Glaucoma Detection using Deep Learning Techniques

Submitted in partial fulfillment of the requirement for the award of Degree of Bachelor of Technology in Computer Engineering Discipline

Submitted To

SVKM's NMIMS,
Mukesh Patel School of Technology Management & Engineering,
Shirpur Campus (M.H.)

Submitted By:

Kushal Jha – 70021119030
Naman Gokharu – 70021119016
Rohan Mathur – 70021119040

Under The Supervision Of:

Prof. Sachin Bhandari
(Department& Designation)

DEPARTMENT OF COMPUTER ENGINEERING
Mukesh Patel School of Technology Management & Engineering

SESSION: 2022-23

CERTIFICATE

This is to certify that the work embodies in this Project entitled **"Glaucoma Detection using Deep Learning Techniques"** being submitted by

Kushal Jha – 70021119030
Naman Gokharu – 70021119016
Rohan Mathur – 70021119040

for partial fulfillment of the requirement for the award of "**Bachelor of Technology** in **Computer Engineering"** discipline to "SVKM's NMIMS, Mumbai (M.H.)" during the academic year 2022-23 is a record of bonafide piece of work, carried out by him under my supervision and guidance in the **"Department of Computer Engineering", MPSTME, Shirpur (M.H.).**

APPROVED & SUPERVISED BY:

Prof. Sachin Bhandari
(Assistant Professor, Computer Engg.)

FORWARDED BY

Dr. N.S. Choubey
H.O.D., Computer
MPSTME, Shirpur Campus

Dr. Venkatadri M.
Associate Dean,
MPSTME, Shirpur Campus

DEPARTMENT OF COMPUTER ENGINEERING
Mukesh Patel School of Technology Management & Engineering

CERTIFICATE OF APPROVAL

The Project entitled **“Glaucoma Detection using Deep Learning Techniques”** being submitted by

Kushal Jha	**– 70021119030**
Naman Gokharu	**– 70021119016**
Rohan Mathur	**– 70021119040**

has been examined by us and is hereby approved for the award of degree **“Bachelor of Technology in Computer Engineering Discipline”**, for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein, but approve the project only for the purpose for which it has been submitted.

(Internal Examiner) **(External Examiner)**

Date: **Date:**

DEPARTMENT OF COMPUTER ENGINEERING
Mukesh Patel School of Technology Management & Engineering

DECLARATION

We,

Kushal Jha

Naman Gokharu

Rohan Mathur

The students of **Bachelor of Technology** in **Computer Engineering discipline**, S**ession: 2022-23, MPSTME, Shirpur Campus,** hereby declare that the work presented in this Project entitled **"Glaucoma Detection using Deep Learning Techniques"** is the outcome of our work, is bonafide and correct to the best of our knowledge and this work has been carried out taking care of Engineering Ethics. The work presented does not infringe any patented work and has not been submitted to any other university or anywhere else for the award of any degree or any professional diploma.

(Kushal Jha with Signature)
SAP ID.: 70021119030
(Naman Gokharu with Signature)
SAP ID.: 70021119016
(Rohan Mathur with Signature)
SAP ID.: 70021119040

Date:

ACKNOWLEDGEMENT

After the completion of this Major Project work, words are not enough to express my feelings about all those who helped me to reach my goal; feeling above this is my indebtedness to The Almighty for providing me this moment in life.

It's a great pleasure and moment of immense satisfaction for me to express my profound gratitude to **Prof. Sachin Bhandari, Assistant Professor**, Computer Engineering Department, MPSTME, Shirpur, whose constant encouragement enabled me to work enthusiastically. Their perpetual motivation, patience and excellent expertise in discussion during progress of the project work have benefited me to an extent, which is beyond expression. Their depth and breadth of knowledge of Computer Engineering field made me realize that theoretical knowledge always helps to develop efficient operational software, which is a blend of all core subjects of the field. I am highly indebted to them for their invaluable guidance and ever-ready support in the successful completion of this project in time. Working under their guidance has been a fruitful and unforgettable experience.

We express my sincere thanks and gratitude to Dr. N.S. Choubey, Head of Department, Computer Engineering Department, MPSTME, Shirpur, for providing necessary infrastructure and help to complete the project work successfully.

We also extend my deepest gratitude to Dr. Akshay Malhotra, Director, SVKM'S NMIMS, Shirpur Campus and Dr. Venkatadri M., Associate Dean, SVKM'S NMIMS, Shirpur Campus for providing all the necessary facilities and true encouraging environment to bring out the best of my endeavors.

We sincerely wish to express my grateful thanks to all members of the staff of computer engineering department and all those who have embedded me with technical knowledge of computer technology during various stages of B.Tech. Computer Engineering.

We would like to acknowledge all my friends, who have contributed directly or indirectly in this Major Project work.

The successful completion of a Major Project is generally not an individual effort. It is an outcome of the cumulative effort of several persons, each having their own importance to the objective. This section is a vote of thanks and gratitude towards all those persons who have directly or indirectly contributed in their own special way towards the completion of this project.

Naman Gokharu
Kushal Jha
Rohan Mathur

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 PURPOSE

The purpose of the "Glaucoma Detection using Deep Learning Methods" project is to develop a reliable, accurate, and efficient system for the early detection of glaucoma using deep learning methods. Glaucoma is a progressive disease that affects the optic nerve, leading to irreversible vision loss if left untreated. Therefore, early detection of glaucoma is crucial to prevent vision loss and improve patient outcomes.

The project aims to develop a deep learning-based system that can accurately detect glaucoma in retinal images. The system will use convolutional neural networks (CNNs) or other deep learning architectures to analyze retinal images and identify signs of glaucoma, such as optic nerve damage or changes in the retina.

To achieve this goal, the project will collect a large dataset of retinal images with and without glaucoma. The dataset will be pre-processed to remove noise and enhance the features that are relevant to glaucoma detection. The deep learning model will be trained on this dataset using supervised learning techniques.

The project's objectives are to develop a deep learning-based model for glaucoma detection, evaluate the performance of the model on a test dataset of retinal images, and compare the performance of the developed model with existing methods for glaucoma detection. The goal is to provide healthcare professionals with a user-friendly interface for the glaucoma detection system, which can help improve the accuracy and efficiency of glaucoma diagnosis.

The successful development of the glaucoma detection system can have significant benefits for patients, as early detection and treatment can prevent irreversible vision loss. The system can also benefit healthcare professionals by providing them with a reliable and efficient tool to diagnose glaucoma, which can lead to better patient outcomes and reduced healthcare costs. Additionally, the project can contribute to the field of medical imaging by demonstrating the potential of deep learning methods in improving the accuracy and efficiency of medical diagnosis.

1.2 SCOPE

The scope of our project is difficult to define clearly. This holds true for all Deep Learning and Artificial Intelligence projects. A project in this domain may find scope of improvements in another project after even little time of its implementation. This is because technology is evolving extremely quickly. Our project may be a practical and applicable solution for a few years if implemented successfully but it may be affected by updates and evolutions in Machine Learning and Artificial Intelligence. The scope of the project includes:

i. Collecting a large dataset of retinal images with and without glaucoma.
ii. Pre-processing the retinal images to remove noise and enhance features.
iii. Developing a deep learning-based model for glaucoma detection, using convolutional neural networks (CNNs) or other deep learning architectures. Training the model on the dataset of retinal images with and without glaucoma.
iv. Evaluating the performance of the developed model on a test dataset of retinal images.
v. Comparing the performance of the developed model with existing methods for glaucoma detection, such as rule-based systems or machine learning models.
vi. Developing a user-friendly interface for the glaucoma detection system, allowing healthcare professionals to input retinal images and obtain a diagnosis.

1.3 OVERVIEW

Glaucoma is the third most common cause of blindness globally with approximately 7% of total cases of blindness worldwide. It is a chronic eye condition that progressively impairs vision and harms the optic nerve. Attacks from glaucoma typically do not occur until the disease is somewhat advanced even though glaucoma is treatable and its progression can be delayed. Glaucoma can be identified early by using digital fundus images (DFI). The possibility of acquiring DFIs in a noninvasive manner, which is suitable for large-scale glaucoma screening, has made DFI a preferred modality. In a glaucoma screening program, an automated system analyzes whether an image exhibits any glaucoma symptoms. The system will send any suspicious images to ophthalmologists for additional evaluation.

CHAPTER 2

LITERATURE SURVEY

Glaucoma is a silent killer of the optic nerve. Using machine learning, an intelligent prediction system is designed to predict glaucoma. Researchers trained the dataset with Linear support vector machine (LSVM), Medium Gaussian support vector machine (MGSVM) and Bayesian Optimization SVM (BOSVM) and predicted the output classification. In this proposed system, BOSVM is incorporated into the local real time diabetic people dataset.

Based on the results obtained, BOSVM appears to be a best fit classifier with the accuracy of 96.6%, area under curve of 0.83 in the training set, and accuracy of 97.4% with 0.943, 0.951 of sensitivity and specificity. The accuracy of 95.4% and 94.3% were achieved using LSVM and MGSVM respectively. [1]

Glaucoma is a chronic, irreversible eye disease that deteriorates vision and quality of life. In this paper, a deep learning (DL) architecture with convolutional neural network for automated glaucoma diagnosis was developed and implemented which can infer a hierarchical representation of images to discriminate between glaucoma and non-glaucoma patterns for diagnostic decisions. The proposed Deep Learning Architecture has six layers, four convolutional layers and two fully connected layers. The output of these layers is then passed to the softmax classifier for glaucoma detection. [1]

Model was implemented using the ORIGA and SCES dataset. The ORIGA dataset with clinical glaucoma diagnosis consists of 168 glaucoma and 482 normal fundus images whereas the SCES dataset contains 1676 fundus images, and 46 images are glaucoma cases. The performance measure includes area under the curve (AUC) of receiver operating characteristic curve (ROC) to evaluate the performance of glaucoma diagnosis. The ROC is plotted as a curve which shows the tradeoff between sensitivity TPR (true positive rate) and specificity TNR (true negative rate). The results show the area under the curve (AUC) of the receiver operating characteristic curve in glaucoma detection at 0.831 and 0.887 in the two databases, much better than state-of-the-art algorithms. [2]

The optic cup and disc can be separated from retinal fundus pictures using the Field of Experts (FoE) model that the researchers suggested in this paper. The model has the subsequent stages, a.) approximate optic disc localization, b) generating probability maps for the disc and, c) segmenting the disc into the cup and rim region. Without using manually crafted features, FoE develops a previous model of the cup and neuro-retinal rim region through filters. Responses from the FoE filter are combined with a Random Forest classifier, which generates probability maps for the test image and then separates the cup from the rim. The performance of this method is superior to the existing State-of-the-Art, according to experimental results.[3]

Researchers have used the DRISHTI-GS dataset, which consists of 50 patient pictures captured with a 30-degree FOV with a resolution of 2896 × 1944. With 40 training images and 10 test images in each fold, researchers used a 5-fold cross validation technique. The f-score, the overlap measure, and the boundary error were utilized as performance measures. [3]

The presence of the cup in the disc is a strong indicator of glaucoma, a method to detect glaucoma was presented here by properly detecting the location of the cup. The disc segmentation was done by thresholding, the vessel segmentation was done using edge detection, and for the cup segmentation it was presented a method that uses the vessels and the cup intensities. A set of Fundus images obtained from the Center of Prevention and Attention of Glaucoma in Bucaramanga. The CDR was calculated by using the following formula CDR=Area of cup/Area of Disc. If the CDR is greater than 0.6 then the patient has Glaucoma. The accuracy for this method was 88.50%. [4]

The student’s t-test and PCA were two feature selection techniques that were compared in this study for the purpose of glaucoma screening using fundus pictures. When compared to independent output from various channels, the proposed improved decision-making strategy on classifier output based on a combination of R, G, B, and Grey components produced a good classification accuracy. The sub-band pictures were split using 2D EwT, and correntropy was employed to extract features. For feature selection and ranking, the student’s t-test and z-Score normalization were used. PCA was used to reduce the dimensionality. Then, RBF is used as the kernel and LS-SVM is used as the classifier.[4]

The final output is strengthened by the proposed feature selection technique and bit-wise OR operation at the classifier's output from digital retinal fundus images. The accuracy of 99% obtained using the bitwise OR operation, which is comparable to the accuracy so far achieved by other studies, makes this technique of medicinal potential. [5]

The glaucoma detection was carried out by making use of the CNN model. In the presented mechanism, a dropout mechanism is also applied. The overall architecture has six layers. The four layers are the convolutional layers, and the last two layers are connected fully. The output obtained from the very last layer is given to the classifier for the detection of the glaucoma disease.[5]

For the ORIGA dataset, the proposed method's AUC obtained was .822, while for the SCES, it had a value of.882. In contrast, the state-of-the-art methodology gives an AUC value of 0.809 for the ORIGA dataset and 0.859 for the SCES dataset. [6]

Initially CNN was used on the raw dataset of images to extract the important features from the image and then these images were trained by the SVM algorithm which classified it into normal and abnormal categories. Confusion matrix was used to measure the performance of the algorithm. It gave the following results- accuracy, specificity and sensitivity of 88.2%, 90.8%, and 85%, respectively which clearly demonstrate that the proposed deep learning method is promising in automatic diagnosis of glaucoma and at considerably lower computational cost. [7]

The area under receiver operating characteristic curve (AUC) of a 10-fold cross validation (CV) was used to evaluate the models. 10-fold CV AUCs of the CNNs were 0.940 for color fundus images, 0.942 for RNFL thickness maps, 0.944 for macular GCC thickness maps, 0.949 for disc RNFL deviation maps, and 0.952 for macular GCC deviation maps. Random Forest combining the five separate CNN models improved the 10-fold CV AUC to 0.963. Therefore, the machine learning system described here can accurately differentiate between healthy and glaucomatous subjects based on their extracted images from OCT data and color fundus images. [8]

A supervised learning method, Support Vector Machine (SVM) was used to classify the retinal fundus images. The classifier will take these images and formation of the presence/ absence of glaucoma by calculating the Cup-to-disc ratio (CDR) value in the respective image. Here different image processing techniques like Image preprocessing, Image enhancement, feature extraction, Image classifications are carried and compared their training accuracy for various algorithms by using various retinal images among the comparisons of results support vector machine performs better accuracy (93%) and computational efficiency of SVM is good. [9]

Literature Review Table:

Paper	Author	Journal & Year of Publication	Methods Used
[1]	Eswari MS, Balamurali S.	International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE 2021	LSVM, MGSVM and SVM (BOSVM)
[2]	Chen X et al.	37th annual international conference of the IEEE engineering in medicine and biology society (EMBC) 2015	The adopted DL structure consists of six layers: four convolutional layers and two fully-connected layers
[3]	Mahapatra D, Buhmann JM	12th International al Symposium on Biomedical Imaging (ISBI) IEEE 2015	Random Forest Classifier
[4]	Divya L, Jacob J.	8th International Conference on Advances in Computing and Communication (ICACC-2018)	Least-squares support-vector machines (LS-SVM)
[5]	Saxena A. et al.	International Conference on Electronics and Sustainable	Convolutional Neural Network

		Communication Systems (ICESC), IEEE 2020	
[6]	Carrillo J et al.	XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA) 2019 IEEE	Cup to Disc Ratio
[7]	Guangzhou An, et al.	Journal of healthcare engineering. 2019	Convolutional Neural Network, Random Forest Classifier
[8]	Baidaa Al-Bander, et al.	14th International Multi-Conference on Systems, Signals & Devices (SSD), IEEE 2017	Convolutional Neural Network, Support Vector Machine
[9]	J. Civet-Masot, et al.	IEEE Access. 2020	Convolutional Neural Network, Ensemble Learning
[10]	Sarkar D, Das S.	InAdvances in Optical Science and Engineering 2017	Cup to Disk ratio
[11]	Rao PV, et al.	2nd International Conference on Nanomaterials and Technologies (CNT 2014)	Multilayer Perceptron (MLP) and Back Propagation (BP) neural network (MLP-BP ANN), Naive Bayes Classifier
[12]	Abbas Q.	(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 6, 2017	Convolutional Neural Network, Deep Belief network(dbn)
[13]	Li L, et al.	In Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition (CVPR), 2019	Attention-based CNN (AG-CNN)

[14]	Thangaraj V, Natarajan V.	In2017 International Conference on Intelligent Support vector machine Computing and Control Systems (ICICCS) 2017 Jun 15 (pp. 394-399). IEEE.	Support vector machine

Table 1: Literature review summary

CHAPTER 3

PROBLEM DEFINITION & PROPOSED SOLUTION

3.1 PROBLEM STATEMENT

The pressure inside the eye, known as the intraocular pressure, keeps the eye hard and almost spherical in shape (IOP). The IOP of a human eye fluctuates with age and typically ranges from 12 mm to 22 mm of mercury. The fragile fibers that make up the optic nerve make it vulnerable to variations in intraocular pressure. Blood pressure changes and reduced blood flow to the nerve fibers both cause injury to the nerves.
This optic nerve is afflicted by the chronic disease glaucoma, which lengthens the optic cup and causes increasing damage to it. It internally comprises a number of disorders with like traits, thus its prompt detection is sufficient to determine the likelihood of the retinal abnormality.

Some of the high-risk factors associated with glaucoma are (i) age (ii) increased eye pressure (iii) thin cornea (iv) hypertensive (v) diabetic patients (vi) history of injuries to eyes (vii) a family history of glaucoma (viii) nearsightedness (ix) cardiovascular diseases (x) steroid consumption (xi) rate of blood pressure (xii) severe anemia. As a result, it is crucial to have a quick and reliable approach for diagnosing the development of glaucoma because most individuals don't realize they have the disease until significant damage has been done.

3.2 PROPOSED SOLUTION

In the proposed solution, we will first remove the unwanted columns from the csv file, and will keep only the columns:

i. Image name, Retinopathy grade, type

After this we will map the five categories to final two categories:

ii. DR (Diabetic Retinopathic), No-DR (Not Diabetic Retinopathic)

Then the images will be preprocessed (shuffled, rescaled, and normalized). Then images will be split to train, test and validation set. After this model will be built using the five different models listed below:

a) CNN
b) ResNet50
c) VGG16
d) CNN+ResNet50
e) CNN+VGG16

Each of the model will be trained on 50 epochs, and their accuracy and loss will be compared.

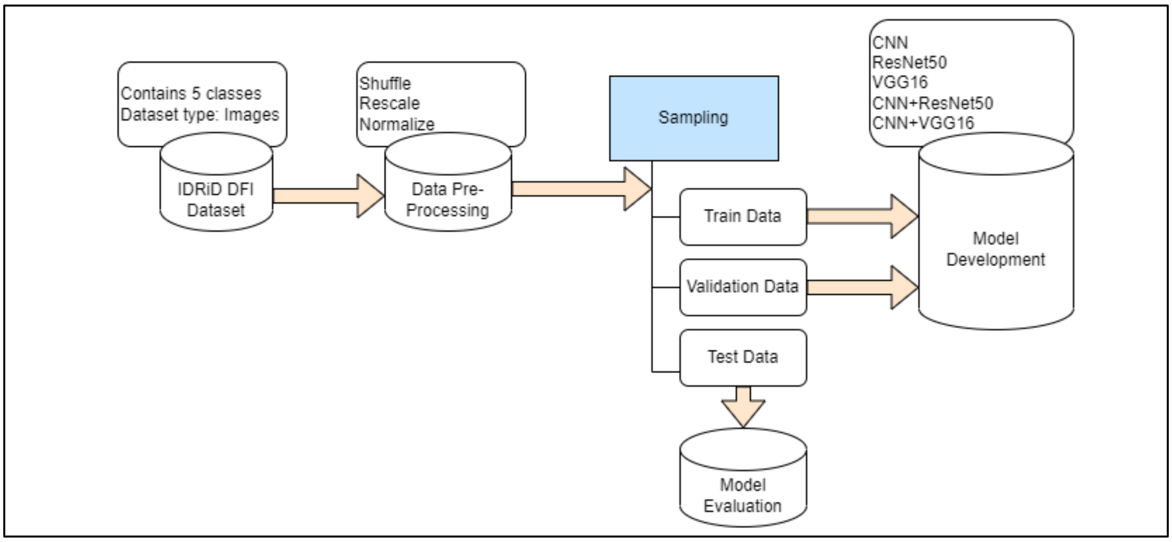


Fig 1: Proposed Solution

CHAPTER 4
DESIGN

4.1 ARCHITECTURAL DIAGRAM

An architectural diagram is a diagram of a system that is used to abstract the overall outline of the software system and the relationships, constraints, and boundaries between components. It is an important tool as it provides an over- all view of the physical deployment of the software system and its evolution roadmap.

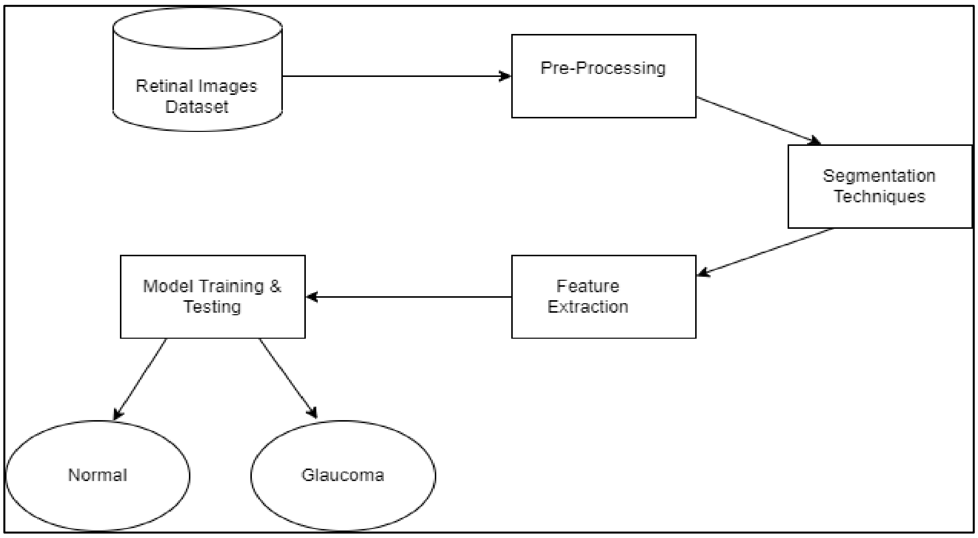


Fig 2: Architectural Diagram

4.2 DATA FLOW DIAGRAM

4.2.1 Level 0 DFD

DFD Level 0 is also called a Context Diagram. It is a basic overview of the whole system or process being analyzed or modeled. It is designed to be an at-a-glance view, showing the system as a single high-level process, with its relationship to external entities.

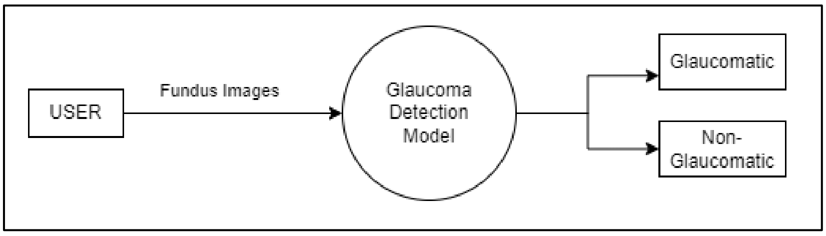


Fig 3: Level 0 DFD

4.1.1 Level 1 DFD

In 1-level DFD, the context diagram is decomposed into multiple bubbles/processes. In this level, we highlight the main functions of the system and breakdown the high-level process of 0-level DFD into subprocesses.

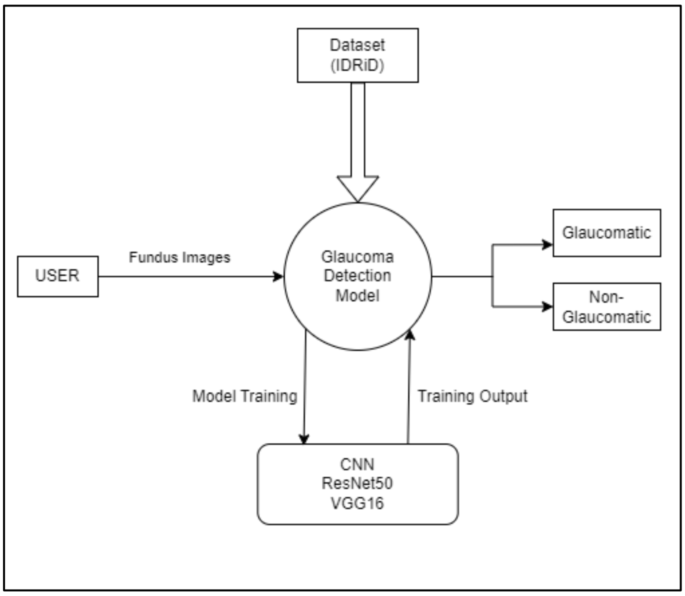


Fig 4: Level 1 DFD

4.3 USE CASE DIAGRAM

A use case diagram at its simplest is a representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved. Use case diagrams are used to gather the requirements of a system including internal and external influences. These requirements are mostly design requirements. Hence, when a system is analyzed to gather its functionalities, use cases are prepared and actors are identified. In this, we have actor and use cases. With the help of actor and use cases the diagram is made and all the functionalities is defined in a very appropriate manner.

Use Case Diagrams can be used for:

a) Requirement Analysis and high-level design
b) Model the context of a system
c) Reverse Engineering
d) Forward Engineering

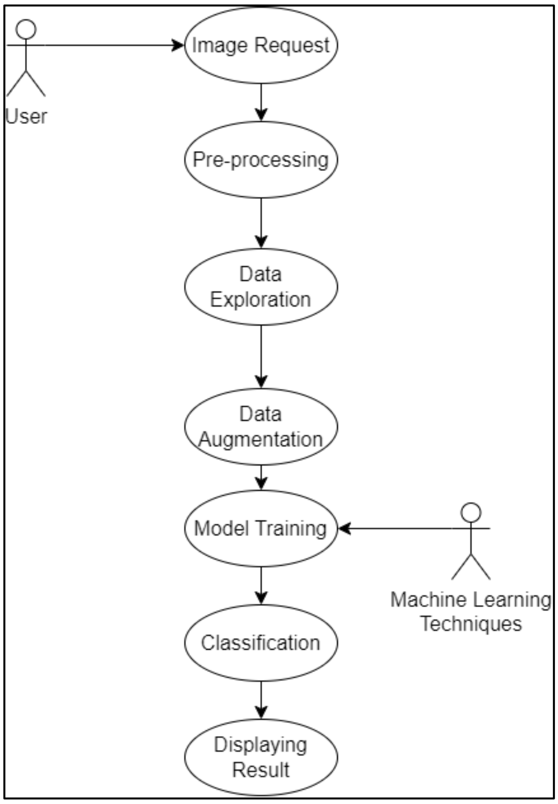


Fig 5: Use Case Diagram

4.4 ACTIVITY DIAGRAM

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all type of flow control by using different elements such as fork, join, etc.

The basic purpose of activity diagrams is similar to other four diagrams. It captures the dynamic behavior of the system. Other four diagrams are used to show the message flow from one object to another but activity diagram is used to show message flow from one activity to another.

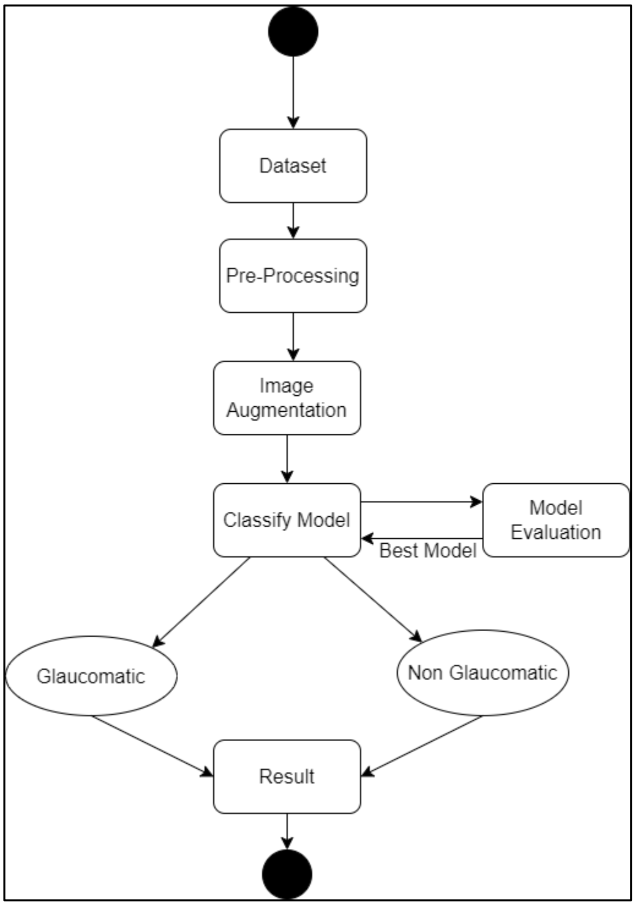


Fig 6: Activity Diagram

4.5 CLASS DIAGRAM

In UML, class diagrams are one of six types of structural diagram. Class diagrams are fundamental to the object modeling process and model the static structure of a system. Depending on the complexity of a system, you can use a single class diagram to model an entire system, or you can use several class diagrams to model the components of a system.

Class diagrams are the blueprints of your system or subsystem. You can use class diagrams to model the objects that make up the system, to display the relationships between the objects, and to describe what those objects do and the services that they provide.

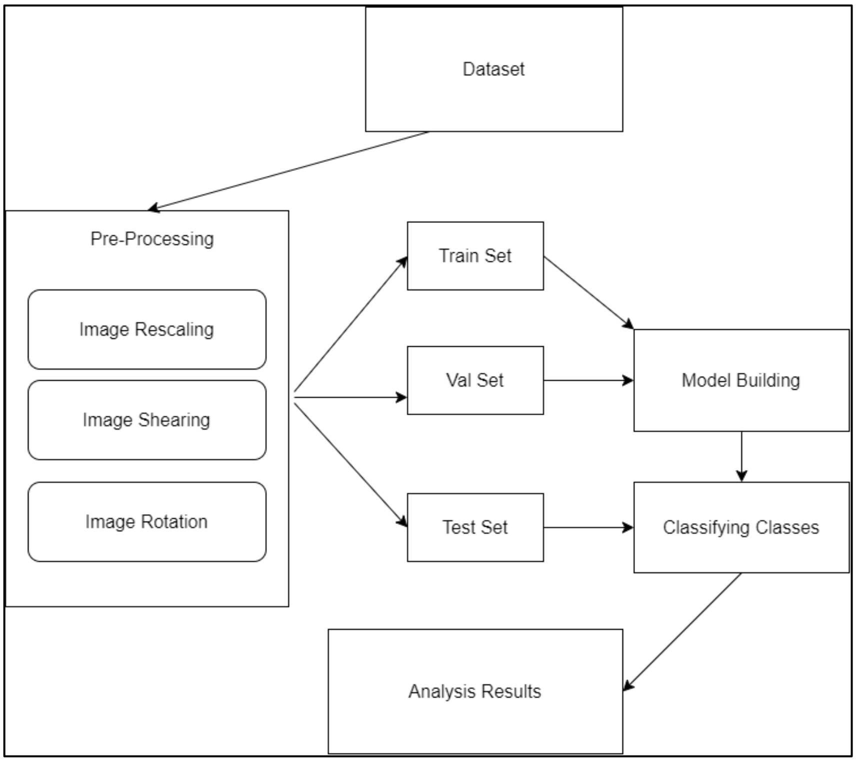


Fig 7: Class Diagram

4.6 SEQUENCE DIAGRAM

A sequence diagram simply depicts interaction between objects in a sequential order. The sequence diagram represents the flow of messages in the system and is also termed as an event diagram. It helps in envisioning several dynamic scenarios. It portrays the communication between any two lifelines as a time-ordered sequence of events, such that these lifelines took part at the run time. In UML, the lifeline is represented by a vertical bar, whereas the message flow is represented by a vertical dotted line that extends across the bottom of the page. It incorporates the iterations as well as branching.

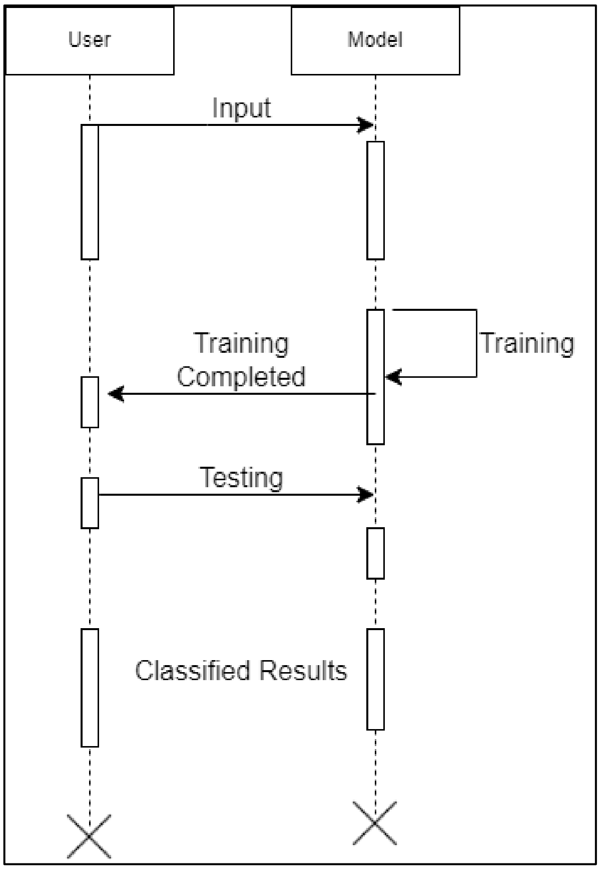


Fig 8: Sequence Diagram

4.7 COMPONENT DIAGRAM

A component diagram is used to break down a large object-oriented system into the smaller components, to make them more manageable. It models the physical view of a system such as executables, files, libraries, etc. that resides within the node.

It visualizes the relationships as well as the organization between the components present in the system. It helps in forming an executable system. A component is a single unit of the system, which is replaceable and executable. The implementation details of a component are hidden, and it necessitates an interface to execute a function. It is like a black box whose behavior is explained by the provided and required interfaces.

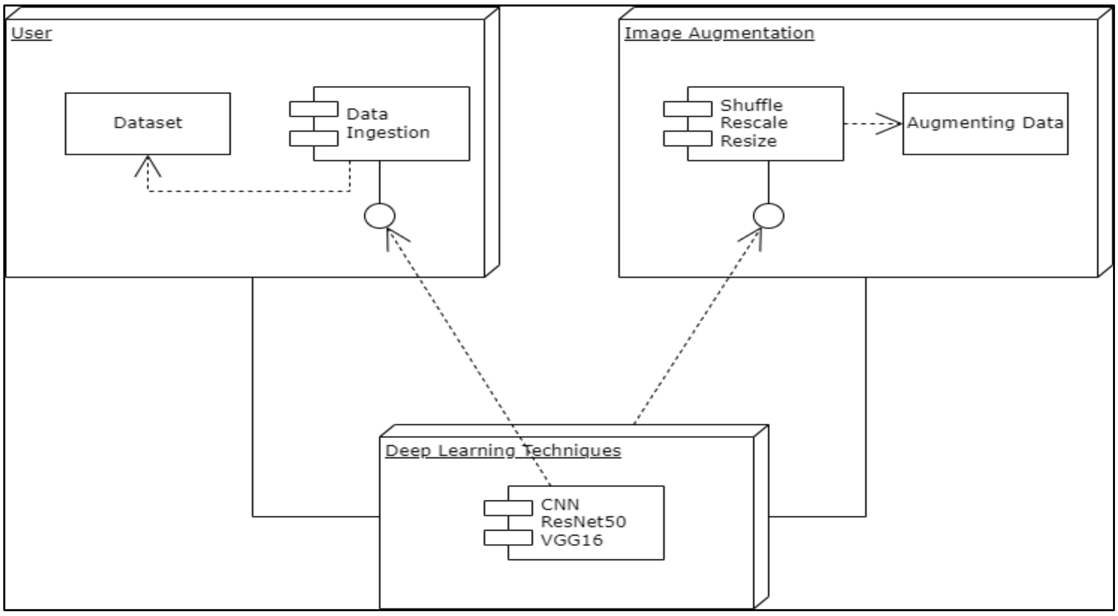


Fig 9: Component Diagram

CHAPTER 6

IMPLEMENTATION

6.1 Dataset

IDRiD (Indian Diabetic Retinopathy Image Dataset), is the first database representative of an Indian population. Moreover, it is the only dataset constituting typical diabetic retinopathy lesions and normal retinal structures annotated at a pixel level. This dataset provides information on the disease severity of diabetic retinopathy, and diabetic macular edema for each image. This makes it perfect for development and evaluation of image analysis algorithms for early detection of diabetic retinopathy.

The dataset consists of:

a) Original color digital fundus images
b) 516 images divided into train set
c) (413 images) and test set (103 images)
d) Groundtruth Labels for Diabetic Retinopathy and Diabetic Macular Edema Severity Grade (Divided into train and test set - CSV File)

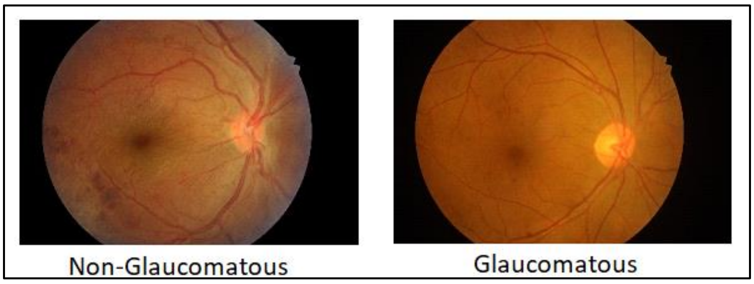


Fig 10: DFI Images (IDRiD Dataset)

Images in the dataset are categorized into 5 different types and their retinopathy grade:

a) No-DR 0
b) Mild 1
c) Proliferate DR 2
d) Moderate 3
e) Severe 4

Where, No-DR (Non-Diabetic Retinopathic) is non-glaucomatous, rest all four types are glaucomatous.

1	Image name	Retinopathy grade	Risk of macular edema
2	IDRiD_001	3	2
3	IDRiD_002	3	2
4	IDRiD_003	2	2
5	IDRiD_004	3	2
6	IDRiD_005	4	0
7	IDRiD_006	4	1
8	IDRiD_007	4	0
9	IDRiD_008	4	2
10	IDRiD_009	3	2
11	IDRiD_010	4	1
12	IDRiD_011	3	1
13	IDRiD_012	3	2
14	IDRiD_013	3	0
15	IDRiD_014	4	2
16	IDRiD_015	4	2
17	IDRiD_016	2	2
18	IDRiD_017	4	2

Fig 11: csv file for training dataset

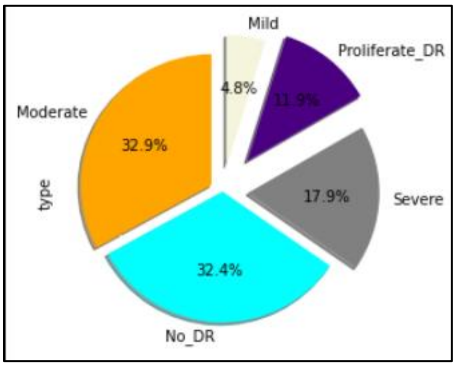


Fig 12: Pie chart displaying % of each type of image category

The figure below shows the count of each major type of the image, i.e., Glaucomatous and Non- Glaucomatous. The four categories of type ‘Glaucomatous’ are merged and mapped to single category, which is ‘DR or Diabetic Retinopathic.’

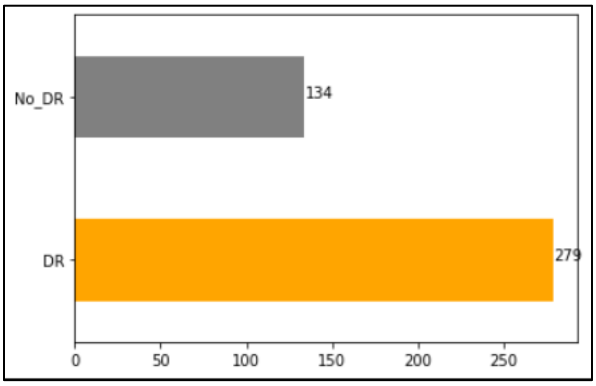


Fig 13: Number of glaucomatous and non-glaucomatous images

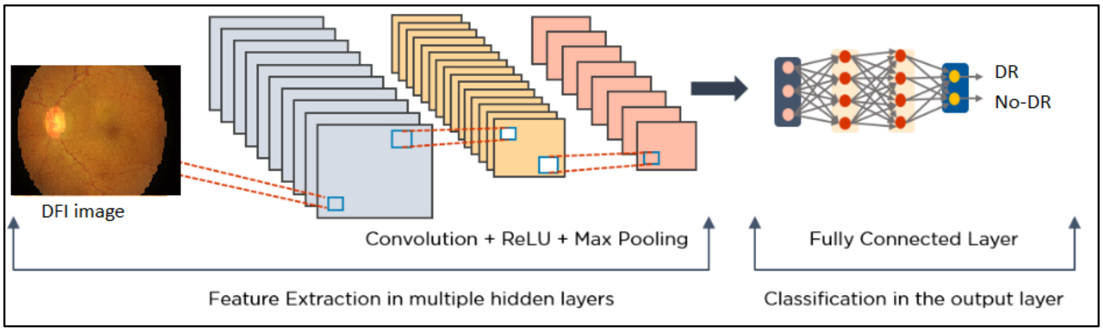


Fig 14: model components

The final model is like the architecture shown in the figure 14. It consists of feature extraction component that contain convolution, activation, and pooling layers, and then followed by the fully connected dense layers, which will have two layers as the output, i.e., DR and No-DR. Five models are used here which differ in the number of these layers mentioned above. These models along with their model summary and architecture are mentioned below:

6.2 Simple CNN

i) A Convolutional Neural Network (CNN) is a type of deep learning model commonly used for image and video processing tasks such as object recognition, classification, segmentation, and detection. It can automatically learn and extract relevant features from the input data without the need for explicit feature engineering.

ii) The basic architecture of a CNN consists of several layers, including convolutional layers, pooling layers, and fully connected layers.

iii) Convolutional layers use filters (also called kernels) to convolve over the input image, creating feature maps that capture relevant patterns and structures. The number and size of the filters can vary depending on the complexity of the task and the input data.

iv) Pooling layers are used to reduce the size of the feature maps while retaining the most relevant information. Max pooling and average pooling are two common pooling techniques used in CNNs.

v) Fully connected layers are used at the end of the CNN to perform classification or regression tasks based on the extracted features.

vi) Overall, CNNs are a powerful tool for image and video processing tasks and have achieved state-of-the-art performance in many domains. They have applications in fields such as computer vision, robotics, and autonomous driving.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 221, 221, 64)      3136

 max_pooling2d (MaxPooling2D  (None, 110, 110, 64)     0
 )

 batch_normalization (BatchN  (None, 110, 110, 64)     256
 ormalization)

 conv2d_1 (Conv2D)           (None, 108, 108, 32)      18464

 max_pooling2d_1 (MaxPooling  (None, 54, 54, 32)       0
 2D)

 batch_normalization_1 (Batc  (None, 54, 54, 32)       128
 hNormalization)

 conv2d_2 (Conv2D)           (None, 52, 52, 16)        4624

 max_pooling2d_2 (MaxPooling  (None, 26, 26, 16)       0
 2D)

 batch_normalization_2 (Batc  (None, 26, 26, 16)       64
 hNormalization)

 flatten (Flatten)           (None, 10816)             0

 dense (Dense)               (None, 32)                346144

 dropout (Dropout)           (None, 32)                0

 dense_1 (Dense)             (None, 2)                 66

=================================================================
Total params: 372,882
Trainable params: 372,658
Non-trainable params: 224
_________________________________________________________________
```

Fig 15: CNN model summary

6.3 VGG16

i. VGG16 is a deep learning model that was introduced in 2014 by the Visual Geometry Group (VGG) at the University of Oxford. It is a convolutional neural network architecture that is commonly used for image classification tasks.

ii. The VGG16 model consists of 16 layers, including 13 convolutional layers and 3 fully connected layers. The convolutional layers have a fixed kernel size of 3x3 and a stride of 1 pixel. Max pooling layers are used to down sample the spatial dimensions of the output feature maps. The model uses the Rectified Linear Unit (ReLU) activation function, which helps to introduce non-linearity into the network.

iii. The VGG16 model is known for its simplicity and high accuracy on image classification tasks. It has been used in a variety of applications, including object detection, image segmentation, and image recognition.

iv. One of the main advantages of the VGG16 model is its transfer learning capabilities. By using pre-trained weights from the ImageNet dataset, which contains millions of labeled images, the model can be fine-tuned for a specific task with much less training data than would be required to train the network from scratch.

v. Overall, the VGG16 model is a powerful tool for image classification and related tasks, and has been widely used in both academic and industrial settings.

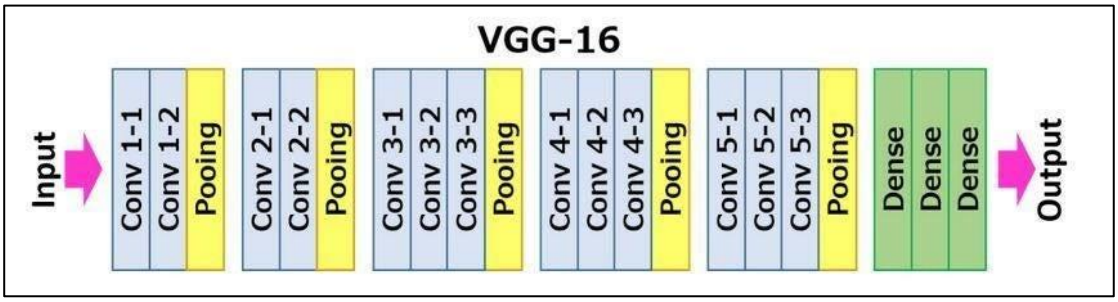
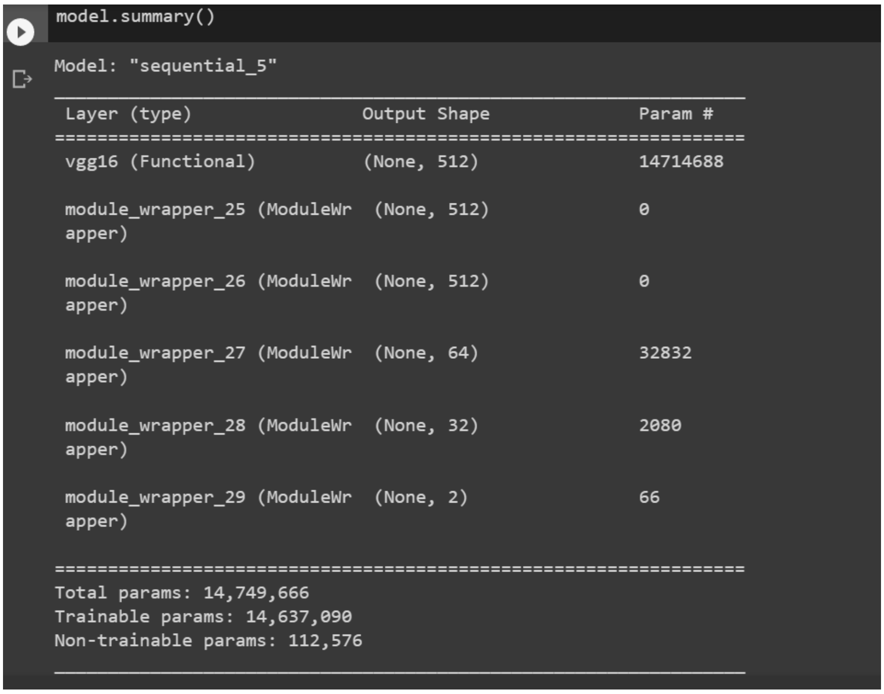


Fig 16: VGG16 model architecture



```
model.summary()

Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 512)               14714688

 module_wrapper_25 (ModuleWr  (None, 512)              0
 apper)

 module_wrapper_26 (ModuleWr  (None, 512)              0
 apper)

 module_wrapper_27 (ModuleWr  (None, 64)               32832
 apper)

 module_wrapper_28 (ModuleWr  (None, 32)               2080
 apper)

 module_wrapper_29 (ModuleWr  (None, 2)                66
 apper)

=================================================================
Total params: 14,749,666
Trainable params: 14,637,090
Non-trainable params: 112,576
_________________________________________________________________
```

Fig 17: VGG16 model summary

6.4 ResNet50

i. ResNet50 is a convolutional neural network (CNN) architecture that was proposed in 2015 by researchers at Microsoft Research. It is part of the ResNet family of models that are known for their ability to train very deep neural networks without suffering from the vanishing gradient problem.

ii. The architecture of ResNet50 consists of 50 layers, with skip connections that allow information to bypass several layers, helping to address the vanishing gradient problem. The network is composed of a series of convolutional layers, followed by a global average pooling layer, and then several fully connected layers that produce the final output.

iii. ResNet50 is often used for image recognition tasks, such as object detection, image classification, and semantic segmentation. It has achieved state-of-the-art performance on a number of benchmark datasets, including ImageNet, which contains over one million images.

iv. One of the key features of ResNet50 is its ability to learn very complex features and representations from raw image data, which makes it particularly effective for tasks that require high levels of accuracy and precision. It is also relatively easy to train and tune, which has made it a popular choice for researchers and practitioners in the deep learning community.

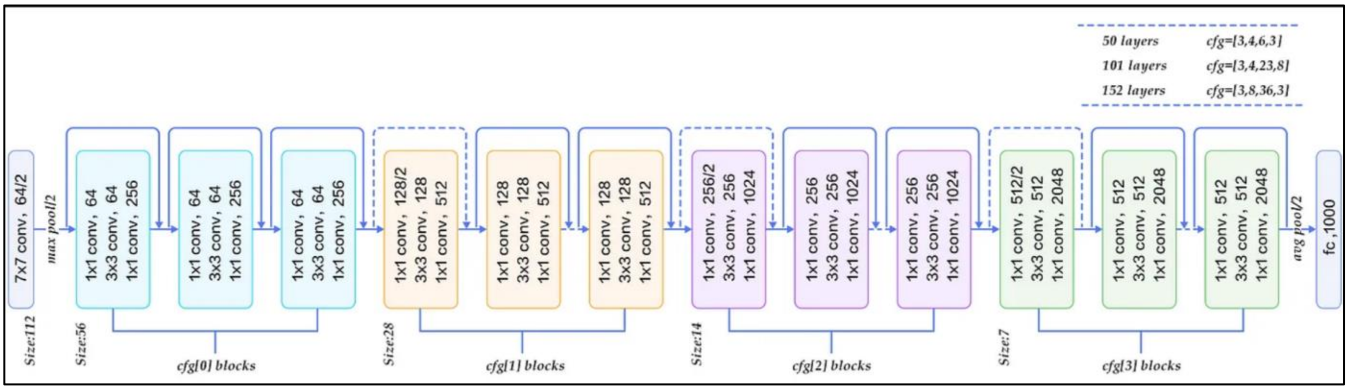
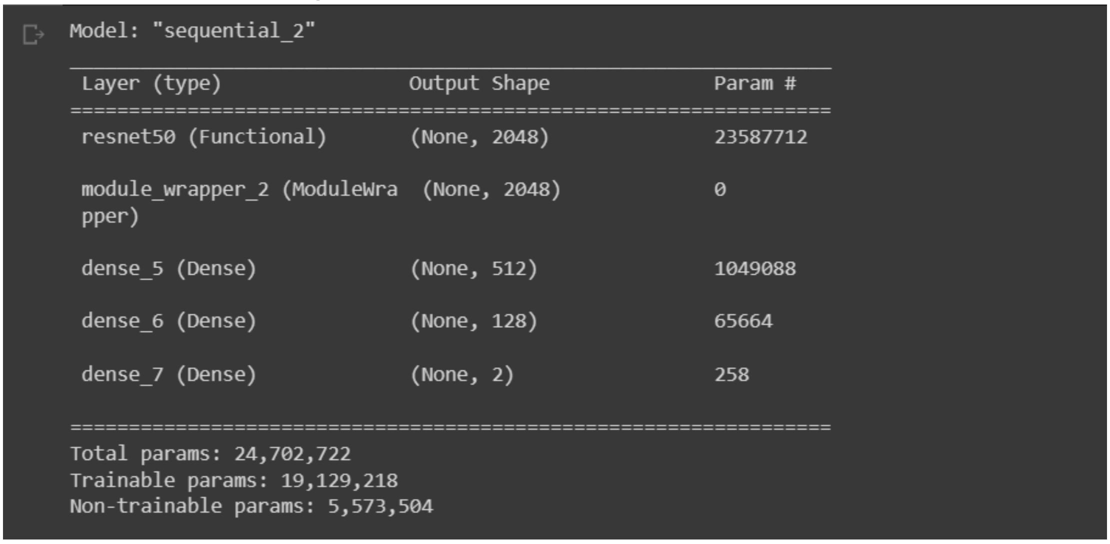


Fig 18: ResNet50 model architecture



Model: "sequential_2"

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 2048)	23587712
module_wrapper_2 (ModuleWrapper)	(None, 2048)	0
dense_5 (Dense)	(None, 512)	1049088
dense_6 (Dense)	(None, 128)	65664
dense_7 (Dense)	(None, 2)	258

Total params: 24,702,722
Trainable params: 19,129,218
Non-trainable params: 5,573,504

Fig 19: ResNet50 model summary

6.5 CNN+ ResNet50

```
Model: "model_8"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_22_input (InputLayer  [(None, 224, 224, 3)]    0
 )

 conv2d_22 (Conv2D)          (None, 221, 221, 16)      784

 max_pooling2d_22 (MaxPoolin  (None, 110, 110, 16)     0
 g2D)

 conv2d_23 (Conv2D)          (None, 107, 107, 3)       771

 max_pooling2d_23 (MaxPoolin  (None, 53, 53, 3)        0
 g2D)

 sequential_26 (Sequential)  (None, 2)                 24702722

=================================================================
Total params: 24,704,277
Trainable params: 19,130,773
Non-trainable params: 5,573,504
_________________________________________________________________
```

Fig 20: CNN + VGG16 model summary

6.6 CNN + VGG16

```
Model: "model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_9_input (InputLayer)  [(None, 224, 224, 3)]    0

 conv2d_9 (Conv2D)           (None, 221, 221, 16)      784

 max_pooling2d_4 (MaxPooling  (None, 110, 110, 16)     0
 2D)

 conv2d_10 (Conv2D)          (None, 107, 107, 3)       771

 max_pooling2d_5 (MaxPooling  (None, 53, 53, 3)        0
 2D)

 sequential_21 (Sequential)  (None, 2)                 15043266

=================================================================
Total params: 15,044,821
Trainable params: 14,932,245
Non-trainable params: 112,576
_________________________________________________________________
```

Fig 21: CNN + VGG16 model summary

CHAPTER 7
RESULT ANALYSIS

For all the models mentioned above, we have calculated the ‘testing accuracy,’ and have plotted accuracy vs epoch line plot below:

7.1 CNN (Test Accuracy – 80.76%)

```
Plotting accuracy versus epoch
The model is being evaluated
52/52 - 14s - loss: 0.4445 - acc: 0.8077 - 14s/epoch - 265ms/step
The accuracy of the model is:
0.807692289352417
```

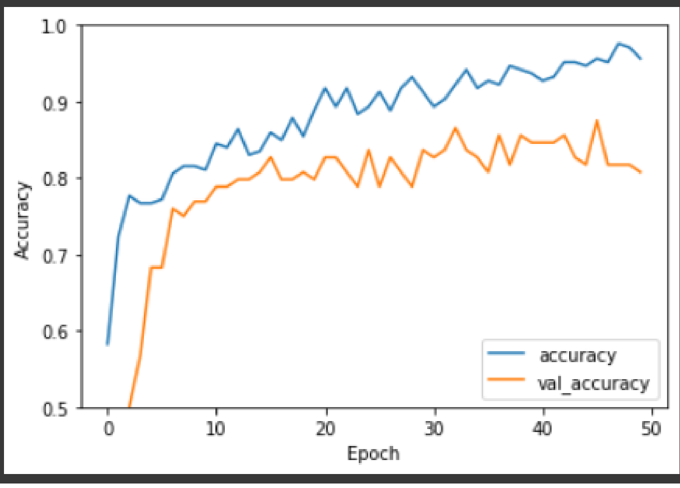


Fig 22: Accuracy and val-accuracy vs Epoch (CNN)

7.2 VGG16 (Test Accuracy – 81.55%)

```
Plotting accuracy versus epoch
The model is being evaluated
4/4 - 17s - loss: 0.6964 - accuracy: 0.8155 - 17s/epoch - 4s/step
The accuracy of the model is:
0.8155339956283569
```

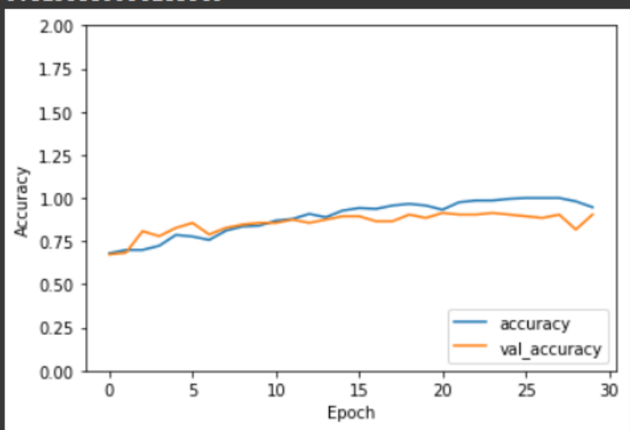


Fig 23: Accuracy and val-accuracy vs Epoch (VGG16)

7.3 CNN+VGG16 (Test Accuracy – 66.01%)

```
Plotting accuracy versus epoch
The model is being evaluated
4/4 - 15s - loss: 0.8545 - accuracy: 0.6602 - 15s/epoch - 4s/step
The accuracy of the model is:
0.6601941585540771
```

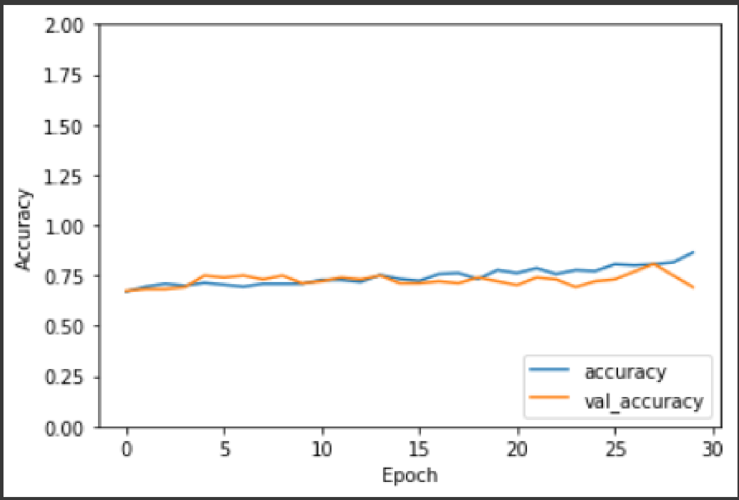


Fig 24: Accuracy and val-accuracy vs Epoch (CNN+VGG16)

7.4 ResNet50 (Test Accuracy – 90.29 %)

```
Plotting accuracy versus epoch
The model is being evaluated
52/52 - 57s - loss: 0.3215 - accuracy: 0.9029 - 57s/epoch - 1s/step
The accuracy of the model is:
0.9029126167297363
```

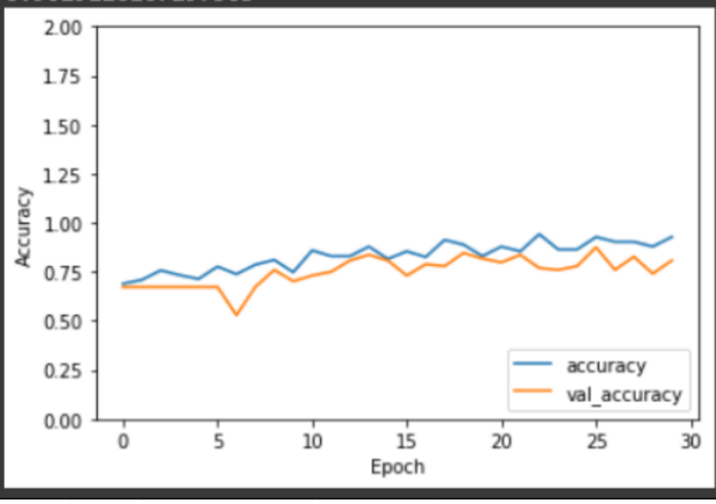


Fig 25: Accuracy and val-accuracy vs Epoch (ResNet50)

7.5 CNN+ResNet50 (Test Accuracy s–66.01%)

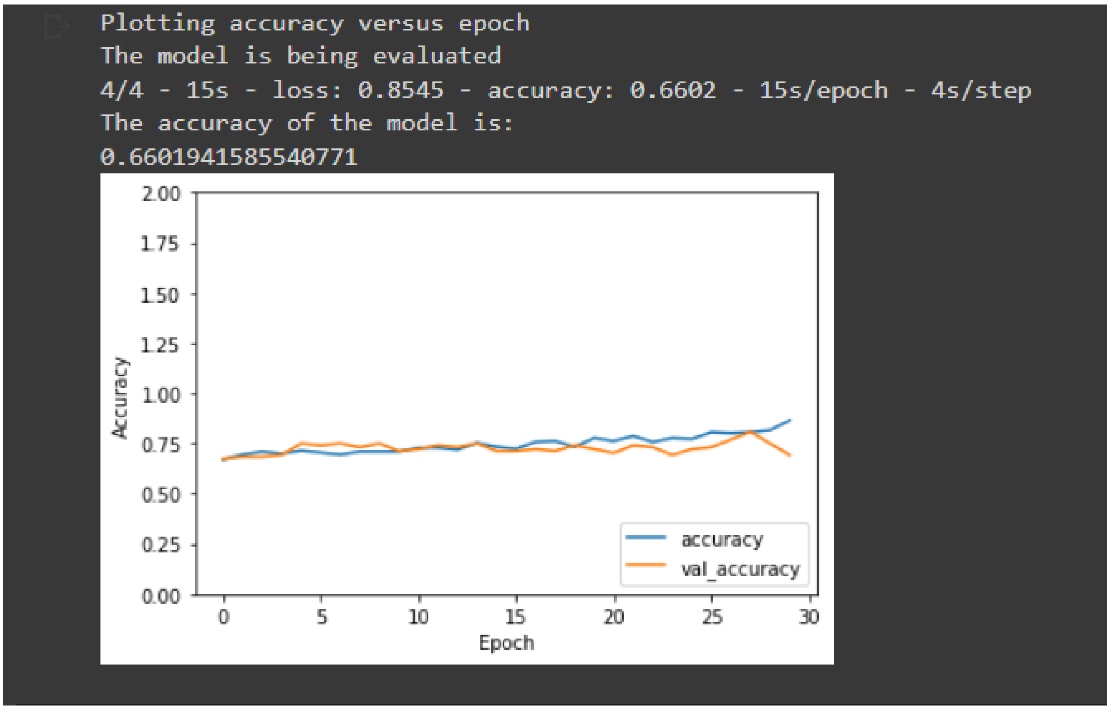


Fig 26: Accuracy and val-accuracy vs Epoch (CNN+ResNet50)

7.6 Tabular comparison of accuracy obtained from the proposed models:

Model Used	Train Accuracy	Test Accuracy
CNN	91.4	80.76
VGG16	88.02	81.05
CNN + VGG16	76.05	66.09
ResNet50	92.18	90.29
CNN + ResNet50	76.2	66.02

Table2: models accuracy comparison

7.7 Tabular comparison of classification accuracy of solutions developed by top-6 teams for grading of DR in IDRiD Disease Grading Challenge with proposed model:

Here we have displayed the comparison tables of the deep learning models we have used in our proposed solution and other teams have also used on the IDRiD dataset in IDRiD Disease Grading Challenge.

CNN

TEAM	Accuracy %
Our Proposed Model	80.76
VRT	59

Table3: CNN accuracy comparison

ResNet50

TEAM	Accuracy %
Our Proposed Model	90.29
LzyUNCC	59
AVSASVA	55

Table4: ResNet50 accuracy comparison

Here we observed that two models, CNN and ResNet50 were common methods that we could compare. We found on comparison that team "VRT" obtained an accuracy of around 59%, whereas we obtained 80.76% accuracy on testing on the same dataset. On the other hand, teams, LzyUNCC, AVSASVA obtained accuracy 59% and 55% through ResNet50 models, whereas we got accuracy around 90.29%.

Thus, we can say that we were successfully able to obtain higher accuracy from the other teams that tested their models on the same dataset.

CHAPTER 8

CONCLUSION & FUTURE WORK

CONCLUSION

In this project we have trained and tested the IDRiD dataset on 5 deep learning models which includes CNN, Resnet50, VGG16, CNN + Resnet50, CNN + VGG16. From the work done in past by the researchers on the same dataset, it was found that they obtained an accuracy of around 75% form the deep learning models they trained and tested. We tried to improve their accuracy by changing the model parameters. We were successfully able to obtain accuracy greater than them. It was observed that resnet50 model gave the highest accuracy compared to all the other models. Interestingly, it was also found that the hybrid models with CNN gave less accuracy as compared to the pretrained transfer learning model and standalone CNN. Thus, it can be concluded from our study that transfer learning models can be used on IDRiD dataset for glaucoma detection efficiently.

FUTURE WORK

Deep learning algorithms such as convolutional neural networks (CNNs) and ResNet50 have shown promising results in the detection of glaucoma from fundus photographs and OCT images. Further research can explore the use of AI for glaucoma detection in different imaging modalities and the development of more efficient and accurate algorithms. New imaging modalities such as swept-source OCT and adaptive optics imaging hold promise for more accurate and earlier detection of glaucoma. The development of portable and affordable devices for glaucoma detection may improve accessibility to eye care and help in the early detection of glaucoma. In summary, there are several promising areas for future research in glaucoma detection, including exploring new imaging modalities, developing multimodal approaches, using AI techniques, developing portable and affordable devices, and conducting longitudinal studies.

CHAPTER 9

REFERENCES

[1] M. S. Eswari and S. Balamurali, "An Intelligent Machine Learning Support System for Glaucoma Prediction Among Diabetic Patients," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 447-449, doi: 10.1109/ICACITE51222.2021.9404635.

[2] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong and J. Liu, "Glaucoma detection based on deep convolutional neural network," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015, pp. 715-718, doi: 10.1109/EMBC.2015.7318462.

[3] D. Mahapatra and J. M. Buhmann, "A field of experts model for optic cup and disc segmentation from retinal fundus images," 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), 2015, pp. 218-221, doi: 10.1109/ISBI.2015.7163853.

[4] Divya L, Jacob J. Performance analysis of glaucoma detection approaches from fundus images. Procedia computer science. 2018 Jan 1;143:544-51.

[5] A. Saxena, A. Vyas, L. Parashar and U. Singh, "A Glaucoma Detection using Convolutional Neural Network," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, 815-820, doi: pp.10.1109/ICESC48915.2020.9155930.

[6] J. Carrillo, L. Bautista, J. Villamizar, J. Rueda, M. Sanchez and D. rueda, "Glaucoma Detection Using Fundus Images of The Eye," 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), 2019, 1-4, doi: pp. 10.1109/STSIVA.2019.8730250.

[7] An G, Omodaka K, Hashimoto K, Tsuda S, Shiga Y, Takada N, Kikawa T, Yokota H, Akiba M, Nakazawa T. Glaucoma diagnosis with machine learning based on optical

coherence tomography and color fundus images. Journal of healthcare engineering. 2019 Feb 18;2019.

[8] B. Al-Bander, W. Al-Nuaimy, M. A. Al-Taee and Y. Zheng, "Automated glaucoma diagnosis using deep learning approach," 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD), 207-210, 2017, doi: pp. 10.1109/SSD.2017.8166974.

[9] J. Civit-Masot, M. J. Domínguez-Morales, S. Vicente-Díaz and A. Civit, "Dual Machine-Learning System to Aid Glaucoma Diagnosis Using Disc and Cup Feature Extraction," in IEEE 8, pp. 127519-127529, 2020, Access, vol. doi: 10.1109/ACCESS.2020.3008539.

[10] Sarkar D, Das S. Automated glaucoma detection of medical image using biogeography based optimization. In Advances in Optical Science and Engineering 2017 (pp. 381-388). Springer, Singapore.

[11] Rao PV, Gayathri R, Sunitha R. A novel approach for design and analysis of diabetic retinopathy glaucoma detection using cup to disk ration and ANN. Procedia Materials Science. 2015 Jan 1;10:446-54.

[12] Abbas Q. Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. International Journal of Advanced Computer Science and Applications. 2017;8(6).

[13] Li L, Xu M, Wang X, Jiang L, Liu H. Attention based glaucoma detection: a large-scale database and CNN model. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2019 (pp. 10571-10580).

[14] Thangaraj V, Natarajan V. Glaucoma diagnosis using support vector machine. In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS) 2017 Jun 15 (pp. 394- 399). IEEE.