TRIBHUVAN UNIVERSITY
INSTITUTE OF SCIENCE AND TECHNOLOGY

Final Year Project Report on

Ocular Disease Detection and Classification Using Convolutional Neural Network and Transfer Learning

For the partial fulfillment of the requirements of the degree of

Bachelor of Science in Computer Science and Information Technology (B.Sc.CSIT)
awarded by Tribhuvan University

Submitted By:

RASHU SHRESTHA (23348/076)

TU Registration No. : (5-2-282-26-2019)

SAFALTA KHANAL (23354/076)

TU Registration No. : (5-2-282-32-2019)

ST. XAVIER'S COLLEGE

Maitighar, Kathmandu, Nepal

Submitted To:

Office of the Dean

Institute of Science and Technology

Tribhuvan University

Kathmandu

March 2024

“Ocular Disease Detection and Classification Using Convolutional Neural Network and Transfer Learning”

[CSC – 412]

A final year project report submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Information Technology awarded by Tribhuvan University.

Submitted by:

Rashu Shrestha(T.U. Exam Roll No. 23348/076)

Safalta Khanal (T.U. Exam Roll No. 23354/076)

Submitted to:

ST. XAVIER’S COLLEGE

Department of Computer Science

Maitighar, Kathmandu, Nepal

March 13, 2024

DEPARTMENT OF COMPUTER SCIENCE

DEPARTMENT OF COMPUTER SCIENCE
ST. XAVIER'S COLLEGE, MAITIGHAR, KATHMANDU

SUPERVISOR RECOMMENDATION

This is to certify that the final year project entitled **“Ocular Disease Detection and Classification Using Convolutional Neural Network and Transfer Learning”** is an academic work completed by **Rashu Shrestha (T.U. Roll No.:23348/076)** and **SafaltaKhanal (T.U. Roll No.:23354/076)** submitted in the partial fulfillment of the requirements forthe degree of Bachelor of Science in Computer Science and Information Technology awardedby Institute of Science and Technology, Tribhuvan University under my guidance and supervision.

The information presented by him/her in the project report has not been submitted earlier to the best of my knowledge,

____

Er. Rajan Karmacharya
Project Supervisor/ Lecturer
Chief Technology Officer
Department of Computer Science
St. Xavier’s College
Date:

CERTIFICATE OF APPROVAL

The undersigned certify that they have read and recommended to the Department of Computer Science, St. Xavier's College for acceptance, a Final year Project Report entitled **"Ocular Disease Detection and Classification Using Convolutional Neural Network and Transfer Learning**" submitted by **Rashu Shrestha (T.U. Roll No.:23348/076)** and **Safalta Khanal (T.U. Roll No.:23354/076)** for the partial fulfillment of the requirement for the degree of Bachelor of Science in Computer Science and Information Technology awarded by Tribhuvan University.

……………………………..
Er. Rajan Karmacharya
Project Supervisor
Chief Technology Officer
St. Xavier's College

……………………………..
Mr. Ganesh Yogi
Internal Examiner
HoD, Department of Computer Science
St. Xavier's College

……………………………..
Mr. Ganesh Yogi
HoD, Department of Computer Science
St. Xavier's College

……………………………..
Mr. Nawaraj Paudel
External Examiner
CDCSIT, TU

ACKNOWLEDGEMENT

We would want to take this chance to express our sincerest gratitude to everyone who helped us throughout our project and supported us to complete this final year project.

It brings us great pleasure to express our sincere gratitude and heartfelt appreciation to our highly respected and esteemed supervisor, **Er. Rajan Karmacharya, Chief Technology Officer, Lecturer, Department of Computer Science**, for his valuable direction, inspiration, and assistance in completing this research. We truly thank him for his helpful recommendations and cooperative attitude during this entire project. We would also like to express our gratitude towards **Mr. Ganesh Yogi**, **Head of Department** for his constant support.

We would also like to thank all our lecturers for their unwavering support and direction. Additionally, we would like to express our sincere gratitude to **Mr. Deepak Bdr. Bohara, Administration Assistant, Kirtipur Eye Hospital and Ophthalmic Study Center** for his support. In the end, we would want to extend our sincere gratitude to all of our friends, and everyone who supported and assisted me in some way during this project.

Rashu Shrestha (T.U. Exam Roll No. 23348/076)

Safalta Khanal (T.U. Exam Roll No. 23354/076)

ABSTRACT

In the face of the increasing global incidence of eye conditions, it is critical to have reliable diagnostic instruments that enable timely intervention and tailored treatment plans. This study explores the application of deep learning techniques to the diagnosis and classification of ocular disorders using medical image analysis. It explores automating the identification and categorization of various pathological disorders, such as glaucoma, cataract, and diabetic retinopathy, using convolutional neural network (CNN) models such as VGG19 and ResNet50 as the foundation. To improve model performance, extensive training on a variety of large-scale datasets is conducted in conjunction with data augmentation methods and preprocessing the fundus images. The evaluation of each algorithm's performance on the testing dataset involves measuring its accuracy, precision, recall, and F1-score metrics followed by training on the training dataset. The experimental results show that the proposed method achieves high accuracy of 89% using VGG19 on validation images. These models can assist doctors in identifying eye disease more precisely, leading to fewer unnecessary biopsies and better patient outcomes. This investigation contributes significantly to the field of medical image analysis advancement, paving the way for the use of deep learning models into ophthalmic diagnostic processes.

Keywords: *Convolutional neural network, Eye disease, Pre-trained model, Transfer learning, Cataract, Glaucoma, VGG19, ResNet50*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AMD	Age related Macular Degeneration
API	Application Programming Interface
CNN	Convolutional Neural Network
CSS	Cascading Style Sheets
DFD	Data Flow Diagram
DL	Deep Learning
FN	False Negative
FP	False Positive
FR	Functional Requirements
GPU	Graphical Processing Unit
GUI	Graphical User Interface
HTML	Hypertext Markup Language
ML	Machine Learning
NFR	Nonfunctional Requirements
NN	Neural Network
ResNet	Residual Network
TN	True Negative
TP	True Positive
VGG	Visual Geometry Group
WHO	World Health Organization

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Introduction

The World Report on visual highlights the serious global difficulties in eye care, with at least 2.2 billion individuals suffering from visual impairment, much of which might have been avoided or is untreated. The study done by the World Health Organization (WHO) emphasizes gaps in the availability and caliber of eye care services, as well as insufficient integration into healthcare systems and a lack of qualified experts. The report, which was created with feedback from specialists around the world, provides information on the frequency of eye disorders as well as suggestions and tactics for enhancing eye care services [1].

Machine and deep learning algorithms are rapidly growing in the dynamic research of medical imaging. Work is being done to improve medical imaging applications by using algorithms designed to find mistakes in disease diagnostic systems, which could lead to the avoidance of unclear medical interventions. Deep learning techniques, in particular, using convolutional networks, have become a specialized methodology for quickly analyzing medical images.

It uses supervised or unsupervised algorithms using some specific standard dataset to indicate the predictions [2].

Deep learning algorithms are a subset of machine learning algorithms, which aim at discovering multiple levels of distributed representations. The Convolutional Neural Networks (CNN) is one of the most notable deep learning approaches where multiple layers are trained robustly. It has been found highly effective and is also the most commonly used in diverse computer vision applications [3].

The eye is a vital sensory organ that allows us to perceive visuals by transmitting signals to the brain, providing organisms with the gift of sight. Their functionality, however, can be compromised by curable diseases. Cataract and Glaucoma are two degenerative eye diseases that can cause vision impairment and even vision loss. If detected and diagnosed at an early stage, both of the diseases can be cured.

Cataract is a medical condition caused by the clouding of lenses. It is primarily caused by protein breakdown in the lens. As cataracts do not affect vision during the early stages, people do not notice any symptoms. With the growth of cataracts, the signs become evident too. The visuals seen by people affected by cataracts have been compared to looking through a frosty or fogged-up window [4].

When fluids accumulate in the front part of the eye, it increases pressure and damages the optic nerve causing glaucoma [5]. Given that the optic nerves are responsible for transmitting signals to the brain, damage caused to it may result in vision loss. While cataract-related blindness can be cured with surgery, the vision loss caused by glaucoma is irreversible [6].

Diabetic retinopathy, another main cause of blindness among the age-related eye disease is caused by diabetes. This highlights how important it is for people with the disease to get frequent eye exams. Blood flow can be impeded by retinal blood vessel damage caused by high blood sugar levels, which can result in edema, leakage, or closure. Furthermore, this illness might promote the growth of asymmetrical new blood vessels in the retina because of the potential for serious vision effects, diabetics should take extra care of their eyes [7].

While most of the projects in this field focus on a single disease, this project aims to develop a model that can classify eye diseases such as cataracts, glaucoma, diabetic retinopathy, and normal eyes. The author’s aim is to aid in the early diagnosis of diseases and enhance the quality of life for those who are affected.

1.2 Problem Statement

The human eye is susceptible to both treatable and incurable diseases like cataracts and glaucoma. They lead to vision impairment, and in some cases, blindness. Both cataracts and glaucoma can be treated if they are detected in the initial stages. However, people are unaware they have these diseases as the initial symptoms are mild. They carry on with their lives, without seeking medical help. Existing AI systems often focus on individual eye diseases in isolation, resulting in fragmented diagnostic workflows. There is a pressing need for a unified AI solution capable of detecting multiple eye conditions simultaneously, thereby enhancing clinical efficiency and accuracy. Early detection of eye diseases is crucial for effective treatment. Existing screening methods often lack sensitivity, leading to late diagnoses. Deep learning algorithms can improve early disease detection rates. By utilizing Deep Learning techniques, we aim to develop a tool that is convenient to the users and can identify cataracts and glaucoma at an early stage.

1.3 Objective

The main objective of this project is to build an automated system that identifies multiple eye diseases among the several input retinal images by developing methods and algorithms to quantify and analyze biomedical data and thus build a model. The objectives of building this model are:

- To develop an eye disease detection and classification model that can differentiate between Cataracts, Glaucoma, Diabetic Retinopathy, and NormalEyes.
- To deploy a web application for classifying specified eye diseases using the deep learning algorithms and compare their accuracies.

1.4 Scope and Limitation

Combining partial differential equations (PDEs) with convolutional neural networks (CNNs) offers a powerful and efficient method for automatically identifying and classifying different eye pathologies. These deep learning models demonstrate the potential for extremely accurate and early identification of eye illnesses through lengthy training on large-scale datasets and the use of data augmentation methodologies to boost model resilience. When such technology is used in ophthalmology, it has the potential to revolutionize diagnostic procedures and enable timely intervention and individualized treatment regimens. This could lead to a notable improvement in patient outcomes and advance the field of eye care worldwide. Even so, there are certain shortcomings with the project's approach, they are:

- Additional system improvements will be required for ophthalmologists to trust the system's performance, interpretability, and reliability.

1.5 Methodology

This project's development process is built on an agile machine learning development methodology. The group members split up the overall task. It is possible to lower risk and accelerate completion by breaking the project up into smaller components.At each iteration, the team will work through the complete software development process, which includes requirements analysis, design, coding, testing, and planning.

The algorithm used in the research is based on deep learning. The foundation of deep learning is software. Since it's a distinct form of program, a different approach is required. Despite the fact that the study calls for several back propagations, the project as a whole is built on an agile methodology. To get the most efficiency, preprocessing data in accordance with model performance is necessary.

1.5.1 Procedure of the project

The procedure of the project is as:

- **Data Collection:** Gather a diverse dataset of eye images, including labeled samples of various eye diseases and healthy eyes.
- **Data Preprocessing:** Standardize image sizes, perform data augmentation, and preprocess images to enhance model generalization.
- **Model Development:** Train and fine-tune deep learning models using the collected dataset, experimenting with different architectures and hyper parameters.
- **Validation and Evaluation:** Assess model performance using cross-validation and a separate validation dataset. Optimize models for high sensitivity and specificity.
- **Real-time Application:** Develop a user-friendly web interface for capturing and processing eye images in real time, integrating the trained models.
- **Interpretability:** Use mapping and to visualize and explain the model's decisions.

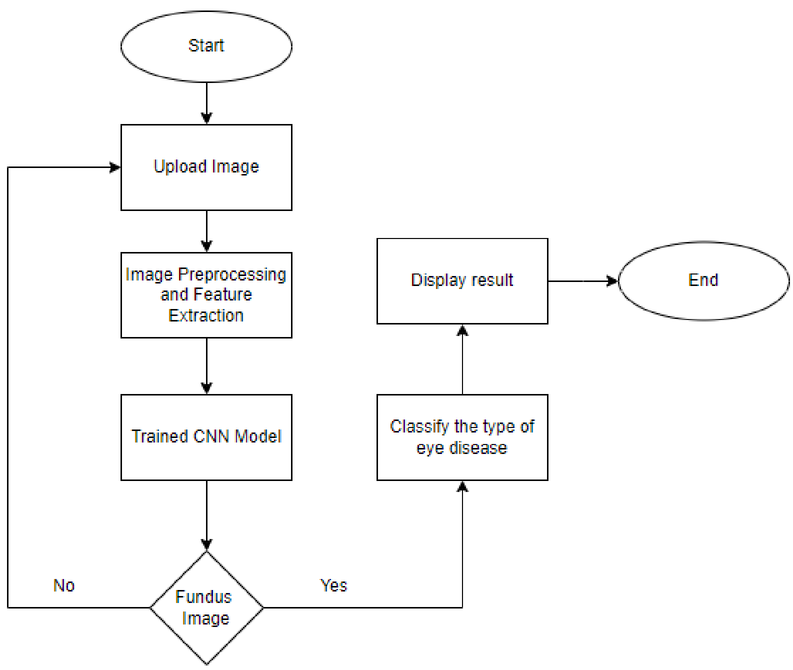


Figure 1: Project Methodology.

1.6 Report Organization

This report is organized in such a way that the introduction, problem statement, objectives, scope, and limits are all included in the first chapter. The project's background and a review of the previous works on the systems in use were covered in the next chapter. System analysis, which includes feasibility and requirement analysis,is covered in the third chapter. Algorithms utilized in the construction of this application are covered in the fourth chapter on system design. In a similar vein, the system'simplementation and testing are covered in the fifth chapter. The sixth chapter includes a conclusion and recommendations for the future. As a result, the report is divided into six chapters that cover every aspect of the system.

CHAPTER 2: LITERATURE REVIEW

2.1 Background Study

2.1.1 Fundus Photography

The retina, macula, optic disc, fovea, and blood vessels are all located in the innermost region of the eye is called fundus. Using a specialized camera that shines light through the pupil, fundus photography takes pictures of this area. Eye experts can efficiently identify, monitor, and treat a variety of eye diseases with the use of these photographs [8].

The neurosensory tissue known as the retina translates the optical images into electrical signals for the brain that can be seen by the means of fundus photography. In order to take a direct snapshot of the retina, this approach entails passing light rays through the pupil. A flash activates when the shutter is released, creating a fundus picture. The ophthalmologists use these images for monitoring, diagnosing, and treating eye problems [9].

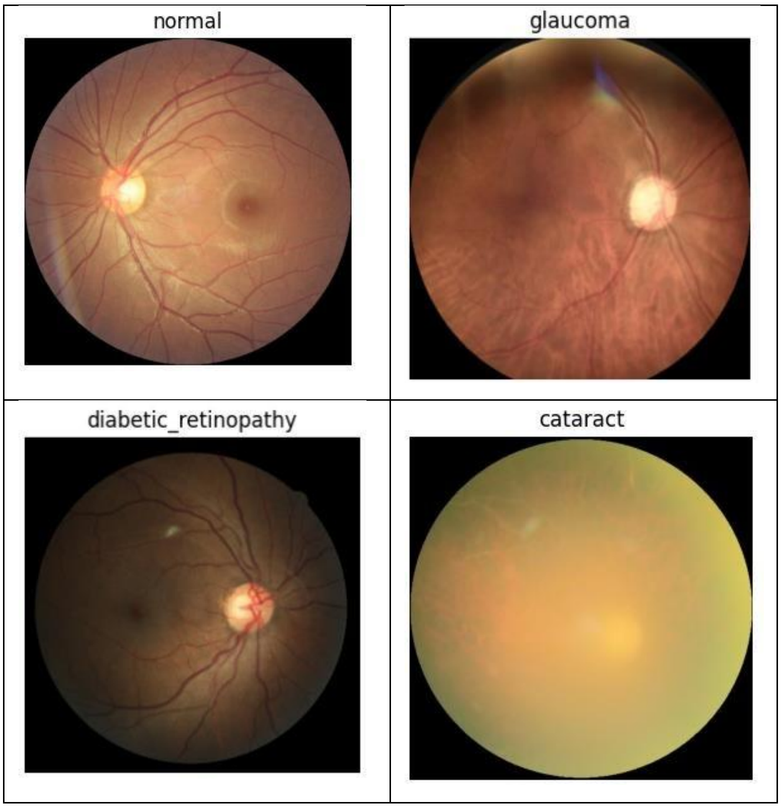


Figure 2: Fundus images showing different diseases

The fundus camera is distinguished by their angle of view, which denotes the lens's optical range. A film image that is 2.5 times larger than life is produced by using a standard angle of view of 30°. With viewing angles ranging from 45° to 140°, wide-angle fundus cameras are able to collect images that cover a larger section of the retina with a lower magnification. Narrow-angle fundus cameras, on the other hand, have a field of vision of no more than 20° [9].

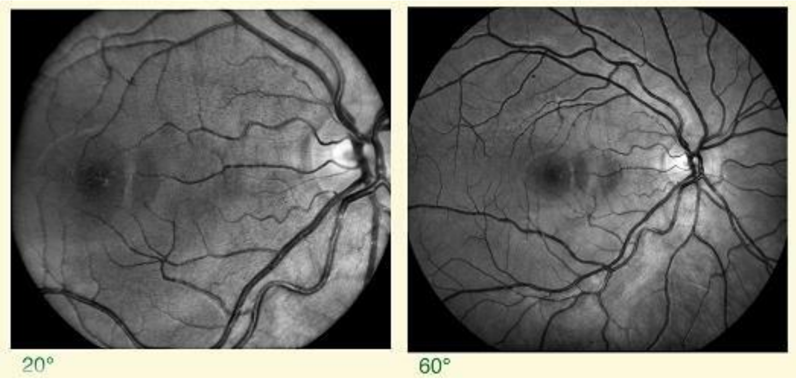


Figure 3: Fundus Photography with different angles of view [8]

2.1.2 Image Preprocessing

Pre-processing is an enhancement of the images that reduces undesired distortions and brings out certain features of the image that are crucial for additional processing. Geometric image transformations (such as rotation, scaling, and translation) are included in this category of pre-processing methods [10].

Building an effective deep learning model needs to carefully consider the input data format. There are several techniques which enhances the pixels of the image, filtering is also done and detection of the blood vessels through these images and comparing it with the recorded data to predict the results according to the statistical data to get a clear vision about the severity of the disease caused via different preprocessing of the images and then integrating them with encoders to convert the data and provide the information in a readable form. There are different methods for preprocessing the retinal photographs of poor quality using deep learning method such as Grayscale conversion, Adaptive Histogram Equalization, Discrete Wavelet Transform, Matched filter and Fuzzy C Means in Segmentation preprocessing [11].

2.1.3 Dataset

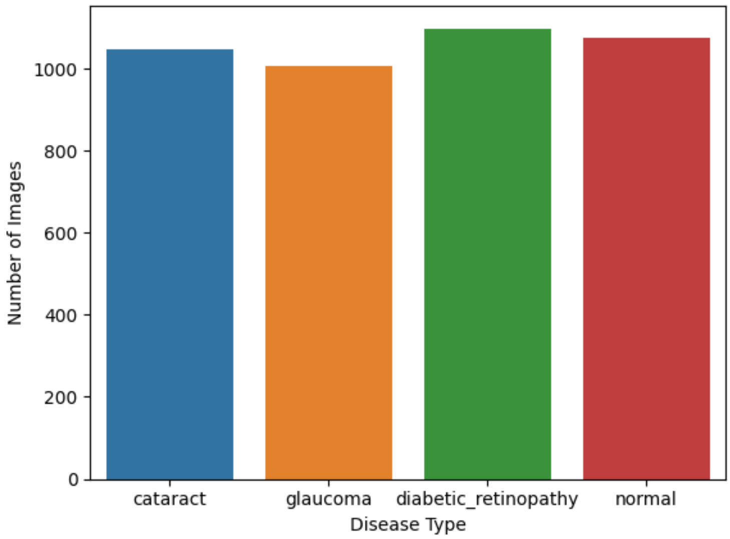
The training and validation dataset utilized in this model, titled "Eye Disease Classification," was obtained from Kaggle This dataset comprises a diverse collection of retinal images categorized into four distinct disease types: Normal, Diabetic Retinopathy, Cataract, and Glaucoma. Each class has approximately 1000 images, contributing to a balanced dataset. The distribution of images across disease types is as follows:

- Diabetic Retinopathy: 1098 images
- Normal: 1074 images
- Cataract: 1038 images
- Glaucoma: 1007 images

These images are sourced from various reputable repositories, including IDRiD (Indian Diabetic Retinopathy Image Dataset), Ocular Recognition, HRF (Human Retina Fundus), and others. The incorporation of images from diverse sources ensures the dataset's richness and relevance in capturing the wide spectrum of eye diseases [12].

The testing dataset was obtained from Kirtipur Eye Hospital and Ophthalmic Study Center of about 20 images that is respective to the training disease type. In 2022, Kirtipur Hospital performed total OPD of 37,524 and total number of performed surgery were 1,560 [13].

Figure 4: Training data distribution among the diseases [12]



2.1.4 Neural Network

Neural network uses interconnected nodes arranged in layers to simulate how the human brain makes decisions which is one of the essential part of artificial intelligence. These nodes, which are also referred to as artificial neurons, analyze data by giving input variables weights and putting them via activation functions. These networks steadily improve accuracy and decrease errors by modifying their weights and biases through supervised learning using labeled datasets. They iteratively optimize their solutions by fine-tuning their parameters through the use of methods like gradient descent and backpropagation.

Deep neural networks can use backpropagation for training, even though they usually have a feedforward architecture. Neural networks are capable of making complex judgments, as demonstrated by real-world applications like image recognition and classification, which are made possible by mathematical ideas like linear regression and activation functions [14].

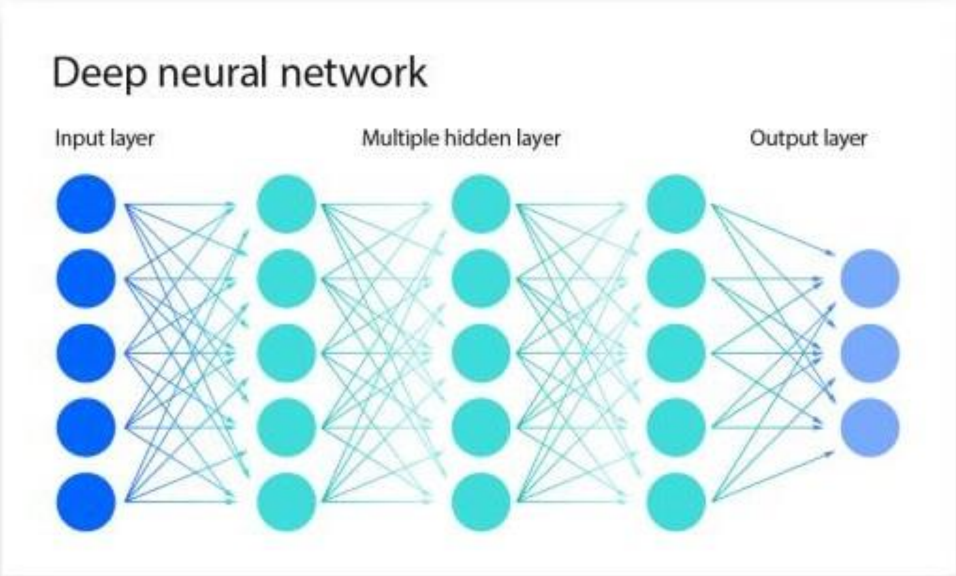


Figure 5: Neural Network architecture [14]

2.1.5 Transfer Learning

In machine learning, transfer learning refers to applying a trained model to a new task. With this method, a machine can improve its capacity to generalize to new issues by using the knowledge it has gained from previous jobs. For example, the knowledge gained during training can be used to identify beverages when a classifier is trained to identify food items in photos. This method that uses information from a trained model to enhance a model's performance on an untrained task. Dealing with limited labeled data or computational resources can make transfer learning very useful. Examples of transfer learning applications are given in a variety of fields, including computer vision and natural languageprocessing. There are different ways to transfer learning, such as feature extraction and fine-tuning. It is crucial to choose the right pre-trained model and fine-tuning technique depending on the particular dataset and situation. Inception-v3, one of the most well-known pre-trained machine learning models, is well-known for having been trained in the

ImageNet "Large Visual Recognition Challenge." Beside pre- trained models like VGG, ResNet also uses transfer learning for feature extraction and classification of diseases [15].

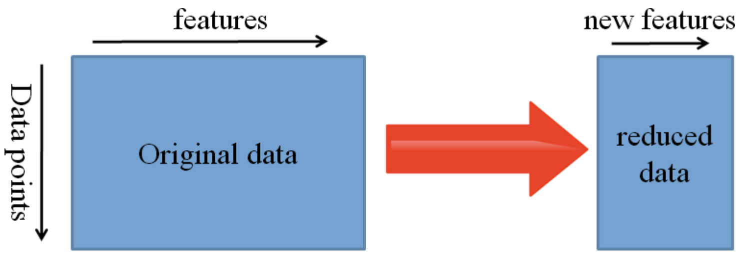


Figure 6: Transfer learning process [15]

2.1.6 Convolutional Neural Network

Convolutional neural networks (CNNs) are among the most popular deep learning techniques, especially for computer vision problems. Convolution layers, pooling layers, and fully connected layers are some of the basic building blocks that CNNs are made of. CNNs are designed to learn the spatial hierarchies of features in grid-patterned data, including photographs. The model's structure is modeled after that of the animal visual cortex. Fully linked layers manage feature mapping in order to classify the features extracted by convolution and pooling layers. CNN's performance in image processing tasks is increased by convolution layers, which use small grid structures called kernels to perform mathematical operations across image positions. Backpropagation and gradient descent are two optimization methods that are used during training to fine-tune parameters like kernels in order to reduce differences between outputs and ground truth labels. More intricate characteristics are gradually revealed by this repeated approach [16].

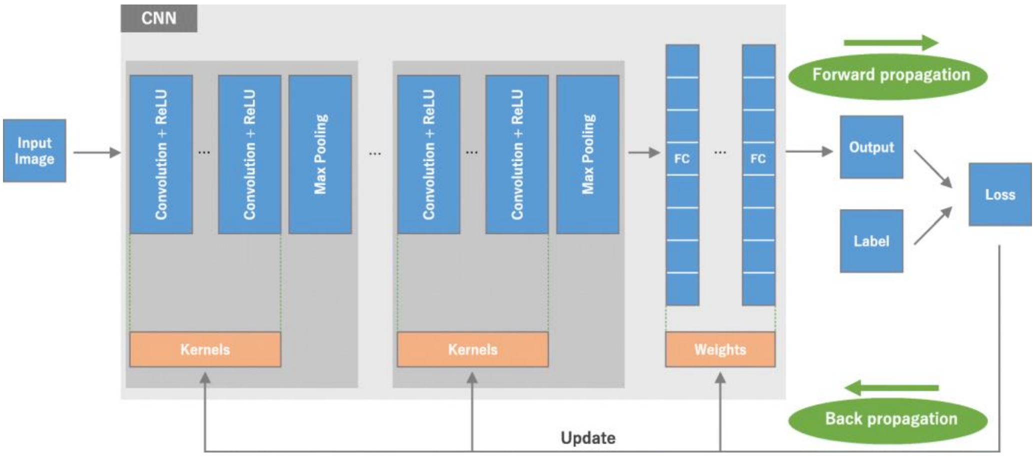


Figure 7: Convolutional Neural Network [16]

I. Convolutional Layer

A convolutional process is a particular kind of linear feature extraction in which a small array of numbers known as a kernel is processed through an input tensor, which is essentially an array of numbers. By calculating the element-wise product between each element of the kernel and the input tensor at each point of the tensor, the output value—also referred to as a feature map is generated at the corresponding position of the output tensor. Next, this product is totaled. This procedure generates a variety of feature maps, each of which reflects a unique property of the input tensors, using a number of kernels. Thus, several kernels can be considered as distinct feature extractors [16].

II. Rectified Linear Unit (ReLU)

An essential component of artificial neural networks is the activation function, which decides when to activate neurons. It establishes a node's output in response to a single or a specified set of inputs. Rectified Linear Unit, or ReLU, is the most commonly used activation function in deep learning. It is defined by the formula $R(x) = \max(0, x)$, where the output is equal to x in non-negative inputs and zero in negative inputs.

The reason for its popularity is because, when compared to other options such as sigmoid or tanh functions, it can accelerate the convergence of gradient descent. Its ability to selectively activate neurons while converting negative inputs to zero prevents excessive simultaneous activation, which simplifies network calculation and is one of its main advantages. Furthermore, ReLU solves the diminishing gradient descent problem and is intrinsically nonlinear because of the nonlinearity introduced by the $\max(0, x)$ operation, even if its definition appears to be linear [17].

III. Pooling Layer

The pooling layer is often used to arrange the layers within the architecture and comes after the convolutional layer. Depending on how the particular requirements of the task at hand change, its inclusion can be made more than once. These pooling layers can be used with popular Deep Learning frameworks such as PyTorch, TensorFlow, and Keras, among others; nevertheless, their efficient use requires a solid understanding of deep learning principles. Max, Min, Average, and Global Pooling are the four primary categories of pooling layers [18].

IV. Fully connected Layer

Fully connected layers (FCLs) in neural networks emphasizes on model building and understanding. FCLs play a crucial role in creating connections between each neuron in one layer and those in the layer below, enabling complex data processing and feature extraction. They are very important in model training and explores ways to improve the effectiveness of model. Along with this, it addresses typical issues with them, like overfitting [19].

2.1.7 Libraries

Libraries are similar to machine learning algorithms; however, they are implemented differently. As opposed to machine learning algorithms, which process machine tools automatically and carry out user-specified tasks, other kinds of software have been developed. Additionally, they carry out various neural networks and operations [20]. The following is a list of some of the libraries that are used in the project:

I. TensorFlow

TensorFlow is an extensive open-source platform especially for machine learning applications created by Google that offers a broad ecosystem that meets a variety of machine learning demands thanks to its adaptable tools and libraries. Because of its multi-level abstractions, users can choose the best strategy for any problem they need to solve. Large-scale machine learning projects are best suited for TensorFlow because of its Distribution Strategy API, which allows distributed training across different hardware configurations without requiring modifications to the model definition. In addition to its adaptability, TensorFlow has interoperability with several contexts, including TensorFlow.js, Extended, and Lite, which makes it possible to create, train, and distribute models across servers, devices, and the internet [21].

II. Keras

TensorFlow's high-level API, Keras, provides a user-friendly and effective interface for addressing machine learning problems, especially those related to contemporary deep learning. Rapid experimentation is prioritized by Keras, which covers every facet of the machine learning workflow, from data preprocessing to hyper parameter adjustment and deployment. With Keras, users may take advantage of GPU clusters, TPUs, and even export models for usage in web browsers or mobile devices which is due to TensorFlow's cross-platform and scalability [22].

III. Scikit-learn

The popular python package Scikit-learn, sometimes known as sklearn, is used for machine learning applications for model selection, dimensionality reduction, clustering, regression, and classification. It offers an extensive range of tools and techniques to increase its utility and adaptability. This library is constructed on top of existing scientific Python libraries like matplotlib, SciPy, and NumPy and is well known for having an intuitive user interface that makes it simple to develop and implement machine learning models [23].

2.2 Literature Review

Number of research has been performed to explore the feasibility of predicting different types of eye disease by examining different machine learning, deep learning, mobile technology as a potential learning system. Most research papers regrading this project are regarding one specific disease and some are multiclass depending on their dataset and objectives.

A deep learning architecture was tested using three-dimensional optical coherence tomography (OCT) images from patients who were sent to an eye hospital in a 2019 study titled "Clinically applicable deep learning for diagnosis and referral in retinal disease." Even with a limited training sample of 14,884 images, the model outperforms skilled doctors in terms of referral recommendations for sight-threatening retinal disorders, according to the authors. Moreover, the model's tissue segmentations serve as a device-neutral representation, guaranteeing accurate referrals across a range of devices. The strategy was evaluated by junior and senior graders who demonstrated a low error rate of 3.4% on crucial referral judgments. This indicates that the method can be applied more broadly across medical imaging techniques. This breakthrough removes earlier obstacles to broader clinical application, with low training data needs and efficacy across different disorders in an actual scenario. This work offers a novel framework that evaluates clinical OCT scans and recommends referrals to clinical specialists [24].

In 2017, research was performed using 494,661 retinal pictures that evaluated a deep learning system (DLS) for the detection of diabetic retinopathy and related eye illnesses. This system demonstrated a high degree of sensitivity and specificity in identifying vision-threatening diabetic retinopathy (AUC 0.958), probable glaucoma (AUC 0.942), referable diabetic retinopathy (AUC 0.936), and age-related macular degeneration (AUC 0.931) after being trained on many datasets collected from 2010 to 2013. Under all settings, sensitivity varied from 90.5% to 100%, while specificities were typically higher than 87%. The system had a strong performance in the first validation dataset, indicating its potential for broad screening in diabetic multiethnic populations. However, the study showed that it is necessary to conduct additional studies to examine how the system can be included in healthcare environments to enhance visual outcomes [25].

A study was performed on a deep learning system that was used to categorize age-related macular degeneration (AMD) phases using color fundus pictures to increase the efficiency of disease staging in 2018. The proposed algorithm performed exceptionally well, outperforming human graders in the Age-Related Eye Disease Study (AREDS) using a large dataset of 120,656 fundus images from study participants. The algorithm was further validated using 5555 images from the Kooperative Gesundheitsforschung in der Region Augsburg (KORA) study. Identifying 84.2% of AMD-positive photos and properly categorizing 94.3% of healthy images, the ensemble of six convolutional neural net designs achieved a quadratic weight of 92%. The study presents how the algorithm might help diagnose AMD earlier, especially in adults [26].

The authors used ImageNet-trained CNNs to assess glaucoma using Fundus images. Architectures such as VGG16, VGG19, InceptionV3, ResNet50, and Xception were employed on databases consisting of a total of 1707 images. Validation strategies such as Cross-Validation and Cross-testing were used. 10-fold cross-validation was performed to obtain the F-score along with accuracy and sensitivity, the average along with standard deviation was determined for all of the architectures. Based on the AUC and CNN parameters, Xception classified glaucoma most accurately. Additionally, the authors have introduced a new clinical database named ‘ACRIMA’. The 705 images in the database including both normal and glaucomatous images are labeled [27].

In 2021, a review paper was proposed by a group of researchers where they discussed the application of deep learning techniques for the detection and classification of diabetic retinopathy (DR) using fundus pictures. This paper compiles research on datasets used for DR detection and classification, deep learning methods, and classification models. For training, testing, and segmentation, the study makes use of retinal fundus picture datasets from sources including Kaggle EyePACS, Kaggle APTOS 2019, Messidor, IDRiD, and DDR, among others, and uses preprocessing methods like normalization, denoising, and contrast enhancement to improve the model's performance. Several deep learning architectures were discussed and it found that the 5-class dataset utilizing Ensemble architecture for the Healthy vs. Diseased category had the greatest accuracy among any classification model with an accuracy of 97.67% and the highest segmentation accuracy was 99.97% [28].

One of the drawbacks in the detection of glaucoma with the use of fundus images is redundancy. As the attention mechanism has the potential to enhance CNN, the authors have proposed an attention-based convolutional neural network (AG-CNN). In addition to establishing a LAG (large-scale attention-based glaucoma) database, human attention maps from ophthalmologists collected through a simulated eye-tracking experiment were integrated into the LAG database. The AG-CNN architecture comprises three subnets. The Attention prediction subnet generates attention maps for fundus images which plays a pivotal role in reducing redundancy. Consisting of both convolutional layers and fully connected layers, the Pathological area localization subnet, uses guided backpropagation to obtain a Visualization map of the pathological area. Used for the binary classification of fundus images, the Glaucoma classification subnet determines if the case is of positive glaucoma or negative glaucoma. With an accuracy of 95.3%, the deep learning model, AG-CNN, demonstrates a notable advancement in glaucoma detection [29].

In 2021, a group of authors created a dataset called Retinal Fundus Multi-Disease Image Dataset (RFMiD), which includes 3200 retinal fundus images that display a variety of pathologies such as retinitis pigmentosa, macular holes, cotton-wool spots, and more. The dataset is accessible for download and is meant for ophthalmology research and diagnostic applications. RFMiD, which is recorded by experienced retinal specialists with three distinct fundus cameras, intends to assist research on the detection of multiple diseases and the creation of broadly applicable models for retinal screening. This paper provides insight regarding different types of pathologies and also provides multiple datasets for further research [30].

A research entitled “Automated detection of mild and multi-class diabetic eye diseases using deep learning” was performed in 2020. This paper addressed the problem of automating the identification of diabetic eye disorders (DED), which include cataracts, diabetic macular edema, glaucoma, and diabetic retinopathy. The goal of the paper is to develop an automated classification system that can recognize both mild and multi-class cases of DED using retinal fundus images. A variety of datasets were pre-trained on convolutional neural network (CNN) models, such as VGG16 and InceptionV3, in conjunction with optimization and fine-tuning methods. The paper creates an automated system that complies with the British Diabetic Association's DED detection guidelines, which is a major advancement in expediting the identification of diabetic eye disorders. The model then compares several CNN architectures and optimization methodologies where the VGG16 model obtains a maximum accuracy of 88.3% for multi-class classification and 85.95% for moderate multi-class classification [31].

A paper on “Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network” in 2020 proposes an automated classification model that can identify diverse ophthalmological disorders using fundus images to prevent blindness by early disease diagnosis. The dataset was obtained from the Ocular Disease Intelligent Recognition (ODIR) database which includes pictures of patients with various ocular disorders. The paper emphasizes multi-class multi-label categorization using Convolutional neural networks (CNNs). The models are pre-trained and refined using the ODIR database, and performance measures are assessed. Out of four pre-trained CNN architectures, the VGG16 model with an SGD optimizer setup produces the best results. This research provides us an overview of how crucial it is to handle fundus images separately to make accurate predictions and offers directions for future studies, such as incorporating explainable AI strategies and investigating cutting-edge data augmentation techniques to improve the model's performance and interpretability [32].

Using the UCI machine learning repository's Diabetic Retinopathy Debrecen dataset, a study was proposed to improve the identification of diabetic retinopathy by combining machine learning and deep learning methods. To reduce dimensionality in the dataset, a

Deep Neural Network (DNN) is used in conjunction with Principal Component Analysis (PCA), and the Harris Hawks Optimization (HHO) algorithm is used to optimize the feature extraction and classification processes. The hybrid model DNN-PCA-HHO outperforms with an accuracy of 97%. In the study, the performance of the algorithm is enhanced while training time is decreased when PCA and HHO are combined. The research recognizes the possible drawbacks of overfitting with low-dimensional datasets which may limit the model's functionality. Regardless, the outcomes show a notable gain over current systems. This study values early identification of diabetic retinopathy and suggests more research in related health fields and high-dimensional data [33].

The study entitled "Robust optic disc and cup segmentation with deep learning forglaucoma detection" addresses the crucial problem of early glaucoma detection using a modified U-Net architecture. This paper presents a strong segmentation method for distinguishing between the optic disc and cup and also combines traditional U-Net decoding layers with the ResNet-34 model as encoding layers. The model was trained usingthe RIGA dataset and performs on par with the state-of-the-art techniques having remarkable dice values of 97.31% for disc segmentation and 87.61% for cup segmentation.This study attains state-of-the-art results by fine-tuning two more databases, the average disc dice values for DRISHTI-GS and RIM-ONE are 97.38% and 96.10%, respectively, while the average cup dice values are 88.77% and 84.45%, respectively. The proposed method's main benefit is it uses pre-trained ResNet and U-Net that allow for quick trainingand reliable results, helping in early detection of glaucoma [34].

In 2022, a paper entitled "Optimized convolution neural network based multiple eye disease detection" was proposed that addresses the need for early diagnosis of age-related eye diseases. This paper focuses on using maximum entropy transformation to improve the quality and data extraction of retinal fundus images that are obtained from online sources. To classify diseases, the study intends to develop a convolutional neural network (CNN) specifically designed for feature extraction through the use of a flower pollination optimization algorithm (FPOA) in conjunction with a Multiclass Support Vector Machine (MSVM) classifier. The Ocular Disease Intelligent Recognition (ODIR) dataset is used to evaluate the proposed CNN-based multiple disease detection (CNN-MDD) model. The paper shows a 7.5% increase in validation accuracy. The rates of 98.30%, 95.27%, 95.21%, and 93.3% were achieved by precision, accuracy, specificity, recall, and F1 score, respectively. According to the study's findings, this technique helps with the automatic diagnosis of a variety of eye conditions and provides timely treatment [35].

For the growth of Automatic clinical diagnosis, the authors have proposed a hybrid model dependent on deep neural networks (DNNs) and support vector machines (SVM). In addition to integrating U-Net for image segmentation, SVM, and Principal Component Analysis (PCA) are used for classification. The EyeNet data set was initially introduced and the proposed model was trained on it. For the segmentation of images, a modified version of U-Net was introduced. The reduced copy and crop process proved to be satisfactory for small-size images. To avoid overfitting Principal component analysis (PCA) is combined with SVM classifier. While 70% of EyeNet was divided for training, 10% was used for validation, and the rest of 20% was used in the testing process. Before being classified by the SVM, the data that is to be trained go through PCA. With an accuracy of 89.73%, the model outperforms other pre-trained DNN models [36].

A study was conducted in 2022 where the effectiveness of mobile health applications for disease screening in the adult population was analyzed. Due to the lack of traditional healthcare services during COVID-19, the authors focused on the user experience and integration of artificial intelligence in smartphones that would be cost-effective and could screen chronic diseases like diabetes through retinal images, cardiovascular disease, and cancer. The authors utilized data from articles published in various databases and followed PRISMA guidelines to examine chronic diseases. Although the sample size and generalizability are limited, the study offers valuable insight into improving healthcare outcomes through mobile technology [37].

A study was conducted in 2021 to create a model that could forecast a patient's visual field (VF) inside the center 10 degrees if they have glaucoma. This research was carried out by using optical coherence tomography (OCT) pictures to train a convolutional neural network (CNN), and then using data from the Humphrey Field Analyzer (HFA) 24-2 test to fine-tune the predictions. The study took place at several Japanese medical institutes and used a dataset made up of 648 eyes from 358 participants, including both normal eyes and those suffering from primary open-angle glaucoma (POAG). Data on macular retinal thickness determined by OCT was used to train two CNN models: VGG19 and ResNet152. The models' mean absolute error (MAE) ranged from 9.4 to 9.5 dB at first then the MAE dropped to an average of 5.5 dB after correcting the predictions using the HFA 24-2 test data. According to the study's findings, OCT pictures and a trained CNN model can forecast HFA 10-2 test outcomes, with modifications made in response to HFA 24-2 test data [38].

With the high incidence of cataract cases in the United States, a study was carried out in 2022 to create an effective smartphone app for cataract detection. The proposed method comprised taking pictures of the eyes with different smartphone models and analyzing them to extract pertinent characteristics that could be signs of cataracts. The photos were preprocessed using techniques like watershed transformation and median filtering to get rid of background objects and noise and characteristics from the lens picture were extracted using a unique luminance transformation approach. Support Vector Machine (SVM) was used for classification which identified cataracts with a sensitivity of 93.75% and specificity of 93.4%, yielding a diagnosis accuracy of 96.6%. The proposed methodology for cataract detection is quick, easy to use, and reasonably priced, which makes it appropriate for usage in distant locations with limited medical resources. This study offers a viable path for telemedicine and remote/at-home monitoring applications in cataract identification [39].

As image quality is pivotal for the classifier's performance, selecting high-quality fundus images is important. The authors use NIQE and PIQE to examine image quality. A threshold point (T) between 5 and 50, based on NIQE and PIQE scores, serves as a judicious criterion for image selection. The selected images are preprocessed for quality improvement. After resizing, green channel extraction, and normalization are applied, the images are converted into frequency spectrograms by 2D DFT for detection and grading of cataracts. The CNN architecture proposed by the authors has Convolutional layers (16,16, 32, 64, 128, 256 filters) which use 3×3 kernels with the same padding. backpropagation and optimization functions are also implemented to minimize cross-entropy loss. Inaddition to its lesser computation time, the proposed architecture has greater accuracy thanother pertained models. The transformative application of 2D DFT to convert preprocessedimages into frequency spectrograms emerges as a distinctive methodology for cataract detection and grading [40].

In 2023, a study proposed a Convolutional Neural Network(CNN) as a method to overcome a drawback of U-Net, which is high memory consumption. The dataset used in the proposed model is of EyeNet featuring 32 types of retinal diseases. The model architecture not only involves ten convolution layers but also undergoes a sequential process. Initially, low-level feature extraction is performed and then moved onto middle and high-level features for effective classification. Compared to other models like Inception3, and AlexNet, the proposed CNN model minimizes the number of layers, contributing to the reduction of training time. To avoid overfitting and enhance the diversity of the dataset, data augmentation was also performed. Saturation, grayscale conversion, flipping, and brightening were part of a six-step process. While the validation loss varied while trained in 10 epochs and 15 epochs, the validation accuracy of 95% remains constant. In addition to achieving high accuracy, the proposed model consumes less memory and can operate with hardware limitations [41].

With the use of deep learning algorithms, this study entitled "Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach" attempts to establish an automatic diagnostic tool for glaucoma, addressing the important requirement for early diagnosis. Four publicly available datasets were used in the study and to improve model resilience and enrich the dataset, preprocessing involved turning training photos into grayscale and using image preprocessing methods. An accuracy of 98.48% was produced on the G1020 dataset. Several datasets were used for testing the model throughout the evaluation to confirm its efficiency and generalizability. The study demonstrates how deep learning algorithms may be used to automatically diagnose glaucoma from retinal scans, with a high degree of efficiency and accuracy [42].

The study entitled "Eye Disease Identification using Deep Learning" addresses the growing prevalence of ocular disorders in the diabetic population and the absence of specialized early detection apps. This paper presents an automated algorithm that analyses observable symptoms to identify eye disorders through the use of deep learning techniques. Four eye conditions; crossed eyes, bulging eyes, cataracts, uveitis, and conjunctivitis are classified by the model. The authors used internet resources to collect the dataset and employed convolutional neural network (CNN) architecture with three convolutional layers in conjunction with digital image processing techniques like segmentation and morphology. Two separate models were introduced to train the dataset where one for photographs of a single eye and another for photos of both eyes. One of the models concentrates on illnesses like crossed eyes and protruding eyes, the former forecasts ailments like cataracts and conjunctivitis/uveitis showed a precision rate of 96% and 92.31%, respectively. The study's findings point to the possible creation of a mobile application and web platform that uses a customized deep-learning model to identify diseases from uploaded eye photos [43].

CHAPTER 3: SYSTEM ANALYSIS

3.1 System Analysis

System analysis is a systematic process that seeks to identify and address issues within a system in order to achieve specific objectives. It collects and evaluates data to find problems and break the system down into its constituent parts. Every component needsto be closely inspected in order to understand the system's construction and functioning. Understanding the system's needs and environment is the main focus of this crucial stage in the software development life cycle. Eventually, system analysis entails disassembling the system into its constituent elements and understanding how they interact in order to accomplish the system's overall goal [44].

3.1.1 Requirement analysis

One of the first things that should be done while developing a program is requirement analysis. Typically, it involved analyzing the current system, gathering data, determining the hardware and software requirements, and so on. It is the process of outlining what users expect from newly developed or changed software. It is sometimes referred to informally in software engineering as requirements gathering or requirements capturing [45]. The basic hardware and software requirements required for this project are:

Hardware Requirements

- A minimum of 7th generation (Intel Core i7 processor)
- 4 GB RAM or Higher
- Internet Connection
- Input device: Keyboard, Mouse
- Output device: Monitor

Software Requirements

- Code Editor: Google Colab (GPU), VS Code
- Web Browser: Chrome/Edge/Firefox/Opera etc.

The functional and non-functional requirements of the project are discussed as below:

I. Functional Requirements

A Functional Requirement (FR) is a description of the services that the software must provide. It provides an explanation of a software system or its parts. All that a function consists of are the inputs, outputs, and behavior of the software system. It can be any particular feature that indicates what kind of task a system is expected to carry out, such as a calculation, data manipulation, business process, user interaction, etc [46].

In this study, the functional requirement of the system is to detect and classify the different types of eye disease when a image is provided by the user and identify whether there is a presence of the disease of the image is normal. Product features or functions must be implemented by developers in order for users to complete their duties. For the development team as well as the stakeholders, it is crucial to make them apparent. Thus, the functional requirements are:

1. A web interface, where users or opthomologists can provide their fundus image file.
2. An eye disease detection and classification system.

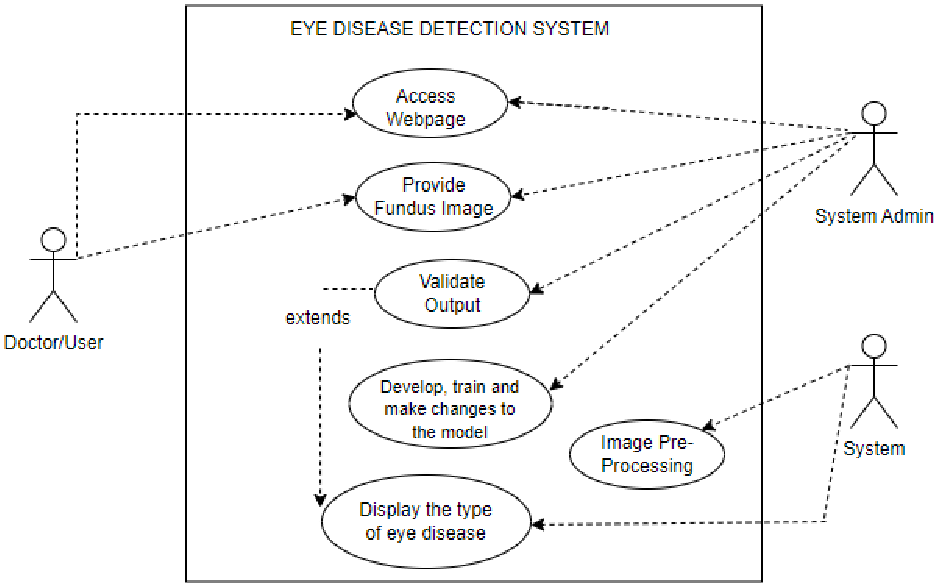


Figure 8: Use Case Diagram for Eye Disease Detection System

The use case diagram illustrates the interaction between system components and the actor in the system and shows how the system responds to the user input.

- **User:** User can access the system through the webpage where they are allowed to upload the fundus images of their eyes.
- **System:** System manipulates the input fundus image that is compatible for the model training.
- **System Admin:** Admin develops and trains the CNN model to detect the type of eye disease as well as validates the output given by the system.

II. Non-Functional Requirements

System characteristics including security, dependability, performance, maintainability, scalability, and usability are defined by Nonfunctional Requirements, or NFRs. They act as limitations or restrictions on how the system is designed across the various backlogs. They guarantee the overall efficacy and usability of the system [47]. The non-functional requirements of the system consist of performance, maintenance, and reliability. The application should identify the provided images in the minimum time possible. Similarly, the system should be maintained so that the cost incurred is less and the classification should be accurate and optimal.

i. Feasibility Analysis

- **Technical Feasibility**

  The evaluation of technical feasibility involves assessing whether it's possible to address a technological challenge and if the proposed solution aligns with the project's goals and constraints [48]. The project is technically feasible; compiles with current technology, including both hardware and software. The computer application is supported by almost all latest computers with minimum hardware and software requirements.

- **Economic Feasibility**

  From an economic perspective, feasibility studies are inquiries conducted to determine whether a product development endeavor would generate profits and be sustainable for a company [48]. The requirement for this project is only of a computer system with working camera or functionality to upload an image and an internet connection with access to the webpage. This will be feasible if a computer system is currently present as other components are not required. So, it is economically feasible.

- **Operational Feasibility**

  Operational feasibility assesses the alignment of a proposed system with the system requirements identified during the requirements analysis phase and its ability to effectively tackle the challenges and capitalize on the opportunities identified in the scope definition and problem analysis stages [49]. The project application can solve the problem of traditional approaches to eye disease detection with better computability by using new model. The use of the algorithm can leverage the task of eye disease detection using various pre-built libraries and models. So, it is operationally feasible.

- **Schedule Feasibility**

  To gauge the feasibility and achievability of a project schedule, it is essential to consider the specific project deadlines and constraints associated with it that is considered to be schedule feasibility [49]. A project's schedule feasibility is assessed as high if it has a high possibility of being finished on time. This project will be completed in schedule to the submitted project proposal. Hence, this system is schedule feasible.

- **Legal Feasibility**

  The implementation of this project does not violate any rules or standards defined by the government of Nepal or any principles internationally. The concepts implemented and literatures defined are noted with their respective references. The project does not violate any national or international laws and is legally feasible.

Figure 9: Gantt Chart

3.1.3 Analysis

I. Data Modeling

The study and characterization of different data types and their interactions inside are important steps in the data modeling process. Data modeling develops visual representations of how data is collected, stored, and used through the use of text, symbols, and diagrams. Organizations may streamline team collaboration, find areas for process optimization, and allocate resources more wisely by recording data kinds, consumption, and management requirements. Improved data retrieval and analytics performance, less errors, and the capacity to establish and monitor KPIs that are in line with corporate goals are all advantages of data modeling, which eventually results in more effective and efficient operations [50].

- **Entity-Relationship Diagram**

An entity relationship (ER) diagram is a type of diagram that displays the relationships between various entities in a system, which can be people, objects, or concepts. ER diagrams expressly focus on the relationships between entities themselves and are commonly used in conjunction with data flow diagrams (DFDs), which illustrate the flow of information between processes or systems [51].

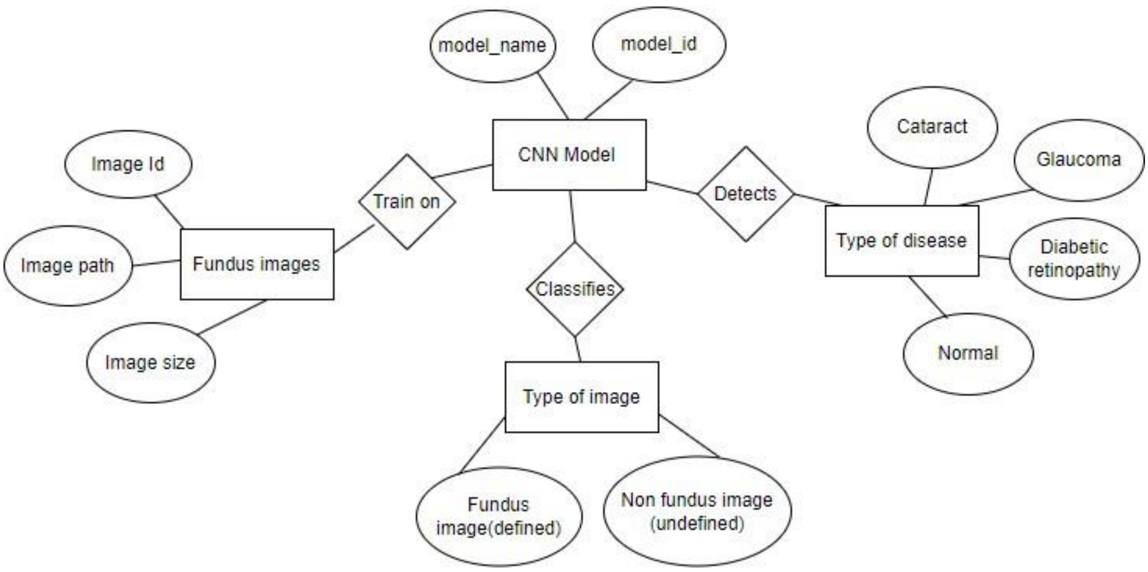


Figure 10: ER diagram

The figure above shows the entity relationship of the proposed model. The entities involved in the system are:

- **CNN model:** The model provided (VGG19 and ResNet50) trains on the dataset provided.
- **Type of image:** The trained model detects if the provided input image is fundus image or not.
- **Fundus image (Dataset):** The dataset contains the image path to the images on which the model is supposed to be trained on. The dimension of the dataset is 224x224.
- **Type of disease:** The model if classifies the dataset is fundus or not then if the provided image is fundus. The image is classified on the basis on the class that are cataract, glaucoma, diabetic retinopathy or normal.

II. Process Modeling

The graphical representation or workflows is called process modeling. Similar to a flow chart, the process's component steps are depicted to provide an end-to-end picture of the tasks involved in the project environment. With the use of a process model, projects can visualize their internal processes and improve their understanding, management, and efficiency. Usually, this is a flexible exercise for ongoing development. Since tasks must be defined and the workflow optimized before it can be automated, process modeling is an essential part of process automation [52].

- **Data Flow Diagrams**

The visual representation of data flow within an information system, or any process or system, is provided by a data-flow diagram. It not only shows the data flow but also lists all of the inputs and outputs for every entity including the process itself involved. Data-flow diagrams, in contrast to flowcharts, do not have control flow components like loops or decision rules. Rather, they only pay attention to data movement representation [53].

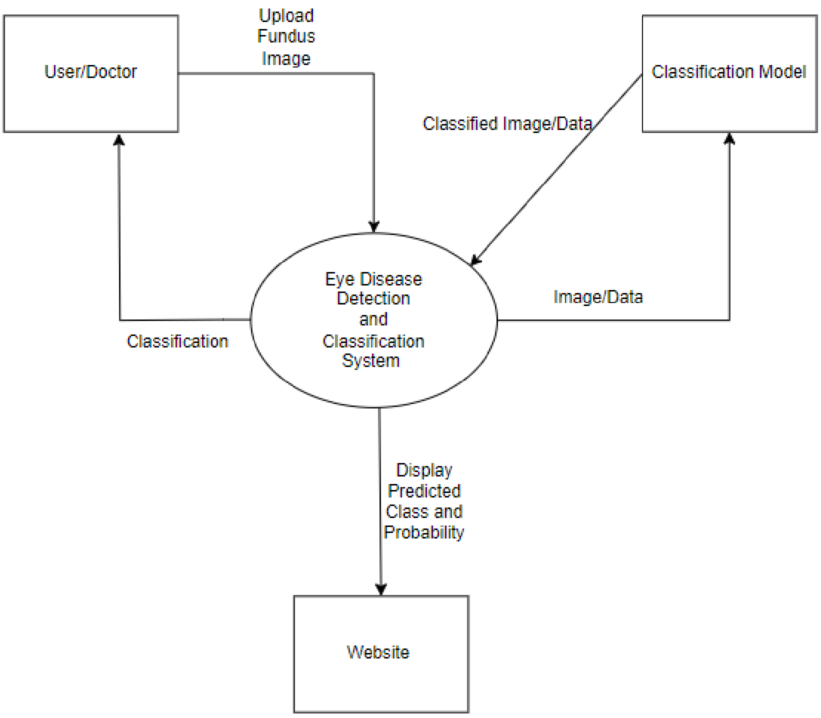


Figure 11: Context Diagram

Above figure represents the context diagram of Ocular Disease Detection and Classification System. It is also called a DFD Level 0 diagram and shows the basic overview of the whole system. It shows the system as a single high-level process. There are basically 3 entities that interact with the system. They are as follows:

- **User/Doctor:** The user uploads the input image into the system and receives a classification based on the model
- **Model:** The system provides the trained model with the input image provided by the user. Then the model classifies and detects the input image based on the trained dataset.
- **Website:** This is an interface between the user and the model so this displays the predicted class of the input image from the model.

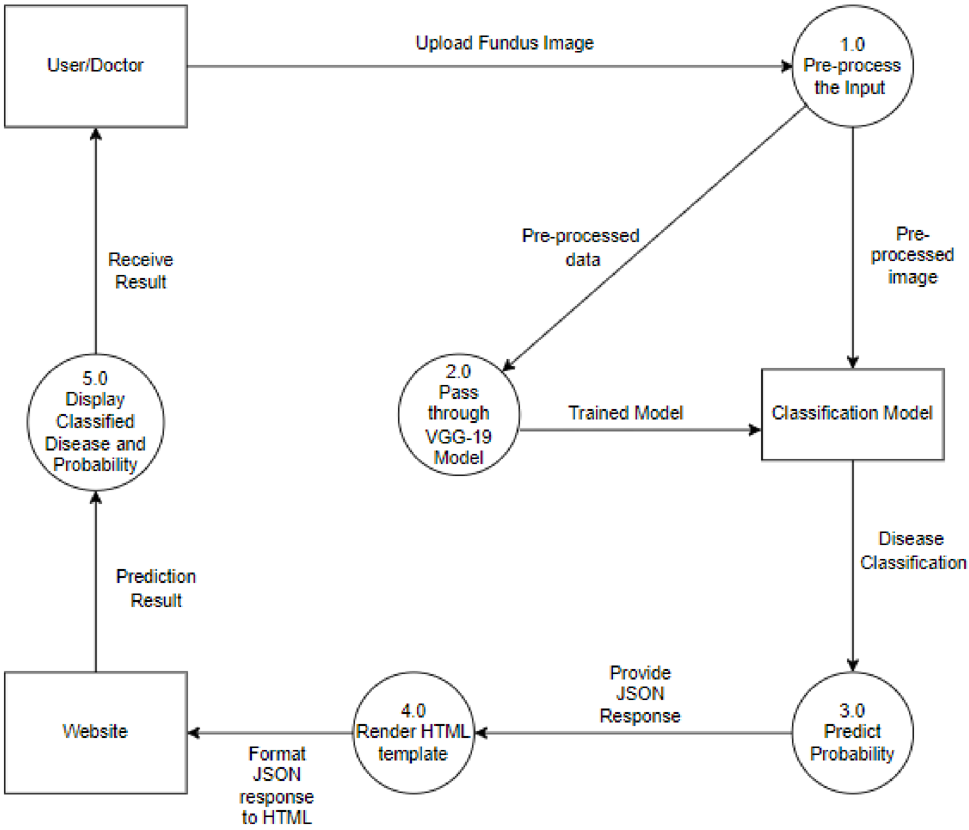


Figure 12: DFD Level 1

The figure above represents the more detailed approach of the context level diagram. It displays all of the system's operations as well as how the entities and processes interact with one another. The level 1 DFD of the proposed system consists of 3 entities: User/Doctor, Classification Model and Website with 5 processes. They are:

- **Pre-process the input:** After the user uploads the input image, the input is preprocessed using augmentation and preprocessing methods such as resizing, rotation, rescaling of the image.
- **Pass through the model:** The preprocessed input is passed through the best model with highest accuracy for further classification.
- **Predict probability:** The classification model classifies the disease based on the training dataset or on the parameters that they are trained on and predicts the probability of the disease on which it has been classified.
- **Render HTML template:** After the model classifies and detects the type of disease, JSON response is sent to render the template that formats the JSON format into HTML format.

- **Display classified disease and probability:** After the conversion into HTML format, the website displays the classified disease with probability of that particular disease.

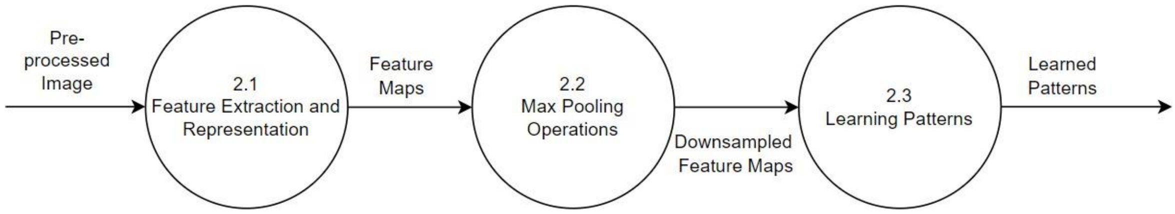


Figure 13: DFD Level 2

The figure above is the extended version of the DFD level 1 where the feature extraction process is shown. This level has 3 processes that finally leads to the pattern for the model to be trained on. The feature is extracted from the preprocessed image and represented. The represented feature is then mapped to perform max pooling operations. The max pooling operation downsizes the input dataset and trains model to learn the pattern. The learned pattern is passed to the model for correct classification and detection of the project.

CHAPTER 4: SYSTEM DESIGN

4.1 System Design

Establishing a system's requirements, such as its structure, connections, components, and layout, is known as system design. It includes every step of the system development process, from outlining requirements to delivering the finished product. System development is the process of creating, enhancing, or changing a system while making necessary adjustments to procedures, methods, and techniques. Effective management of the design methodology and system prerequisites requires a methodical approach. Physical design, which describes input and output processes, and logical design, which depicts conceptual data flow, are the two primary components of the process that are usually involved. Determining components, architecture, modules, interfaces, data, and other factors to satisfy certain organizational or commercial objectives and requirements is known as systems design [54].

4.1.1 System Flowchart

A system flowchart illustrates how decisions impact the data flow and how various system components interact to function as a whole [55].

The figure below shows the system flowchart of the proposed model. When the user input the image, the system loads the image, preprocesses the image and extracts the features of the image which the trained on the model and detected if the input image is fundus image or not. If the image is not fundus image, the result is undefined. If the image is a fundus image, the model classifies and displays the result in the user interface.

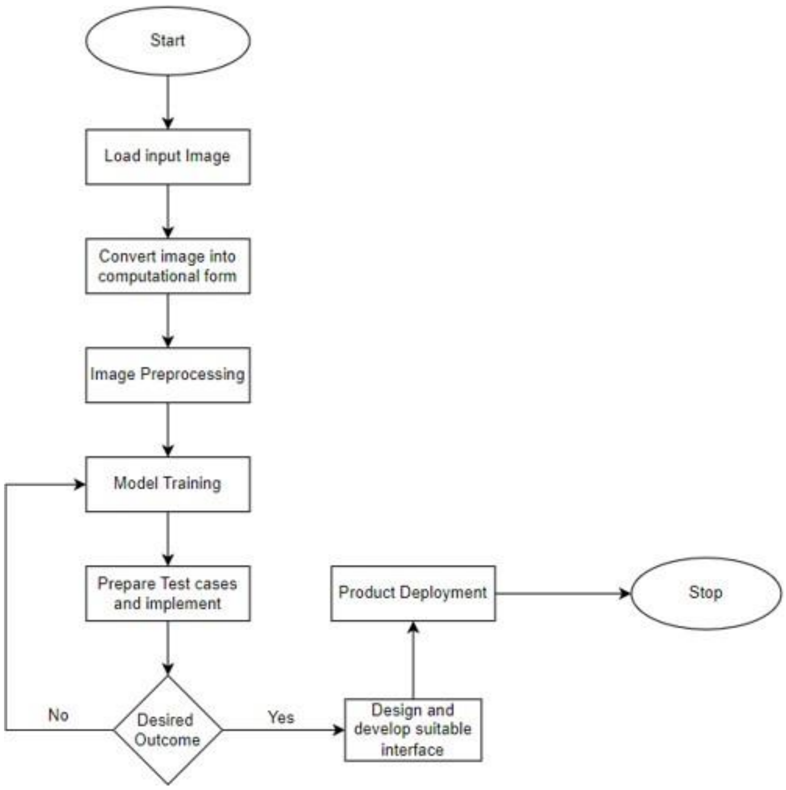


Figure 14: System Flowchart

4.1.2 Interface Design

Through the web interface, users can upload fundus images to detect and classify the diseases. Users can learn about the different diseases through the slideshow too.
Implemented through flask, the frontend is built with HTML, CSS, and JavaScript.

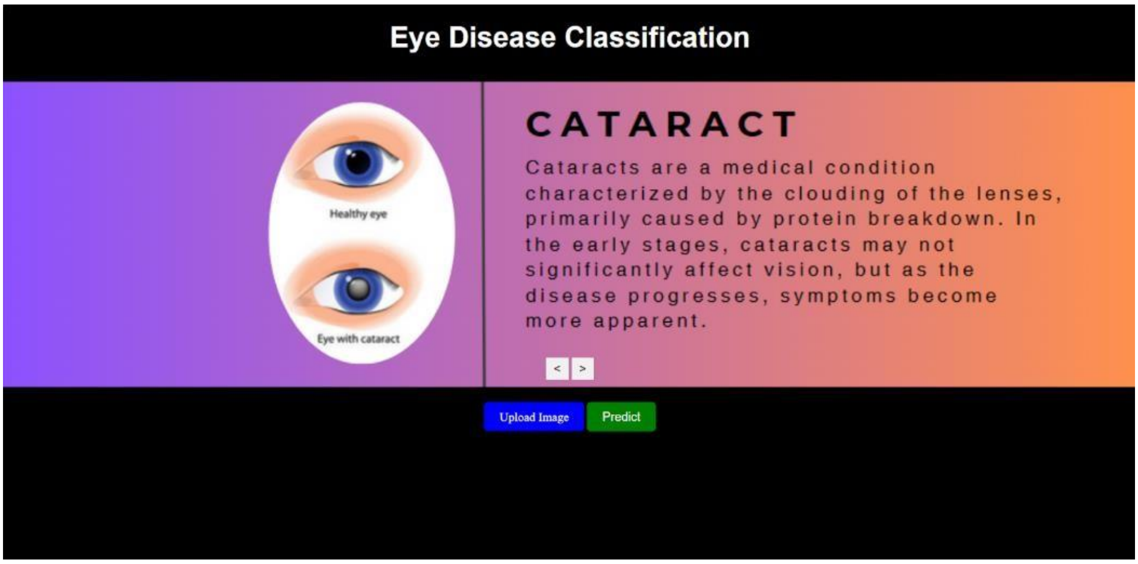


Figure 15: Index Page Design

4.2 Algorithm Details

4.2.1 System Algorithm

I. Convolutional Neural Network (CNN)

Convolutional neural networks (CNNs) are a particular class of deep learning algorithms intended for image processing and recognition applications. Due to their ability to automatically extract hierarchical feature representations from raw input images, CNNs require less setup than other classification models. They set themselves apart by identifying significant objects and characteristics in images using convolutional layers, which employ filters to find local patterns. CNNs are modeled after the connection system of the visual cortex in the human brain and are very good at capturing spatial relationships and patterns in images. By arranging numerous convolutional and pooling layers, CNNs are able to comprehend complex features, improving accuracy in tasks like segmentation, object detection, and picture categorization.

This neural network implements the principles of algebra; the deep learning models extracts features from images such that they can be identified correctly [56].

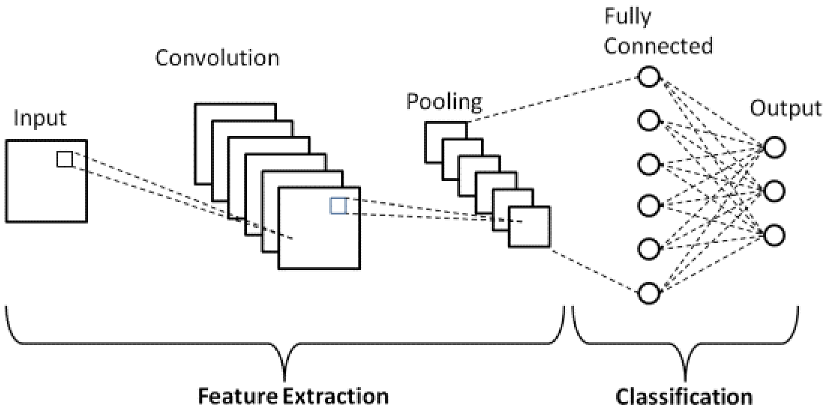


Figure 16: Basic CNN architecture [57]

- **Convolutional Layer**

Convolutional Layer is the fundamental component of the CNN. By utilizing small filters also known as kernel, it is able to scan images and identify specific characteristics. The filter creates feature maps that represent identified characteristics like edges or textures by computing dot products with pixel values as it passes through the image.

It's crucial to set up three critical hyper parameters that affect the output volume size before training a neural network. First, the depth of the output is determined by the number of filters used, each of which creates a unique feature map. Second, a smaller output is produced by a larger stride value, which indicates the distance the kernel travels across the input matrix. Last but not least, there are three types of zero-padding that can be used when filters don't fit the input image: full padding (which increases output size by adding zeros to the input border), same padding (which ensures output size matches input), and valid padding (dropping the last convolution if dimensions don't align). These hyper parameters have a significant impact on the architecture of the network and ultimately its performance [58].

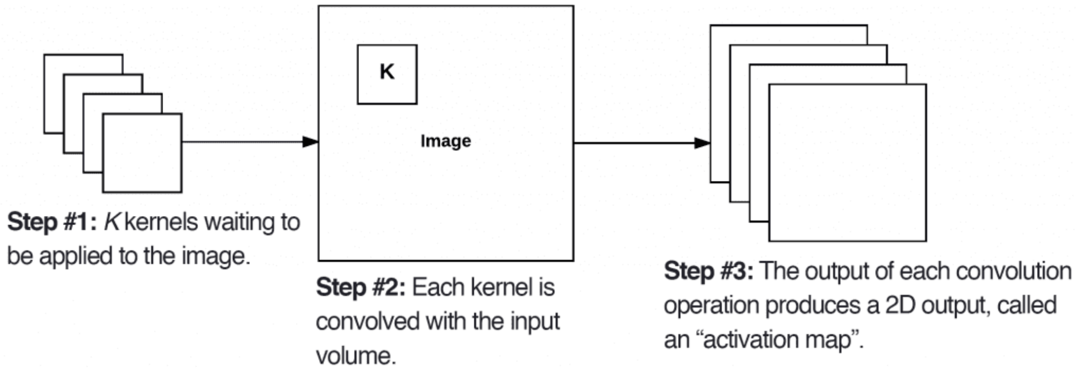


Figure 17: Convolutional Layer Steps [59]

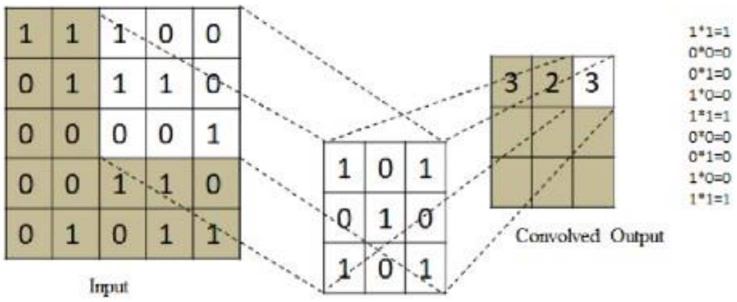


Figure 18: Kernel with size of 3*3 [59]

- **Pooling Layer**

The pooling layer reduces input dimensionality but retains vital information. Down sampling is commonly used in pooling, such as average pooling, which use averages, or max pooling, which keeps maximum values within a frame. Through the reduction of parameters and computations, down sampling promotes efficiency and generality.

Max pooling and average pooling are two popular pooling techniques used in neural networks. By choosing the highest value from each pool, max pooling preserves the feature map's most noticeable elements and produces an image that is crisper than the original.

Conversely, average pooling smooths the image while preserving its key characteristics by calculating the average value of each pool while keeping the feature map's average features [61].

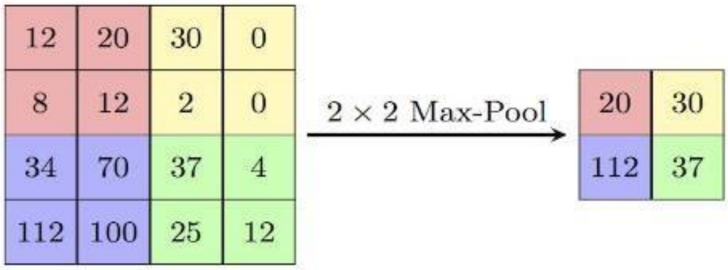


Figure 19: 2x2 max pooling method [61]

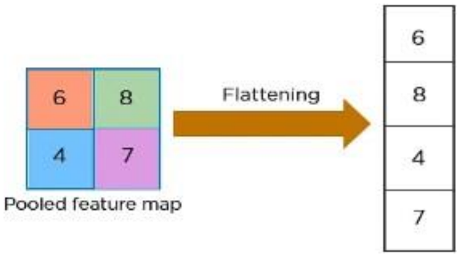


Figure 20: Flattening of a pooled feature map [61]

- **Activation Layer**

To introduce necessary non-linearity and allow the network to record non-linear correlations between input and output data, the activation function layer is important. It makes complex relationships and patterns in the data easier to model. The activation function, which is normally applied to each neuron's output, processes the weighted sum of inputs to produce an output that is then sent to the layer below.

Because of its propensity to produce sparse representations of input data and its computational efficiency, ReLU is frequently used in CNNs. By converting input into a meaningful data representation, this activation function acts as the "brain" of the CNN, returning the input value if it is positive and zero otherwise. It remains an essential part of CNN architecture [62].

Softmax Activation Function:

Softmax is used to convert numbers or logits into probabilities. A vector (let's say v) containing the probabilities of every potential result is the result of a Softmax. Vector v's probability add up to one for every potential result or class [63].

Mathematically,

$$S(y)_i = \frac{\exp(y_i)}{\sum_{j=1}^{n} \exp(y_j)}$$

Where y is an input vector to a softmax function with n elements for n classes, y_i is the i-th element of input vector, n is the number of classes and $\sum_{j=1}^{n} \exp(y_j)$ is the normalization term to ensure the values of output term.

Exp (y_i) denoted that the standard exponential function applies to y_i.

- **Fully Connected Layer:**

The fully connected layer uses extracted features to classify images. Neurons in this layer connect to every neuron in the subsequent layer, integrating extracted features and mapping them to specific classes. Fully connected layers are arranged in a way that strikes a balance between generalization and computational performance [64]. A neuron applies linear transformation to input vector through weights matrix as:

$$y_{jk}(x) = f\left(\sum_{i=1}^{n_H} w_{jk} x_i + w_{j0}\right)$$

Where W is weights matrix, x is the input vector and W_o is the bias term added to a linear function.

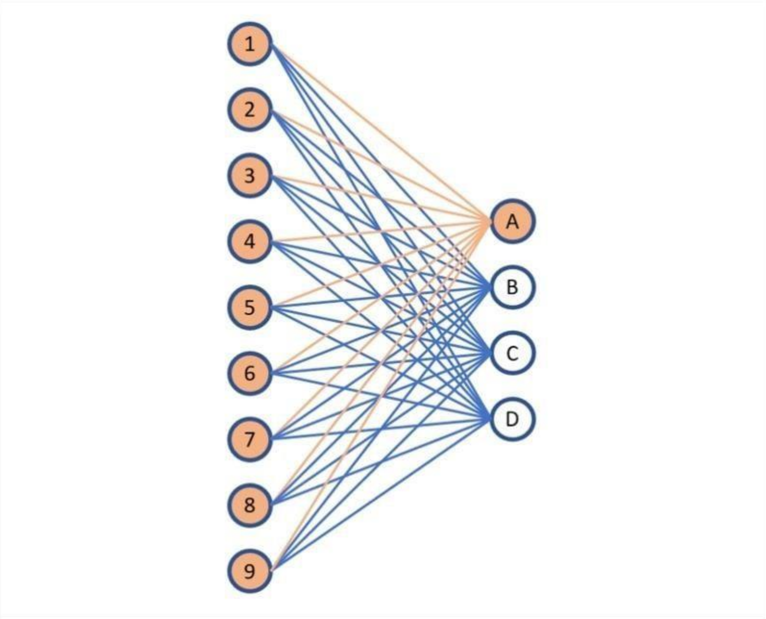


Figure 21: Fully connected layer [64]

- **Normalization Layer:**

It is a layer of preprocessing where continuous features are normalized. Inputs will be scaled and shifted by this layer into a distribution with a standard deviation of 1 and a center of 0 values. To do this, the mean and variance of the data are precomputed, and (input - mean) / sqrt(var) is called at runtime. There are different types of normalization but batch normalization is used in this project [65].

A popular method for normalizing data across the batch dimension in convolutional neural networks (CNNs) is batch normalization that calculates the mean and variance for every feature over all instances in the batch, adds epsilon for numerical stability, and then normalizes the results by deducting the mean and dividing by the standard deviation. Although this normalizing procedure can improve generalization and expedite training, it may present difficulties in sequence models and other cases with small or variable batch sizes. If normalization is found to be unneeded, scale and shift factors, "gamma" and "beta," respectively, are introduced after normalization to enable the layer to adjust [66].

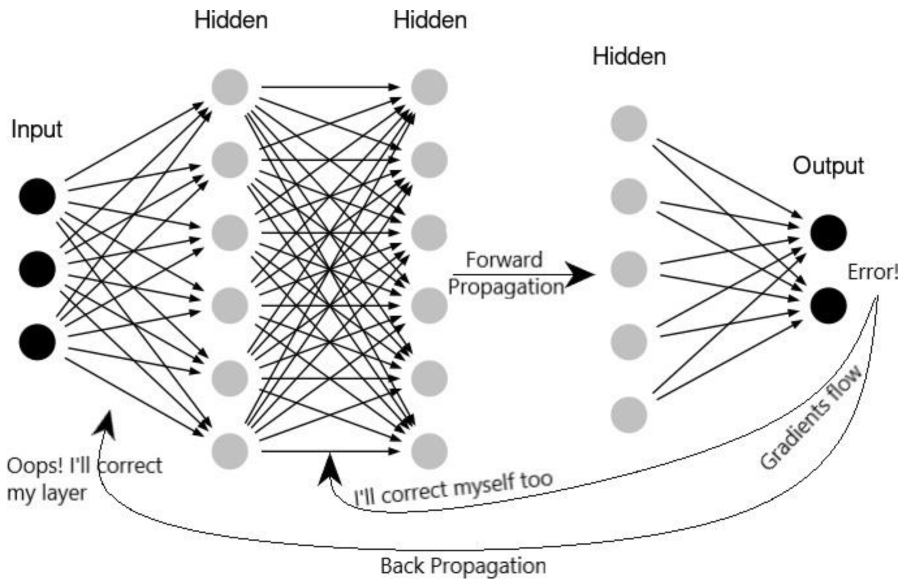


Figure 22: Normalization process layer [67]

- **Dropout Layer:**

A key regularization technique in neural networks is the dropout layer, which is primarily used to prevent overfitting by periodically turning off neurons during training. By promoting the discovery of robust features in the dataset, it improves the model's capacity to generalize and reduces the possibility that the model may memorize training data. This method ensures computing effectiveness and ease of implementation by providing efficient regularization without complex designs or computational loads. Furthermore, by preventing overfitting to unimportant noise, it strengthens the model's resistance to noisy data. To sum up, the dropout layer is a useful tool in deep learning that goes beyond traditional regularization techniques [68].

- **Dense Layer**

A dense layer, sometimes referred to as a completely connected layer, is a basic building block of neural networks in which dense connections are formed between every neuron in the layer and every neuron in the layer above it. It is appropriate for capturing complicated relationships in data because of this connectedness structure. Following convolutional layers in image classification tasks, Dense Layers are essential for processing the recovered features and generating predictions. Each neuron in the Dense Layer calculates the weighted sum of its inputs and applies an activation function to generate an output by combining weights and biases. Subsequent layers receive these outputs and process or classify them further [69].

A. VGG-19

19 layers deep Convolutional Neural Network, the VGG-19 model layers can be sub divided into: 1 SoftMax layer, 16 convolution layers, 5 MaxPool layers, and 3 Fully connected layers. Additionally, it uses 3x3 convolutional filter throughout the network [70]. The VGG-19 model has been implemented to train the model as:

I. **Model Initialization**
   - Weights are initialized with the pre-trained ImageNet weights.
   - The `include_top` parameter is set to `False` to exclude the fully connected layers at the top of the network.

II. **Freezing Layers**
   - All layers of the VGG-19 model are frozen to prevent them from being trained again during the training process.

III. **Building the Model**
   - A new Sequential model (`model`) is created.
   - The VGG-19 model is added as the first layer of the Sequential model.

IV. **Flattening Layer**

- A Flatten layer is added after the VGG-19 model to convert the output into a 1D array, preparing it for the fully connected layers.

V. **Fully Connected Layers**

- A Dense layer with softmax activation is added as the output layer with 8 units corresponding to the number of classes.

VI. **Model Compilation**

- The model is compiled with the Adam optimizer and sparse categorical crossentropy loss function.

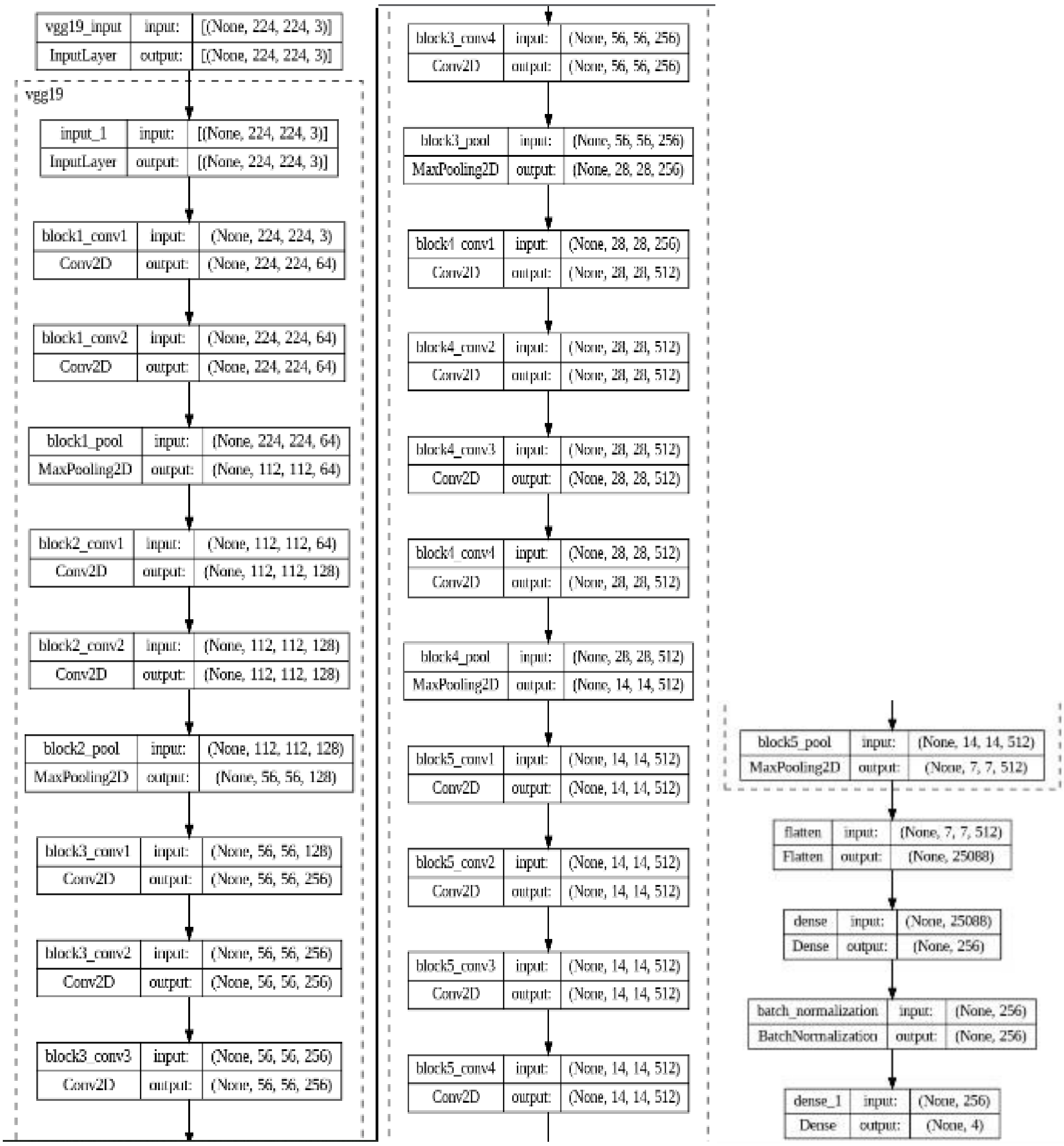


Figure 23: VGG19 Model architecture

B. ResNet-50

Also known as Residual Network, it consists of 50 layers, out of which 48 layers are convolutional layers, one is MaxPool layer, and the remaining layer is the average pool layer. With the utilization of residual blocks, ResNet tackles the issue of vanishinggradients in deep neural networks [71]. ResNet-50 has been implemented to train the modelas:

I. **Model Initialization**
   - Weights are initialized with the pre-trained ImageNet weights.
   - The `include_top` parameter is set to `False` to exclude the fully connected layers at the top of the network.

II. **Freezing Layers**
   - All layers of the ResNet-50 model are frozen to prevent them from being trained again during the training process.

III. **Building the Model**
   - A new Sequential model (`model2`) is created.
   - The ResNet-50 model is added as the first layer of the Sequential model.

IV. **Flattening Layer**
   - A Flatten layer is added after the ResNet-50 model to convert the output into a 1D array, preparing it for the fully connected layers.

V. **Fully Connected Layers**
   - A Dense layer with softmax activation is added as the output layer with 8 units corresponding to the number of classes.

VI. **Model Compilation**
   - The model is compiled with the Adam optimizer and categorical cross entropy loss function.

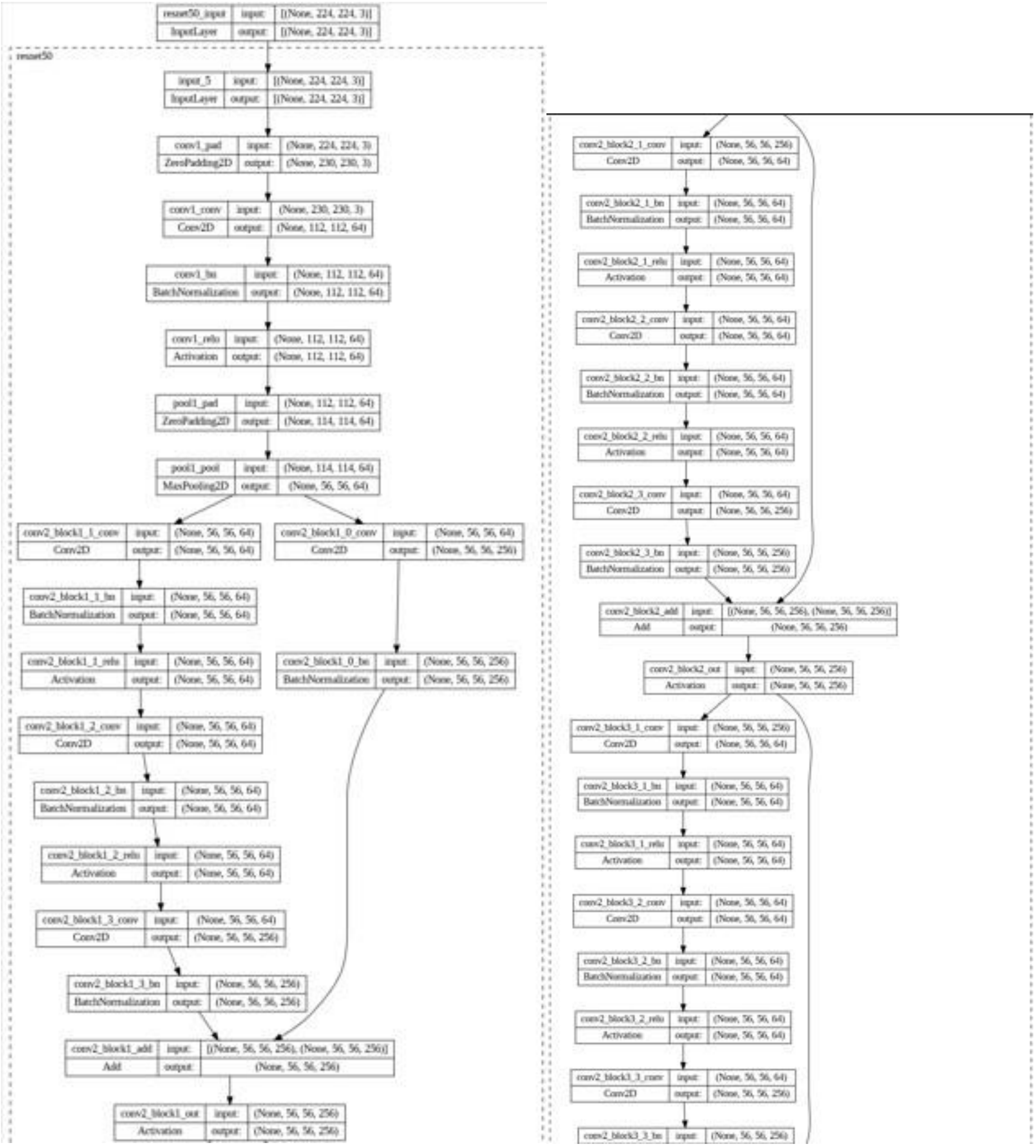


Figure 24: ResNet-50 Model

- **Data Training**

After the architecture of the deep learning model, the model was trained by assembling it. This important phase makes use of a particular function in the Keras framework that is intended to set up the model for training. The optimizer, loss function, and metrics are among the crucial arguments that this function includes and which determine important parts of the training process. The compile function's optimizer parameter chooses the optimization technique that controls how the model's parameters are updated during training. 'Adam' was chosen in this case due to its ability to change learning rates and effectively modify model weights.

- **Loss function**

  This function assess how well the algorithm represents the data in the dataset. A bigger value will be produced by the loss function if all of the forecasts are incorrect. It will produce a lower number if the dataset is fairly decent. It is the backpropagation function is and calculates the value that a model should aim to minimize while being trained [72].

- **Softmax Loss**

  For multiclass classification tasks, categorical cross entropy, also known as softmax loss, is a popular technique that combines cross-entropy loss and softmax activation. Convolutional Neural Networks can be trained to generate probability distributions over N classes for each input image by utilizing this loss function. The neural network's raw outputs are softmax activated during multiclass classification, producing a vector of predicted probabilities for each of the input classes. Typically, in a one-hot encoded label multiclass classification scenario, only one element in the target vector is non-zero and only the positive class is included in the loss computation. For these kinds of multiclass classification problems, categorical cross entropy is especially made for assessing the difference between expected class probabilities and actual class labels [73].

$$J(w,b) = -\frac{1}{m_{train}} \sum_{i=1}^{m_{train}} \sum_{j=1}^{C} y_j^{(i)} \cdot log\left(f\left(w, b\left(x^{(i)}\right)\right)_j\right)$$

  where

  $J\ (w,\ b)$ = Loss function,

  w and b are weights and biases, $mtrain$ as the number of training examples, $yj\ (i)$ as the true label for class j in the $i\ th$ example, $x\ (i)$ as the input, and $f\ (w,\ b\ (x\ (i)$)) j as the predicted probability for class j given the input and model's parameters

- **Model summary**

  The model summary for VGG19 provides an overview of the sequential architecture, detailing layer parameters and a total of 26449220 parameters, of which 6424324 were trainable and 20024896 were non trainable.

Similarly, the model summary of ResNet50 provide an overview of the sequential architecture detailing layer parameters a total of 49280132 parameters of which 6424324 were trainable and 53632 were non trainable.

Layer (type)	Output Shape	Param#
vgg19(Functional)	(None,7,7,512)	20024384
flatten(Flatten)	(None,25088)	0
dense(Dense)	(None,256)	6422784
batch_normalization (BatchNormalization)	(None,256)	1024
dense_1(Dense)	(None,4)	1028
Total params:26649220 Trainable params: 6424324 Non-trainable params: 20024896		

Table 1: Model Summary for VGG19

Layer (type)	Output Shape	Param#
Resnet50(Functional)	(None,7,7,2048)	23587712
flatten(Flatten)	(None,100352)	0
dense(Dense)	(None,256)	25690368
batch_normalization (BatchNormalization)	(None,256)	1024
dense_1(Dense)	(None,4)	1028
Total params:49280132 Trainable params: 6424324 Non-trainable params: 53632		

Table 2: Model Summary for ResNet 50

- **Data fitting**

  After the training, in order to define data, and parameters the following components were defined as:

i. Training Data: The dataset used to train the models
ii. Batch Size: The number of samples used in each iteration during training.
iii. Epochs: Number of times the model iterates over the dataset.
iv. Validation Data: Benchmark to evaluate model performance.
v. Verbosity: This controls the level of details shown during training.

vi. Callbacks: Functions called during training at defined points.

vii. Model Checkpoint: This saved the best model based on validation accuracy.

viii. Early Stopping: This stops the training if there's no improvement in validation accuracy after a certain number of epochs.

4.2.2 Project Algorithm

Step 1: Initialize.

Step 2: User accesses the website.

Step 3: User uploads the fundus image by clicking the upload button.

Step 4: User clicks on the predict button to initiate disease detection and classification.

Step 5: The integrated model evaluates the uploaded image, computes the class probabilities, and identifies the disease class with the maximum probability.

Step 6: The system displays the classification result alongside the probability on the website to the user.

Step 7: End.

CHAPTER 5: IMPLEMENTATION AND TESTING

5.1. Implementation

The eye disease classification system developed from the previously mentioned requirements and designs is implemented in this phase. The tools used and the implementation details are mentioned as follows:

5.1.1. Tools Used

The following hardware and software tools are used to develop the diabetic retinopathy detection model:

Hardware tools

- 4 GB RAM or higher.
- 1 GHz or faster processor.
- Input device: Keyboard, Mouse
- Output device: Monitor

Software Tools

The following software tools are used to develop the eye disease classification system:

I. **Python**

Used in Machine Learning, software engineering and web development, Python is a versatile programming language. As it is easy to understand, it is said to be beginner friendly. It's wide range of libraries makes it an excellent aid for developers [74].

In fact, Python has been used as the primary language in this project. Implementation of libraries such as Tensorflow, NumPy, Pandas, Matplotlib, and Seaborn has been done to train and test the machine learning model.

II. **Google Colab**

Google Colab is a game-changer in the fields of machine learning and data science because it offers easy, setup-free access to powerful computing resources like GPUs and TPUs. Its combination with GitHub and Google Drive makes it simple to organize the projects and work with others. Colab facilitates access to cutting-edge technologies like TensorFlow, enabling anybody to learn, experiment, and develop in the field of machine learning—regardless of whether they are a researcher, developer, or educator [75]. The model implemented in the project has been trained and tested through Google Colab.

III. Google Drive

Created by Google, Google Drive is a flexible cloud-based platform for collaboration and storage. It gives users the ability to store data safely on the cloud, making it possible to view them from anywhere at any time and to share them with others with ease [76]. Drive's cross-platform capabilities and the storage provided by the college was used to handle datasets and project models in a single central folder.

IV. Flask

Without the need for extra libraries or code files, Flask, a Python online application framework, enables users to interact with Python code so that machine learning models directly from web browsers [77]. It was the best option for implementing the deep learning model of the project on the local host because of its simplicity and ease of setup.

V. VS Code

Produced by Microsoft, Visual Studio Code (VS Code), is a popular source-code editor. In addition to being compatible with multiple programming language, it is also free. It is compatible with Linux, macOS, and Windows and provides features like version control system integration and debugging [78]. VS Code was used to write code in Flask for backend functionality, HTML, CSS, and JavaScript for frontend development, as well as to incorporate the machine learning model developed in Google Colab.

VI. HTML

Used to structure content on web pages, HTML, is the building block of an website. It wraps various content sections in elements to give them a specific appearance and behavior. These elements have the ability to alter text style, font size, and generate links [79]. Flask was used to create user interface template using HTML

VII. CSS

Cascading Style Sheets, or CSS for short, is a language used to style elements expressed in markup languages such as HTML. It facilitates the division of a webpage's visual design from its content. The interaction between HTML and CSS is essential; although HTML establishes the structure of the website, CSS determines how it looks overall. Keeping all of the styling in one file with CSS makes it simpler to create and update the design [80]. In order to improve the web application's user experience, the website used combined templates with CSS.

VIII. JavaScript

An interpreted language, JavaScript is used to add interactive features to websites. It works well with HTML and is easy to use on web pages. Likewise, it's cross platform functionality makes it portable. Due to its portability and seamless integration with languages such as HTML, CSS and frameworks such as Flask, the project was implemented on JavaScript to enhance the functionality of the website [81].

IX. Draw.io software: Draw.io also known as diagrams.net, is a free online application for making flowcharts and diagrams. Users can create, edit, and share a variety of graphical representations, including flowcharts and DFDs, using it [82]. To craft flowcharts, context diagrams, data flow diagrams, use cases, and class diagrams, the authors utilized Draw.io. These diagrams have been illustrated to represent the conceptual aspects of the project.

X. Microsoft Word:

Microsoft Word, a word processing application within the Microsoft Office suite, enables users to create documents ranging from simple to complex [83]. Microsoft Word was essential for work documentation during the project phases. The project ensured that every section of this report was accurate and comprehensive by following the necessary formats and criteria. Furthermore, extensive research was done and cited, which added insightful material to the report.

XI. ClickUp

ClickUp is a feature-rich productivity platform created to support group brainstorming, planning, coordination, and teamwork on a range of assignments, from product design to process documentation. Even while using one tool for several tasks can be difficult, ClickUp was created expressly to meet this demand. Initially established to save time, ClickUp has maintained its leadership in productivity by providing adaptable job management tools [84]. This tool was used inorder to plan, coordinate and design the project details and information between the team members that made the project a success.

5.1.2. Implementation Details of Modules

I. For Pre-processing of the image

 # Function to extract features using VGG19

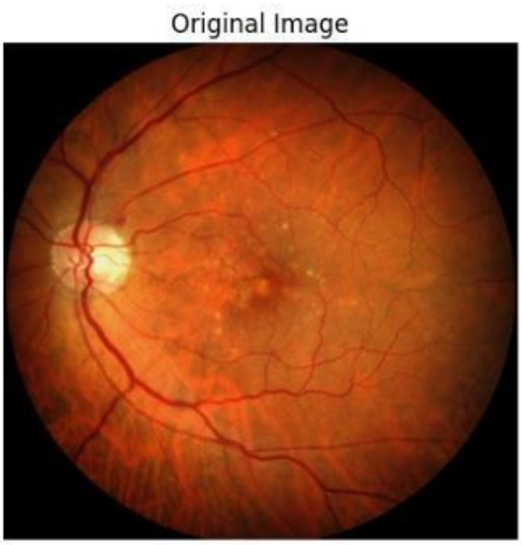
```
def extract_features_vgg19(img_path):
    img = preprocess_image_vgg19(img_path)
    features = vgg19_model.predict(img)
    return features
```

 # Function to enhance image

```
def enhance_image(img):
    img = cv2.convertScaleAbs(img, alpha=1.2, beta=0)
    return img
```

 # Function to preprocess image with enhancement techniques

```
def preprocess_image(img_path):
    img = cv2.imread(img_path)
    img = cv2.resize(img, (IMG_WIDTH, IMG_HEIGHT))
    img = cv2.GaussianBlur(img, (3, 3), 0)
    img = enhance_image(img)
    return img
```



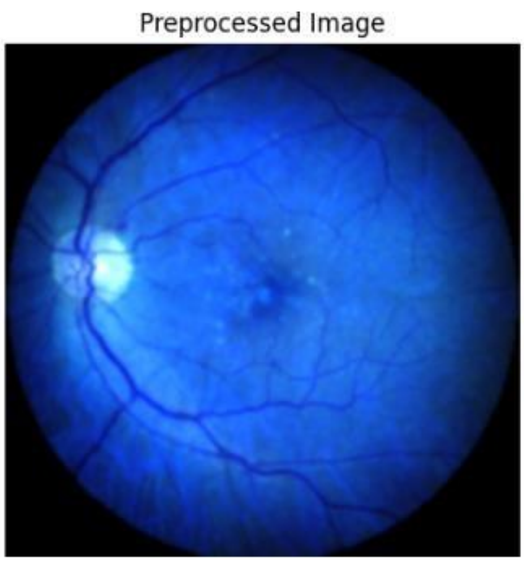


Figure 25: Preprocessed image

II. For result analysis generating confusion matrix Code Snippet

```
import random
import matplotlib.pyplot as plt
```

 # Define the class labels

```
class_labels = ["glaucoma", "cataract", "normal",
"diabetic_retinopathy"]
# Get a batch of data from the valid_data generator
images, labels = next(valid_data)
# Replace these variables with actual data
y_pred = model.predict(images) # Assuming you have a trained
model
plt.figure(figsize=(12, 7))
for i in range(20):
    sample_index = random.randint(0, images.shape[0] - 1)
    image = images[sample_index]
    category_index = labels[sample_index].argmax()
    pred_category_index = y_pred[sample_index].argmax()
    label = class_labels[category_index]
    pred_label = class_labels[pred_category_index]
    plt.subplot(4, 5, i + 1)
    plt.imshow(image)
 # Change the font size for labels
  plt.xlabel("Actual: {}\nPrediction: {}".format(label,
pred_label), fontsize=9) # Change the fontsize as needed
plt.tight_layout()
# Save the plot as an image
plt.savefig('Model Eval.png')
plt.show()
```

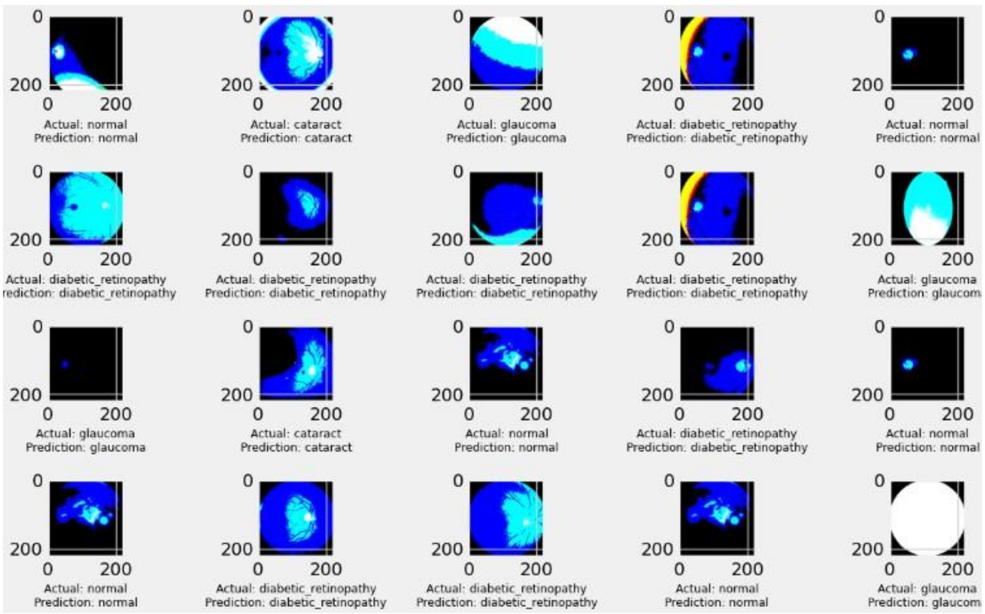


Figure 26: Model Evaluation for dataset

The figure above shows the actual and predicted type of eye disease for the 20 validation set using the proposed model.

5.2. Testing

5.2.1. Test Cases for Unit Testing

Test Case 1:

Test Objectives: Test for starting GUI

Expected Output: Successfully open the application

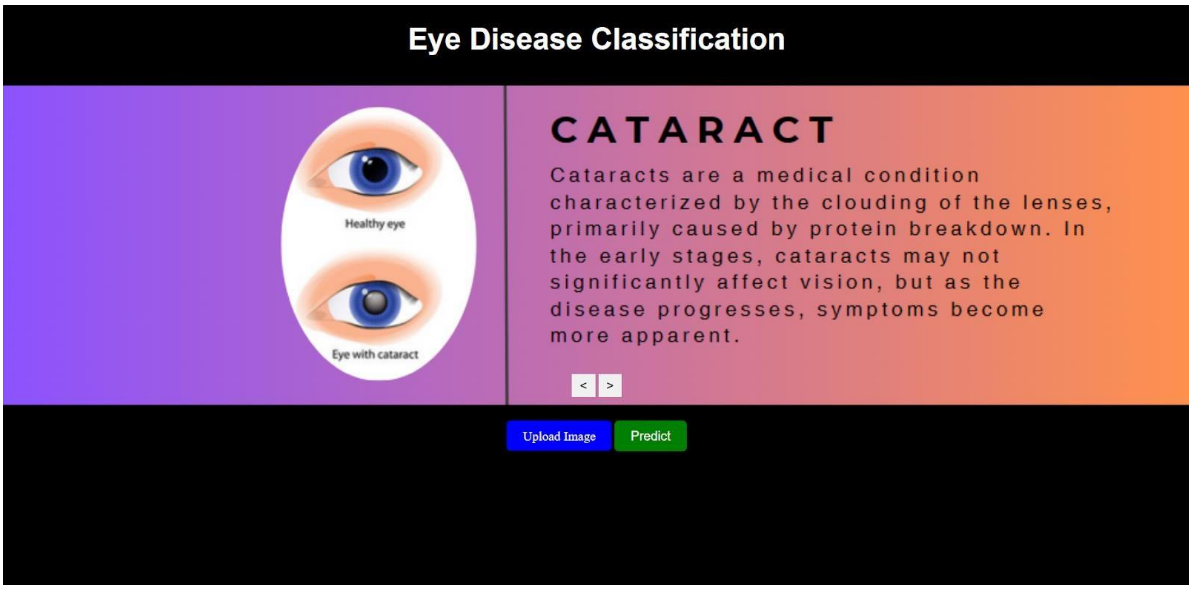


Figure 27: Index Page for web application

Test Case 2:

Test Objectives: Test for uploading image

Expected Output: Successfully upload an image and store it into uploads folder

Figure 28: Upload and Predict Buttons

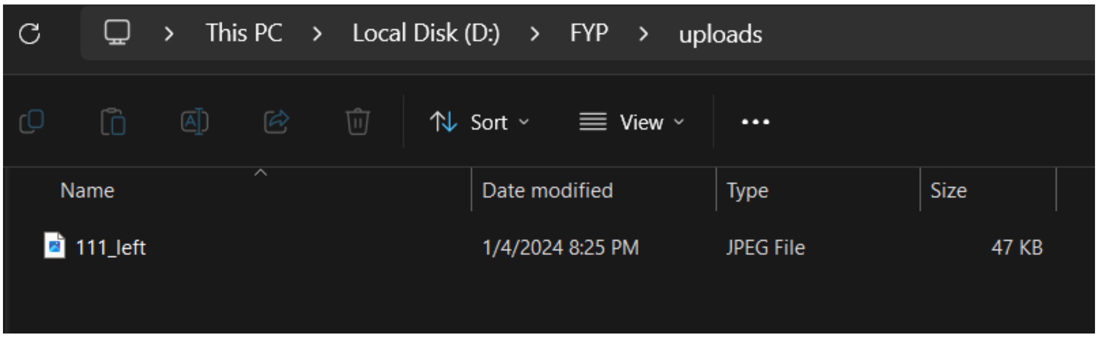


Figure 29: Directory Screenshot

Test Case 3:

Test Objectives: Test for displaying uploaded image

Expected Output: Successfully preview the raw image to the user

Figure 30: Image Preview

Test Case 4:

Test Objectives: Test for displaying the output.

Expected Output: Successfully classify the disease

Figure 31: Disease Classification

5.2.2. Test Cases for System Testing

System Testing ensures that the entire software product works seamlessly within the larger computer-based system. Its goal is to check if the system meets all specifications from start to finish. As software is just one part of a bigger system, System Testing involves various tests to thoroughly assess how well it integrates with other software and hardware components.

Test Case 1:

Test Objectives: Test for Diabetic Retinopathy Classed Image

Expected Output: Successfully classify as Diabetic Retinopathy

Figure 32: Diabetic Retinopathy Image Testing

Test Case 2:

Test Objectives: Test for Glaucoma Classed Image

Expected Output: Successfully classify as Glaucoma

Figure 33: Glaucoma Image Testing

Test Case 3:

Test Objectives: Test for Cataract Classed Image

Expected Output: Successfully classify as Cataract

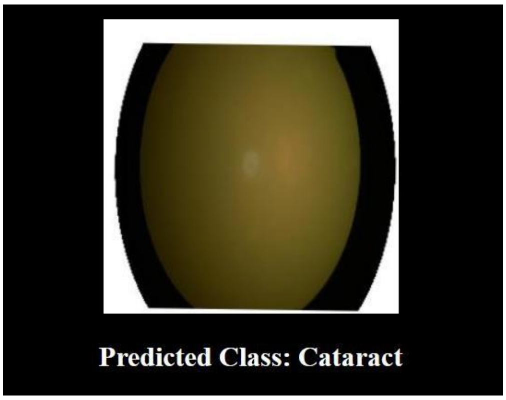


Figure 34: Cataract Image Testing

Test Case 4:

Test Objectives: Test for Normal Classed Image

Expected Output: Successfully classify as Normal

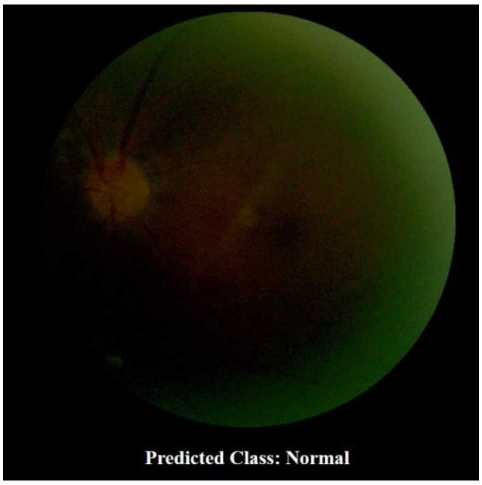


Figure 35: Normal Image Testing

Test Case 5:

Test Objectives: Test for Non-Fundus Image

Expected Output: Successfully classify as undefined

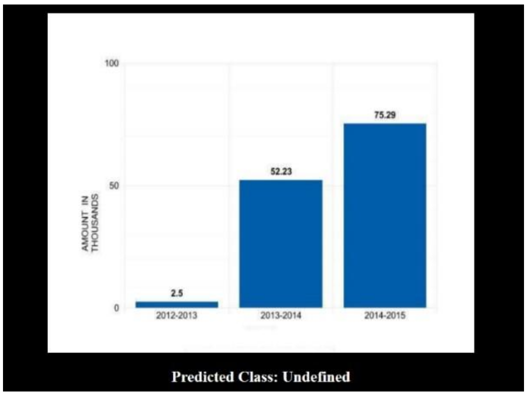


Figure 36: Non-Fundus Image Testing

5.3 Result Analysis

5.3.1 Training Process Overview

The training process was comprehensively analyzed through visualizations generated using for ResNet-50 and VGG-19 models. A layout was formed to show the training and validation accuracy and loss of both of the models. For VGG19, the best epoch is shown at 2 for loss whereas best epoch is shown at 10 for accuracy. Similarly, for the ResNet 50 training and validation loss shows best epoch at 6 and for accuracy it comes out to be 6 as well.

Figure 38: Comparing the training and validation loss and accuracy for VGG19

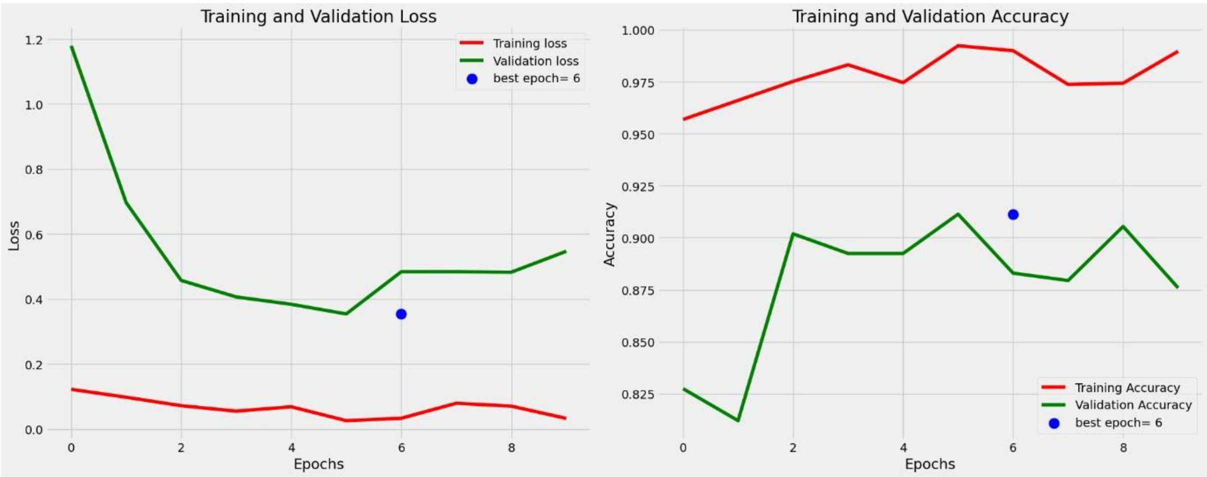


Figure 39: Comparing the training and validation loss and accuracy for ResNet50

5.3.2 Confusion Matrix

The confusion matrix is a matrix used to determine the performance of the classification models for a given set of test data. It can only be determined if the true values for test data are known. The matrix itself can be easily understood, but the related terminologies may be confusing. Since it shows the errors in the model performance in the form of a matrix, hence also known as an error matrix [85].

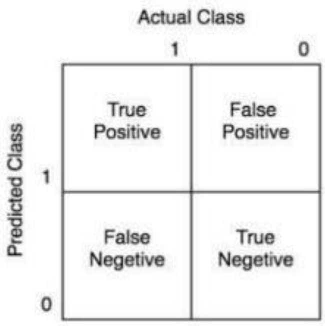


Figure 40: Confusion matrix labeling

The confusion matrix below show that the VGG19 correctly classifies most of the cataract images whereas incorrectly classifies most of the normal images. For ResNet50, the model classified most of the cataract but cannot classify most of the diabetic retinopathy properly.

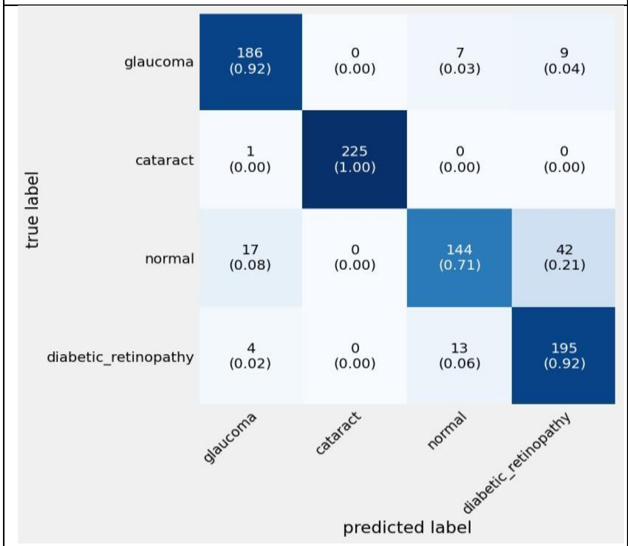
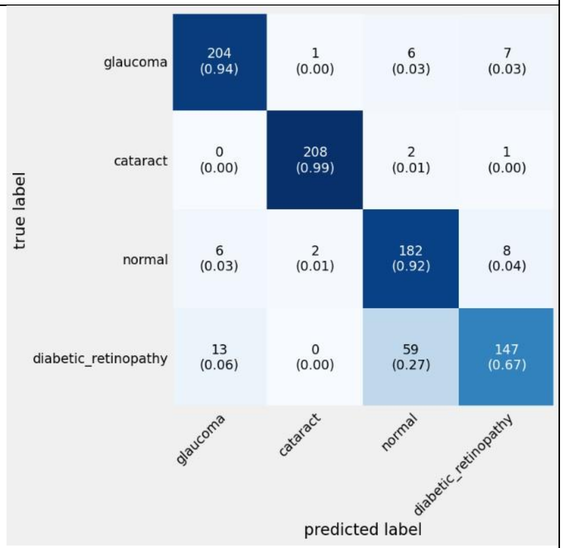
VGG19 Model	ResNet50 Model
	

Figure 41: Confusion Matrix for the proposed models.

5.3.3 Classification Report

The model's accuracy, recall, F1, and support scores are shown in the classification report visualizer. A color-coded heat map and numerical scores are integrated into the report to facilitate simpler interpretation and problem detection. The precision metric calculates the percentage of all positive forecasts that are true positive predictions, or cases that are actually positive. Recall is a metric that counts the percentage of positive instances that were really predicted to be positive. The F1-score, which is a harmonic mean of precision and recall, is helpful when evaluating models when both precision and recall are crucial. The evaluation involves calculating the accuracy, precision, and F1-measure of the model.

Accuracy= $\dfrac{TP+TN}{TP+FP+TN+FN}$

Precision= $\dfrac{TP}{TP+FP}$

Recall= $\dfrac{TP}{TP+FN}$

F1-Measure = 2×Precision×RecallPrecision+Recall

Where

True positive (TP): Prediction is +ve

True negative (TN): Prediction is -ve

False positive (FP): Prediction is +ve

False negative (FN): Prediction is –ve

The authors used these metrics for evaluation [86].

	Precision		Recall		F1-score		Support	
	VGG19	**ResNet50**	**VGG19**	**ResNet50**	**VGG19**	**ResNet50**	**VGG19**	**ResNet50**
Glaucoma	0.89	0.91	0.92	0.94	0.91	0.93	202	218
Cataract	1.00	0.99	1.00	0.99	1.00	0.99	226	211
Diabetic Retinopathy	0.88	0.73	0.71	0.92	0.78	0.81	203	198
Normal	0.79	0.90	0.92	0.67	0.85	0.77	212	219
Accuracy					**0.89**	**0.88**	843	846
Macro Average	0.89	0.88	0.89	0.88	0.89	0.87	843	846
Weighted Average	0.89	0.89	0.89	0.88	0.89	0.87	843	846

Table 3: Classification report for the proposed models

5.3.4 Findings

The analysis of the VGG-19 model reveals an 89% overall accuracy. Precision is high for 'cataract' (100%) and 'glaucoma' (89%), but slightly lower for 'normal' (88%) and 'diabetic_retinopathy' (79%). Recall rates are strong for 'cataract' (100%) and 'glaucoma' (92%), but lower for 'normal' (71%) and 'diabetic_retinopathy' (92%). F1-scores, balancing precision and recall, are highest for 'cataract' (1.00) and 'glaucoma' (0.91), and slightly lower for 'normal' (0.78) and 'diabetic_retinopathy' (0.85). 'Cataract' is the most frequent class (226), while 'normal' is the least (203). The macro and weighted average metrics demonstrate balanced performance across classes, indicating strong overall model performance, particularly in identifying 'cataract' and 'glaucoma', with room for improvement in 'normal' and 'diabetic_retinopathy' classifications.

CHAPTER 6: CONCLUSION AND FUTURE ENHANCEMENTS

6.1 Conclusion

The proposed project was able to use retinal fundus images for the precise detection and classification of different eye diseases. Convolutional Neural Networks (CNNs) model was used as the project's main tool, first concentrating on cataract identification. This served as the groundwork for extending the project's model to cover a wider spectrum of conditions, including glaucoma, diabetic retinopathy, and normal cases. To ensure reliable training data, the "Eye Disease Classification" dataset from Kaggle was used, which contains a balanced selection of images across different disease categories - Normal, Diabetic Retinopathy, Cataract, and Glaucoma. With data from reliable sources such as Ocular Recognition, IDRiD, and HRF, this dataset offered an accurate representation of wide range of eye conditions. Additionally, testing was conducted using fundus images from Kirtipur Eye Hospital, further validating the effectiveness of the model.

The models VGG19 and ResNet50 were trained on the respective datasets and the best model was used to predict and classify the eye disease. VGG19 outperformed with an accuracy of 89% whereas ResNet50 had an accuracy of 88% which might be due to algorithms. The result depicts that both of these models can be used to classify and detect disease. The best model is then saved and integrated with flask as a backend and hosted onto a webpage for better user interaction.

6.2 Future Enhancements

These algorithms struggled to process the dataset due to lack of huge dataset. Further image preprocessing may be employed as a consequence for more precise and accurate prediction. For various results, the datasets can be expanded and other class may be added into the system for multiple diseases. Overall, utilizing properly preprocessed and trained dataset with appropriate deep learning tools can be used to predict eye disease by using fundus images, which can help to enhance the algorithm. The accuracy can further be enhanced by the use of preprocessing tools and by increasing the number of training datasets.

There are several potential enhancements that could be made. Some possibilities include:

1. Implementing Optical Coherence Tomography (OCT) images to detect cataract.
2. Expanding the disease classes and include diseases such as AMD, Hypertension, Pathological Myopia, etc.
3. Enable real-time diagnosis.

REFERENCES

[1] "World Report on vision," World Health Organization, https://www.who.int/publications/i/item/9789241516570 (accessed Mar. 10, 2024).

[2] J. Latif, C. Xiao, A. Imran, and S. Tu, "Medical Imaging using Machine Learning and Deep Learning Algorithms: A Review," in *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, IEEE, Jan. 2019, pp. 1–5. doi: 10.1109/ICOMET.2019.8673502.

[3] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016, doi: 10.1016/j.neucom.2015.09.116.

[4] *Cataracts* (2023) Mayo Clinic. Available at: https://www.mayoclinic.org/diseases-conditions/cataracts/symptoms-causes/syc-20353790

[5] Boyd, K. (2023) *What is glaucoma? symptoms, causes, diagnosis, treatment*, American Academy of Ophthalmology. Available at: https://www.aao.org/eye-health/diseases/what-is glaucoma#: ~:text=Glaucoma%20is%20a%20disease%20that, eye%2C%20damaging%20the%20optic%20nerve.

[6] Gordon, M. (2023) *Glaucoma vs cataracts: What's the difference?* Guy's and St Thomas' Private Healthcare. Available at: https://guysandstthomasprivatehealthcare.co.uk/glaucoma-vs-cataracts-whats-the-difference/.

[7] K. Boyd, "Diabetic retinopathy: Causes, symptoms, treatment," American Academy of Ophthalmology, https://www.aao.org/eye-health/diseases/what-is-diabetic-retinopathy.

[8] D. Turbert, "Fundus" *American Academy of Ophthalmology*,https://www.aao.org/eye-health/anatomy/fundus

[9] Fundus Photography Overview - *ophthalmic photographers' society,* https://www.opsweb.org/page/fundusphotography.

[10] M. Sonka, V. Hlavac, and R. Boyle, "Image pre-processing," in *Image Processing, Analysis and Machine Vision*, Boston, MA: Springer US, 1993, pp. 56–111. doi: 10.1007/978-1-4899-3216-7_4.

[11] *"What is the purpose of image preprocessing in deep learning?* Available:

https://www.isahit.com/blog/what-is-the-purpose-of-i 45 mage-preprocessing-in-deep-learning

[12] Eye_diseases_classification. (n.d.). *Kaggle: Your Machine Learning and Data Science Community*. https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseasesclassification

[13] "Kirtipur Eye Hospital," *Nepal Netra Jyoti Sangh | Eye Hospital*, https://nnjs.org.np/eye-hospital-detail/kirtipur-eye-hospital (accessed Mar. 10, 2024).

[14] "What is a neural network?," IBM, https://www.ibm.com/topics/neural-networks (accessed Mar. 13, 2024).

[15] "What is transfer learning? exploring the popular deep learning approach.," Built In, https://builtin.com/data-science/transfer-learning (accessed Mar. 11, 2024).

[16] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

[17] "Relu and sigmoid activation functions in a neural network - shiksha online," *shiksha*, https://www.shiksha.com/online-courses/articles/relu-and-sigmoid-activation-function/ (accessed Mar. 10, 2024).

[18] "Pooling in a CNN: Pooling Layers explained," *KnowledgeHut*, https://www.knowledgehut.com/blog/data-science/pooling-layer (accessed Mar. 10, 2024).

[19] "Fully connected layer vs. convolutional layer: Explained," *Built In,* https://builtin.com/machine-learning/fully-connected-layer.

[20] "What Is Machine Learning and How Does It Work?" *Simplilearn,* February 21, 2023. Available: https://www.simplilearn.com/tutorials/machinelearning-tutorial/what-is-machine-learning

[21] Doç. Dr. Serkan Savaş "Deep Learning Libraries," Medium, https://medium.com/yapay-zeka-makine-%C3%B6%C4%9Frenmesi-derin-%C3%B6%C4%9Frenme/deep-learning-libraries-6a503227479.

[22] "Keras: The high-level API for tensorflow : Tensorflow Core," *TensorFlow*, https://www.tensorflow.org/guide/keras (accessed Mar. 10, 2024).

[23] "What is Sklearn?," *Domino Data Lab*, https://domino.ai/data-science-dictionary/sklearn.

[24] J. J. Titano *et al.*, "Automated deep-neural-network surveillance of cranial images

for acute neurologic events," *Nat. Med.*, vol. 24, no. 9, pp. 1337–1341, Sep. 2018, doi: 10.1038/s41591-018-0147-y.

[25] D. S. W. Ting *et al.*, "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes," *JAMA*, vol. 318, no. 22, p. 2211, Dec. 2017, doi: 10.1001/jama.2017.18152.

[26] F. Grassmann *et al.*, "A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography," *Ophthalmology*, vol. 125, no. 9, pp. 1410–1420, Sep. 2018, doi: 10.1016/j.ophtha.2018.02.037.

[28] A. Diaz-Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: an extensive validation," *Biomed. Eng. Online*, vol. 18, no. 1, p. 29, Dec. 2019, doi: 10.1186/s12938-019-0649-y.

[29] N. Tsiknakis *et al.*, "Deep learning for diabetic retinopathy detection and classification based on fundus images: A review," *Comput. Biol. Med.*, vol. 135, p. 104599, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104599.

[29] L. Li *et al.*, "A Large-Scale Database and a CNN Model for Attention-Based Glaucoma Detection," *IEEE Trans. Med. Imaging*, vol. 39, no. 2, pp. 413–424, Feb. 2020, doi: 10.1109/TMI.2019.2927226.

[30] S. Pachade *et al.*, "Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research," *Data*, vol. 6, no. 2, p. 14, Feb. 2021, doi: 10.3390/data6020014.

[31] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, "Automated detection of mild and multi-class diabetic eye diseases using deep learning," *Heal. Inf. Sci. Syst.*, vol. 8, no. 1, p. 32, Dec. 2020, doi: 10.1007/s13755-020-00125-5.

[32] N. Gour and P. Khanna, "Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network," *Biomed. Signal Process. Control*, vol. 66, p. 102329, Apr. 2021, doi: 10.1016/j.bspc.2020.102329.

[33] N. Gundluru *et al.*, "Enhancement of Detection of Diabetic Retinopathy Using Harris Hawks Optimization with Deep Learning Model," *Comput. Intell. Neurosci.*,vol. 2022, pp. 1–13, May 2022, doi: 10.1155/2022/8512469.

[34] S. Yu, D. Xiao, S. Frost, and Y. Kanagasingam, "Robust optic disc and cup segmentation with deep learning for glaucoma detection," *Comput. Med. Imaging*

Graph., vol. 74, pp. 61–71, Jun. 2019, doi: 10.1016/j.compmedimag.2019.02.005.

[35] P. Glaret subin and P. Muthukannan, "Optimized convolution neural network based multiple eye disease detection," *Comput. Biol. Med.*, vol. 146, p. 105648, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105648.

[36] C.-H. H. Yang *et al.*, "A Novel Hybrid Machine Learning Model for Auto-Classification of Retinal Diseases," Jun. 2018, [Online]. Available: http://arxiv.org/abs/1806.06423

[37] J. C. Moses, S. Adibi, N. Wickramasinghe, L. Nguyen, M. Angelova, and S. M. S. Islam, "Smartphone as a Disease Screening Tool: A Systematic Review," *Sensors*, vol. 22, no. 10, p. 3787, May 2022, doi: 10.3390/s22103787.

[38] S. Asano *et al.*, "Predicting the central 10 degrees visual field in glaucoma by applying a deep learning algorithm to optical coherence tomography images," *Sci. Rep.*, vol. 11, no. 1, p. 2214, Jan. 2021, doi: 10.1038/s41598-020-79494-6.

[39] B. Askarian, P. Ho, and J. W. Chong, "Detecting Cataract Using Smartphones," *IEEE J. Transl. Eng. Heal. Med.*, vol. 9, pp. 1–10, 2021, doi: 10.1109/JTEHM.2021.3074597.

[40] S. Yadav and J. K. P. S. Yadav, "Automatic Cataract Severity Detection and Grading Using Deep Learning," *J. Sensors*, vol. 2023, pp. 1–17, Jun. 2023, doi: 10.1155/2023/2973836.

[41] A. Nawaz, T. Ali, G. Mustafa, M. Babar, and B. Qureshi, "Multi-Class Retinal Diseases Detection Using Deep CNN With Minimal Memory Consumption," *IEEE Access*, vol. 11, pp. 56170–56180, 2023, doi: 10.1109/ACCESS.2023.3281859.

[42] A. Shoukat, S. Akbar, S. A. Hassan, S. Iqbal, A. Mehmood, and Q. M. Ilyas, "Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach," *Diagnostics*, vol. 13, no. 10, p. 1738, May 2023, doi: 10.3390/diagnostics13101738.

[43] A. M. S. Sushma K Sattigeri, Harshith N, Dhanush Gowda N, K A Ullas, "EYE DISEASE IDENTIFICATION USING DEEP LEARNING," *Int. Res. J. Eng. Technol.*, vol. 9, no. 7, p. 5, 2022.

[44] "System Analysis and Design Tutorial"Accessed on: 1 February, 2023. [Online]. Available: https://www.tutorialspoint.com/system_analysis_and_design/index.html

[45] *"requirements analysis (requirements engineering)"* [Online]. Available: https://www.techtarget.com/searchsoftwarequality/definition/requirement s-analysis

[46] *“What is a Functional Requirement in Software Engineering?”* January 21, 2023 Available: https://www.guru99.com/functionalrequ irement-specification-example.html

[47] *“Nonfunctional Requirements”* Available: https://www.scaledagileframework.com/nonfunctional-requirements

[48] K. Bause, A. Radimersky, M. Iwanicki, and A. Albers, *“Feasibility studies in the product development process,*” Procedia CIRP, vol. 21, pp. 473–478, 2014. doi:10.1016/j.procir.2014.03.128

[49] A. rodriguez, *“Whitten & Bentley (2007) System Analysis and design methods - 7th edition,”* Academia.edu,https://www.academia.edu/8787830/Whitten_and_Bentley_2007_System_Analysis_and_Design_Methods_7th_Edition.

[50] P. B. Team, “What is Data Modeling?: Microsoft power bi,” What is Data Modeling? | Microsoft Power BI, https://powerbi.microsoft.com/en-us/what-is-data-modeling/.

[51] “What is an entity relationship diagram (ERD)?,” Lucidchart, https://www.lucidchart.com/pages/er-diagrams#:~:text=Make%20an%20ERD-,What%20is%20an%20ER%20diagram%3F,each%20other%20within%20a%20system.).

[52] What is process modeling? 6 essential questions answered, https://www.bizagi.com/en/blog/process-modeling-and-mapping/what-is-process-modeling-6-essential-questions-answered#:~:text=Process%20modeling%20is%20the%20graphical,context%20of%20the%20business%20environment..

[53] “Data-Flow Diagram,” Wikipedia, https://en.wikipedia.org/wiki/Data-flow_diagram (accessed Mar. 11, 2024).

[54] ”What is systems design? definition of systems design, systems design meaning,” The Economic Times, https://economictimes.indiatimes.com/definition/systems-design (accessed Mar. 11, 2024).

[55] About Jessie Strongitharm Jessie Strongitharm is a Content Marketer & Writer at Venngage. A quick-witted wordsmith whose passions lie in strategic storytelling (see also: excessive alliteration) and A. J. Strongitharm, “What is a flowchart? use cases, templates & design tips,” Venngage,

[55] R. Awati, “What is convolutional neural network? - Definition from WhatIs.com,” *SearchEnterpriseAI,* Sep. 2022.

https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network

[56] "Convolutional neural network: Benefits, types, and applications," Datagen, https://datagen.tech/guides/computer-vision/cnn-convolutional-neural-network/#

[57] "Basic CNN architecture: Explaining 5 layers of Convolutional Neural Network," upGrad blog, https://www.upgrad.com/blog/basic-cnn-architecture/ (accessed Mar. 13, 2024).

[58] "What are convolutional neural networks?," IBM, https://www.ibm.com/topics/convolutional-neural-networks#:~:text=The%20convolutional%20layer%20is%20the%20core%20building%20block%20of%20a,filter%2C%20and%20a%20feature%20map (accessed Mar. 13, 2024).

[59] A. Rosebrock, "Convolutional Neural Networks (cnns) and layer types," PyImageSearch, https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/

[60] "Papers with code - max pooling explained," Explained | Papers With Code, https://paperswithcode.com/method/max-pooling (accessed Mar. 10, 2024).

[61] T. A. Team, "Introduction to pooling layers in CNN," Towards AI, https://towardsai.net/p/l/introduction-to-pooling-layers-in-cnn (accessed Mar. 13, 2024).

[62] P. Potrimba, "What is a convolutional neural network?," Roboflow Blog, https://blog.roboflow.com/what-is-a-convolutional-neural-network/#:~:text=Activation%20Function%20Layer&text=It%20takes%20in%20the%20weighted,Sigmoid.

[63] Medium, https://towardsdatascience.com/softmax-activation-function-how-it-actually
works-d292d335bd78. (accessed Mar. 13, 2024).

[64] V. Rastogi, "Fully connected layer," *Medium*, https://medium.com/@vaibhav1403/fully-connected-layer-f13275337c7c (accessed Mar. 10, 2024).

[65] K. Team, "Keras documentation: Normalization Layer," Keras, https://keras.io/api/layers/preprocessing_layers/numerical/normalization/. (accessed Mar. 13, 2024).

[66] P. S, “Understanding batch normalization, layer normalization and group normalization by implementing from scratch,” LinkedIn, https://www.linkedin.com/pulse/understanding-batch-normalization-layer-group-implementing-pasha s#:~:text=Layer%20normalization%20computes%20the%20mean,RNNs%20or%20in%20transformer%20models.

[67] Rajagopal, “Batch normalization - speed up neural network training,” Medium, https://medium.com/@ilango100/batch-normalization-speed-up-neural-network-training-245e39a62f85

[68] “What is the dropout layer?,” Data Basecamp, https://databasecamp.de/en/ml/dropout-layer-en (accessed Mar. 13, 2024).

[69] G. Dumane, “Introduction to convolutional neural network (CNN) using tensorflow,” Medium, https://towardsdatascience.com/introduction-to-convolutional-neural-network-cnn-de73f69c5b83#:~:text=Dense%20Layer%20is%20simple%20layer,source (accessed Mar. 13, 2024).

[70] “VGG-19 Convolutional Neural Network - an overview | ScienceDirect Topics,” *www.sciencedirect.com*. https://www.sciencedirect.com/topics/computer-science/vgg-19-convolutional-neural-network (accessed Mar. 10, 2024).

[71] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.

[72] “Introduction to loss functions,” DataRobot AI Platform, https://www.datarobot.com/blog/introduction-to-loss-functions/ (accessed Mar. 13, 2024).

[73] “Cross entropy loss: Intro, applications, code,” V7, https://www.v7labs.com/blog/cross-entropy-loss-guide#:~:text=Categorical%20Cross%20Entropy%20is%20also,N%20classes%20for%20each%20image.. (accessed Mar. 13, 2024).

[74] “Why Use Python for AI and Machine Learning?” Accessed on: 10 March, 2024. [Online]. Available: https://steelkiwi.com/blog/python-for-ai-and-machine-learning/

[75] “Google Colab,” *www.linkedin.com*. https://www.linkedin.com/pulse/google-colab-dipti-goyal (accessed Mar. 10, 2024).

[76] “What’s so great about Google Drive? And why should my students be using it?”September 25, 2015. Accessed on: 20 February, 2024. [Online]. Available:https://www.cultofpedagogy.com/google-drive-for-students/

[77] “Model Deployment using Flask” 10 March, 2024. Accessed on: 17 February, 2023. [Online]. Available:https://towardsdatascience.com/model-deployment- using-flask-c5dcbb6499c9

[78] “Visual Studio Code - ArchWiki,” *wiki.archlinux.org*. https://wiki.archlinux.org/title/Visual_Studio_Code (accessed Mar. 10, 2024).

[79] “HTML basics” Accessed on: 10 March, 2024. [Online]. Available:https://developer.mozilla.org/enUS/docs/Learn/Getting_started_with_the_web/HTML_basics

[80] “What Is CSS and How Does It Work?”January 3, 2023. Accessed on: 15 February, 2024. [Online]. Available:https://www.hostinger.com/tutorials/what-is-css

[81] Javascript Tutorial - Tutorialspoint,” *Tutorialspoint.com*, 2019. https://www.tutorialspoint.com/javascript/index.htm

[82] “draw.io”June 02, 2020. Available: https://www.computerhope.com/jargon/d/drawio.html

[83] “Introduction to Microsoft Word”,Accessed on: 18 February, 2024. [Online]. Available:https://app.myeducator.com/reader/web/1204b/lesson1/mx1ix/

[84] H. Parker, “What is clickup used for and how does it work?,” ClickUp, https://clickup.com/blog/what-is-clickup-used-for/#:~:text=ClickUp%20is%20an%20all%2Din,ClickUp%20was%20built%20for%20this (accessed Mar. 13, 2024).

[85] *“Confusion Matrix in Machine Learning”* Available: https://www.javatpoint.com/confusion-matrix-in-machine-learning

[86] G. Varoquaux and O. Colliot, “Evaluating machine learning models and their diagnostic value,” *Machine Learning for Brain Disorders*, pp. 601–630, 2023. doi:10.1007/978-1-0716-3195-9_20

APPENDICES

Ocular Disease Detection and Classification Using Convolutional Neural Network and Transfer Learning

ORIGINALITY REPORT

1%

SIMILARITY INDEX

PRIMARY SOURCES

1	elibrary.mec.edu.om Internet	112 words — 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 1%
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 10 WORDS

नेपाल नेत्रज्योति संघ
कीर्तिपुर आँखा अस्पताल तथा अध्ययन केन्द्र
(Kirtipur Eye Hospital and Ophthalmic Study Centre)

VISION 2020
THE RIGHT TO SIGHT

PAN No: 301127031

Ref. No. 311-080-081

7th March 2024

To,

The Chief Technology Officer,

Department of Computer Science

St. Xavier's Collage, Maitighar, Kathmandu.

Respscted Sir/Ma'am,

We would like to thank you for reaching out to us regarding your final year project on "Ocular Disease Detection and Classification using Convolutional Neural Network and Transfer Learning." We appreciate your interest in utilizing our resources for academic research purposes.

We understand the importance of access to relevant data for your project. So we are providind the data for your students.

We wish you and your students the best of luck with your research endeavors and commend your dedication to advancing knowledge in the field.

7th March 2024

Deepak Bdr. Bohara
Administration Assistant
Kirtipur Eye Hospital & Ophthalmic Studies Centre.

प्रशासन शाखा
Administration Department

ताजाल्चा, तिनकुने, कीर्तिपुर, काठमाण्डौ ।
फोन नं. : ४३३१८२३, ४३३२०८९
Email: kirtipureyehospital@nnjs.org.np Wedsite: www.nnjs.org.np

Name of the Project: Ocular Disease Detection and Classification using Convolutional Neural Network and Transfer Learning

Members: Rashu Shrestha,Safalta Khanal

S.No.	Name of the Authors	Title of the Article	No. of Citations	Name of the Journal	Impact Factor of the Journal	Published Year	Problem Statement	Study Area / Context / Geographical Area of Study	Data Used/ Sources of Data	Methods Used (How has this study been done?)	Technology Used/ Implementation	Result/Finding	Validation/Verification/Evaluation Method used	How much is the accuracy?	Gap/Shortcoming/Limitation/Enhancement
1	Jeffrey De Fauw, Joseph R Ledsam, Bernardino Romera-Paredes, Stanislav Nikolov, Nenad Tomasev, Sam Blackwell, Harry Askham, Xavier Glorot et al.	Clinically applicable deep learning for diagnosis and referral in retinal disease	2133	Nature Medicine	87.244	2018	To overcome the increasing volume and complexity of diagnostic imaging, with the goal of achieving or surpassing human expert performance in clinical settings.	The study area of this paper is centered around Moorfields Eye Hospital NHS Foundation Trust, a tertiary referral center in London, UK.	The clinical data used for the training, validation and test sets were collected at Moorfields Eye Hospital,UK in de-identified format.	The study utilized 15,877 OCT scans from 7,981 patients to develop a deep learning framework for automated diagnosis of eye disorders, employing a classification network for referral recommendations, retinal pathology detection, and a 3D U-Net architecture for segmentation, generating multiple hypotheses for each scan.	Optical Coherence Tomography (OCT) scans alongside segmentation and classification networks,	The study highlights divergent training requirements between devices in achieving expert performance on referral decisions, yet affirms the efficacy of the developed deep learning framework for making critical referral decisions in OCT imaging for eye diseases.	The paper evaluates the method's performance based on referral decisions and retinal morphology, employing three junior graders with oversight from a senior retinal specialist to assess a subset of OCT scans during validation and assessment, while adhering to standard grading criteria for OCT evaluation and having access to follow-up data.	The paper suggests incorrect referral decisions in 4 of the 116 cases, a total error rate of 3.4%. Due to the small number of cases in the new test set, this is not significantly different to the error rate of 5.5% on device type 1 (P(4 out of 116 < 55 out of 997) = 0.774.	In future work, researchers can address a much wider range of medical imaging techniques, and incorporate clinical diagnoses and tissue types well outside the immediate application demonstrated here.
2	Daniel Shu Wei Ting, Carol Yim-Lui Cheung, Gilbert Lim, et al	Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes	1716	JAMA Network (Journal of the American Medical Association)	120.7	2017	The study evaluates a deep learning system's performance in detecting referable and vision-threatening diabetic retinopathy, possible glaucoma, and age-related macular degeneration in multiethnic populations	Singapore National Diabetic Retinopathy Screening Program (SIDRP)	The study utilizes retinal images from patients with diabetes enrolled in the Singapore National Diabetic Retinopathy Screening Program (SIDRP) between 2010 and 2013,.	The deep learning system employed a convolutional neural network trained on retinal images to recognize features of referable diabetic retinopathy, possible glaucoma, and AMD, enabling classification of unseen images after exposure to diverse examples during training.	Deep learning systems, utilizing convolutional neural networks, outperformed professional graders in assessing diabetic retinopathy from retinal images captured in JPEG format	In the primary validation dataset, the deep learning system exhibited high sensitivity and specificity for detecting referable and vision-threatening diabetic retinopathy, possible glaucoma, and age-related macular degeneration..	The primary analysis aimed to determine the deep learning system's performance in detecting referable and vision-threatening diabetic retinopathy compared to experienced nonmedical professional graders, using a retinal specialist as the reference standard, followed by external validation across 10 diverse cohorts of participants with diabetes.	The deep learning system achieved high accuracy in detecting referable diabetic retinopathy (AUC: 0.936, sensitivity: 90.5%, specificity: 91.6%),	The study's training set lacked retinal specialists' grading, while the primary validation dataset included their assessments; however, external datasets had varied assessments by different professionals, underscoring limitations in accurately identifying diabetic macular edema from fundus photographs
3	Felix Grassmann, Judith Mengelkamp, Caroline Brandl, Sebastian Harsch, Martina E. Zimmermann, Birgit Linkohr, Annette Peters, Iris M. Heid, Christoph Palm, Bernhard H.F. Weber	A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography	460	American Academy of Opthamology	13.7	2018	Age-related macular degeneration (AMD) is a common threat to vision. While classification of disease stages is critical to understanding disease risk and progression, several systems based on color fundus photographs are known.	The research obtains fundus images from Zeiss and Topcon cameras, totaling 120,656 images, along with manual gradings, from the Genotypes and Phenotypes database for analysis.	The algorithm's performance was evaluated on 5555 fundus images from the population-based KORA study.	The proposed deep learning classification strategy involves preprocessing color fundus images, training multiple independent convolutional neural networks (CNNs) to optimize an evaluation metric iteratively, building a model ensemble with a random forest algorithm.	The study trained 13 classes using various convolutional deep learning architectures including AlexNet, GoogLeNet, VGG, Inception-V3, ResNet-101, and Inception-ResNet-V2.	The study trained 6 CNNs on 86,770 AREDS fundus images, achieving improved accuracy and weighted kappa (63.3% and 92.1% respectively).	Validation is performed on a cross-sectional, population-based study. 21867 fundus images datasets were used to evaluate the performance of various methods to compute the network ensemble. Using the predicted class probabilities of the 6 CNN models, the random forest using 1000 trees were used on the training dataset.	6 different neural net architectures predicted the 13 classes in the AREDS test set with an overall accuracy of 63.3%. The algorithm detected 84.2% of all fundus images with definite signs of early or late AMD. Overall, 94.3% of healthy fundus images were classified correctly.	The model can be improved by incorporating images from population-based surveys that contain other phenotypes to enable a CNN that can be used to prescreen patients, to aid diagnosis in day-to-day patient care, or both.
4	Tao Li, Wang Bo, Chunyu Hu, Hong Kang, Hanruo Liu, Kai Wang, Huazhu Fu	Applications of deep learning in fundus images: A review	223	Medical Image Analysis	10.9	2021	This research aims to address the challenge of limited high-quality labeled data in fundus image analysis by exploring solutions to enhance the performance of deep learning models for early screening of eye diseases.	The study includes utilizing images from various sources such as UK Biobank, and hospitals in India. The area of study focuses on the applications of deep learning in fundus images for early screening of eye diseases.	Fundus images from various sources such as EyePACS, Inoveon, AREDS, UK Biobank, and hospitals in India.	The study collects and analyzes 143 papers from databases including DBLP, ScienceDirect, JAMA Network, and Investigative Ophthalmology & Visual Science, focusing on deep learning applications in fundus images from January 2016 to August 2020 to identify trends and suggest future directions for early screening of eye diseases.	The technology used in this paper revolves around deep learning methods, specifically Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and Residual Networks.	The paper suggests that deep learning techniques, including CNNs, GANs, and Residual Networks, hold potential for enhancing the diagnosis and grading of ophthalmic diseases from fundus images, while addressing current limitations and proposing solutions to improve model.	The study evaluates deep learning models' effectiveness in early screening of eye diseases using performance metrics like accuracy, sensitivity, specificity, and AUC, while also discussing their generalization to real-world data and comparing their diagnostic performance with human experts.	The accuracy values range from around 76% to over 98% in different studies	The study highlights the need for simultaneous improvement in fine lesion segmentation and disease diagnosis, particularly emphasizing the lesser focus on AMD diagnosis compared to other eye diseases, alongside discrepancies in reported data and the necessity for enhanced data management and reporting practices.

5	Andres Diaz-Pinto, Sandra Morales, Valery Naranjo,Thomas Köhler, Jose M. Mossi, and Amparo Navea	CNNs for automatic glaucoma assessment using fundus images: an extensive validation	172	BioMedical Engineering OnLine	3.9	2019	The paper addresses the challenge of improving glaucoma assessment by utilizing ImageNet-trained CNNs, aiming to overcome limitations associated with handcrafted features and segmentation methods in existing algorithms.	Medical Image Analysis for Glaucoma Assessment	The study utilizes five public databases comprising 1707 fundus images, as well as a newly introduced ACRIMA database with 705 labeled images.	ImageNet-trained CNNs, specifically VGG16, VGG19, InceptionV3, ResNet50, and Xception, are employed for automatic glaucoma assessment. The approach involves fine-tuning and testing these models.	The implementation involves training and testing the CNN models using raw pixel intensities from fundus images. The technology includes deep learning frameworks for CNNs and associated software tools.	The Xception architecture yields significant improvements, achieving an average AUC of 0.9605, specificity of 0.8580, and sensitivity of 0.9346. The ACRIMA database is introduced as the largest public database for glaucoma diagnosis.	The study employs extensive validation using cross-validation and cross-testing strategies, comparing results with previous works in the literature. The evaluation is conducted on publicly available glaucoma-labeled databases.	The proposed approach demonstrates high accuracy, with an average AUC of 0.9605, specificity of 0.8580, and sensitivity of 0.9346.	The introduction of the ACRIMA database enhances data availability but may pose challenges in terms of generalizability to diverse populations. Further, the paper does not provide information on the computational resources required for training and testing.
6	Nikos Tsiknakis, Dimitris Theodoropoulos, Georgios Manikis, Emmanouil Ktistakis, Ourania Boutsora, Alexa Berto, Fabio Scarpa, Alberto Scarpa. Dimitrios, Fotiadis. Kostas Marias	Deep learning for diabetic retinopathy detection and classification based on fundus images: A review	140	Computers in Biology and Medicine	7.7	2021	Diabetes mellitus, affecting 463 million people globally and potentially rising to 700 million by 2045, presents a significant public health concern, with diabetic retinopathy (DR) being the most prevalent eye disease among at least one third of diabetics.	The review paper studies the different study done on deep learning algorithms, classfication models, detection models for diabetic retinopathy detection and classififaction under different datasets.	The study utilizes retinal fundus image datasets from various sources, including Kaggle EyePACS, Kaggle APTOS 2019, Messidor & Messidor 2, IDRiD, DDR, , and HRF, for training, testing, and segmentation tasks.	This review comprehensively examines the application of deep learning methods throughout the diabetic retinopathy detection pipeline, encompassing image preprocessing techniques	Some of the technologies listed in this review study are ResNET, AlexNet, VGG16, GoogleNET, InceptionV3, InceptionV4 etc.	This article presents the current state of research regarding the application of deep learning in diagnosing diabetic retinopathy and provides insights and future research directions on what models have been applied	Performance metrices such as sensitivity, specificity, precision, AUC are utilized to evaluated the performance of the methods along with their datasets. Besides segmentation also has een used to find the sensitivity, specificity, IoU etc.	The article reviews various classification models, with the highest accuracy (97.67%) achieved by an Ensemble architecture on the Healthy vs Diseased, 5-class dataset, while the lowest accuracy (45%) was attained using the Inception V4 architecture on the same dataset.	Further improvements can be done regarding performance, interpretability and trustworthiness from ophthalmologists.
7	Liu Li; Mai Xu; Hanruo Liu; Yang Li; Xiaofei Wang; Lai Jiang; Zulin Wang; Xiang Fan; Ningli Wang	A Large-Scale Database and a CNN Model for Attention-Based Glaucoma Detection	133	IEEE Xplore	3.557	2020	The problem addressed in the paper is the inefficiency of existing approaches in removing high redundancy in fundus images for glaucoma detection, which can reduce the reliability and accuracy of glaucoma detection.	The study area and context of the paper is in the field of glaucoma detection using fundus images. The paper specifically focuses on proposing an attention-based convolutional neural network (CNN) for glaucoma detection, called AG-CNN.	Creation of the LAG database, which consists of 5,824 fundus images labeled with positive or negative glaucoma	The study involves the design and implementation of a deep learning model (AG-CNN) with attention prediction, pathological area localization, and glaucoma classification subnets. The model uses attention maps to focus on salient regions and visualize CNN features for glaucoma detection.	It uses convolutional layers, fully connected layers, and multi-scale building blocks in the AG-CNN architecture. It also implements the down-sampling and normalization of attention maps and visualization of pathological areas through guided backpropagation.	AG-CNN method outperforms other methods in glaucoma detection, as evidenced by accuracy, sensitivity, specificity, F2-score, and AUC metrics.	Use of attention prediction loss, feature visualization loss, and glaucoma classification loss for end-to-end training has been implemented.	Over LAG database: 95.3% Over RIM-ONE database: 85.2%	The AG-CNN model's reliance on weakly supervised training for predicting attention maps may compromise its ability to accurately highlight salient regions for glaucoma detection, while the study overlooks potential challenges or biases associated with using fundus images for glaucoma detection, including variations in image quality and acquisition techniques.
8	Ling-Ping Cen, Jie Ji, Jian-Wei Lin, Si-Tong Ju, Hong-Jie Lin, Tai-Ping Li, Yun Wang, Jian-Feng Yang, Yu-Fen Liu, Shaoying Tan, Li Tan, Dongjie Li, Yifan Wang, Dezhi Zheng, Yongqun Xiong, Hanfu Wu, Jingjing Jiang, Zhenggen Wu, Dingguo Huang, Tingkun Shi, Binyao Chen, Jianling Yang, Xiaoling Zhang, Li Luo, Mingzhi Zhang	Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks	133	Nature Communications	16.6	2021	Limitations in detecting retinal fundus diseases using deep learning technology include inadequate image data for training, high intra-class variation, low inter-class variations, overlaps between classes, and challenges in accurately locating lesion areas like exudates, hemorrhages, and cotton-wool spots.	The geographical area of study includes data collected from multiple hospitals in China, as well as data from the United States and public datasets.	The data sources include the Joint Shantou International Eye Center (JSIEC), the Lishui Eye Disease Randomized Controlled Trial System (LEDRS), the Early Treatment Diabetic Retinopathy Study (ETDRS), and the Eye Picture Archive Communication System (EyePACS) in the United States.	A deep learning platform was developed utilizing convolutional neural networks trained on a diverse dataset from global hospitals and public sources, incorporating multi-label classification, dynamic data resampling, and model ensembling for improved accuracy and robustness, with interpretability ensured through modified Class Activation Maps and DeepShap, showing high efficiency in triaging retinal fundus diseases, especially in primary healthcare settings.	Multi-label classification, data resampling, and model ensembling were employed to handle complexities and improve accuracy. Interpretability tools like modified CAM and DeepShap were integrated, enabling clinicians to understand the decision-making process. The platform, trained on diverse datasets, demonstrated high efficiency in detecting fundus diseases.	The proposed platform demonstrated outstanding performance in the classification of 39 different forms of retinal fundus illnesses of referable diabetic retinopathy (DR). When evaluated on a variety of datasets from different hospitals and ethnicities, it showed good generalization ability and robustness across subclasses within big classes.	Algorithm testing and implementation details scrutiny were done as a verification process. This study assessed the platform's performance across multiple datasets, measuring metrics such as accuracy, sensitivity, and specificity. Evaluation encompassed comprehensive testing, focusing on detecting diabetic retinopathy and other fundus diseases while assessing generalization across diverse datasets, ensuring the platform's robustness and generalizability.	__________	The paper acknowledges limitations including insufficient image data for training, challenges in accurately detecting lesions, particularly DR1, and the reliance solely on fundus photography for diagnosis. The model has difficulty in interpreting deep learning decisions and highlights data availability issues, suggesting avenues for future research and improvement in retinal disease detection.

9	Samiksha Pachade,Prasanna Porwal,Dhanshree Thulkar,Manesh Kokare,Girish Deshmukh,Vivek Sahasrabuddhe,Luca Giancardo, Gwenolé Quellec, andFabrice Mériaudeau	Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research	118	Data	3	2021	The world faces difficulties in terms of eye care, including treatment, quality of prevention, vision rehabilitation services, and scarcity of trained eye care experts. Early detection and diagnosis of ocular pathologies to prevent visual impairment, focusing on rare sight-threatening diseases are some of the problems addressed in this paper.	The geographical area of study for this paper is Nanded, Maharashtra, India, specifically at the Eye Clinic, Sushrusha Hospital, and the Center of Excellence in Signal and Image Processing at SGGS Institute of Engineering and Technology.	The data used in this paper were raw and manual annotations acquired from three different digital fundus cameras: TOPCON 3D OCT-2000, Kowa VX-10 α, and TOPCON TRC-NW300, centered on the macula or optic disc that were collected from an eye clinic in Maharashtra, India.	The study was conducted by acquiring retinal images using three different digital fundus cameras centered on the macula or optic disc. The fundus images were captured with the patient sitting upright and a specific distance between the lenses and the examined eye using a non-invasive fundus camera. The images were labeled by two ophthalmologists independently based on the subjects' clinical records and visual fields and consulted when differences in diagnostic assessments were observed.	Digital fundus cameras such as TOPCON 3D OCT-2000, Kowa VX-10 α, and TOPCON TRC-NW300.	The results of the research include the creation of datasets containing images of various retinal pathologies. Additionally, the	The labeling of retinal images by two ophthalmologists independently based on the subjects' clinical records and visual fields. Adjudicated consensus for the labels was obtained through consultation when differences in diagnostic assessments were observed	————	The lack of explicit mention of the accuracy of the model developed in the research and the study focused on rare sight-threatening diseases that may limit the generalizability of the findings to more common ocular pathologies.
10	Rubina Sarki, Khandakar Ahmed, Hua Wang and Yanchun Zhang	Automated detection of mild and multi-class diabetic eye diseases using deep learning	114	Health Information Science and Systems	5.017	2020	Manual analysis of retinal fundus images for detecting diabetic eye diseases is time-intensive, while deep learning, though accurate, faces challenges in mild and multi-class classification, especially in early-stage impairments, which constitute a significant risk of permanent visual impairment in diabetes cases.	The paper focuses on mild and multiclass Diabetic Eye Disease (DED) identification using experimental evaluation of various techniques for improving classification.	The study utilizes open-access platforms like DRISHTI-GS, Messidor, and Messidor-2, alongside publicly accessible retina datasets, with Messidor-2 of the Diabetic Eye Disease (DED) algorithm, Drishti-GS comprising 101 retinal images, and the retina dataset Github containing 100 cataract images in the cataract dataset.	The study employs pretrained CNNs such as VGG16 and InceptionV3 on ImageNet for classifying retinal fundus images in the DED dataset, utilizing methods like under-sampling and over-sampling to mitigate misclassification, while considering both mild multi-class and multi-class classification with four and five classes, respectively.	The studies conducted used Python, Keras library, TensorFlow as a back-end. The default ADAM was the optimiser.	The study compared pretrained CNN models through fine-tuning, evaluated efficiency gain or loss in accuracy, identified top-performing architectures with different optimizers, and focused on mild and multi-class DL algorithms for automated detection of Dry Eye Disease, achieving sensitivity within the range of 85% to 98%.	Various performance improvement techniques were employed, i.e., fine tune, optimization, and contrast enhancement. The range of parameters used across CNN depends on the total hidden layers present on every system model. The possible classifcation enhancement of the DED detection task assessed as a result of the models' proposed customization options. The optimization techniques used for the experiments were RMSprop, SGD, Adagrad, Adam.	Maximum accuracy of 88.3% obtained on the VGG16 model for multi-class classifcation and 85.95% for mild multi-class classifcation.	DED lesion segmentation for enhancing the identifcation of mild disease and moving to more complex and advantageous identifcation of multi-grade diseases.
11	Neha Gour, Pritee Khanna	Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network	101	Biomedical Signal Processing and Control	5.1	2020	Early detection of ophthalmological diseases is a crucial step for the prevention of blindness in patients. Most of the research in the field of ophthalmological image analysis is confined to some selected set of diseases.	The study area of this research is in the field of ophthalmological image analysis, specifically focusing on the automated multi-class multi-label transfer learning-based CNN for the detection of ocular diseases using fundus images	The data used in this research is from the Ocular Disease Intelligent Recognition (ODIR) database, which consists of fundus images of both eyes of patients with various ocular diseases	The study developed a CNN-based system for ocular disease detection, fine-tuning pre-trained models on the ODIR database and evaluating performance metrics such as F1-score and AUC. Training included 100 epochs with early stopping based on validation accuracy, selecting the best-performing model for testing.	The technology used in this research included the implementation of four state-of-the-art pre-trained CNN architectures for finetuning on the Ocular Disease Intelligent Recognition (ODIR) database.	The study highlighted the importance of treating fundus images individually for proper prediction of ocular disease labels and showed that the concatenated input approach led to overfitting. The activation maps correctly classified images of patients with various ocular diseases, providing insights into the working of the architectures.	The models were evaluated online on the Grand Challenge competition website for the ODIR database, and the testing performance was assessed based on F1-score and AUC parameters.	The accuracy of the two-input VGG16 architecture with SGD optimizer in the study was 87.16% for validation accuracy	Enhancements for future research could involve incorporating explainable AI techniques to provide insights into the model's decision-making process . Exploring advanced data augmentation strategies and conducting a more extensive benchmarking against the latest deep learning architectures could further improve the model's performance and robustness.
12	Masum Shah Junayed, Md Baharul Islam, Arezoo Sadeghzadeh, Saimunur Rahman	CataractNet: An Automated Cataract Detection System Using Deep Learning for Fundus Images	64	IEEE Access	3.9	2021	Cataract is one of the most common eye disorders that causes vision distortion. Accurate and timely detection of cataracts is the best way to control the risk and avoid blindness.	In this paper, a novel deep neural network, namely CataractNet , is proposed for automatic cataract detection in fundus images.	In this research, the high-resolution fundus (HRF) image, fundus image registration (FIRE) dataset, ACHIKO-I fundus image dataset, Indian diabetic retinopathy image dataset (IDRiD), etc are used	. The developed CataractNet focused on investigating different layers, activation function, loss function, and optimization algorithms for minimizing the computational cost without sacrificing the model accuracy.	All the experiments are carried out in a computer with the following properties: core i9-10850K CPU, 64GB RAM, and NVIDIA Geforce RTX 2080 super GPU with 3.60 GHz. Image pre-processing, augmentation, and the CNN-based model are all, implemented	An automated cataract detection system, namely CataractNet, based on lightweight deep learning. It focused on investigating different layers, activation function, loss function, and optimization algorithms for minimizing the computational cost.	Various evaluation metrics such as Recall/Sensitivity, Precision, Specificity, F-Score, and Matthews Correlation Coefficient (MCC) are employed to evaluate our model and five pre-trained deep learning models.	Comparing with five pre-trained CNN models, i.e., MobileNet, VGG-16, VGG-19, Inception-v3, and ResNet-50, the CataractNet achieved competitive performance in terms of accuracy (99.13%).	The proposed method can not discriminate the three types of age-related cataracts (nuclear cataracts, cortical cataracts, and PSCs). Besides, it was only proposed for cataract detection and not for grading or finding its exact location, which can be helpful for ophthalmologists.

13	Sadaf Malik,Nadia Kanwal, Mamoona Naveed Asghar,Mohammad Ali A. Sadiq, Irfan Karamat, andMartin Fleury	Data Driven Approach for Eye Disease Classification with Machine Learning	59	Applied Science	2.7	2019	This study states the need for improved methods of diagnosing eye diseases using artificial intelligence and machine learning techniques.	The context of the paper is focused on the development of a data-driven approach for classifying eye diseases using machine learning algorithms. The research primarily revolves around the analysis of diagnostic data related to eye diseases, utilizing machine learning techniques to improve the accuracy of disease classification and prediction.	The data used in the study includes information such as patient symptoms, clinical observations, age, and illness history. The data is derived from electronic health records and structured in a standardized format to facilitate accurate disease diagnosis and prediction using machine learning algorithms.	The research involves developing a comprehensive framework to standardize diagnostic data for disease prediction using machine learning techniques, employing algorithms like Decision Tree, Random Forest, Naive Bayes, and Neural Networks, structured hierarchically by medical professionals with diagnoses relying on ICD-10 coding, and possessing self-learning capabilities for assimilating fresh classifications for symptoms and diagnoses.	This study employs Decision Tree, Random Forest, Naive Bayes, and Neural Network algorithms to analyze patient data and predict disease diagnosis based on various features.	The proposed framework for classifying eye diseases using machine learning algorithms performs well with a prediction rate of over 90% when utilizing tree-based methods. By analyzing patient data based on various features such as age, illness history, and clinical observations, the system demonstrates high accuracy in predicting disease diagnosis.	The proposed model divides the dataset into training and testing sets with a 70:30 split, utilizes 10-fold cross-validation to validate classification algorithms, and evaluates performance using statistical measures including kappa statistics, RMSE, accuracy, precision, recall, and AUC of ROC graphs.	Decision Tree: 85.81% Naive Bayes: 81.53% Random Forest: 86.63% Neural Network: 85.98%	This study presents innovative contributions in data modeling and disease classification but still there are areas for improvement in terms of standardization, data size, and generalizability to enhance the applicability of the findings in broader medical contexts.
14	Nagaraja Gundluru,Dharmendra Singh Rajput,Kuruva Lakshmanna,Rajesh Kaluri,Mohammad Shorfuzzaman,Mueen Uddin,and Mohammad Arifin Rahman Khan	Enhancement of Detection of Diabetic Retinopathy Using Harris Hawks Optimization with Deep Learning Model	52	Computational Intelligence and Neuroscience	3.12	2022	The study addresses the critical issue of diabetic retinopathy, a severe health concern affecting individuals across various age groups. The primary problem is the early detection of diabetic retinopathy to prevent vision loss and potential complications such as retinal detachment and glaucoma-induced blindness.	The study focuses on diabetic retinopathy detection	The Diabetic Retinopathy Debrecen dataset from the UCI machine learning repository is employed in this study. This dataset contains 1151 instances and 20 attributes, capturing essential information for diabetic retinopathy analysis.	The study utilizes a combination of deep learning and machine learning techniques. Specifically, a Deep Neural Network (DNN) is employed, coupled with Principal Component Analysis (PCA) for dimensionality reduction. Additionally, the Harris Hawks Optimization (HHO) algorithm is used for optimizing the classification and feature extraction process.	The study leverages Python for experimental implementation.	The DNN-PCA-HHO model outperforms other prominent ML hybrid models in terms of specificity, precision, accuracy, recall, and sensitivity. The application of PCA alone on DNN and other ML algorithms minimizes training time but slightly degrades performance metrics. The combination of PCA and HHO improves ML algorithm performance while reducing training time.	The HHO algorithm is applied to optimize an objective function relevant to diabetic retinopathy detection. The objective function involves parameters related to the classification model and feature extraction.	DNN-PCA-HHO achieves the highest accuracy at 97%. DNN and SVM-PCA-HHO also show high accuracy, with values of 96.7% and 90.3%, respectively. KNN has the lowest accuracy at 59.4%.	While the results are promising, the study acknowledges the potential limitation of overfitting, particularly in low-dimensional datasets. This suggests that the model's performance might be constrained under certain conditions. The study encourages further research in high-dimensional data and similar health disciplines.
15	Moahmmed Rashid Ahmed, Adil Deniz Duru, Osman Nuri Uçan, Oğuz Baya	An Expert System to Predict Eye Disorder Using Deep Convolutional Neural Network	44	Academic Platform Journal of Engineering and Science	——	2020	The problem stated in the paper is the complexity of detecting the severe eye disease, glaucoma, using biomedical information systems. This disease can lead to significant damage to the optic nerves, yet there is currently no automated expert system utilizing deep learning techniques for its detection.	The context of this paper is focused on developing an expert system for detecting eye disorders, specifically Glaucoma, using Deep Convolutional Neural Network (DCNN) technology	Kaggle	The project aims to enhance early glaucoma detection through a MATLAB-implemented Deep Convolutional Neural Network (DCNN) expert system, conducted in phases including fundus image processing, feature extraction using DCNN, data organization for analysis, model construction, data partitioning, and model training.	This paper utilizes a Deep Convolutional Neural Network (DCNN) in MATLAB. Image processing techniques such as Grayscale, B&W, Complement, Robert, Resize, and Power Transform are employed. The system is trained on a Kaggle dataset using Neural Network Back Propagation.	The study achieved Deep Convolutional Neural Network (DCNN) achieving 92.78% accuracy in Glaucoma detection from processed fundus images,	Through dataset splitting, training-validation-testing iterations, and accuracy evaluation, the study ensures the model's reliability and effectiveness in early diagnosis of eye disorders.	The accuracy achieved in this paper for detecting Glaucoma using the proposed model is 92.78%.	The paper emphasizes the need for expert systems in this detection glaucoma. Limitations include data availability constraints, potential generalization issues, and the scope of the expert system. Enhancements proposed in the paper involve expanding the dataset, optimizing the model, and conducting clinical validation studies to validate real-world effectiveness.
16	Shuang Yu, Di Xiao , Shaun Frost , Yogesan Kanagasingam	Robust optic disc and cup segmentation with deep learning for glaucoma detection	29	Computerized Medical Imaging and Graphics	7.422	2019	The problem addressed in the paper is the early detection of glaucoma, which is a leading cause of irreversible vision loss worldwide.	The study focuses on general problem of glaucoma detection	The model was trained on the RIGA dataset, and tested on DRISHTI-GS and RIM-ONE datasets. Additionally, fine-tuning was performed on two databases, DRISHTI-GS, and RIM-ONE.	The paper describes the use of a modified U-Net architecture that combines a pre-trained ResNet-34 model as encoding layers with classical U-Net decoding layers for optic disc and cup segmentation.	The technology used involves the combination of pre-trained ResNet and U-Net architectures. The model was implemented for segmentation tasks related to optic disc and cup.	The model achieved promising results with an average dice value of 97.31% for disc segmentation and 87.61% for cup segmentation on the RIGA dataset. It also demonstrated comparable performance on DRISHTI-GS and RIM-ONE datasets without re-training.	Fine-tuning was performed on two databases, and performance metrics such as dice values were used.	The reported average dice values are 97.31% for disc segmentation and 87.61% for cup segmentation on the RIGA dataset. Fine-tuned models achieved 97.38% disc dice and 88.77% cup dice on DRISHTI-GS, and 96.10% disc dice and 84.45% cup dice on the RIM-ONE database.	————

17	P. Glaret subin, P. Muthukannan	Optimized convolution neural network based multiple eye disease detection	24	Computers in Biology and Medicine	7.7	2022	People in the world are visually challenged mostly due to the age-related eye diseases such as age-related macular degeneration (AMD), cataract, diabetic retinopathy (DR) and glaucoma. These diseases lead to blindness if not diagnosed at an early stage.	This study focuses on multiclass eye disease detection.	The proposed CNN-based multiple disease detection was tested with the online dataset, namely Ocular Disease Intelligent Recognition (ODIR).	The pre-processed images were fed to a convolution neural network which was optimized using a flower pollination optimization algorithm (FPOA) for feature extraction. Hyperparameters were optimized using FPOA for training the CNN. The CNN output was fed to a Multiclass Support Vector Machine (MSVM) classifier for the classification of the type of disease.	CNN was used train the images and MSVM classifier classifed the dataset.	The study found that pre-processing of fundus image using maximum entropy transformation, which provided better quality and more information relating to the image. The optimization of CNN hyperparameters using FPOA, which produced results with greater accuracy in the diagnosis of multiple diseases and classification of multiple eye diseases using a multiclass SVM classifier.	7.5% increased validation accuracy compared to the non-optimized CNN.	The proposed model performance was analysed with the other optimized models which yielded the best performance in terms of precision, accuracy, specificity, recall, and F1 score of 98.30%, 95.27%, 95.21%, and 93.3%, respectively.	This paper used only limited dataset for only age related eye diseases. Further classes can be added for multiclass multilabel classification of these type of diseases.
18	C.-H. Huck Yang, Jia-Hong Huang, Fangyu Liu, Fang-Yi Chiu, Mengya Gao, Weifeng Lyu, I-Hung Lin M.D, Jesper Tegner	A Novel Hybrid Machine Learning Model for Auto-Classification of Retinal Diseases	20	Semantic Scholar	2	2018	The problem addressed in this study is the limited accessibility to specialized medical expertise in diagnosing retinal diseases, particularly in remote areas. The aim is to develop an automated system that can assist in the diagnosis of such diseases, potentially bridging the gap in healthcare access.	The focus of this study is on the automatic diagnosis of retinal diseases.	The study uses a new clinical retina label collection that incorporates images and text information of 32 different retinal diseases classes.	The study employs a hybrid approach combining support vector machines (SVM) and deep neural networks (DNNs) for automatic diagnosis. The integration of SVM and DNNs is detailed, and the EyeNet model is introduced as a result of this hybridization.	The technology used in this study includes support vector machines, deep neural networks, and specifically, the EyeNet model.	The developed hybrid model, EyeNet, achieves a diagnosis accuracy of 89.73%.	The dataset is split into three subsets: 70% for training, 10% for validation, and 20% for testing.Cross-entropy loss is used for evaluating the training process of the U-Net model.	The developed hybrid model, EyeNet, achieves a diagnosis accuracy of 89.73%.	The study hints at potential future directions, such as Visual Question Answering based on retinal images, and highlights the potential application of the model in remote rural areas.
19	Jeban Chandir Moses, Sasan Adibi, Nilmini Wickramasinghe, Lemai Nguyen, Maia Angelova, andSheikh Mohammed Shariful Islam	Smartphone as a Disease Screening Tool: A Systematic Review	17	Sensors	4	2022	COVID-19 has restricted hospital visits for screening and other healthcare services resulting in the disruption of screening for cancer, diabetes, and cardiovascular diseases.	This paper is focused on evaluating mobile health (mHealth) applications for disease screening in the adult population	The data used in this study were obtained from articles published in online databases such as Medline Complete, Web of Science, Embase, and Proquest . Research articles on health and medical research published..	The study was conducted through a systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.	Smartphone technologies, including built-in sensors and wireless capabilities, were utilized for disease screening and monitoring, showing potential for cost-effective and high-quality results	The review of nine studies highlights the potential of smartphone applications in improving healthcare outcomes for chronic diseases, emphasizing user acceptance and technology usability.	The study utilized a systematic review following PRISMA guidelines to evaluate the effectiveness of smartphone applications for disease screening and monitoring, highlighting the utility of smartphones with built-in sensors and wireless capabilities as cost-effective tools for this purpose.	————	The study had small sample sizes in the studies reviewed and the possibility of bias as only one reviewer performed article screening and data extraction . The limited timeframe i.e. upto September 2020 can potentially miss newer developments. Lack of standard models for evaluating user perception of information systems, hindering meta-analysis
20	Shotaro Asano, Ryo Asaoka, Hiroshi Murata, Yohei Hashimoto, Atsuya Miki, Kazuhiko Mori, Yoko Ikeda, Takashi Kanamoto, Junkichi Yamagami, Kenji Inoue	Predicting the central 10 degrees visual field in glaucoma by applying a deep learning algorithm to optical coherence tomography images	12	Scientific Report	4.6	2021	The study aims to predict visual field (VF) in the central 10 degrees for glaucoma patients using a convolutional neural network (CNN) trained with optical coherence tomography (OCT) images. The challenge is to enhance prediction accuracy by adjusting values with Humphrey Field Analyzer (HFA) 24-2 test results.	The research involves multiple medical institutions in Japan, including the University of Tokyo Hospital, Inouye Eye Hospital, JR Tokyo General Hospital, Hiroshima Memorial Hospital, Osaka University Hospital, Kyoto Prefectural University of Medicine Hospital, and Oike-Ganka Ikeda Clinic.	The training dataset includes 648 eyes from 358 subjects, with OCT and HFA 10-2 tests conducted within a 3-month period. The data encompass 90 normal eyes and 558 eyes with primary open-angle glaucoma (POAG). Institutional databases and patient consent facilitated data usage.	Two CNN models, VGG19 and ResNet152, were employed. OCT-measured macular GCC, RNFL, and OS + RPE thicknesses were used for training. Transfer learning, utilizing pre-trained models on the ImageNet dataset, addressed the need for a large training dataset.	Convolutional Neural Networks (CNNs), specifically VGG19 and ResNet152, were implemented for the predictive modeling. Optical Coherence Tomography (OCT) provided retinal thickness data, and transfer learning techniques leveraged pre-trained models.	The models demonstrated prediction accuracy, with mean absolute error (MAE) values initially around 10 dB. Significantly improved accuracy was achieved by adjusting predictions using TD values from the innermost four points of the HFA 24-2 test.	The validation method involved testing the performance of the VGG and ResNet models using a separate testing dataset. Evaluation was done using the mean absolute error (MAE) approach.	Initial MAE values were around 10 dB, but after adjustment with HFA 24-2 test data, the MAE reduced to an average of 5.5 dB, indicating improved accuracy.	A potential limitation is the relatively larger prediction errors compared to a previous study, possibly due to a wider range of disease stages in the testing dataset. The study suggests the need for further work to enhance prediction accuracy, potentially through a larger training dataset and continued advancements in deep learning techniques.

21	Behnam Askarian; Peter Ho; Jo Woon Chong	Detecting Cataract Using Smartphones	9	IEEE Journal of Translational Engineering in Health and Medicine	3.4	2021	The significance of cataracts lies in their potential to lead to blindness if not identified and addressed promptly. Therefore, there is a critical need for early detection methods to mitigate the progression of cataracts and prevent vision impairment.	The study focuses on the feasibility of detecting cataracts using smartphones.	Eye images were captured using iPhone X, iPhone 6, and iPhone 11 Pro smartphone cameras. Images from an Axis Scientific 7-Part Human Eye model were used to emulate healthy and diseased eyes.	Images were preprocessed using a median filter and watershed transformation to remove background noise. A novel luminance transformation algorithm was applied to extract lens image features. Support Vector Machines (SVM) were used as the classification method.	A re-targetable application platform is mentioned for implementing the proposed method on any smartphone.	The proposed method achieved a diagnosis accuracy of 96.6%, specificity of 93.4%, and sensitivity of 93.75% for detecting cataracts.	10-fold cross-validation and SVM were employed for the classification task. Statistical hypothesis tests were conducted to validate the difference between healthy and diseased luminance values.	The proposed method achieved an accuracy of 96.6% for detecting cataracts.	Further clarification is needed on the re-targetable application platform mentioned for implementation on various smartphones.
22	Sunita Yadav and Jay Kant Pratap Singh Yadav	Automatic Cataract Severity Detection and Grading Using Deep Learning	5	Journal of Sensors	1.9	2023	Cataracts represent a significant global cause of vision loss, and timely detection is critical for preventing further impairment. Existing challenges include inadequate medical care and high treatment costs, hindering timely intervention for affected individuals.	The study area is not explicitly mentioned, but it seems to be applicable globally. The context is the need for improved and accessible cataract diagnosis, especially in regions with limited access to qualified ophthalmologists.	The study utilizes fundus images sourced from various open-source databases available on the internet.	The proposed method involves a combination of deep learning (DL) and 2D-discrete Fourier transform (DFT) applied to fundus images. A quality selection module, preprocessing steps, and augmentation techniques are incorporated. Convolutional Neural Network (CNN) is employed for feature extraction, and a softmax classifier is used for classification.	The implementation is carried out using MATLAB R2019a, including Image Processing, Neural Network Toolboxes, and Deep Learning Toolboxes.	The proposed system outperforms existing methods, achieving a four-class accuracy of 93.10%. Comparative analyses demonstrate superior performance against standard SVM, AlexNet-softmax, VGGNet-softmax, and ResNet-softmax. The CNN-based approach, incorporating 2D-DFT spectrograms, proves more effective in feature extraction and classification.	The performance evaluation includes metrics such as accuracy, sensitivity, specificity, precision, and F1-score. The study employs a set of criteria based on true positive, true negative, false positive, and false negative classifications.	The proposed method achieves an accuracy of 93.10%.	The study acknowledges limitations in training and testing on a limited dataset. Real-time and larger datasets are suggested for future evaluations.
23	Asif Nawaz; Tariq Ali; Ghulam Mustafa; Muhammad Babar; Basit Qureshi	Multi-Class Retinal Diseases Detection Using Deep CNN With Minimal Memory Consumption	2	IEEE Access	3.9	2023	This research addresses the high memory and CPU consumption challenges associated with existing techniques, particularly the U-Net Segmentation	The study is centered around medical image classification, specifically the identification and classification of retinal diseases.	The research employs the EyeNet dataset, which comprises 32 classes of retinal diseases.	The proposed approach involves the development and evaluation of a convolutional neural network (CNN) model. The model is trained on the EyeNet dataset using Python and Keras.	The implementation is carried out in Python using the Keras library. The research leverages an Intel Core i5-7200CPU at 2.70GHz for the implementation and Google GPU for training.	The proposed CNN model demonstrates better performance in terms of memory management and accuracy compared to U-Net Segmentation.	The model is evaluated based on precision, recall, and accuracy. The validation is performed on a standard benchmark dataset, and the results are compared across different numbers of epochs.	The proposed CNN model achieves an accuracy of 95% on the EyeNet dataset.	Absence of a clear strategy for addressing evolving challenges in data quality and bias management.
24	Ayesha Shoukat, Shahzad Akbar, Syed Ale Hassan, Sajid Iqbal, Abid Mehmood, and Qazi Mudassar Ilyas	Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach	2	Diagnostics	3.9	2023	Early detection of glaucoma is critical to prevent vision loss, but the current manual assessment methods are skill-intensive, time-consuming, and often lead to delayed diagnoses, particularly in the elderly population.	The study focuses on glaucoma detection using deep learning techniques.	Four publicly available datasets are used: G1020, DRISHTI-GS, ORIGA, and RIM-ONE.	Image Processing has been done for conversion of training images to greyscale for better clarity. Techniques such as flipping, rotation, cropping, and scaling has been applied to increase the dataset's size and enhance model robustness. Similarly, ResNet-50 architecture was used.	The research employs deep learning techniques, specifically using the ResNet-50 architecture for glaucoma detection.	The proposed model achieved promising results, with an accuracy of 98.48%, sensitivity of 99.30%, specificity of 96.52%, AUC of 97%, and an F1-score of 98% on the G1020 dataset.	The evaluation method involves testing the model on multiple datasets (G1020, DRISHTI-GS, ORIGA, and RIM-ONE) to validate its generalizability and effectiveness.	The accuracy of the proposed model on the G1020 dataset is reported as 98.48%.	The study primarily focuses on greyscale fundus images, potentially limiting its applicability to multimodal imaging approaches.
25	Sushma K Sattigeri, Harshith N., Dhanush Gowda N., K. A. Ullas, Aditya M. S.	Eye Disease Identification using Deep Learning	2	International Research Journal of Engineering and Technology (IRJET)	8.226	2022	The increasing rate of diabetes worldwide has raised people's vulnerability to age-related vision issues. On sites like Google, there aren't many specialised apps for early eye disease detection	The proposed model focuses on visually observable symptoms to analyze and categorize four eye diseases: crossed eyes, bulging eyes, cataracts, uveitis, and conjunctivitis.	The authors collected information on cataract, bulging eyes, and crossed eyes symptoms from Kaggle and small portion from the internet.	The study utilizes digital image processing techniques and convolutional neural networks with three layers, each comprising 3x3 filters and linear activation functions, for segmentation and morphology in deep learning methods.	The model is put together using the Adam optimizer. Some core libraries like OpenCV, keras, TensorFlow, pandas, NumPy are also used in this paper.	The study builts a deep learning model which classifies between normal eye and a diseased eye. The model will be able to guide the users to know their eye condition (for specified eye diseases namely crossed eye, bulged eye, conjunctivitis and cataract)..	In this model, disease classes including cataract, uveitis/conjunctivitis, crossed eyes, and bulging eyes are compared, with two different models trained on images depicting one eye and two eyes respectively, wherein the latter predicts crossed eyes and bulging eyes while the former forecasts cataract and conjunctivitis/uveitis.	For the dataset with images of a single eye, the authors were able to achieve a precision of 96%, and for the dataset with images of both eyes, a precision of 92.31% was able to be achieved.	Based on the paper, it might be possible to create a web-based platform and mobile application that uses a customised Deep Learning model to recognise diseases or problems with the eyes that are visible from the outside using uploaded eye images.