the ResNet50 model in terms of classifying "Normal" with 373 true positives and "Cataract" with 353 true positives. It also performs similarly to the ResNet50 model in the other classes with a minimal number of misclassification where it has a small number of false positives for "Glaucoma" and "Diabetic Retinopathy." DensNet121 on the other hand, has a somewhat lower performance than both ResNet models. While it has a relatively high number of true positives for "Normal" with 370 and for "Cataract" with 354 classifications, it still struggles more with the "Glaucoma" and "Diabetic Retinopathy" classes, showing higher incorrect positive and incorrect negative rates. This suggests that DensNet121 has some trouble in distinguishing between images belonging to these specific diseases.

VGG16 has also demonstrated competitive results notably with the "Normal" and "Cataract" classes being well-classified, with 362 and 356 true positives respectively. However, it has shown a confusion between the "Glaucoma" and "Diabetic Retinopathy" classes with some noticeable misclassifications, especially in "Glaucoma" where 19 images were misclassified as "Normal." VGG19 provides the lowest performance among the models, notably in the "Glaucoma" class, where there are significant misclassifications of 33 false positives and 13 false negatives. It has also shown some struggles with "Diabetic Retinopathy" which indicates that it might have a higher generalization error for the dataset.

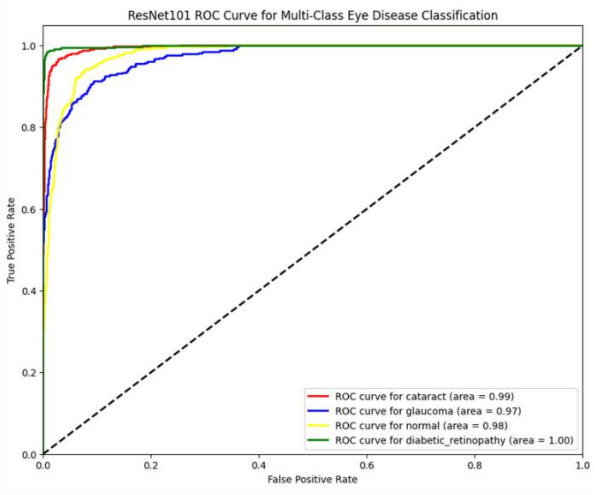
4.3 General Discussion of Results

The overall performance of the models generally indicates that deeper, more complex architectures like ResNet50 and ResNet101 provide the best classification accuracy, precision, and recall. Both models have shown high accuracy in distinguishing between the four classes, especially in the Normal and Cataract classes since it demonstrated a relatively low misclassification rate. This is also confirmed from the confusion matrices that demonstrated that these models maintain fewer false positives and false negatives. DensNet121 and VGG16 also have shown a solid performance, however they exhibit more challenges particularly with distinguishing between Glaucoma and Diabetic Retinopathy. DensNet121's slightly lower performance can be because of its fewer training parameters and possible higher generalization error, considering that it performs worse in certain categories, especially with Glaucoma. VGG19 has shown a lower overall accuracy despite having a deeply layered structure and appears less effective for the classification task at hand. This is further demonstrated by its relatively high misclassification rates in Glaucoma, having the highest rate of false positive and false negative predictions. This may generally indicate that VGG19 faces issues with the fine-grained classification of certain classes despite its deep architecture.

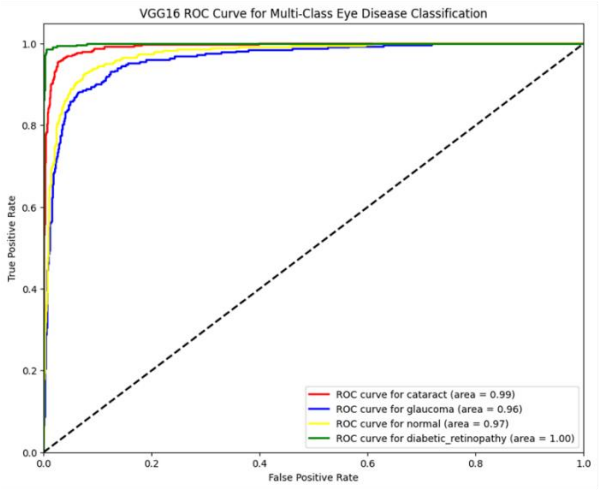
From the results obtained by the performance metrics, it can be deduced that ResNet50 and ResNet101 stand out with the highest performance on test and validation datasets, which makes them the most reliable models for this four-class eye disease classification task. The AUC values for these models further demonstrate their strong performance, by highlighting their ability to correctly classify input images in terms to their most probable true class.



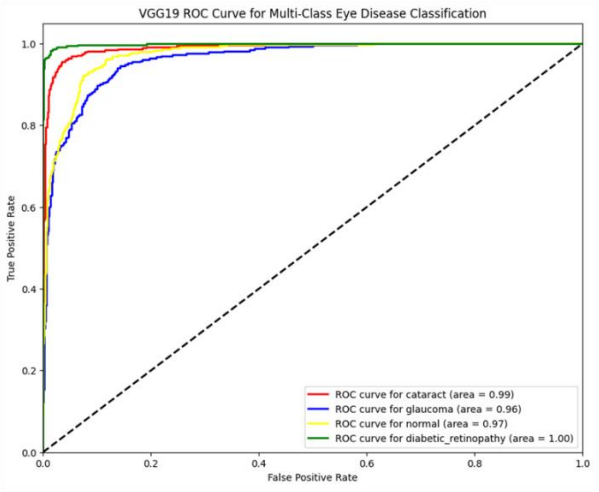
(a)



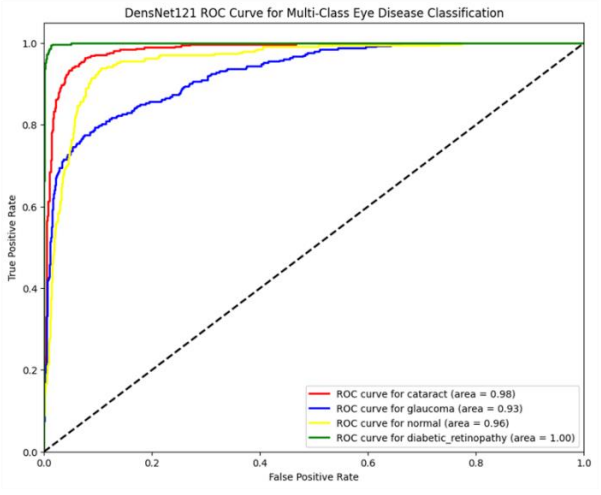
(b)



(c)

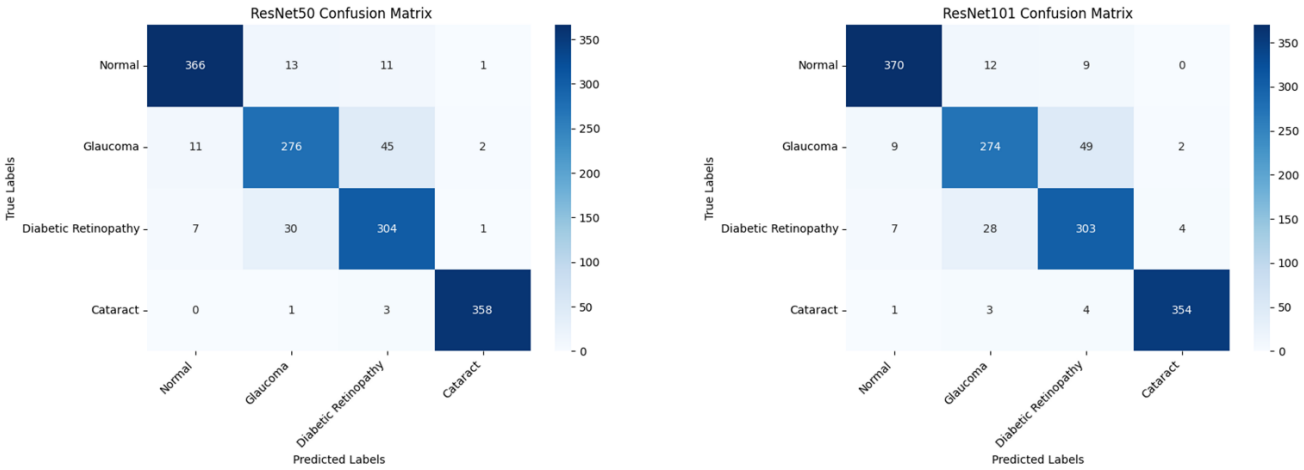


(d)



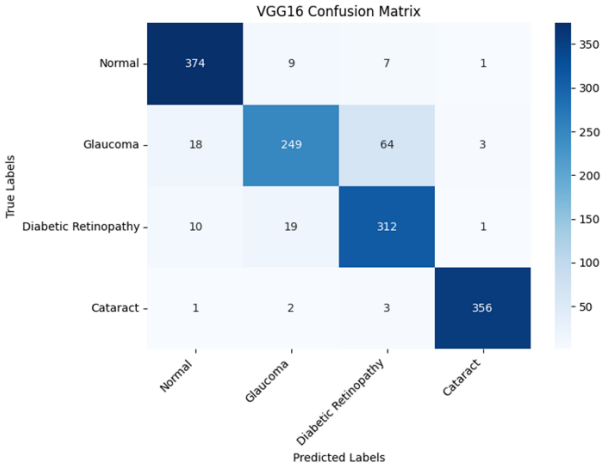
(e)

Figure 4.1 ROC curves for each CNN model; where (a) represents ROC curve of ResNet50, (b) ResNet101, (c) VGG16, (d) VGG19, and (e) DensNet121.

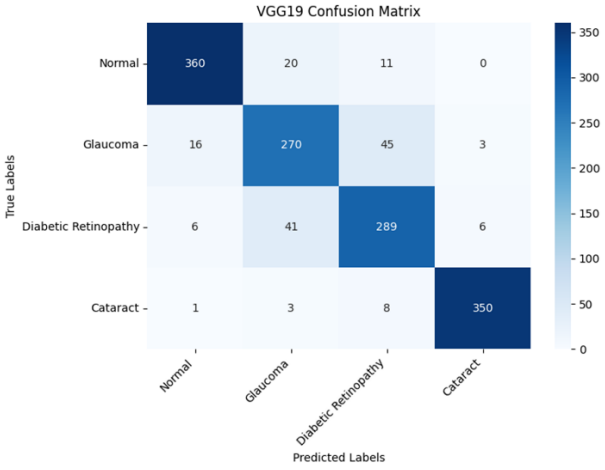


(a)

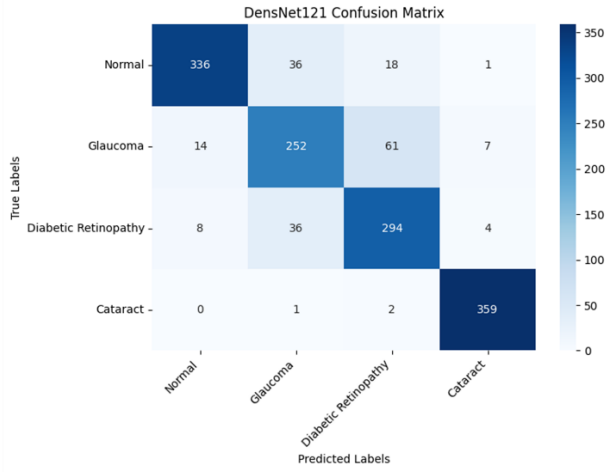
(b)



(c)



(d)



(e)

Figure 4.2 Confusion matrix for each CNN model; where (a) represents the confusion matrix of ResNet50, (b) ResNet101, (c) VGG16, (d) VGG19, and (e) DensNet121.

4.4 Implementation of Results into Web App

4.4.1 Introduction to WebApp

In this work, a web application ‘I-SCANNER’ was developed to facilitate the usage and complement the deep learning model of this study for the eye disease classification task. This web app is developed to assist healthcare professionals with the detection and diagnosis of ocular pathologies. The utilisation of a command-line interface (CLI) to predict eye diseases can be a tedious process considering that it requires typing commands for each prediction. This web-based application simplifies the whole process by providing an intuitive and user-friendly dashboard where images can be uploaded by a user and interact with the model in a seamless way to promote an easy interpretation. In addition to allowing users to get a prediction of eye disease that may be represented in an uploaded retinal image, I-SCANNER also allow users to securely register or login to access their personalized account as displayed in figures 4.3 and 4.4. This framework also saves all the predicted results of each patient automatically for future analysis and reference, making it much more manageable, accessible, and user friendly.

This web application combines the usage of different technologies including Flask, TensorFlow/Keras, JavaScript and others to create a smooth experience to predict eye diseases. It uses an application programming interface (API) to connect the backend models with the frontend interface, which allows users to see and work with the predictions visually.

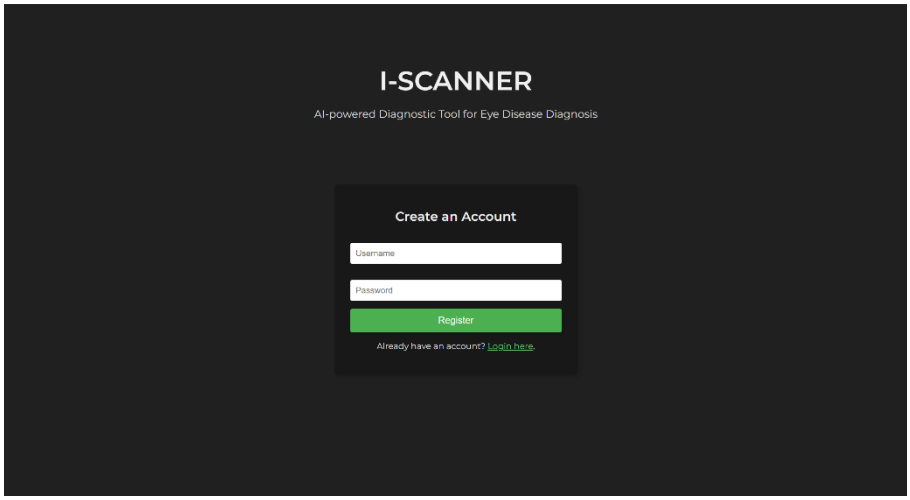


Figure 4.3: Figure showing the registration feature of I-SCANNER

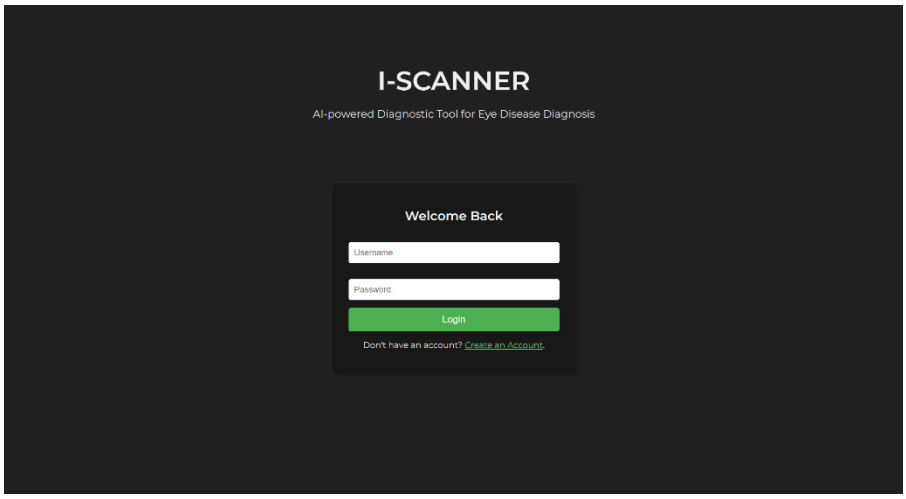


Figure 4.4: Figure showing the login feature of I-SCANNER

4.4.2 The API

The model for eye disease classification was built using Python along with a popular framework used around machine learning, called TensorFlow. To keep things simple and compatible, the models were saved as Keras .h5 files without compression or format changes ensuring the API can access them easily without running through compatibility issues.

The API was developed using Flask, a widely used Python platform for building web applications and APIs, TensorFlow/Keras which allowed the loading of the trained CNN model, and Werkzeug Security to provide a secure password and authentication for users. The frontend sends a request carrying image data to the backend server that processes the image and generate the predictions. These results are stored in a database along with relevant metadata in a database for future refence. To optimize efficiency, the input image size is adjusted to 224×224 pixels in order to reduce processing time and ensure a quick generation of response.

4.4.3 The Frontend

The frontend operates as the interface for users navigating the web application and displays elements such as buttons, text, and images. It dynamically updates the content using JavaScript depending on the predicted results that were received from the API. HTML The application begins with a login page where users can authenticate themselves to gain access. This ensures that data is secured and prevents any unauthorised access that can compromise

the system or disrupt API calls. The frontend technologies that were used include HyperText Markup Language (HTML) and Jinja2 to create a dynamic template, Cascading Style Sheets (CSS) which was used to ensure a clean and appealing stylistic design to the web application, and JavaScript was used to ensure a good management and handling of file uploads and updating results.

After registering and the logging in, users are then directed to a personalized dashboard where they can upload medical images that are the sent to the server where the AI model analyses them, figures 4.5 and 4.6 display the dashboards where users can upload and analyse retinal images. Predictions and the confidence scores (probabilities) of each class are then displayed, and the class with the highest confidence is highlighted as the primary prediction that is most likely represented in the uploaded image as shown in figure 4.7. This design enables users to easily navigate the application in an efficient way. These details and results are then saved to the database for future reference.

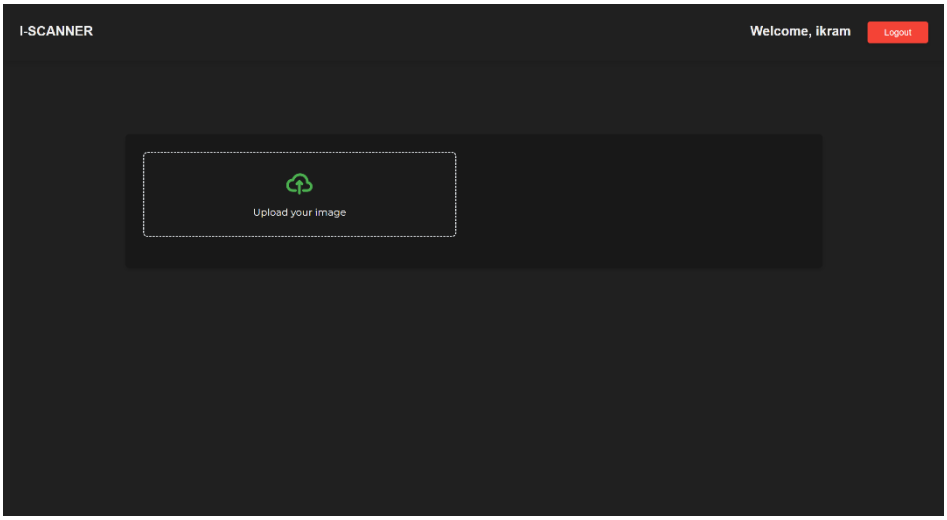


Figure 4.5: Figure displaying the dashboard for image uploading

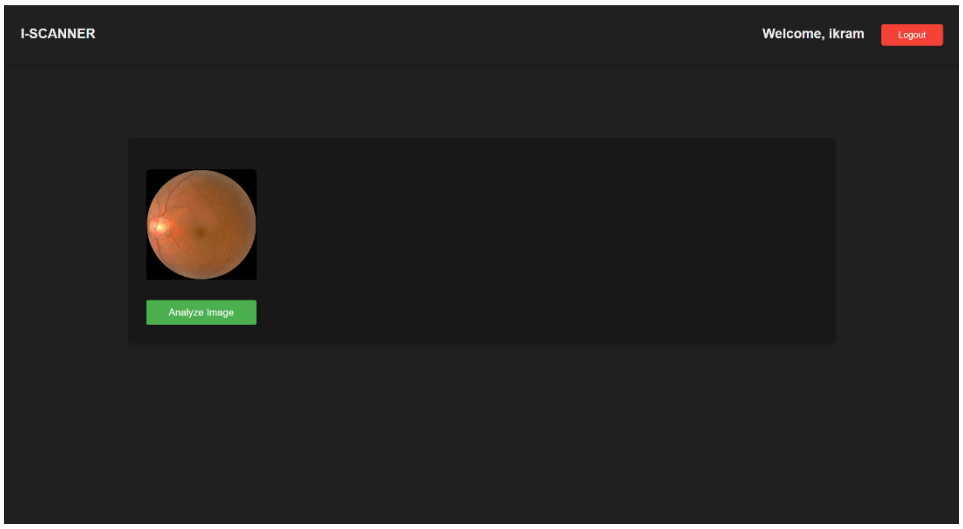


Figure 4.6: Figure displaying the dashboard for image analysis

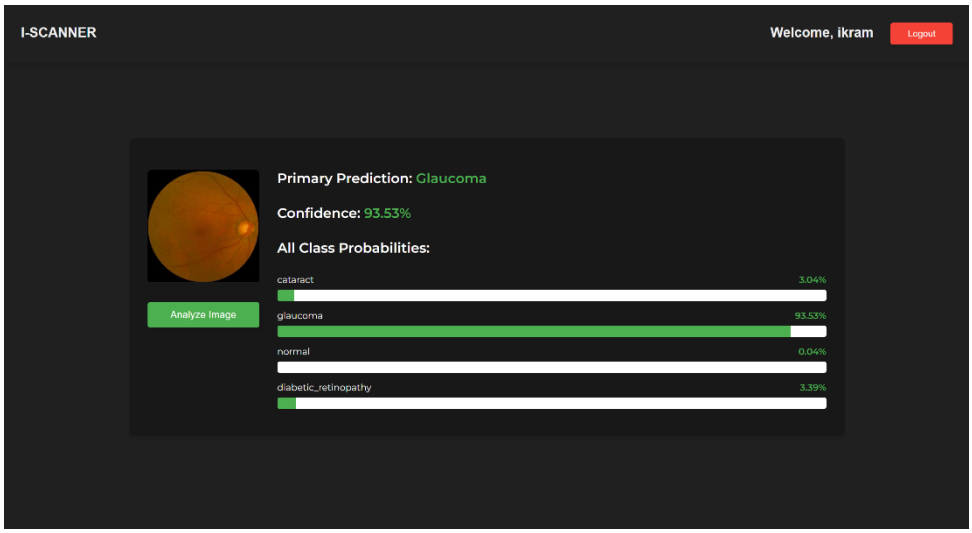


Figure 4.7: Shows the primary disease prediction and its confidence along with all class probabilities

4.4.4 The Database

The I-SCANNER web app uses SQLite, which is a reliable SQL-based database to manage and store data efficiently. This database allows for quick searches and effectively handles powerful join calls through tables enduring the application runs smoothly.

Each prediction is stored with a unique identifier for each patient. This setup helps users track their previous predictions. Each user is made sure to have their own account to ensure that all predictions and settings are personalized. This guarantees privacy and creates a better overall user experience making this WebApp practical and powerful as a tool to assist medical professionals for automated eye disease classification and analysis.

CHAPTER V
LIMITATIONS AND FUTURE WORK

5.1 Limitations and Future Work

A major limitation that was observed in the proposed modality is the imbalanced performance of the classification models across different disease classes. While most models achieved high accuracies in distinguishing normal fundus images and cataract cases, they still found some difficulties to distinguish between the glaucoma and diabetic retinopathy categories. This confusion is mainly due to the complexity and similarity of the visual features in these two classes. Glaucoma and diabetic retinopathy may also present overlapping characteristics with other conditions, such as subtle vessel abnormalities or optic nerve changes, and this can reveal to be harder for the model to distinguish. This emphasises the need for a more refined approach like using class-specific data augmentation or more advanced feature extraction techniques to enhance the classification accuracy for these challenging categories.

Another notable issue was overfitting, especially in VGG19 and DenseNet121. These models achieved high training accuracies exceeding 90%, however their performance on the testing sets was lower in comparison, and this also indicates that the models memorized training data patterns rather than generalizing to unseen data. The overfitting can also be due to the high complexity of these architectures considering that they include a significant number of parameters. To solve this problem, further optimization such as simplifying the models, using less layers, and more fine-tuning may be required to achieve better generalization.

There is also the issue of limited dataset size and variety; even though the dataset that was used had over 4,217 fundus images, the size may still be considered limited especially when training deep learning models that require vast amounts of data to operate efficiently. Additionally, the images may not represent all the variations seen in real-world clinical cases, such as differences in imaging devices, lighting conditions, or patients' demographics (e.g., age, ethnicity, clinical history). This limitation could affect the model's ability to generalize to actual clinical environments. Therefore, a bigger and more diverse dataset would help with improving the robustness and applicability of the proposed models.

Another disadvantage is that this work focused solely on coloured fundoscopy images, which are commonly used and moderately accessible as a diagnosis modality in eye care. However, they don't always provide a complete picture of the eye condition and capture all the relevant pathological features needed to achieve an accurate diagnosis. Other imaging procedures, such as OCT or fluorescein angiography may show additional and more explicit details of the retinal structure that could improve diagnosis accuracy. Including these imaging techniques and using a multimodal dataset in future works, could make the models more versatile and reliable in clinical practice.

Despite the promising overall performance, the models have shown a non-negligible rate of misclassifications, especially in the glaucoma and diabetic retinopathy classes. These misclassifications may be due to multiple factors including insufficient feature representation, noise in the dataset, or overlap in features between classes. These missteps indicate the need for further refinement of the models. Improvements such as focusing on critical regions and areas of the images, or using more advanced deep learning pre-trained architectures, or including some expert-labelled regions to better guide the learning process of the models.

By addressing these limitations and proposing potential solutions, future studies could significantly improve the models' effectiveness, robustness, and clinical applicability, which may greatly contribute to more accurate and reliable automated eye disease diagnosis.

CHAPTER VI
CONCLUSION

6.1 Conclusion

In summary, the ResNet50 and ResNet101 models emerge as the top-performing models, which shows and reinforces their effectiveness in classifying medical images of ocular diseases with high accuracy and low misclassification rates. On the other hand, DenseNet121, VGG16, and VGG19 have shown some weaknesses regarding the classification of the more challenging classes like Glaucoma and Diabetic Retinopathy. These results may suggest that deeper architectures with more layers, like the ResNet models, provide better domain adaptation and generalization for medical image classification tasks. Despite the limitations of the other models, all the architectures show overall promising results in terms of evaluation metrics, highlighting their potential for utilization in automated diagnostic systems in healthcare. The best performing model was integrated into a web application 'I-SCANNER' that was specifically developed and designed to assist ophthalmologists in the detection and classification of ocular pathologies into normal, cataract, glaucoma, and diabetic retinopathy. The I-SCANNER platform aims to facilitate the accessibility to diagnostic tools when it comes to eye diseases by contributing to the early detection and therefore, early treatment of these burdensome conditions that can greatly affect life quality.

REFERENCES

Abbas, Q., Ibrahim, M. E., & Jaffar, M. A. (2018). Video scene analysis: An overview and challenges on deep learning algorithms. *Multimedia Tools and Applications*, *77*(16), 20415–20453.

Abràmoff, M. D., Garvin, M. K., & Sonka, M. (2010). Retinal Imaging and Image Analysis. *IEEE Reviews in Biomedical Engineering*, *3*, 169–208. IEEE Reviews in Biomedical Engineering. https://doi.org/10.1109/RBME.2010.2084567

Ahmad, A., Mansoor, A. B., Mumtaz, R., Khan, M., & Mirza, S. (2014). *Image processing and classification in diabetic retinopathy: A review*. 1–6.

Alam, K. N., & Khan, M. M. (2021). CNN Based COVID-19 Prediction from Chest X-ray Images. *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 0486–0492. https://doi.org/10.1109/UEMCON53757.2021.9666508

Alsaih, K., Lemaitre, G., Rastgoo, M., Massich, J., Sidibé, D., & Meriaudeau, F. (2017). Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *BioMedical Engineering OnLine*, *16*(1), 68. https://doi.org/10.1186/s12938-017-0352-9

Amoaku, W. M., Ghanchi, F., Bailey, C., Banerjee, S., Banerjee, S., Downey, L., Gale, R., Hamilton, R., Khunti, K., & Posner, E. (2020). Diabetic retinopathy and diabetic macular oedema pathways and management: UK Consensus Working Group. *Eye*, *34*(Suppl 1), 1–51.

Borrás, T. (2012). Advances in Glaucoma Treatment and Management: Gene Therapy. *Investigative Ophthalmology & Visual Science*, *53*(5), 2506–2510. https://doi.org/10.1167/iovs.12-9483o

Budai, A., Bock, R., Maier, A., Hornegger, J., & Michelson, G. (2013). Robust vessel segmentation in fundus images. *International Journal of Biomedical Imaging*, *2013*, 154860. https://doi.org/10.1155/2013/154860

Burlina, P. M., Joshi, N., Pekala, M., Pacheco, K. D., Freund, D. E., & Bressler, N. M. (2017). Automated Grading of Age-Related Macular Degeneration From Color Fundus Images Using Deep Convolutional Neural Networks. *JAMA Ophthalmology*, *135*(11), 1170. https://doi.org/10.1001/jamaophthalmol.2017.3782

Chakrabarti, A., Singh, S., & R, K. (2000). Phacoemulsification in eyes with white cataract1. *Journal of Cataract & Refractive Surgery*, *26*(7), 1041–1047. https://doi.org/10.1016/S0886-3350(00)00525-3

Cheng, Y., Ma, M., Li, X., & Zhou, Y. (2021). Multi-label classification of fundus images based on graph convolutional network. *BMC Medical Informatics and Decision Making*, *21*(2), 82. https://doi.org/10.1186/s12911-021-01424-x

Chuck, R. S., Dunn, S. P., Flaxel, C. J., Gedde, S. J., Mah, F. S., Miller, K. M., Wallace, D. K., & Musch, D. C. (2021). Comprehensive Adult Medical Eye Evaluation Preferred Practice Pattern®. *Ophthalmology*, *128*(1), P1–P29. https://doi.org/10.1016/j.ophtha.2020.10.024

Cohen, L. P., & Pasquale, L. R. (2014). Clinical Characteristics and Current Treatment of Glaucoma. *Cold Spring Harbor Perspectives in Medicine*, *4*(6), a017236. https://doi.org/10.1101/cshperspect.a017236

Diaz-Pinto, A., Morales, S., Naranjo, V., Köhler, T., Mossi, J. M., & Navea, A. (2019). CNNs for automatic glaucoma assessment using fundus images: An extensive validation. *BioMedical Engineering OnLine*, *18*(1), 29. https://doi.org/10.1186/s12938-019-0649-y

Dipu, N. M., Shohan, S. A., & Salam, K. M. A. (2021). Ocular Disease Detection Using Advanced Neural Network Based Classification Algorithms. *Asian Journal For Convergence In Technology (AJCT) ISSN -2350-1146*, *7*(2), Article 2. https://doi.org/10.33130/AJCT.2021v07i02.019

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fung, T. H., Patel, B., Wilmot, E. G., & Amoaku, W. M. (2022). Diabetic retinopathy for the non-ophthalmologist. *Clinical Medicine*, *22*(2), 112–116. https://doi.org/10.7861/clinmed.2021-0792

Galveia, J. N., Travassos, A., Quadros, F. A., & da Silva Cruz, L. A. (2018). Computer aided diagnosis in ophthalmology: Deep learning applications. *Classification in BioApps: Automation of Decision Making*, 263–293.

Gour, N., & Khanna, P. (2021). Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network. *Biomedical Signal Processing and Control*, *66*, 102329. https://doi.org/10.1016/j.bspc.2020.102329

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (n.d.). *Deep learning for visual understanding: A review*.

Gus, P. I., Kwitko, I., Roehe, D., & Kwitko, S. (2000). Potential acuity meter accuracy in cataract patients. *Journal of Cataract & Refractive Surgery*, *26*(8), 1238. https://doi.org/10.1016/S0886-3350(00)00409-0

H. S, R. (2017). *Prevalence and Risk Factors in Primary Open Angle Glaucoma of Patients Attending Ophthalmology OPD at Kims Hubli* [M.S., Rajiv Gandhi University of Health Sciences (India)]. https://www.proquest.com/docview/2866083516/abstract/E8C99CE536974659PQ/1

He, H., Boyd-Graber, J., Kwok, K., & Hal Daumé, I. I. I. (2016). Opponent Modeling in Deep Reinforcement Learning. *Proceedings of The 33rd International Conference on Machine Learning*, 1804–1813. https://proceedings.mlr.press/v48/he16.html

He, J., Li, C., Ye, J., Qiao, Y., & Gu, L. (2021a). Multi-label ocular disease classification with a dense correlation deep neural network. *Biomedical Signal Processing and Control*, *63*, 102167. https://doi.org/10.1016/j.bspc.2020.102167

He, J., Li, C., Ye, J., Qiao, Y., & Gu, L. (2021b). Self-speculation of clinical features based on knowledge distillation for accurate ocular disease classification. *Biomedical Signal Processing and Control*, *67*, 102491. https://doi.org/10.1016/j.bspc.2021.102491

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). *Densely Connected Convolutional Networks*. 4700–4708. https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html

Islam, M. M., Yang, H.-C., Poly, T. N., Jian, W.-S., & (Jack) Li, Y.-C. (2020). Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis. *Computer Methods and Programs in Biomedicine*, *191*, 105320. https://doi.org/10.1016/j.cmpb.2020.105320

Jayaram, H., Kolko, M., Friedman, D. S., & Gazzard, G. (2023). Glaucoma: Now and beyond. *The Lancet*, *402*(10414), 1788–1801. https://doi.org/10.1016/S0140-6736(23)01289-8

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., … Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, *172*(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Kingma, D. P. (2014). Adam: A method for stochastic optimization. *arXiv Preprint arXiv:1412.6980*.

Lam, D., Rao, S. K., Ratra, V., Liu, Y., Mitchell, P., King, J., Tassignon, M.-J., Jonas, J., Pang, C. P., & Chang, D. F. (2015). Cataract. *Nature Reviews Disease Primers*, *1*(1), 1–15. https://doi.org/10.1038/nrdp.2015.14

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. Proceedings of the IEEE. https://doi.org/10.1109/5.726791

Lee, D. A., & Higginbotham, E. J. (2005). Glaucoma and its treatment: A review. *American Journal of Health-System Pharmacy*, *62*(7), 691–699. https://doi.org/10.1093/ajhp/62.7.691

Li, F., Wang, Y., Xu, T., Dong, L., Yan, L., Jiang, M., Zhang, X., Jiang, H., Wu, Z., & Zou, H. (2022). Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. *Eye*, *36*(7), 1433–1441. https://doi.org/10.1038/s41433-021-01552-8

Li, Z., Keel, S., Liu, C., He, Y., Meng, W., Scheetz, J., Lee, P. Y., Shaw, J., Ting, D., Wong, T. Y., Taylor, H., Chang, R., & He, M. (2018). An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs. *Diabetes Care*, *41*(12), 2509–2516. https://doi.org/10.2337/dc18-0147

Lim, R. (2022). The surgical management of glaucoma: A review. *Clinical & Experimental Ophthalmology*, *50*(2), 213–231. https://doi.org/10.1111/ceo.14028

Madsen-Bouterse, S. A., & Kowluru, R. A. (2008). Oxidative stress and diabetic retinopathy: Pathophysiological mechanisms and treatment perspectives. *Reviews in Endocrine and Metabolic Disorders*, *9*(4), 315–327. https://doi.org/10.1007/s11154-008-9090-4

Maitra, M., & Chatterjee, A. (2006). A Slantlet transform based intelligent system for magnetic resonance brain image classification. *Biomedical Signal Processing and Control*, *1*(4), 299–306. https://doi.org/10.1016/j.bspc.2006.12.001

Meng, X., Xi, X., Yang, L., Zhang, G., Yin, Y., & Chen, X. (2018). Fast and effective optic disk localization based on convolutional neural network. *Neurocomputing*, *312*, 285–295. https://doi.org/10.1016/j.neucom.2018.05.114

ODIR-2019—Grand Challenge. (2020). Grand-Challenge.Org. https://odir2019.grand-challenge.org/

Ouda, O., AbdelMaksoud, E., Abd El-Aziz, A. A., & Elmogy, M. (2022). Multiple Ocular Disease Diagnosis Using Fundus Images Based on Multi-Label Deep Learning Classification. *Electronics*, *11*(13), Article 13. https://doi.org/10.3390/electronics11131966

Patz, A., Fine, S., Finkelstein, D., Prout, T., Aiello, L., Bradley, R., Briones, J. C., Myers, F., Bresnick, G., de Venecia, G., Stevens, T. S., Wallow, I. H. L., Chandra, S. R.,

Norton, E., Blankenship, G., Harris, J., Knobloch, W., Goetz, F., Ramsay, R. C., … Rand, L. (1978). Photocoagulation Treatment of Proliferative Diabetic Retinopathy: The Second Report of Diabetic Retinopathy Study Findings. *Ophthalmology*, *85*(1), 82–106. https://doi.org/10.1016/S0161-6420(78)35693-1

Pektaş, M. (2023). Performance Analysis of Efficient Deep Learning Models for Multi-Label Classification of Fundus Image. *Artificial Intelligence Theory and Applications*, *3*(2), 105–112.

Porta, M., & Bandello, F. (2002). Diabetic retinopathy. *Diabetologia*, *45*(12), 1617–1634. https://doi.org/10.1007/s00125-002-0990-7

Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., & Meriaudeau, F. (2018). *Indian Diabetic Retinopathy Image Dataset (IDRiD)* [Dataset]. IEEE Dataport. https://doi.org/10.21227/H25W98

Powers, D. M. W. (2020). *Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation* (No. arXiv:2010.16061). arXiv. https://doi.org/10.48550/arXiv.2010.16061

Rajasekaran, A., & Indirani, G. (2021). *IOT BASED GLAUCOMA DETECTION AND CLASSIFICATION USING FUNDUS IMAGES-COMPARISON*. 2249–6661.

Reguant, R., Brunak, S., & Saha, S. (2021). Understanding inherent image features in CNN-based assessment of diabetic retinopathy. *Scientific Reports*, *11*(1), 9704. https://doi.org/10.1038/s41598-021-89225-0

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R. P., de Berker, A., Ganguli, S., Gillon, C. J., Hafner, D., Kepecs, A., Kriegeskorte, N., Latham, P., Lindsay, G. W., Miller, K. D., Naud, R., Pack, C. C., …

Kording, K. P. (2019). A deep learning framework for neuroscience. *Nature Neuroscience*, *22*(11), 1761–1770. https://doi.org/10.1038/s41593-019-0520-2

Rocha, K. M., Nosé, W., Bottós, K., Bottós, J., Morimoto, L., & Soriano, E. (2007). Higher-order aberrations of age-related cataract. *Journal of Cataract & Refractive Surgery*, *33*(8), 1442–1446. https://doi.org/10.1016/j.jcrs.2007.03.059

Ryu, G., Lee, K., Park, D., Park, S. H., & Sagong, M. (2021). A deep learning model for identifying diabetic retinopathy using optical coherence tomography angiography. *Scientific Reports*, *11*(1), 23024. https://doi.org/10.1038/s41598-021-02479-6

Sahu, M. K. (2024). A review on glaucoma: Causes, symptoms, pathogenesis & treatment. *Journal of Clinical Research and Ophthalmology*, *11*(1), 001–004.

Shabbir, S., Tariq, A., & Akram, M. U. (2013). A Comparison and Evaluation of Computerized Methods for Blood Vessel Enhancement and Segmentation in Retinal Images. *International Journal of Future Computer and Communication*, 600–603. https://doi.org/10.7763/IJFCC.2013.V2.235

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (No. arXiv:1409.1556). arXiv. https://doi.org/10.48550/arXiv.1409.1556

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Song, E., Sun, H., Xu, Y., Ma, Y., Zhu, H., & Pan, C.-W. (2014). Age-Related Cataract, Cataract Surgery and Subsequent Mortality: A Systematic Review and Meta-Analysis. *PLOS ONE*, *9*(11), e112054. https://doi.org/10.1371/journal.pone.0112054

Sriram, R., .s, G., Madhusudanan, J., Perumal, I., Venkatesan, V., & Murugesan, G. (2015). *A Study on Context-aware Computing Framework in Pervasive Healthcare* (p. 5). https://doi.org/10.1145/2743065.2743104

Sultan, A. S., Elgharib, M. A., Tavares, T., Jessri, M., & Basile, J. R. (2020). The use of artificial intelligence, machine learning and deep learning in oncologic histopathology. *Journal of Oral Pathology & Medicine*, *49*(9), 849–856. https://doi.org/10.1111/jop.13042

Supuran, C. T. (2019). The management of glaucoma and macular degeneration. *Expert Opinion on Therapeutic Patents*, *29*(10), 745–747. https://doi.org/10.1080/13543776.2019.1674285

Tayal, A., Gupta, J., Solanki, A., Bisht, K., Nayyar, A., & Masud, M. (2022). DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases. *Multimedia Systems*, *28*(4), 1417–1438. https://doi.org/10.1007/s00530-021-00769-7

Vasavada, A., Singh, R., & Desai, J. (1998). Phacoemulsification of white mature cataracts. *Journal of Cataract & Refractive Surgery*, *24*(2), 270–277. https://doi.org/10.1016/S0886-3350(98)80210-1

Wang, J., Yang, L., Huo, Z., He, W., & Luo, J. (2020). Multi-Label Classification of Fundus Images With EfficientNet. *IEEE Access*, *8*, 212499–212508. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3040275

Wang, W., & Lo, A. C. Y. (2018). Diabetic Retinopathy: Pathophysiology and Treatments. *International Journal of Molecular Sciences*, *19*(6), Article 6. https://doi.org/10.3390/ijms19061816

Wang, Z., Keane, P. A., Chiang, M., Cheung, C. Y., Wong, T. Y., & Ting, D. S. W. (2022). Artificial Intelligence and Deep Learning in Ophthalmology. In N. Lidströmer & H.

Ashrafian (Eds.), *Artificial Intelligence in Medicine* (pp. 1519–1552). Springer International Publishing. https://doi.org/10.1007/978-3-030-64573-1_200

Wilkinson, C. P., Ferris, F. L., Klein, R. E., Lee, P. P., Agardh, C. D., Davis, M., Dills, D., Kampik, A., Pararajasegaram, R., & Verdaguer, J. T. (2003). Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, *110*(9), 1677–1682. https://doi.org/10.1016/S0161-6420(03)00475-5

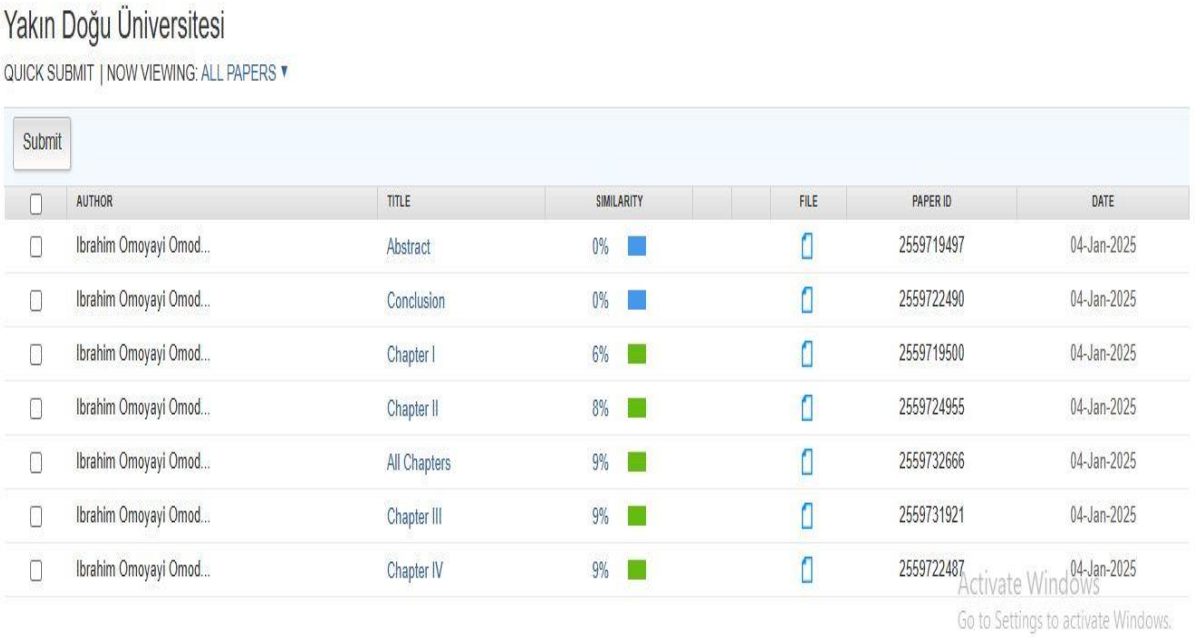
Wong, T.-Y., & Bressler, N. (2016). Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening. *JAMA*, *316*. https://doi.org/10.1001/jama.2016.17563

Yannuzzi, L. A., Ober, M. D., Slakter, J. S., Spaide, R. F., Fisher, Y. L., Flower, R. W., & Rosen, R. (2004). Ophthalmic fundus imaging: Today and beyond. *American Journal of Ophthalmology*, *137*(3), 511–524. https://doi.org/10.1016/j.ajo.2003.12.035

Zhang, W., Zhong, J., Yang, S., Gao, Z., Hu, J., Chen, Y., & Yi, Z. (2019). Automated identification and grading system of diabetic retinopathy using deep neural networks. *Knowledge-Based Systems*, *175*, 12–25. https://doi.org/10.1016/j.knosys.2019.03.016

APPENDIX X

Similarity Report



Yakın Doğu Üniversitesi

QUICK SUBMIT | NOW VIEWING: ALL PAPERS ▾

Submit

AUTHOR	TITLE	SIMILARITY	FILE	PAPER ID	DATE
Ibrahim Omoyayi Omod...	Abstract	0%		2559719497	04-Jan-2025
Ibrahim Omoyayi Omod...	Conclusion	0%		2559722490	04-Jan-2025
Ibrahim Omoyayi Omod...	Chapter I	6%		2559719500	04-Jan-2025
Ibrahim Omoyayi Omod...	Chapter II	8%		2559724955	04-Jan-2025
Ibrahim Omoyayi Omod...	All Chapters	9%		2559732666	04-Jan-2025
Ibrahim Omoyayi Omod...	Chapter III	9%		2559731921	04-Jan-2025
Ibrahim Omoyayi Omod...	Chapter IV	9%		2559722487	04-Jan-2025