Title: THE PREDICTION AND CLASSIFICATION ON TRACHOMA DISEASE USING TRANSFER LEARNING METHOD

Principal Investigator: Dr.Sasikumar Perumal, Associate Professor, Dept.of AI&ML, Sphoorthy Engineering College, Hyderabad

Co-Investigators:

1. Dr.Girithar Akulla, Professor, Dept.of AI&ML, Sphoorthy Engineering college, Hyderabad
2. Dr.J.Paul Mansingh, Professor , School of VAIAL, VIT, Vellore, Tamilnadu
3. Dr.P.Srinivas, Professor,Dept of CSE Malla Reddy Institute of Engineering & Technology,Hyderabad
4. Mrs. S.Indhumathi, Assistant Professor , Dept.of AI&ML, Sphoorthy Engineering college, Hyderabad
5. Mr.Rakesh Reddy, Assistant Professor ,Dept.of Cyber Security, Sphoorthy Engineering college, Hyderabad
6. Mr.T.Manigandan, Assistant Professor ,Dept.of Cyber Security, Sphoorthy Engineering college, Hyderabad
7. Mr.S.T.Saravanan, Assistant Professor ,Dept.of Cyber Security, Sphoorthy Engineering college, Hyderabad

ABSTRACT

In many of the most impoverished and rural parts of Asia, Central and South America, Africa, Australia, and the Middle East, trachoma is hyper endemic. Approximately 1.9 million people have loss of vision or impaired vision as a result of it. In this research work, we develop a model that detect and classify of trachoma and other conjunctiva disease using transfer learning approach.

There have been numerous studies conducted to automate the detection of trachoma diseases on image processing, by using different algorithms. However, almost the existing studies concentrate on only detecting trachoma using machine learning and deep learning. This requires human intervention that can achieve its designed purpose and large labeled datasets for training. Also, no studies have been done to distinguish trachoma from other eye diseases (no differential diagnosis of eye disease has been done).

In Image processing , Normalize an image to a typical size during image preprocessing and use different image filter technique. To extract the region of interest, segmentation is employed. Transfer learning techniques for feature extraction automatically extract features from photos. Pre-trained convolutional neural networks are used in feature learning. We employed Softmax for classification, and sigmoid was utilized to group patients into distinct classes (trachomatous scarring, trachomatous intense, trachomatous follicular, and normal conjunctiva). This study effort consists of a review and experiments on various accessible transfer learning models, namely Yolov3 and Yolov5 for detection, and VGG16, MobileNet v3, and Xception for classification.

The suggested framework is put into practice on the Google colaboratory platform using Keras (using Tensor Flow as a backend) in Python. The neural network is fitted with the extracted features over 50 epochs, with 80% of the data used for training, 10% for validation, and 10% for testing and 0.0001 learning rate. Generally from this work the ophthalmologists can easily classify trachoma severity which has related but different features and easily identify other conjunctiva disease to apply treatments according to the level of severity without wrong diagnostic decisions and time-consuming.

Keywords: Transfer learning, Trachoma, image processing, deep learning.

TABLE OF CONTENTS

Contents page

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYM

Adam	Adaptive learning rate optimization algorithm
BCE	Binary Cross-Entropy
CCE	Categorical Cross-Entropy
CNN	Convolutional Neural Networks
CO	Corneal Opacity
FHF	Fred Hollows Foundation
GPU	Graphics Processing Unit
IDE	Integrated Development Environment
RELU	Rectified linear unit
RGB	Red-Green-Blue
ROI	Region of Interest
TDC	Trachoma Detection and Classification
TF	Trachomatous inflammation Follicular
TI	Trachomatous inflammation Intense
TT	Trachomatous Trichiasis
VGG	Visual Geometry Group
WHO	World Health Organization
WSGI	web Server Gateway Interface
YOLO	You Only Look Once

CHAPTER ONE

INTRODUCTION

1.1 Background

If eye illnesses are not treated in the early stages, they can cause partial or even total blindness. Early detection of some eye conditions can stop vision loss (Akram, 2020). The eye condition known as trachoma is brought on by an infection with the Chlamydia trachomatis bacteria. It is transmitted by eye-seeking insects, sharing towels or cloths, and direct human contact with infected eye and nasal secretions. The illness typically affects underprivileged areas where residents have little access to basic medical treatment, clean water, and sanitary facilities. The World Health Organization (WHO) estimates that trachoma poses a risk to over 190 million people across 41 nations. An estimated 1.9 million people worldwide suffer from trachoma, resulting in half a million becoming permanently blind. (WHO, 2018).

In many of the most impoverished and rural regions of Africa, Asia, Australia, Central and South America, and the Middle East, trachoma is hyper endemic. Roughly 1.4% of blind people worldwide are affected by it. African nations have extremely high rates of trachoma prevalence. For example, in Algeria, the national prevalence of trachoma in schools was 26%; in Burkina Faso, the prevalence was 26.9% (TF/TI); in Ghana, the overall prevalence was 16.1% (TF/TI); in Kenya, the prevalence was 30% (TF/TI); and in Chad, the prevalence of active trachoma (TF/TI) among children under the age of ten was 29.7%. (Berhane *et al*., 2014).

The Ethiopian national prevalence of trachoma is very high for children in the age group 1-9 years are 40.14%. Considerable regional variations are observed in the occurrence of active trachoma; the highest prevalence is in Amhara (62.6%), Oromia (41.3%), SNNP (33.2%), Tigray (26.5%), Somali (22.6%) and Gambella (19.1%). The rural prevalence of active trachoma is almost fourfold compared to the urban (42.5% rural Vs 10.7% urban). The national prevalence of trachomatous trichiasis (TT) is 3.1% with the highest prevalence in Amhara regional state (5.2%) (Berhane *et al*., 2014).

Trachoma has many impacts in a different way of social life. It can cause blindness, loss of eyesight, social status loss, stigma, and financial hardship for individuals, families, and communities. An estimated US$ 3–6 billion is lost in productivity annually due to trachoma worldwide (WHO, 2021).

The World Health Organization divides trachoma grading into five stages: The presence of five or more follicles in the upper tarsal conjunctiva indicates the presence of stage 1 trachomatous inflammation-follicular (TF). Follicles are white, grey, or yellowish patches that are less pigmented than the conjunctiva around them. Stage 2: Intense Trachomous Inflammation (TI): The conjunctiva on top of the tarsus is red, rough, and thickened. Normal visibility of the blood vessels is obscured by follicles or a widespread inflammatory infiltrate. Stage 3: trachomatous scarring (TS): The tarsal conjunctiva develops white lines, bands, or patches as a result of the disappearance of the follicles. Stage 4: trachomatous trichiasis (TT): The eyelashes scrape on the cornea, causing ulcerations and persistent inflammation. This is caused by numerous scars that cause the eyelid edge, usually the upper lid, to bend inwards (entropion). Stage 5: corneal opacity (CO): The cornea progressively becomes less transparent, impairing vision and eventually resulting in blindness (WHO, 2021).

Early-stage differential diagnosis any chronic allergic conjunctivitis, such as toxic conjunctiva, viral conjunctiva, atopic keratoconjunctivitis, giant papillary response, and vernal keratoconjunctivitis (VKC), is included in trachoma. In the upper tarsal conjunctiva, trachoma was distinguished from the other conditions by the presence of five or more follicles. The blood vessels, which are typically visible, are hidden by a diffuse inflammatory infiltrate or follicles, and scars appear as white lines, bands, or patches in the tarsal conjunctiva. The upper tarsal conjunctiva is red, rough, and swollen. Whereas Cobble stoning of superior tarsal conjunctiva or limbal with HornerTrantas dots, corneal implication, itching, mucous discharge in Vernal Keratoconjunctivitis. Lid eczema, superior and inferior tarsal papillae, photophobia, conjunctiva cicatrization, severe itching in Atopic Keratoconjunctivitis. Giant papillae of the superior tarsal conjunctiva, itching, eye discomfort in Giant papillary reaction(Beatriz Vidal Villegas & Jose Manuel Benitez-del, 2021).

Machine learning, a recent development in artificial intelligence, allows for the automatic extraction of information from images through the development of algorithms. Deep learning is the term for an algorithm that uses a neural network with multiple hidden layers. In the field of image classification, where feed-forward convolutional neural networks (CNNs) are utilized for automatic picture categorization, deep learning can be applied. The issue with CNN training has been addressed by the introduction of a unique domain called transfer learning, which allows the weights of a prior network that was trained on a different huge dataset to be utilized in place of starting the network weights from scratch. CNN image classification has a lot of potential applications in the medical profession, but its primary limitation is the sheer volume of training images. This is where transfer learning can be useful. (Abdel Hamid Kandel, 2021).

One of the most effective machine learning models for treating unlabeled data after learning from labeled data is a neural network. For neural networks to perform better, a substantial amount of training data is required. That being said, achieving such requirements in real-time is not required because the only way to obtain tagged data is through a laborious and error-prone manual process. (Hanshu, 2020).

The process of moving knowledge from one domain to another is called Transfer Learning. Adapting features from the source problem and applying them to a new but related target problem is the approach of machine learning (Satheshkumar, 2019). It is not necessary to train the domain from scratch in order to transfer learn the model in the target. As a result, there is a much smaller need for massive training data, and training in the target domain requires less computing power.

There is pioneer researches work is done to detect and classification of eye diseases. However, all research is focused and the detection and classification of glaucoma(Hyeonsung *et al*., 2021), diabetic retinopathy(Inas Al-kamach, 2019), and in a lesser extent to cataract(Masum *et al*., 2021). To the best of this study's ability, research has been done on deep convolutional neural networks for the detection and grading of trachoma only in four (Normal, TS, TT, and CO) cases, but the remaining TF and TI were not addressed and they did not work differential diagnosis.

In addition they used deep learning since new deep learning model must be trained from start, which takes a lot of data, powerful computing power, and hours or even days of training time. Because there are frequently very few images available, Deep Learning techniques are very challenging to apply in the context of medical imaging. The process of gathering and annotating medical image takes a lot of effort and skill.

Generally to overcome the above problem the researcher proposed a transfer learning model to detect and classify three stages of trachoma disease (Scar, TI and TF) , Normal conjunctiva and other eye disease of upper tarsal conjunctiva like viral conjunctiva allergic conjunctiva and toxic conjunctiva. This research focus on only on upper tarsal conjunctiva eye disease because of it has the same region of interest and similar features. But TT and CO have different features which easy to identify from other.

1.2 Motivation

The most frequent infectious cause of blindness in the world is Trachoma. It mostly affects Asia and Africa, two of the world's poorest continents. Loss of vision, blindness, social status loss, and stigmatization are possible outcomes. Detecting and diagnosing the disease needs a highly professional person which is time-consuming and costly for the patients in the manual system. One way to avoid this problem is early detection and treatment of the diseases through the current technology. In previous works researchers used machine learning and deep learning which required both a large volume of data and strong computing power computers. Transfer learning simplifies the complex feature extraction problem and needs a small amount of data. In current technology, by using transfer learning, we can diagnose and detect different eye diseases which have very similar features with small data and low computation costs. This research work decreases the time taken for both patient and ophthalmologists and also reduces the mistake of ophthalmologists at the time of diagnosis.

In general, the goal of this research is to evaluate and assess the effectiveness of the transfer learning model, which enables the diagnosis and classification of trachoma (Scar, TI and TF) from other eye disease and normal conjunctiva by using tiny datasets and limited computational resources.

1.3 Statement of the problem

The giving of a "second opinion" can result in an incorrect diagnosis due to limits caused by a large number of overlapping structures and cases, distractions, fatigue, and limitations with the human visual system (Kaustubh et al., 2020). The 2020 WHO study states that trachoma affects public health in 44 countries worldwide. It is the cause of blindness or vision impairment for about 1.9 million people. In endemic areas, 137 million people are at risk of blindness due to trachoma. It is the cause of 1.4% of blindness worldwide. Africa remains the continent suffering from the worst consequences and the one under the strictest control. Trachoma is the second most common cause of blindness in Ethiopia. In certain regions of the nation, over 76 million people suffer from trachoma. More than 9 million children between the ages of one and nine have active trachoma, while 1.3 million adults over the age of fifteen have TT. Over 27 million people in Oromia reside in areas that have been confirmed to be endemic, according to the Global Trachoma Mapping Project's (GTMP) data on active trachoma in children aged 1 to 9 years (Dedefo Tuke et al., 2023).

In an effort to stop the disease's spread, it needs diagnosis and treatment at an early stage using current technology. Manually detecting and diagnosing these diseases and identifying this disease from other eye disease like allergic viral and toxic conjunctiva need highly trained medical professionals called ophthalmologists. For this reason the patient wait professional expert for two or three weeks even for months which are time consuming and budget consuming for patient. Therefore by using this system any person can identify and classify the disease and take the treatment based recommendation automatically given by the system.

Numerous studies have been conducted employing various methods to automate the identification of trachoma illnesses on image processing. However, almost the studies concentrate on only detection of trachoma using machine learning which needs human intervention only capable of what they are designed for(Akram, 2020). And also they used deep learning which needs big labeled data for training where data is a great challenge in the image processing area and time-consuming for train model(Belesti, 2019), (Masum *et al*., 2021), (Juan *et al*., 2021),(Matthew *et al*., 2019). Several studies have been done to detection and classification of trachoma, for trichiasis (Juan *et al*., 2021), active trachoma (Two-grade TF and TI) (Matthew *et al*., 2019), four grade (Normal, TS, TT, and CO) cases (Belesti, 2019) .

All the above researches uses machine learning and deep learning technique, they did not use transfer learning. And also the researcher did not use differential diagnosis of trachoma at early stage. Therefore The aim of this research was designed and developed transfer learning models for the detection and classification of three of trachoma stages (Scar, TI and TF) and other conjunctiva eye disease(allergic conjunctiva, viral conjunctiva and toxic) and normal conjunctiva using a small dataset to solve the above problem.

1.4 Research questions

- How to collect images of trachoma and other disease tarsal conjunctiva classified according to its stage to prepare datasets?
- How to develop a model of the suggested system
    - How to compare the different models of transfer learning to identify the most accurate model?
- How to test and assess the proposed system using test datasets?

1.5 Objective

1.5.1 General Objective

Main objective is to use transfer learning to create a more accurate and efficient model for the identification and categorization of trachoma illness.

1.5.2 Specific Objectives

In order to fulfill this study's Primary /Main objective, the following particular objective are executed in this work:

- ❖ To create datasets, pictures of tarsal conjunctiva illness, including trachoma, should be gathered and categorized based on stage.
- ❖ To review different related work to identify the problem and propose a solution.
- ❖ To design a transfer learning model for detecting and classification trachoma's stages from other eye disease.
- ❖ To compare the different models of transfer learning to identify the most accurate model.
- ❖ To assess the model's performance by utilizing the testing dataset.

1.6 Significance of the Study

This study has the following significances.

- ❖ enabling ophthalmologists to easily grade trachoma which has related but different features
- ❖ enabling ophthalmologists to apply treatments according to the level of severity
- ❖ enabling ophthalmologists to identify trachoma disease from other eye disease
- ❖ prevent wrong diagnostic decisions at a time of treatment and diagnosis
- ❖ Speed up the detection and diagnosis process and hence reduce the time and effort of patients and ophthalmologists
- ❖ Improve patient care and reduce the workload of ophthalmologists.

1.7 Scope and Limitation

This study's primary objective is to use a transfer learning approach to develop, build, and model automatic trachoma illness detection and classification as well as differential diagnosis. However, TT and CO the latter two phases of trachoma, are not included in this research project.

It can be clearly distinguished from other stages of trachoma because to its unique features. Furthermore, medication is not a part of what we do.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter presents a comprehensive analysis of the concepts and literature pertaining to the diagnosis and classification of trachoma. Literature on the concept of trachoma diagnosis and the approaches used for diagnosis are discussed. We begin with a brief introduction of trachoma diagnosis and classification according to its severity including other eye disease and trachoma control strategy. And also the concept of computer vision, image classification, the steps in digital image processing and image enhancement technique like image filter technique and image segmentation technique are discussed. Finally, a detailed description of Machine Learning, Deep Learning, Image Net and Transfer Learning and its models are presented.

2.2 Overview of Medical Diagnosis Application

The global shortage of physicians, the aging and burnt-out physician population, and the growing need for long-term care have resulted in a serious manpower problem in the healthcare sector. A global shortage of about 17 million healthcare workers exists in addition to an aging workforce. Sleep disorders and burnout among healthcare workers are on the rise due to the rising patient base and physician shortage. Artificial intelligence and technology can easily fix these problems and socioeconomic disparities. Artificial intelligence might speed up medical diagnosis and make administrative tasks easier, which would relieve workers of some tasks(Sharma, 2021).

In most developing countries, the number of people dying from different diseases has gone up since there aren't enough medical experts. There is never going to be a quick fix for the medical expertise gap. In order to receive more diagnosis and treatment, patients were required by existing medical practice to contact specialists. Other medical professionals might not have the training or experience necessary to treat some high-risk disorders, creating a burden on the staff. However, it typically takes a few days, weeks, or even months to receive treatment. It's possible that the illnesses have progressed by the time the patients see the doctor. Patients with high-risk diseases may have to endure their suffering for the remainder of their lives because most of them can only be treated in their early stages(Wan Ishak & Siraj, 2019).

Impaired vision results in a significant reduction in living quality and might even have an impact on household welfare. Services are scarce, and societies in Sub-Saharan Africa don't seem to be ready to deal with these kinds of disability. Up until today, global health interventions aimed at reducing vision impairment have received little attention. Globally, the most prevalent causes of vision impairment include trachoma, diabetic retinopathy, age-related macular degeneration, glaucoma, cataract surgery, and uncorrected refractive error(Genet, 2022).

The Medical Decision-Support System is a computer application meant to assist healthcare providers in making clinical decisions. The system uses medical data and knowledge domains to diagnose patients' problems and suggest appropriate courses of action for individual patients. Furthermore, the system offers support to both patients and medical professionals. The technology reduces iatrogenic illness and medical errors, boosts patient compliance, and enhances the standard of medical decision-making(Wan Ishak & Siraj, 2019).

2.3 Prevalence and Diagnosis of Trachoma

In Ethiopia, there is an extremely high prevalence of trachomatous trichiasis in those over the age of 15, as well as active trachoma in youngsters. There are notable disparities between rural and urban areas in terms of active trachoma cases and gender variations in cases of trachomatous trichiasis. These findings highlight the gender gap in access to preventive and treatment services as well as the overall inequity in the population's availability of clean water and sanitary facilities. Taking into account Ethiopia's population in 2006, it was projected that over 1.2 million adults aged 15 and above had trachomatous trichiasis (TT), while over 9 million children had active trachoma. In comparison to the urban population, the prevalence of active trachoma in rural areas is four times higher (42.5% vs. 10.7%). (Berhane *et al*., 2014).

Trachoma, encompassing both active trachoma and trachomatous trichiasis, is primarily found in certain parts of the nation (Amhara, Oromia, SNNPR), which house a significant portion of the population. Due to their high rural population density and unsanitary circumstances, these areas are more likely to transmit trachoma. According to a study conducted in central Ethiopia, the amount of active trachoma is influenced either directly or indirectly by altitude and sanitary conditions (Berhane et al., 2014).

Careful examination of the tarsal conjunctiva, lashes, cornea, and eversion of the upper lid are necessary when examining clinical indications of trachoma. When examining clinical signs, ophthalmologists employ a slit lamp camera (if one is available) (Belesti, 2019). For the purpose of diagnosing and evaluating trachoma, WHO has created a straightforward five-sign grading system (WHO, 2021). The phases of trachoma as per the WHO simplified grading system criteria are listed below.

Trachomatous Inflammation – Follicular (TF): is detected by the presence of five or more follicles (each of which is at least 0.5 mm in diameter) on the upper tarsal conjunctiva. The presence of follicles is believed to be a sign of active trachoma. It is most commonly found in 3-5-year old children. Follicles are round swellings or spots that are lighter than the surrounding conjunctiva, appearing white, grey, or yellow. Follicles have rounded edges that are not sharply defined.

Intense **Trachomous inflammation (TI)** is characterized by a conspicuous thickening of the upper tarsal conjunctiva due to inflammatory infiltration or follicles, which obscures over 50% of the normal deep tarsal arteries. The tarsal conjunctiva appears red, gritty, and thickened in cases with severe TI. The thicker conjunctiva covers a large number of follicles, either entirely or partially.

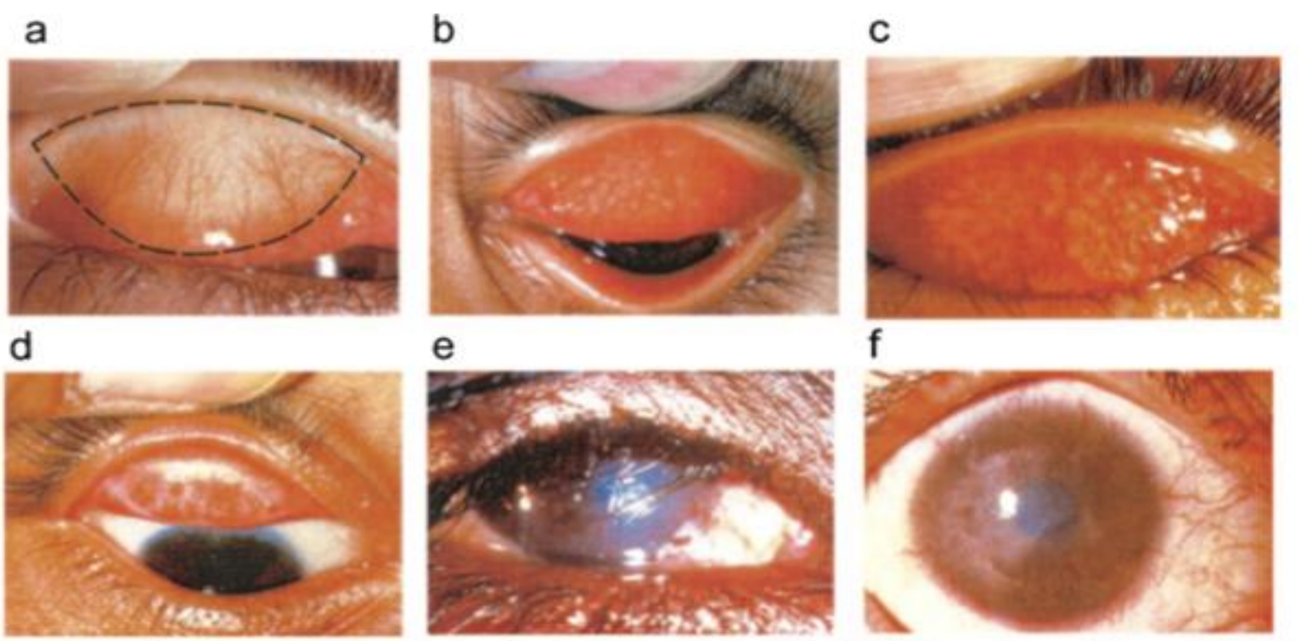
Easy-to-see scars on the upper tarsal conjunctiva are a hallmark of **Tragomatous Scarring (TS).** It is identified by the tarsal conjunctiva's white fibrosis sheets, bands, and lines. Trichiasis and dry eye syndrome are linked to TS, which is a sign of inflammatory diseases in the past. Unlike follicles, which are spherical, scars have sharp edges. Scars have straight, angular, or feathery edges and a shiny, fibrous look. It is important to distinguish between diffuse inflammatory thickening and scarring, particularly diffuse fibrosis, as the latter may conceal the tarsal blood vessels (WHO, 2021).

The condition known as trachomatous trichiasis (TT) is defined as having at least one ingrown lash touching the eyeball or as having recently removed in-turned lashes. Vision can be restored if TT is corrected on time (Who, 2013).

When corneal opacity obscures a portion of the pupil margin, it is known as corneal opacity (CO). It's the trachoma blindness stage. Pannus, infiltration, and epithelial vascularization are examples of opacity. Since corneal opacities significantly impair vision, it is advisable, if at all possible, to measure visual acuity as well (WHO, 2021).

Table 1 trachoma grade and their signs

	Grade		Signs
Infectious	TF	Trachomatous inflammation-Follicular	Presence of 5 or more follicles of >0.5mm in diameter on the upper tarsal conjunctiva
	TI	Trachomatous inflammation - Intense	Presence of pronounced inflammatory thickening of the upper tarsal conjunctiva obscuring more than half of the normal deep tarsal vessels
Non-infectious	TS	Trachomatous conjunctival Scarring	Presence of easily visible scars on the upper tarsal conjunctiva
	TT	Trachomatous Trichiasis	Presence of at least one in-grown eyelash touching the eyeball, or evidence of recent removal of in-turned lashes
	CO	Corneal Opacity	Presence of corneal opacity blurring part of the pupil margin



WHO simplified system. (a) Normal conjunctiva, showing area to be examined. (b) Follicular trachomatous inflammation (TF). (c) Intense trachomatous inflammation (TI) (and follicular trachomatous inflammation). (d) Conjunctival scarring (TS). (e) Trichiasis (TT). (f) Corneal opacity (CO). Reproduced with the permission of the World Health Organization.

Figure 1 : The normal tarsal conjunctiva and the five stage of trachoma

2.4 Trachoma Transmission

Ocular C. trachomatis is shared between people either directly or indirectly by eye-seeking flies or by infected nasal or ocular secretions that land on skin or clothing (fomites). Young children are the main source of energy for other kids and caregivers, mostly moms. Animal reservoirs for C. trachomatis are not known to exist. Research has indicated that nasal and ocular discharge serve as an efficient means of spreading C. trachomatis, particularly in households with limited access to latrines or clean water, or living in cramped quarters. Since frequent face washing of children has been demonstrated to lower transmission rates, infection control measures such as improved sanitation, access to clean water, and health education are critical (Ivanov, 2014).

2.5 Trachoma Control: SAFE Strategy

The World Health Organization recommends using the "SAFE" technique based on the epidemiology of trachoma and its risk factors. According to Paul et al. (2006), this approach entails surgery for TT patients, antibiotics to treat the infection's community pool, face washing, and environmental modifications to stop transmission. The World Health Organization has approved the comprehensive SAFE strategy as a means of eradicating trachoma. surgery for those whose blindness is an immediate risk Antibiotic treatment to address individual active cases and lower the infection reservoir in the community Enhanced hygiene and cleanliness of the face to lessen transmission Enhancements to the environment to enhance living conditions so that trachoma cannot be maintained and spread through the environment The foundation of the fight to eradicate blinding trachoma consists of these four elements.

The success of a trachoma control program depends on the presence of all four elements. That is, delivering surgery, prescribing antibiotics, encouraging good cleanliness, and enhancing the environment all require equal attention. Merely providing surgery and antibiotic therapy without implementing long-term improvements in sanitation and hygiene will simply treat the symptoms of the illness rather than its underlying causes. In addition to the S and A components of the SAFE method, a program needs to include the F and E components for the long-term control of trachoma (Paul & Laura, 2006).

2.5.1 Surgery

The typical treatment advantage of surgery might serve as a foundation for the legitimacy of preventive measures. A quick and affordable surgical technique that turns eyelashes away from the eye to stop more corneal scarring has been approved by the WHO. Local anesthetic surgery can be performed by skilled nurses and ophthalmic assistants. It takes about two weeks to complete training. The actual operation takes about fifteen minutes, and the procedure has an eighty percent long-term success rate. The community level is the ideal setting for the provision of surgical services. Trichiasis is most common in women between the ages of 20 and 60. It's possible that these ladies are unaware of available treatment alternatives or do not have easy access to them. Thus, outreach services and health education are needed for the surgical component of the SAFE strategy (Wondu et al., 2010).

2.5.2 Antibiotics

(TF or TI) Antibiotics cure active disease. The SAFE strategy's antibiotic component treats the infection pool that is present in particular demographics in an effort to reduce community transmission. Tetracycline eye ointment served as the antibiotic component of the SAFE strategy prior to Pfizer starting to donate Zithromax® for the treatment of trachoma. Tetracycline, however, does not address the extra ocular reservoirs of infection linked to the disease's recurrence as a topical preparation, and its six-week treatment regimen resulted in low adherence (Wondu et al., 2010).

2.5.3 Facial Cleanliness

The availability and usage of water are significant determinants in the spread of trachoma. More than the other SAFE approach elements, facial cleanliness depends on health education. In order to influence people to alter their behavior, health educators must support the creation and distribution of persuasive messaging. Innovative techniques to modify the conduct of kids and their caregivers have been tried, like the "child to child" strategy, which involves utilizing older kids to shape the behavior of younger kids and school programs to teach the next generation of caregivers. Control and prevention actions must be planned, devised, and implemented with the cooperation of the community, especially women, since they are frequently the primary caregivers for the children from the beginning. This is because behavioral changes can be challenging to attain and maintain (Wondu et al., 2010).

2.5.4 Environmental Improvement

Numerous environmental factors can either prevent or promote trachoma, and improving the environment is a component of the larger process of community development. The management of human waste is crucial, in addition to expanding access to and usage of water. Water supply and home latrine construction are two aspects of global trachoma elimination projects that aim to meet these needs in the places where the disease is most prevalent. This component's breadth is so wide that cooperation outside the health sector is essential. Enhancements in water and sanitation facilities will not only help decrease trachoma but also have long-term consequences. For this reason, it's critical to coordinate any environmental initiatives with the Ministries of Sanitation and Water as well as other development organizations working on related projects. These advancements need to be socially acceptable, politically viable, and technically solid in order to make an impact (Wondu et al., 2010).

2.6 Why Trachoma Classification (Grading)

Research studies and field surveys employ trachoma grading systems to standardize diagnosis. Anthony and colleagues (2018). It was planned for the WHO simplified system to coexist with the FPC system, which it was designed to be a scaled-down version of. For "less experienced observers" conducting "population-based surveys or for the simple assessment of the disease at the community level," Thylefors et al. deemed the condensed scheme appropriate for use. It offers a great deal less detail than the FPC scale. Nonetheless, the streamlined approach has gained widespread acceptance and is currently utilized extensively in research, community evaluation, and program oversight by both non-specialists and ophthalmologists. According to Anthony et al. (2018), the system mandates that the examiner evaluate a subject for the presence or absence of each of the five signs.

The WHO simplified grading system, which was created expressly to enable quick evaluation of the frequency and severity of disease within a population, is utilized by the majority of trachoma initiatives. Individual trachoma diagnosis was not the planned use for it (Heathcote R Wright et al., 2007).

When a person has four huge follicles or several follicles that are less than 0.5 mm in diameter, the simplified WHO grading system determines that person "does not have trachoma" even though their clinical condition may be compatible with the disease. Because of the oversimplified method, there is a possibility that some people who test positive for follicular disease but do not receive a follicular disease grade actually have trachoma symptoms. For study on the microbiology of trachoma, the more comprehensive WHO grading system may be preferable to the more straightforward method because it would essentially remove this uncertainty (Heathcote R Wright et al., 2007).

2.7 Other Conjunctiva Disease

Pink eye, often known as conjunctivitis, is an infection of the conjunctiva. The thin layer covering the inner surface of the eyelids and the white of the eye is called the conjunctiva. It creates mucus to cover and lubricate the eye's surface. Inside the thin layer are small blood vessels. These blood vessels get bigger and more pronounced when it gets irritated. The eye will look furious and red or pink due to this inflammation.

VKC is a persistent, bilateral ocular allergy that can cause anything from modest vision loss to a life-threatening condition. It is more prevalent in tropical regions, mainly in the Middle East, West Africa, and the Mediterranean Sea. Before the age of ten, VKC typically manifests itself in boys, but females are more likely to experience a later onset. The condition tends to worsen for a number of years before beginning to diminish in the late teens.

The reversible inflammatory syndrome known as giant papillae formation on the upper tarsal conjunctiva (GPC) is caused by the tarsal conjunctiva becoming hypersensitive to sutures, prosthetics, or contact lens wear. Papillae frequently go away when these external chemicals are removed, leading some researchers to define the condition as a response to a mechanical trauma rather than primarily an allergic reaction.

A bilateral chronic inflammatory illness of the eyelids and ocular surface is known as atopic keratoconjunctivitis (AKC). AKC is caused by a combination of immunological responses mediated by cytokines originating from T helper 1 (Th1) and T helper 2 (Th2) lymphocytes, as well as other inflammatory cells, and chronic immunoglobulin (Ig) E-induced mast cell degranulation(1,2). Notably, patients with AKC have a large propria that contains eosinophil, which are never seen in normal tissues and show higher levels of activation markers on their surface (3). Disease often starts in the second through the fifth decades of life (4). The most common symptom is itching, which can also include bleeding, redness, impaired vision, photophobia, mucus discharge, and pain. Seasonal variations may cause persistent itching and other symptoms. Clinical manifestations include papillae in the tarsal conjunctiva, hyperemia of the conjunctiva and episcleral vessels, and concurrent blepharitis.

2.8 Computer vision

Because the human visual system is capable of perceiving three-dimensional structures with sufficient detail, including shape, appearance, and color, it is easy for humans to describe the objects they see in the outside world. The goal of computer vision research is to replicate or characterize objects in pictures or videos using computers, just as the human visual system can (Ying, 2020).

According to Ying (2020), computer vision is defined as "the automatic extraction, analysis, and understanding of useful information from a single image or a sequence of images" by The British Machine Vision Association and Society for Pattern Recognition (BMVA). Numerous computer vision applications have been created in a variety of domains, such as biometrics, autonomous navigation, surveillance, medical imaging, and fingerprint identification. Still, the bulk of applications comprise a number of standard computer vision tasks, including object detection, object categorization, and object recognition (Ying, 2020).

The goal of computer vision research is to develop algorithms capable of producing sophisticated comprehension from pictures and movies. Deep learning has emerged as the most effective method for a variety of computer vision problems in recent years. Deep learning makes it possible to combine a wide range of operations to meet the unique requirements of various computer vision tasks. Back-propagation can be used to train the combined system end-to-end, which frequently produces superior results than conventional systems with individually trained stages (Xie, 2019).

2.9. Digital Image Processing

A subfield of computer vision known as "digital image processing" involves turning images into an array of numbers or pixels or processing digital images with a computer to create an improved image from which valuable information can be extracted (Zewdu, 2020). The term "digital image processing" describes how digital computers handle digital images. Enhancing the original image's visual appeal for human viewers and preparing it for storage, transmission, and representation for independent machine perception are the two primary goals of the processing. Numerous sectors of application, including industrial inspection, astronomy, meteorology, medicine, satellite imaging, remote sensing, and law enforcement, heavily rely on digital image processing (Fatihah, 2016).

A prominent and quickly expanding field of computer science and engineering applications is digital image processing. Technological advancements in digital imagery, computer processing, and mass storage devices are driving its expansion. Because digital technologies are more affordable and auditable than analog imaging, fields that have historically used analog imaging are now transitioning to them. Prominent instances include the fields of medicine, photography, videography, remote sensing, and security surveillance (Maereg, 2020).

Three main categories can be used to categorize the objectives of digital image processing activities (Maereg, 2020).

- ✓ The first step in image preprocessing is to take an image as input and produce an image of higher quality that is prepared for image analysis;
- ✓ Second, image analysis, where measurements or dimensions are the result and an image is the input.
- ✓ Ultimately, image comprehension comprises measurements and descriptions of images as input, verification, identification, and recognition of images as images, and standard descriptions of images as output.

2.11 Image Segmentation

The process of separating objects from the background is called segmentation. The process of isolating an image from its background is known as image segmentation. According to Abhishek and Kirti (2017), there are four primary methods for segmenting images: region-based techniques, threshold techniques, edge detection techniques, and connectivity-preserving relaxation methods.

The process of splitting an image into several segments that are largely/perceptually homogeneous with respect to desired attributes like color, texture, etc. is called image segmentation. In order to provide a description or classification of the image, image segmentation is commonly used to detect objects, estimate the borders of an image, eliminate undesired parts from the image, compress and edit images, or alter and visualize the data. This method is frequently applied, particularly in the analysis of medical images (Erol, 2018).

2.12 Feature Learning

In order to create classifiers or other predictors for task solving, feature learning, often referred to as representation learning, tries to acquire efficient representations that capture relevant and underlying information of the data. A generic term for this process is "feature learning," which highlights how a learning algorithm may be used to optimize the way data is represented in order to complete a given job. To determine the best possible representation of data for a task, a variety of feature-related procedures, such as feature extraction, feature construction, and feature selection, may be engaged in the feature learning process (Ying, 2020).

1. Feature Selection: Variable selection and attribute selection are other names for feature selection. Choosing a subset of pertinent features from the initial huge set of features is the work of feature selection. To represent the data or domain, a significant number of features are frequently used. However, some aspects may be redundant and unnecessary features that don't help with data representation or task solution, like classification. This problem can be solved by feature selection, which chooses a subset of pertinent features. The lowest subset that is required and adequate to represent the data should ideally be chosen. The benefits of feature selection include lowering the dimensionality of the data, accelerating learning, streamlining the learned model, and/or improving performance.

2. Feature Extraction: The process of obtaining new, useful characteristics to represent the raw data through a functional mapping is known as feature extraction. Certain domains—like text, video, and images—cannot be adequately represented by their original, raw data since they lack information. A set of features is frequently extracted to represent an image while solving a job in these areas, such as picture classifications, and the task can be completed using the derived features. In order to convert the raw data into a set of features, the extraction of features may employ or discover certain equations or principles. For example, determining an image's histogram and using it as a feature. While feature selection seeks to reduce the current feature space, feature extraction has the ability to alter the representation of the data by introducing a new feature space. Feature extraction has the ability to decrease the dimensionality of the data by extracting a limited number of features.

3. Feature Construction: The goal of feature creation is to create new, high-level features that represent the data by building upon the existing features. Typically, feature construction uses mathematical equations to determine correlations between the existing features in order to develop new ones. To represent the data, the created features can often be added to or used in place of the original features. In the former scenario, another method for reducing dimensionality is feature construction. Feature building, as opposed to feature extraction, in the latter scenario enlarges the feature space.

2.13 Image Classification

The process of classifying photographs involves giving them a predetermined label depending on the information they contain. Many tasks, such as medical image classification, remote sensing image classification, face classification, texture classification, and biological image classification, fall within the category of image classification. Image categorization is a difficult problem because of the wide differences in background, illumination, viewpoint, scale, distortion, and occlusion between photos (Ying, 2020)..

The process of classifying photos into sequential labels is known as image classification, and it is a crucial field in many domains, including medicine. Deep learning algorithms, which may be thought of as employing autonomous algorithms that can learn by themselves how to differentiate between diverse picture classes, can discover essential features of images without the need for manual feature engineering (Abdelhamid Kandel, 2021).

In machine learning, classification is a common supervised learning job. The process of assigning pre-defined classes to a collection of cases or samples is known as classification. There are two stages to the categorization procedure overall. Training is the first stage in which a classifier is taught, and Testing is the second stage in which class labels are predicted using the learned classifier on unseen data. A classifier is taught, or learned, from cases with class labels during the training phase. The term "training set" refers to the group of examples utilized in this stage. By creating models and/or modifying parameters, the training process uncovers crucial information and guidelines within the training set.

Getting a classifier (model and/or classifier parameters) that can perform optimally (e.g., in terms of classification accuracy) on the training set is the main objective of training. The trained classifier is used to predict the class labels for data (instances) that were not encountered during the training phase. This technique is called testing. The test set, which consists of a collection of cases from the same issue domain as the training set, is frequently used to assess how well the learnt classifier performs (Ying, 2020).

2.14 Transfer learning over Deep learning and Machine Learning Approach

2.14.1. Machine Learning Technique

The field of machine learning (ML) seeks to develop new algorithms for prediction based on input data. ML uses training data to create generic models that can identify patterns in test (fresh) data and determine whether they are present or absent. When it comes to images, training data can be either labeled or unlabeled images, areas, or pixels. High-level and low-level patterns are also possible. A high-level pattern might be the existence or absence of a disease in a medical imaging, whereas a low-level pattern might be a label for a pixel during segmentation. With a training set of picture-label pairings, the problem in this instance is image classification (Erol, 2018).

As a subfield of artificial intelligence, machine learning gives machines the ability to learn from a variety of learning algorithms and use what they have learned to make intelligent decisions similar to those made by humans. However, because machines can only do what they are intended to do—that is, work for what they are designed to do—they will always remain machines and occasionally require human intervention (Zewdu, 2020). We therefore require both transfer and deep learning.

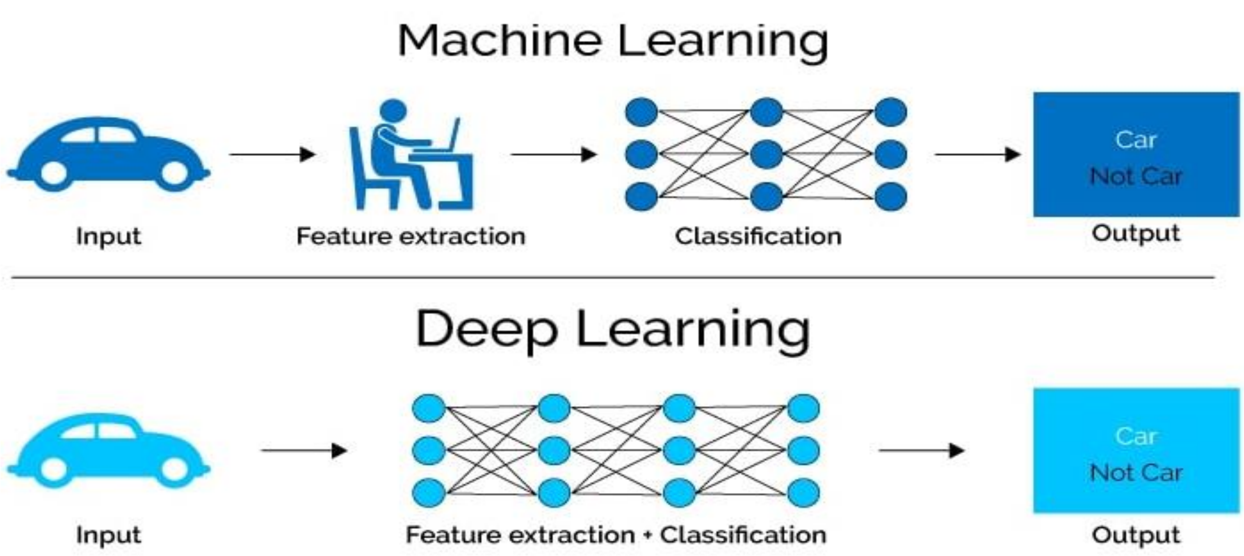


Figure 2 Machine Learning Vs Deep Learning(Sajja & Hemantha, 2019)

ML can be divided into four main categories according on the objectives to be met with it (Boussaada, 2021).

Supervised learning: The reason supervised learning got its moniker is that data scientists teach algorithms what conclusions to draw by acting as guides. It's comparable to how a pupil picks up fundamental math skills from an instructor. Labeled data with the right answers to be anticipated from the algorithm's output are necessary for this kind of learning. In tasks involving regression and classification, supervised learning has shown to be quick and accurate. Predicting the category output value that allows the data to be divided into several classes is the process of classification. Several use cases exist for classification, including figuring out the weather, recognizing whether an email is spam, and classifying distinct animal species after training on an appropriately labeled collection of photos that include the species and other distinguishing traits.

Regression is a kind of problem where it's necessary to forecast a continuous response value, such the price of stocks or homes. In order to forecast the output values for new data based on those relationships that it learned from the prior datasets, it models the relationships and dependencies between the input features and the target prediction output.

Unsupervised Learning: On the other hand, unsupervised learning is more in line with what some experts refer to as true artificial intelligence, which is the idea that a computer can learn to recognize intricate patterns and processes without human supervision. This method is especially helpful when the data does not contain Targets and the experts are unsure of what to look for in the data. It's important to note primary and independent component analysis, association rules, and k-means clustering under the numerous applications of unsupervised machine learning.

K-means clustering is a problem type in which items that are comparable are grouped together. Similar to classification, although this time there are no labels involved; instead, the system will cluster the data based only on what it can determine from the data itself. Clustering news stories and articles based on their type and content would be one application for this. Although this kind of machine learning makes it possible to solve issues that humans would not often take on, its application is not as widespread as that of supervised learning because of its complexity and implementation challenges.

Semi-supervised Learning: Up until now, all of the supplied data has been labeled either with the intended output or not at all. Semi-supervised machine learning combines the two approaches. The cost of labeling is often expensive in real scenarios, and labeling huge datasets makes the process laborious and time-consuming. Furthermore, oversupplying labeled data may introduce biases from humans into the model. Despite the fact that the network is unaware of the unlabeled data, it provides valuable insights about the target group parameters. This leads to the conclusion that unlabeled data can save time and money during the model's construction while simultaneously improving accuracy. Semi-supervised machine learning has potential applications in voice recognition, genetic sequencing, and webpage classification. In some situations, data scientists have access to vast amounts of unlabeled data, and it would take an excessive amount of time to classify it all.

A comparison between these three machine learning models can be established for the same use case, such as classification, using the data gathered thus far. - Supervised classification: Based on the initial labels supplied, the algorithm will categorize the different webpage kinds. - Unsupervised clustering: The program searches for traits and patterns that aid in categorizing webpages. - Semi-unsupervised classification: Using the labeled data to identify the various categories of webpages, the algorithm uses the unlabeled data to draw boundaries between the various types of webpages and to search for other types that may not be included in the labeled data.

Reinforcement Learning: After supervised and unsupervised learning, reinforcement learning is the third primary type of machine learning. It is made up of five essential parts: the agent, the environment, the state, the action, and the reward. Through maximizing its interaction with the environment, reinforcement learning aims to minimize risk and maximize reward. By investigating the surroundings and going through all of the potential states, the RL algorithm—also referred to as the agent—will progressively get better. The agents will automatically find the optimal conduct in order to optimize performance.

2.14.2. Deep Learning Technique

A sophisticated family of machine learning algorithms known as "deep learning" makes use of multiple successive layers. Every layer takes as input the output from the layer before it. There are three types of learning processes: semi-supervised, supervised, and unsupervised. Because the model automatically extracts features during training, deep learning does not require separating feature extraction and classification. It is employed in numerous scientific fields, including bioinformatics, speech recognition, natural language processing, image processing, and image restoration. According to individual definitions, deep learning is a subset of machine learning (Zewdu, 2020).

According to Zewdu (2020), machine learning automatically learns from experience and develops predictions based on data, whereas deep learning involves numerous layers of transformation. Artificial intelligence is the provision of an intelligent system. This network of hidden layers processes input values until they eventually converge at the output layer. The output layer is where the outcome is predicted. An artificial neural network with many hidden layers of neurons, an input layer, and an output layer is called a deep learning network. Therefore, rather than referring to the depth of knowledge, deep learning gets its name from the successive layering of data representation.

Deep neural networks that perform orders of magnitude better than previous networks have been trained thanks to recent advancements in hardware computing power, big dataset accessibility, and reliable training algorithms. Recently, deep learning has shown great promise, particularly in problems involving speech recognition, image classification, and natural language processing (Joseph Kimani, 2021).

Training the deep network weights from scratch in each of these scenarios necessitates large datasets (hundreds of thousands of photos) and a significant amount of time. Due to these prerequisites, 24 deep learning methods are particularly difficult to implement in the context of medical imaging, as there are usually few available photos. An enormous amount of time and experience are needed for medical picture annotation. This is where transfer learning comes into play, enabling the usage of an architecture that has already been trained and adapted to images in the same domain (Abdelhamid Kandel, 2021).

2.14.2.1 Convolutional Neural Networks (CNN)

Several layers work together to create convolutional neural networks, which are used to classify images. The following are the layers that make up the CNN architecture (Sajja & Hemantha, 2019).

1. **Input Layer:** Raw images are received by this layer and sent to later levels for feature extraction.

2. **Convolution Layer:** The convolution layer comes after the input layer. Several filters are applied to photos in this layer in order to extract features from the images. During the testing stage, these features are used to calculate the matches.

3. **Rectified Linear Unit (Rectified-Linear Unit):** This layer comes after the convolution layer or **ReLU**. This layer facilitates quicker and more efficient training by substituting zero (0) for the convolution layer's negative number.

4. Pooling: The pooling layer receives the extracted characteristics. Large photos are captured by this layer, which then lowers them and the parameters to retain crucial information. It keeps each window's maximum value intact.

5. Fully Connected Layer: The last layer is fully connected, in which labels with categories are created from high-level filtered images.

6. Softmax Layer: The output layer is immediately in front of this layer. Each class receives the decimal probabilities from this layer. These probabilities are in decimals, ranging from 0 to 1. The final two phases are referred to as classification stages, and the first four are referred to as feature extraction stages. The final output layer in multi-class classification applications is the softmax layer. The name comes from the softmax activation function, which assigns levels between 0 and 1, adding up to 1, based on a number of projected classification scores as input. This produces an output class with a high probability value (Hanshu, 2020).

2.14.3 Transfer Learning Technique

Using transfer learning techniques is the greatest way to reduce CNNs' weaknesses. For more effective and reliable deep learning model training, transfer learning—also referred to as fine tuning—is frequently used on a network that has already been pre-trained on a sizable dataset (such as ImageNet). Through transfer learning, new models can be tested using relatively smaller training images by using the original parameters (convolution weights) from a model learned on big datasets. For the purpose of training and optimizing models for object detection, freely accessible datasets like ImageNet offer tagged image data of common objects (Hanshu, 2020).

There are numerous benefits to transfer learning. Firstly, it reduces computing time by using the data from the previous training phase rather than creating a new model from scratch. Secondly, it expands upon the understanding it gained from earlier models, and thirdly, transfer learning helps when the new training dataset is not very large. The domains of computer vision, audio classification, and natural language processing stand to benefit greatly from transfer learning (Abdelhamid Kandel, 2021). One of the most often used methods in the fields of deep learning and computer vision for moving knowledge from one domain to another is transfer learning.

When processing capacity is constrained, users can employ pretrained weights from another domain through transfer learning (Erol, 2018).

When applied to machine vision problems, deep learning and machine learning approaches frequently run into issues with the quantity and quality of data. We look into the transfer learning method in an effort to solve this issue. As a result, a training model created from scratch might not adequately capture the diversity of a class. Not to mention, training a network on millions of data points takes a lot of processing power and time, even with access to a sizable dataset. One solution to this issue is to share the neural network model, which has actually "seen" similar or even different data before, rather than the data itself. We refer to this type of strategy as transfer learning.

According to Andrzej et al. (2020), it describes a procedure where a model is first trained on an issue that is comparable to the problem being solved (albeit on significantly different data) and then applied to a different task. There are two primary methodologies within the network-based transfer learning category:

• Feature extraction involves training a new classifier atop the basic model that has already been trained. This approach trains only the final, fully linked layer, leaving the weights learned by the convolution layers unaltered. It is a quick, easy, and nevertheless quite efficient method of utilizing prefabricated architecture.

• Fine-tuning: this involves modifying one or more convolution layers in addition to retraining the fully linked layer. In this solution, we train the final few layers of the original model as well as the newly-added classifier by unlocking some of its layers. The starting point is the weights from the initial training. Unlocked convolution layers are just modified for a new purpose; they are not learned from scratch. The model's performance may be enhanced by this technique, however overfitting may occasionally result. It takes extra time as well.

When selecting which kind of transfer learning to apply to a new dataset, there are two primary things to take into account. These variables include the dataset's size and degree of similarity to the original datasets used to train the repurposed models (Joseph Kimani, 2021). Large datasets can be fine-tuned without running the risk of overfitting, but small new datasets should not be fine-tuned to prevent overfitting. According to Joseph Kimani (2021), the numerous scenarios for new datasets might be summed up as follows.

❖ A little, newly created dataset bears resemblance to the pre-trained ConvNet dataset. As a result, the higher-level characteristics included in the ConvNet will hold significance in the new dataset. Using the pre-trained ConvNet as a feature and solely training a new linear classifier—freezing all other trainable layers—would be the optimal course of action in this situation.

❖ A sizable new dataset that closely resembles the pre-trained ConvNet dataset: in this situation, fine-tuning the new dataset using a pre-trained ConvNet would be the optimal course of action.

A fresh, little dataset distinct from the ConvNet dataset used for training it would be acceptable to utilize the ConvNet that has already been trained as a feature and to train merely a linear classifier because the dataset is minimal. Instead of training the classifier from the top of the network, which has features that are more data-specific, the optimal course of action would be to fine-tune the linear classifier from previous activation levels in the network because the dataset is different.

□ConvNet's can be trained from scratch using a sizable fresh dataset that differs from the pre-trained dataset; however, in practical application, it is still preferable to initialize the network using weights from a pre-trained model and fine-tune across the entire network. In addition to the extensive dataset, a new network can be created and trained from start.

The first thing to keep in mind in this situation is that transfer learning is not a novel idea that is exclusive to deep learning. Using a methodology based on transfer learning principles differs significantly from the conventional way of developing and training machine learning models (Bhargavi et al., 2020).

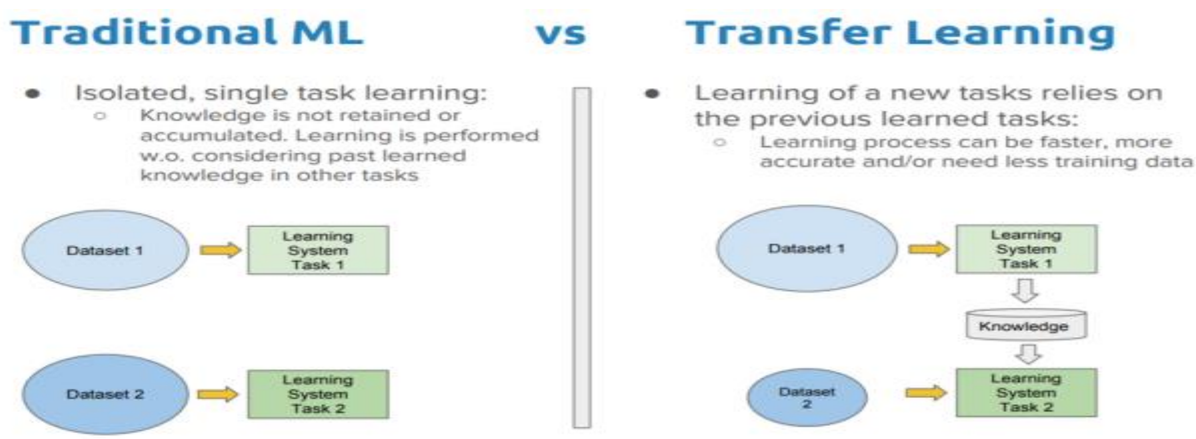


Figure 3:Differences between traditional Machine Learning and Transfer Learning(Bhargavi *et al*., 2020)

2.14.3.1 Image Net

A dataset of over 14 million photos, carefully categorized, arranged hierarchically into 22,000 object categories, and made available as open-source material is part of the Image Net project. The dataset was created with the intention of advancing computer vision research and development. This dataset's imagery serves as the foundation for the Image Net Large Scale Visual Recognition Challenge. The ILSVRC challenge aims to train a model that can correctly identify one of 1,000 common-object classes from an input image. For this challenge, a portion of the Image Net dataset—roughly 1.2 million images for training, 100,000 for testing, and 50,000 for validation—was made available.

It's now common practice to compare computer vision classification algorithms using the Image Net challenge. Convolutional neural network models and other deep learning approaches have topped the challenges' scoreboard since 2012 (Joseph Kimani, 2021).

2.14.3.2Transfer Learning Models

In the field of deep learning, quality and an explosion of quick development may have been observed because of GPU resources and the Image Net database. It led to an enormous manufacturing of several models, including Inception and Alex Net. vgg, ResNet, Dense Net, and exception (Andrzej et al., 2020).

Alex Net: Alex Net is a basic model used for feature extraction that consists of five convolutional layers followed by max pooling layers. The network uses three fully linked layers with Softmax activation for the classification procedure. For improved training performance, non-saturating ReLU activation functions are utilized. There are 60 million parameters in total, and there are 650 000 neurons. Researchers have been inspired by Alex Net to train their architectures using GPU resources.

VGG-16 and VGG-19: In the localization & classification tracks, VGG-16 and VGG-19 secured the top two positions, respectively. The unique aspect of those models was their utilization of tiny convolution filters (3 x 3), which, when paired with the GPU cluster's capability, enabled the network's depth to be increased to 16 or 19 layers. These days, the VGG model's straightforward architecture and strong generalization capacity make it one of the greatest for transfer learning in image recognition applications. With over 134 million parameters, the VGG-16 model features 16 trainable layers, including fully connected, dropout, and convolutional layers in addition to max pooling.

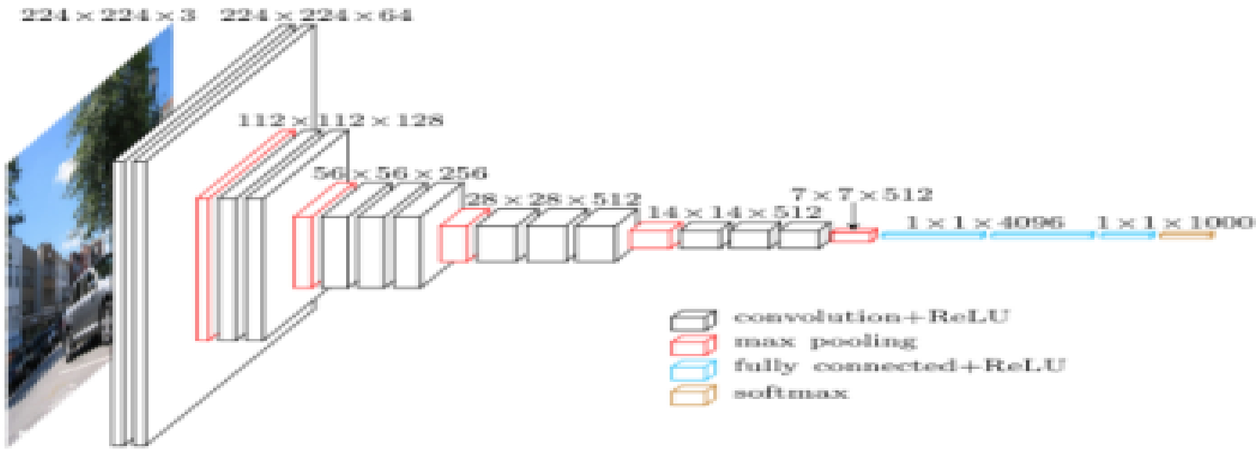


Figure 4: Visualization of VGG Architecture(Joseph Kimani, 2021)

Inception: The primary objective (and the most obvious method) was to expand the network's size, both width and depth. But this straightforward approach had significant disadvantages, such as the possibility of overfitting or—more significantly—a sharp rise in the amount of computer power required. Transitioning from fully connected to sparsely connected architectures was the suggested methodology. Consequently, the fundamental concept underlying the Inception model is to join multiple layers in a parallel fashion, resembling a block, rather than piling them on top of each other.

ResNet: A series of stacked layers make up a residual network, or ResNet. This architecture's defining property is a shortcut that connects the input and output directly at each tier. Remaining block is the layer that is combined with this shortcut. With its 152-layer model, this network was victorious in the 2015 Image Net classification test. It was still less complex than VGG-19 even though it was eight times deeper.

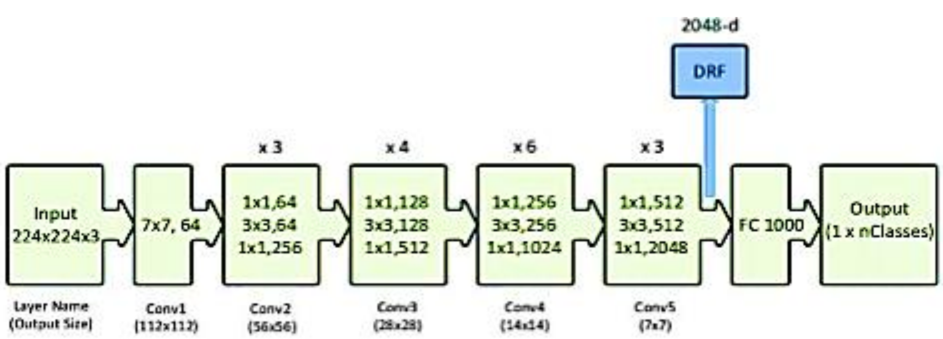


Figure 5 : ResNet architecture(Kevser & Ferhat, 2020)

Xception: The data flow in the Xception architecture is composed of three stages: entering flow, middle flow (which is repeated eight times), and exit flow. The 36 convolutional layers in the design are organized into 14 modules. On the Image Net dataset, Xception performs somewhat better than Inception V3.

Dense Net: The dense net establishes feed-forward connections between each layer and all other layers. Numerous benefits of this type of approach have been demonstrated. First, higher parameter efficiency: Dense Net, which had 20 million parameters, produced results that were on par with ResNet, which had 40 million. Second, Dense Net has enhanced gradient flow and information flow, which helps with strong feature propagation, training, and solving vanishing-gradient problems. Lastly, it's important to note that regularizing the effect of dense connections lowers overfitting.

YOLO: The You Only Look Once (YOLO) approach gets its name from the fact that an image only needs to cross the network once. The processing speed of this real-time approach is 45 frames per second. A "S" x "S" grid divides the image. Bounding boxes, along with associated class probability and offset values, are obtained within each grid. The bounding box is chosen to locate the object if the probability is greater than the threshold. As of right now, there are five versions: YOLO v1 through v5. The majority of the updating focuses on increasing accuracy (Kaili Yu, 2020).

2.15 Role of Transfer Learning in Medical Image Processing

Numerous applications have provided ample evidence of the benefits of using supervised deep learning algorithms for medical imaging tasks. Large volumes of data are needed to construct such algorithms. Unfortunately, access to such large volumes of annotated data presents the biggest obstacle for machine learning techniques in the medical field (Satheshkumar, 2019).

Deep learning requires a lot of labeled data for training in order to prevent overfitting, since all other machine learning algorithms operate by taking the training database of that training, processing new input, and making choices. A Large volumes of tagged images are needed for training purposes in deep learning architectures and computations in order to produce successful image categorization. Large volumes of medical data require time-consuming, highly qualified experts for annotation.

In contrast, transfer learning uses a pre-trained CNN architecture and is able to draw conclusions on its own without the requirement for a large amount of training data, a powerful GPU, a long processing time, or a lot of CPU power. For this reason, we employed the transfer learning strategy, which improves our output.

2.16 Related Work

This case discusses a thorough examination of several research on the identification and categorization of ocular disorders through the use of image processing, machine learning, and deep learning methods. It includes every aspect of eye disease, including diabetic retinopathy, glaucoma, cataracts, and trachoma. Ultimately, this study will highlight the frequent gaps found in the examined publications and how we address them.

2.16.1 Detection and classification of Trachoma

A deep learning CNN model was suggested by Belesti (2019) to diagnose and classify trachoma. They only examined four cases in their research: Normal, TS, TT, and CO. Additionally, the data they used for the previously described situations was not very large. Using empirically determined coordinates from the dataset, they divide the ROI. However, if the ROI is identified automatically, it will be far more effective and efficient. For training and testing, the model's diagnostic accuracy for trachoma detection and grading was 98% and 97.9%, respectively. Compared to the state-of-the-art models, their model was smaller and trained more quickly.

We take into account this gap and work to create or construct a transfer learning model that encompasses the final two trachoma cases for detection and classification.

In 2019, Matthew et al. Convolutional neural networks are suggested for use in the evaluation of eyelid photos. For categorization, they employed 222 TI and 477 TF pictures. Out of which, one hundred pictures were utilized for testing. This suggests that 15% is used for testing and 85% of the dataset is used for training. An image with dimensions of 128 x 128 pixels is fed into the model. A network including three max-pooling layers, three fully connected layers, and seven convolutional layers—each with a filter size of three—has been implemented.

Ultimately, for TF cases, they obtained 92% sensitivity, 48% specificity, and 70% accuracy; for TI cases, they obtained 98% sensitivity, 72% specificity, and 85% accuracy. Because the suggested model only uses the micro (rather than the full) version of VGG, it performs poorly and is limited to the detection of TF and TI situations. To improve performance, no validation process was included. Additionally, the classification of photos with normal eyelids was not taken into account by the suggested system.

Juan et al.'s proposed image sequence analysis using GRU and attention for the classification of Trachomatous trichiasis was published in 2021. In their work, they suggest a technique to decrease the time needed to educate community-based screeners and increase screening accuracy, hence narrowing the assessment gap between specialists and non-experts in TT. Images gathered as part of the Maximizing Trichiasis Surgery Success Trial (MTSS) are used by them. For categorization, they choose a set of 1,706 high-quality pictures (996 TT, 709 non-TT) that surround the eyelid region. For 1,113 photos, our grader annotated the regions of the eye's sclera, cornea, and upper eyelid.

These are two slightly distinct image sets that overlap by 455 photos with TT grading and manual segmentations. With an accuracy of 91%, sensitivity of 92%, and specificity of 87%, the attention-based gated deep learning networks in conjunction with an area identification network can identify TT, demonstrating the practicality of these techniques.

Using machine learning approaches, Akram (2020) suggested an automated system for identifying eye diseases based on the visual content of facial photos. This technique automatically separates the eye portion from the frontal facial image by dividing the facial components. The seven eye diseases—cataracts, trachoma, conjunctivitis, corneal ulcer, ectropion, per orbital cellulitis, and vitamin A deficiency—are analyzed and categorized using the suggested technique. They conclude that the DCNN model performs better than the SVM model based on their experimental findings. Additionally, they contrast their approach with a few other current approaches. When compared to previous approaches, their method exhibits better accuracy. Our DCNN model has an average accuracy rate of 98.79%, 97% sensitivity, and 99% specificity.

2.16.2 Detection and classification of Cataract

Age-related cataracts, cataracts in children, and secondary cataracts are the three types of cataract causes. They can be classified as nuclear cataracts (NC), cortical cataracts (CC), or posterior subcapsular cataracts (PSC) based on where the lens opacity is located.

An Automatic Cataract Detection System Using Deep Learning for Fundus Images is developed by Masum et al. in 2021. Investigating various layers, activation functions, loss functions, and optimization techniques for reducing computational costs without compromising model accuracy was the main goal of the created Cataract Net.

Cataract Net performed competitively when compared to five pre-trained CNN models: ResNet-50, Inception-v3, Mobile Net, VGG-16, and VGG-19. Their model fared better in terms of accuracy (99.13%), precision (99.08%), recall (99.07%), specificity (99.17%), MCC (98.23%), and f1-score (99.07%) than the most advanced cataract detection methods currently available. The ophthalmologists were able to use Cataract Net to more precisely and promptly detect cataract illness because of its high accuracy and efficiency in terms of cost and time. The deep learning model used in this work requires a lot of labeled data to train, which takes time.

2.16.3 Detection and classification of glaucoma

One of the main causes of blindness, glaucoma affects about 64.3 million adults worldwide, or 3.54% of the total adult population. By 2020, it is anticipated that this would have climbed to 76 million. From an economic perspective, the illness has a significant financial impact on both individuals and society, and these costs rise as the illness gets worse (Shalaka & John, 2020).

The Deep Learning Ensemble Method for Classifying Glaucoma Stages Using Fundus Photographs and Convolutional Neural Networks was proposed by Hyeonsung et al. in 2021. This work created and assessed a deep learning ensemble technique to automatically classify glaucoma stages according to severity. In comparison to the best single CNN model, which has an accuracy of 85.2% and an average area under the receiver operating characteristic of 0.950, the proposed method performs significantly better in classifying glaucoma stages, with an accuracy of 88.1% and an average area under the receiver operating characteristic of 0.975.

2.16.4 Detection and classification of Diabetic Retinopathy

Diabetic retinopathy, sometimes referred to as diabetic eye disease, is the result of diabetic retinal degeneration. Up to 80% of patients with diabetes who have had the condition for 20 years or longer are affected by this systemic disease. It is believed that diabetic retinopathy is the primary cause of blindness worldwide (Iyyanar & Parthasarathy, 2020).

Classification of Diabetic Retinopathy Using Pre-Trained Deep Learning Models is developed by Inas, 2019. They put forth an architectural model with varying depths. The suggested model's reduced architectural depth allows it to perform better than the conventional CNN that was created from the ground up. This is also because the experiment's train data set was too large, which made the model labor-intensive and inefficient. Their refined pre-trained model is trained using an Image Net dataset. To observe the behavior of the suggested methods, they modify the parameters, compare it in terms of other domains, and freeze some layers in the experiment.

The researcher reviewed the use of image processing, machine learning, and deep learning approaches for the automated detection of eye illnesses that are relevant to our investigation in this chapter. The researcher made an effort to identify and clarify each work's approach, performance, and limitations as well as how the gaps will be filled up. As far as we are aware, studies have been done to identify and categorize trachoma in cases of trichiasis (TT), active trachoma (TI and TF), and normal, TS, TT, and CO cases. Transfer learning was not used in any of the aforementioned studies; instead, machine learning and deep learning techniques were used. Furthermore, because the researchers did not employ differential diagnosis of trachoma in its early stages, they were unable to distinguish trachoma from other eye diseases.

In order to address the aforementioned issue, the study's goal was to design and develop transfer learning models for the detection and classification of three different stages of trachoma (Scar, TI, and TF), as well as other conjunctival diseases (allergic, viral, and toxic), and normal conjunctiva using a small dataset. For the reasons mentioned above, we included another eye disease from the upper tarsal conjunctiva in this study because it shares characteristics with trachoma disease and is challenging to identify. We also developed a transfer learning model to address the issue with the trachoma disease detection and classification in earlier research.

Table 2: Summary of Related Work

Author	Title	Algorithms	Gaps/Limitation	Performance
Ashrafi Akram(2020)	An Automated Eye Disease Recognition System From Visual Content of Facial Images Using Machine Learning Techniques	DCNN and SVM	Machine Learning Techniques is Used Which is Time Consuming & Need Human Intervention	Accuracy 98.79% with sensitivity of 97% and specificity of 99%
Matthew *et al* (2019)	Sensitivity and Specificity of Computer vision Classification of Eyelid Photographs for Programmatic Trachoma Assessment	Use mini-VGG network architecture Softmax for classification	Works for only two cases: TF and TI Better if they are compared with the normal cases	92% sensitivity, 48% specificity, and 70% accuracy
Masum *et al*., (2021)	An Automated Cataract Detection System Using Deep Learning for Fundus Images	CNN	Proposed for Cataract Detection & Not for Grading or Finding Its Exact Location, Which Can Be Helpful For Ophthalmologists	accuracy of 99.13%
Inas Al_Kamachy (2019)	Classification of diabetic retinopathy using pre-trained deep learning models	CNN	Making connections between a model & a mobile camera by taking pictures	InceptionResNetV2=68 nceptionV3=68 VGG_16=57
Hyeonsung *et al*., (2021)	Deep Learning Ensemble Method for Classifying Glaucoma Stages Using Fundus Photographs and Convolutional Neural Networks	CNN	Low performance	accuracy of 85.2%

CHAPTER THREE

RESEARCH METHODOLOGY

3.1. Overview

The research approach that provides the means to address the research topic is thoroughly covered in this chapter. The set of processes, strategies, instruments, and methodologies utilized to carry out research and arrive at the study's conclusion is known as research methodology. Choosing a suitable research methodology is a crucial stage in the research process.

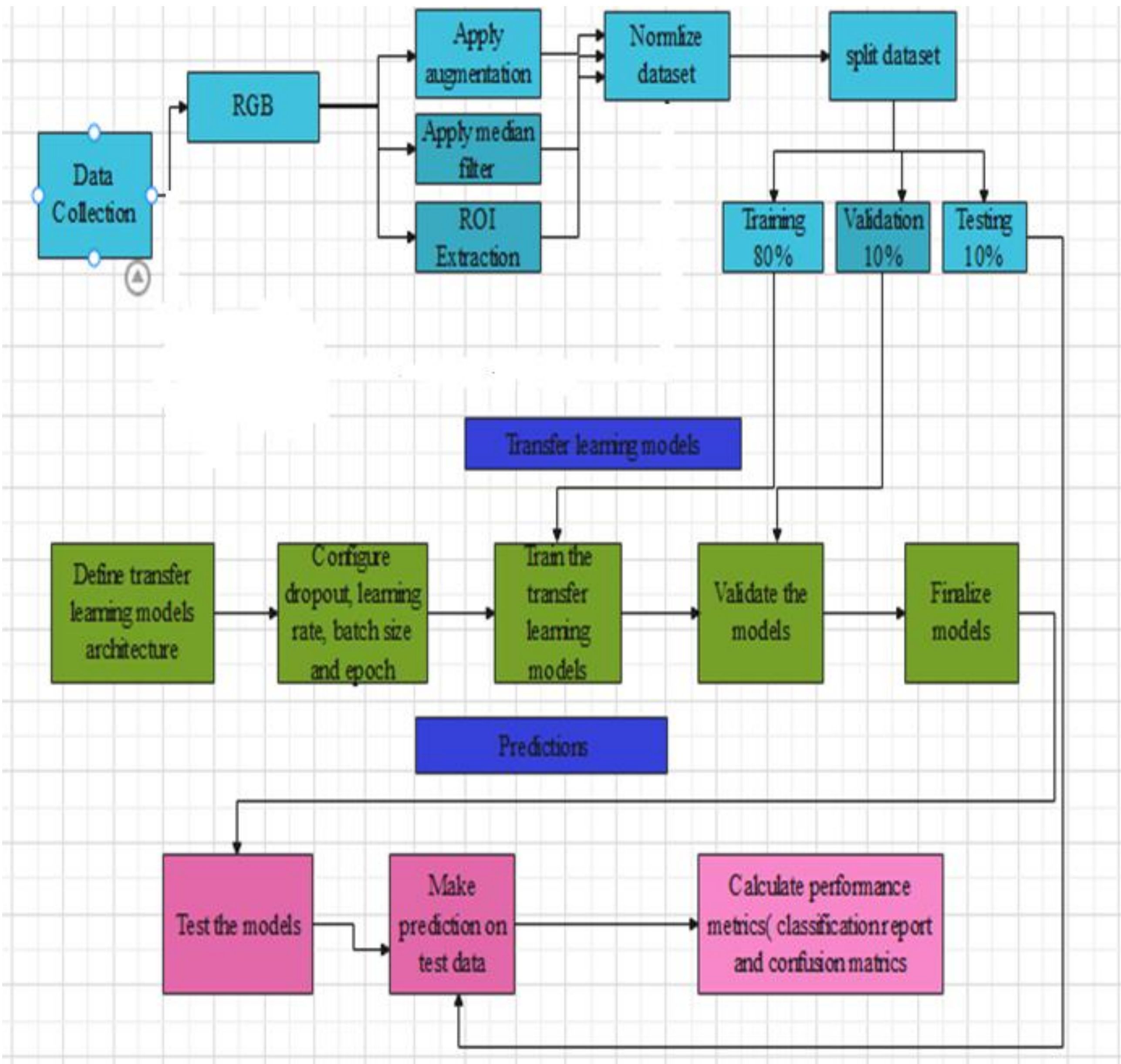


Figure 6: Research flow diagram

To accomplish the goals and objectives of the research as well as provide accurate and dependable results, a methodology that combines techniques and approaches that are most appropriate for the study is selected (Zewdu, 2020). As a result, we go into great length in this chapter about the research process employed, from the conception of the problem to the conclusion.

3.2. Problem Formulation

This section highlights the primary scientific flaw in the identification and categorization of trachoma diseases and investigates potential remedies. In order to learn more about the issue and the significance of the solution, a variety of literatures are reviewed. We examined and assessed a large number of research studies that have been conducted to identify and categorize different kinds of eye diseases in order to determine the approach and methodology employed in doing so, and we found the gaps. This study's primary goal is to distinguish trachoma disease phases from other eye conditions. This is accomplished by training a model based on transfer learning with primary and secondary picture data that is gathered from various sources.

3.3. Literature Review

To comprehend the concepts of image processing, transfer learning, trachoma disease, and how they applied to solve the related problem in the previous research, search articles, theses, dissertations, and books in various databases such as Google Scholar, IEEE, PubMed, Wiley, ACM, Science Direct, SciPub, Springer Link, and Web of Science. The most recent and up-to-date literature was reviewed in order to meet this research goal. The gap in the earlier research was found during the literature review process and used as the foundation for the suggested solution. This is the crucial and essential material where every potential research-related reference and publication is looked into and examined.

3.4 Data source

Since the medical image for the classification of trachoma disease is not available in a publicly accessible database. Trachoma team surveyors gather clinical sign photos from the field for this study, and some expert professional data is also gathered. The image includes photographs of the upper tarsal conjunctiva that show normal eye, diseased eye in each of its stages (TF, TI, TS), and other eye diseases.

3.5 Image Preprocessing

Several image preparation procedures are applied to images in order to improve visualization. A network may extract more distinct and conspicuous elements from images once they are brighter and crisper (Sarki, 2021). Every image in the dataset was preprocessed in this study in accordance with the specifications of the deep neural network at every step.

The input size applied to the model and the picture input size used in training and testing must match. The image was shrunk to 224x224x3 in this instance. The goal of the pre-processing approaches is to prepare the image for additional processing. We employed a median filter approach for this study. Since the median filter preserves image edges while eliminating noise, it is superior to other filtering techniques. The median filter turns all of the image's pixels, and its close neighbors are used to determine whether or not it represents his surroundings.

3.6. Image Augmentation

The technique of creating extra data from the RGB (original dataset) in order to increase the number of training data points in a dataset is known as data augmentation. Big data is necessary for a neural network to be trained effectively. When the trained network generalizes poorly and the training data is not sufficiently varied, parameters are compromised. By performing geometric alterations on pre-existing photos while maintaining their labels, data augmentation provides a solution to this issue. By obtaining additional information from the training dataset, the augmentation step helps the network prevent overfitting (Molla, 2020).

Using augmentation techniques is one option to address the issue of the size of the dataset. A reliable method that has shown promise in computer vision models is image augmentation.

Expanding the number of datasets is crucial. It avoids the overfitting issue and aids the network in learning more intricate features from the data.

Before training the model, a variety of data augmentation techniques were applied to the original pictures dataset in this work. In this instance, the original photos have been cropped, rotated, moved, and brightened.

Brightness: As brightness is neither a geometric nor a synthetic change, it is an augmentation technique that is difficult to classify into the preceding categories. The images' fundamental structure will alter when their brightness varies. Since the self-supervised algorithm's whole functioning will depend on the representations brightness enhancement offers, its application may have a greater detrimental impact on categorization. In this instance, brightness is helpful in representing the disease's parameters at every stage.

Shear: Extends the picture along one of the axial planes, such as the x- or y-axis. The highest shear that is utilized to maintain class preservation is ±20◦.

Cropping: By resizing a central portion of each image, cropping can be applied as a useful processing step for image data with mixed height and width dimensions. Furthermore, random cropping can be applied to provide a result that closely resembles translations.

Rotation: To enhance rotation, rotate the image in either direction on an axis that ranges from 1° to 359°. The rotation degree parameter plays a major role in determining the safety of rotation augmentations. A 180° rotation range, a 25% shifting range, a 40% zoom range, and arbitrary horizontal and vertical image flipping were all employed.

Table 3: Setting used for the augmentation operation

Transformation	Setting
Vertical Shift	-0.4 to 0.4
Horizontal Shift	-0.3 to 0.3
Rotation	-40 to 40
Brightness Range	0.5, 2

3.7. Segmentation

Picture segmentation divides a picture into different sections according on how similar different aspects are to one another. Parts with similar characteristics are grouped together. Segmentation makes it simple to study the image (P., 2017). Not all of the image's content contains the distinguishing traits. Only the image's follicles for follicular and intense white lines, bands, and sheets of fibrosis indicative of an upper tarsal conjunctiva scar are of importance to us in this investigation. Therefore, we employed the oval shape from the upper tarsal conjunctiva to extract the ROI from various phases (Normal, TF, TI, TS, and other eye illness).

3.8 Feature Extraction

An essential first step in developing an image classification system is feature extraction. The process of identifying significant and representative information from a picture that characterizes the image data and lowers the dimensionality of the image data is known as feature extraction. The efficiency of the picture features has a major impact on how well the classification system performs (Ying, 2020).

Following the preparation and segmentation steps, the image's diseased area is extracted. For image processing, this trachoma illness area is considered an intersection. Subsequently, additional traits are derived from the disease symptoms in order to identify the different stages of the disease. The different picture features that are supplied as input to the classifier are kept after the feature extraction stage. Using features, the classifier quickly divides the visual input into several classes. In this instance, pre-trained models extract several aspects of the trachoma disease according to their stages, such as the presence or absence of blood vessels and follicles for the first two stages of the disease (intense and follicular), and the existence of scarring for scarring. And in some diseases, giant papillae, lid eczema, and cobble stoning of the superior tarsal conjunctiva or limbal with Horner-Trantas spots.

3.9 Classification

According to Errol (2018), classification is the process of grouping objects together based on shared traits, attributes, or other predetermined qualities.After learning the distinguishing characteristics, classification is completed. Feature learning involves multiple layers layered on top of one another. We classified trachoma into distinct classes (normal, TF, TI, TS, and other eye diseases) using three pre-trained CNN architectures: VGG16, Mobile Net v3, Xception, and sigmoid and softmax activation function.

3.10 Detection

In order to detect if an eye is diseased with trachoma, infected with another eye illness, or uninfected, we employed two pretrained Yolov3 and Yolov5 in this investigation. The model in this work can identify and categorize data into three groups.

3.11 Model Selection and Building method

Yolov3 and Yolov5, which were pre-trained on the COCO dataset, were chosen for usage in this work in order to detect diseases. Additionally, three pre-trained convolutional neural networks—xception, vgg16, and mobile net v3—that were made accessible using the Keras library and pre-trained on the Image Net dataset were chosen for the disease classification task. Based on factors including model accuracy, model loss, model size, and training time, the researcher chose those models. Following the model's selection, we use two transfer learning techniques to develop them: layer freezing and fine-tuning. After comparing the models' performances on the objective of classifying trachoma disorders, the network with the greatest performance was chosen to be the basis for a web application that would detect and categorize the disease.

3.12. Material and Tools

3.12.1. Software Tools

To determine which software tool is best for implementing the transfer learning algorithm for trachoma picture categorization and detection, an investigation into available software tools and their libraries is carried out. The researcher utilized Python as a programming language together with the Tensor Flow and Keras libraries on Google Collaboration as software tools to create the transfer learning algorithm.

Tensor Flow is the most well-known and quickly developing transfer learning framework at the moment. It is an open-source, free library created by Google. It may be accessed on any desktop computer running Windows, macOS, or Linux, as well as on mobile devices running iOS and Android and in the cloud as a service (Zewdu, 2020).

Python-written Keras is a high-level neural network API that operates on top of Tensor Flow using either Microsoft Cognitive Toolkit (CNTK) or Theano. Most notably, it includes transfer learning models like VGG16, ResNet, exception, and Inception, which were utilized in the experiment. It is also relatively easy to construct a model for and to extend using Python (Zewdu, 2020). Generally, we utilized Visual Studio and the Flask Python Framework to create web applications that identify and diagnose trachoma based on stages and to change code

3.13 Performance Enhancement

Optimizer Selection: To lower the loss function, the neural net nodes' parameters are automatically adjusted throughout the training phase. Furthermore, the optimizer that is used has a significant impact on the size of the parameter adjustment. Regularization and Learning Rate are the two most important weights used to assess the Optimizer's success (Sarki, 2021). One of the often used techniques for optimization is the Adam optimizer. It blends the advantages of the RMSProp and Adagrad algorithms (Mandali, 2022). Thus, Adam is the optimization method applied in this Research Work.

Learning rate: It regulates the amount of weight to update during back propagation and the speed at which the learning process is carried out. Selecting the appropriate learning rate for the experiment was difficult (Zewdu, 2020). It is employed to specify the neural network's learning rate. Weight and bias values are updated during back propagation. The learning rate aids in the modification of their values (Mandali, 2022). We observed in our experiment that training takes longer when the learning rate is too modest than when it is bigger. However, the model performs better when we provide a smaller value than when we use a bigger learning rate. In our instance, a learning rate of 0.0001 was used for the experiment.

Loss function: Since categorical cross-entropy, or CCE, is used for multi-class classification, we used it to address the kind of problem we are attempting to solve in the transfer learning model. The loss function is directly related to activation functions that are used in the output layer, the last fully connected layer of the model. We chose CCE in our experiment even though there are other loss functions like Mean Squared Error (MSE) and Binary cross-Entropy (BCE).

Activation function: It is employed in the computation of the subsequent layer's neuronal value. Numerous activation functions exist, including Sigmoid, Softmax, Leaky Rectified Linear Unit (Leaky ReLU), Rectified Linear Unit (ReLU), and Linear (Mandali, 2022). We employed both softmax and sigmoid activation in this investigation, but since our photos were classified into multiple classes, softmax performed better than sigmoid.

Number of epochs: The number of times the network sees the entire training set during training is known as the epoch count. is the total number of times the training dataset runs through the network or model in the experiment, both forward and backward (Zewdu, 2020). The model in our study was trained utilizing a range of epochs, from 10 to 50. We observed during training that the model overfits between the training error and validation error when we select an epoch that is too little or large.

Batch size: It is used to provide the total number of input rows that will be processed concurrently. It is employed to combine the processing of many inputs. In order to update the parameters, a single set of training input data is broken into smaller subsets known as mini-batches. The total number of training input data we pass into training is present. However, the model is trained over multiple epochs using the entire dataset, split up into mini-batches, to avoid any loss in accuracy or generalizability (Mandali, 2022). For our investigation, 32 is the ideal batch size.

Table 4: Summary of Hyper-Parameters Used During Model Training

Models	Image Size	Optimizer	Batch Size	Learning Rate	Number of epochs	Loss Function	Activation Function
Vgg16	224*224	Adam	32	0.0001	50	CCE	Softmax and sigmoid
xception	224*224	Adam	64	0.0001	50	CCE	Softmax and sigmoid
Mobile Net V3	224*224	Adam	64	0.0001	50	CCE	Softmax and sigmoid

3.14. Performance Evaluation Metrics

We evaluated our model using a confusion matrix, f1-score, precision, recall, and accuracy, among other performance evaluation criteria.

Precision: All projected positive cases divided by the number of correctly discovered positive instances is implied. Keep in mind that this figure shows the percentage of all correctly identified positive cases among all actual positive cases.

Accuracy: One of the most obvious metrics is the total number of cases that were properly identified. The F1-score, which is the harmonic mean of Precision and Recall, is a more accurate measure of cases that were incorrectly classified than the Accuracy Metric.

Another assessment metric used to characterize a classification task's performance is a confusion matrix. A confusion matrix can be used to calculate precision, recall, and accuracy.

CHAPTER FOUR

SYSTEM DESIGN

4.1 Introduction

This chapter focuses on the functions and design of the suggested work. Python 3.10.5 was used to develop the complete system. Pip install is used to install the Flask framework and Flask extension package, which are needed by the whole system. Installing Flask, Flask-Script, Flask-Bootstrap, werkzeug, and other components is necessary.

4.2 System Functions

Both functional and non-functional requirements apply to this system. Functional requirement: The doctor can upload an image from a local file, view the submitted image, and see the anticipated result. The user may also ask questions about the disease and access a wealth of information on the system and trachoma sickness.

Non-function requirement: Efficiency and Usability: The system's high degree of efficiency is ensured by extensive testing and iterative user experience flow modifications. Performance and Scalability: The Firebase platform is developed for both performance and scalability, and the system was created with scalability in mind. Simple to use: There is no need to wait for a specialist because everyone can easily utilize the system. This may entail using the least amount of time and effort possible when browsing the system's contents.

4.3 System Architecture

The user interface and front-end interaction are mostly handled by the client, which gives the user an operator interface to choose the identified images on the interface and get the results of detection and prediction.

The server's main duties include answering the client's request, calling functions, and providing the Corresponding results.

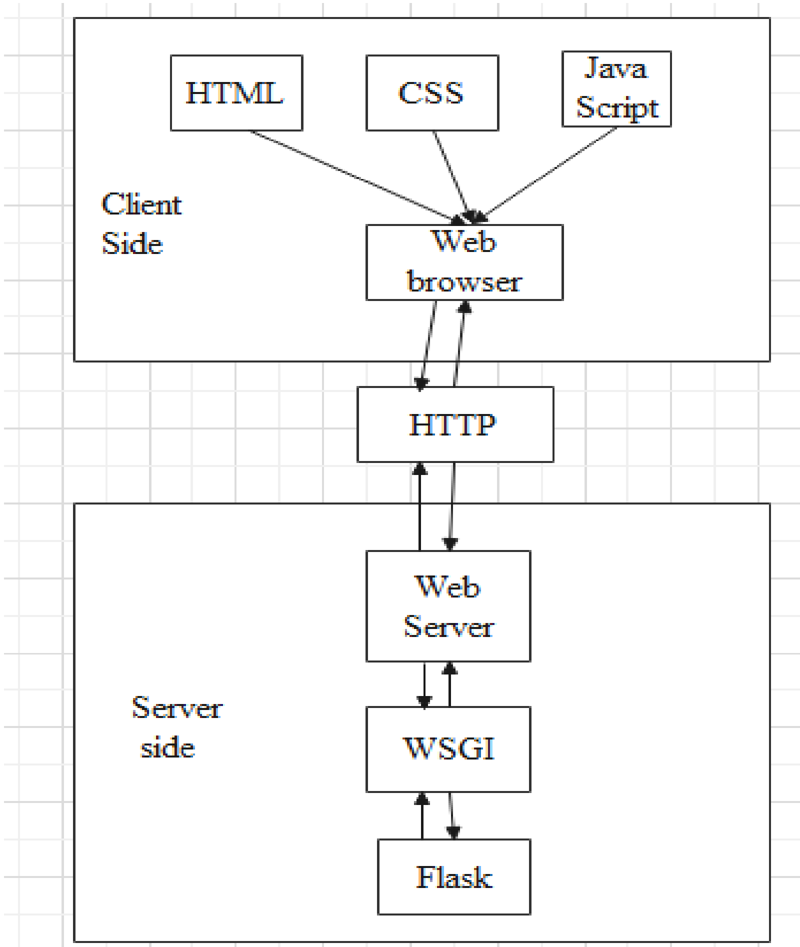


Figure 7: Architecture design

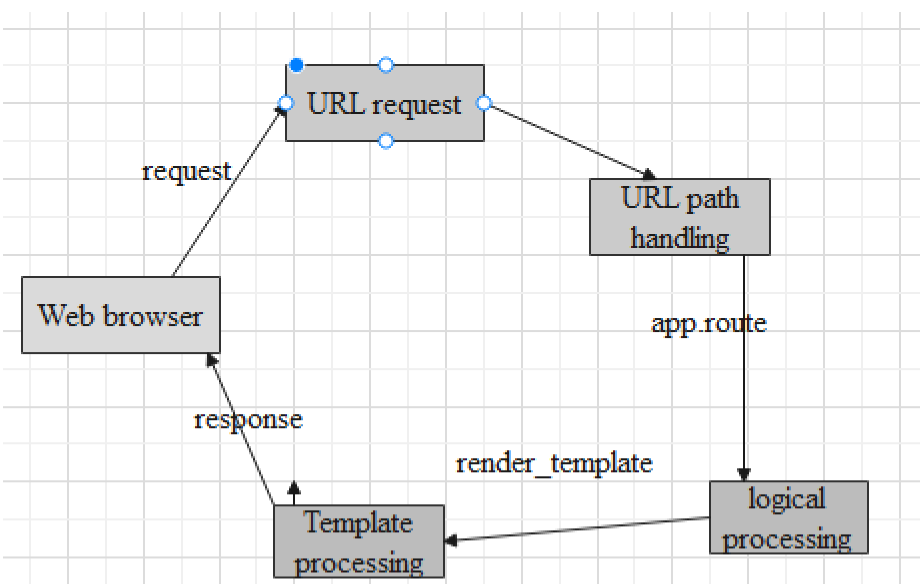


Figure 8: System internal response process

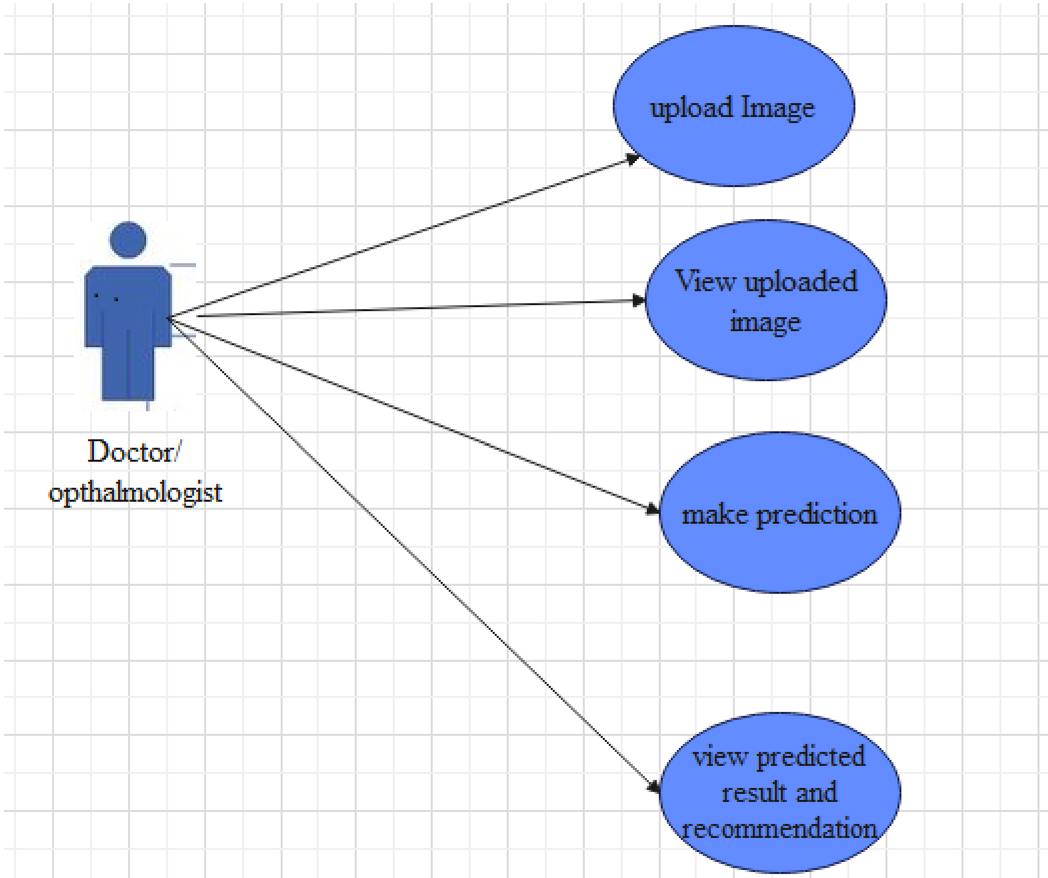


Figure 9: Use Case Diagram for Trachoma Detection and Classification System

4.4 Front-End Implementation

Flask front-end template uses Jinja2 template engine. After rendering a given template, the program's render template method computes any variables or functions in the inherited template to produce an HTML page that is complete.

4.5 Back- End Implementation

Run pip after installing Python 3.10.5, create a flask project and install a flask. This module is mostly responsible for creating and configuring Flask instances. During startup and joining the route, the system will use a few other components to post requests for file uploads. Next, the picture path needs to be saved.

4.6 Summary

The advancement of trachoma detection, classification, and differential diagnosis by various transfer learning techniques are the primary topics of this thesis. Several techniques are employed in this work, including picture augmentation and image filtering to remove image noise. In this thesis, the detection and classification of trachoma images is done using transfer learning models, and the performance of several models is compared in terms of both detection and classification.

Transfer learning, which has a high feedback accuracy and efficiency, was chosen through experimental analysis to enhance the system's classification performance. Essential data augmentation is used in the image augmentation method to resize the image vertically or horizontally, hence expanding the data set. This is a crucial step in increasing the model's classification rate. System design and implementation: in this instance, a web application is built using the Flask framework.

Since most stages have a similar region of interest, the experimental issues are overfitting and underfitting in training, loss of validation or accuracy, and extraction of region of interest. According to the testing results, the mobile net v3 model outperforms xception and vgg16 in terms of classification accuracy. In comparison to the two models, it weighed less and trained more quickly. In the future, researchers can also attempt comparing the speed, flexibility, and adaptability of various frameworks using others, such as Django, Tornado, and others.