Automatic Diabetic Retinopathy Detection Using Convolutional Neural Network

[1]Aravinda H, [2]Dinesh M, [3]Neha M Patil, [4]Srinivas Halvi

[1,2,3]VIII Semester, Department of Medical Electronics, DSCE, Bengaluru
[4]Professor, Department of Medical Electronics, DSCE, Bengaluru

Abstract: **Diabetic Retinopathy is the most common eye disease of diabetic patients which occurs due to high glucose levels leading to the damage of small blood vessels in the retina. About 80% of all patients who have had diabetes for 10 years or more are affected, leading to progressive vision loss. It is highly significant to prioritize eye check-up for chronic diabetic patients and to maintain their blood sugar levels under control so that they can prevent their eyes from getting damaged beyond cure. In this paper, we are proposing a Machine Learning technique using a Convolutional Neural Network which will classify the eye images ranging from No DR to Proliferative DR i.e., last stage of diabetic retinopathy. This system utilizes color fundus photographs as input. A pre-trained CNN model called Inception V3 is deployed. A dataset consisting of 1002 images is used for training this supervised learning model. A visual attention layer is added as the top layer of the pre-trained model focusing on the weighted features helping in image classification. Images will be classified into 5 categories i.e., No DR, Mild DR, Moderate DR, Severe DR, Proliferative DR. Images are taken using a +20D lens placed above a dilated eye using a smartphone camera and are transferred to the Raspberry Pi using an android application. The application will use Bluetooth to transfer the files and this image data is passed to the image classification model.**

This technique assists in the visualization and identification of abnormal structures such as lesions and exudates. A powerful, efficient, and user-friendly way of real-time image classification is offered by this technique.

Keywords **- Diabetic Retinopathy, Machine Learning, Convolutional Neural Network, Inception V3, Raspberry Pi.**

I. INTRODUCTION

A. The Eye

Eyes are paired sense organs located in the sockets or cavity of the skull called the orbits. These are provided with a cushion of fat for frictionless movement and protection from the shock. Eyes are supplied with the second cranial nerve called the optic nerve.

The wall of the eyeball consists of three layers as mentioned below:

1)External layer

This layer is composed of the sclera which is a dense connective tissue that maintains the shape of the eye and protects it. The anterior portion of the external layer is called the cornea which absorbs oxygen from the air and helps to focus light waves as they enter the eye.

2)Middle layer

The middle layer consists of the choroid layer, which has many blood vessels and has a bluish appearance. It is thin over two-thirds of the posterior part of an eyeball. This layer gradually thickens in the anterior part to form a ciliary body (sensory neuron in the retina). The visible colored portion of the eyes formed by the continuous forward formation of the ciliary body forming a pigmented and opaque structure is called the iris. The eyeball possesses a transparent crystalline lens held in place by the ligaments attached to the ciliary body. A pupil is an aperture in front of the lens and is surrounded by the iris.

3)Inner layer

The inner layer of the eye contains the retina. The retina contains three layers of cells called ganglion cells, photoreceptor cells, and bipolar cells. The photoreceptor cells are of two types, rods and cones. The rods and cones contain light-sensitive proteins called photopigments. The scotopic vision is a function of the rods while the daylight and color vision are the functions of cones. The rods include rhodopsin protein (visual purple) that contains a vitamin A derivative. The cone cells include porphyropsin, iodopsin, and cyanopsin containing photopsin protein.

At a point medial to and slightly above the posterior pole of the eyeball, the retinal blood vessels enter the eye while the optic nerves leave the eye. It is called a blind spot because of the absence of photoreceptor cells in this region.

A yellowish pigmented spot called macula lutea with a central pit called fovea is present at the posterior pole of the eye, lateral to the blind spot where the visual resolution is maximum.

The void in between the lens and cornea is the aqueous chamber. This chamber contains a thin watery fluid. The vitreous chamber is the space occupied between the retina and the lens and is filled with a transparent gel called the vitreous humour.

B. Diabetic Retinopathy

Diabetic retinopathy is caused by damage to the blood vessels within the tissue at the retina. Poorly controlled blood sugar is a risk factor. Early symptoms of DR include floaters, blurriness, dark areas of vision, and difficulty perceiving color, even blindness can occur in some cases. Diabetic retinopathy often has no early warning signs. Even macular edema, which causes rapid vision loss, has no warning signs for some time. In general, an individual with macular edema will probably have blurred vision, making it hard to try to do things like reading or drive. In some cases, the vision might recover or worsen during the day.

Progressive stages Of Diabetic Retinopathy

1. Stage I: The first stage of DR, called the Non-Proliferative Diabetic Retinopathy (NPDR) has no symptoms. Patients might not notice any irregularities and will have 20/20 vision. Fundus photographs are often used for objective documentation of the fundus findings, during which microaneurysms (microscopic blood-filled bulges within the artery walls) are often seen.
2. Stage II: The second stage of DR is called Moderate NPDR. These patients will have haemorrhages or Microaneurysms in one to three retinal quadrants; i. cotton wool spots, ii. hard exudates, iii. venous beading. There is a 12% to 27% risk that those suffering from moderate NPDR can develop proliferative diabetic retinopathy (PDR) within a year.
3. Stage III: The third stage of DR is called the Severe NPDR. These patients will have intra-retinal haemorrhages, venous beading in two or more quadrants, or an Intraretinal Microvascular Abnormality (IRMA) in one or more quadrants. Patients having severe NPDR have a 52% risk of developing Proliferative DR within a year, and hence, it is important to have close observation along with regular check-ups to prevent them from getting blind.
4. Stage IV: The final stage of DR is Proliferative Diabetic Retinopathy. This is the advanced stage of the disease that includes the formation of additional new abnormal and fragile blood vessels within the retina. These patients exhibit either neovascularization of the disc, vitreous or pre-retinal haemorrhage, which can lead to vision loss and possibly blindness.

Macular edema is a condition in which blood vessels leak their contents into the macular region of the eye, and can occur at any stage of NPDR. Its symptoms are blurred vision and darkened or distorted images that aren't equivalent in both eyes. Ten percent (10%) of diabetic patients will have vision loss which is associated with macular edema.

Clinicians identify DR by the presence of lesions associated with vascular abnormalities caused by the disease. Methods that are in practice for screening of diabetic retinopathy include direct and indirect ophthalmoscopy, stereoscopic color film fundus photography, mydriatic or nonmydriatic digital color, and monochromatic photography. Traditionally, ophthalmologists have screened for diabetic retinopathy by dilating the pupil and performing indirect ophthalmoscopy, during which the whole retina is examined. This method of screening is possible where access to eye care is sufficient. While this traditional approach is effective, its resource demands are high.

Manual differential diagnosis of DR images remains a challenge due to three main reasons:

- Depends on the experience and professional background of the ophthalmologist;
- The work is tedious and time-consuming; and
- Challenging for the human eye to distinguish subtle changes in tissues.

The expertise and equipment required are over-occupied as the number of diabetes cases in local populations is high and automated DR detection is most needed in this situation. As the number of people with diabetes continues to grow, the infrastructure needed to prevent blindness also has to be improved.

C. Image Classification By CNN

The urge for a comprehensive and automated method of DR screening has long been recognized. Using image classification, pattern recognition, and machine learning; and with color fundus photographs as input, this project aims to be an automated detection system to detect the scale of diabetes and to prevent blindness due to DR. Fundus image classifications can be accomplished by using several different techniques of imaging modalities and machine learning among which deep learning has a greater ability to differentiate healthy and DR images. Deep convolutional neural networks (DCNNs) have the potential to improve diagnostic efficiency and increase the level of interobserver agreement in the classification of fundus images of varying degrees of DR.
Based on the nature of the task, classification, regression, and segmentation-based models are the three major canonical deep learning models that are used.

II. PREVIOUS WORKS AND PRESENT DEVELOPMENTS

The classification of DR images has been accomplished by using several different techniques of imaging modalities and machine learning.

- M. T. Esfahan et al. [1] used ResNet34 CNN in their study to classify DR images into normal or DR images. ResNet34 is a pre-trained CNN architecture on the ImageNet database. A set of image pre-processing techniques is applied to improve the quality of images. The image pre-processing included the Gaussian filter, weighted addition, and image normalization. The image number was 35000 images and its size was 512 × 512 pixels. They reported an accuracy of 85% and a sensitivity of 86%.

- The work by V. Gulshan et al. [2] introduced a method to detect DR and diabetic macular edema (DME) using CNN model. They used Messidor-2 and eyepacs-1 datasets that contain 1748 images and 9963 images, respectively, to test the model. These images are first normalized, and the diameter was resized to 299 pixels wide before feeding them to the CNN. They trained 10 CNNs with the pre-trained Inception-v3 architecture with a various number of images, and the final result was computed by a linear average function. The images were classified into referable diabetic macular edema, moderate or worse DR, severe or worse DR, or fully gradable. They obtained a specificity of 93% in both datasets and 96.1% and 97.5% in sensitivity for the Messidor-2 and eyepacs-1 datasets, respectively; however, they did not explicitly detect non-DR or the five DR stage images.
- T. Shanthi and R. Sabeenian [3] detected the DR stages of the Messidor dataset using a pre-trained architecture called Alexnet. The images are resized and the green channel was extracted before being fed into the CNN. This CNN achieved an accuracy of 96.35. Unfortunately, the work does not detect the lesions in the images and only one dataset and architecture were used to test their method.
- The study of H. Jiang et al. [4] integrated three pre-trained CNN models, namely, Inception V3, Inception-Resnet-V2, and Resnet152 to classify their dataset as referable DR or non-referable DR. In CNNs training, Adam optimizer was used to update their weights. These models were integrated using the Adaboost algorithm. The dataset of 30,244 images was resized to 520 x 520 pixels, enhanced, and augmented before being fed to the CNNs. The work obtained an accuracy of 88.21% and an area under the curve (AUC) of 0.946.
- B. Harangi et al. [5] integrated the available pre-trained AlexNet and the hand-crafted features to classify the five DR stages. The CNN was trained and tested by the IDRiD. 90.07% accuracy was obtained. Unfortunately, the work does not detect the lesions in the images and only one dataset was used to test their method.

III. PROBLEM STATEMENT

1. Timely detection of DR can help in slowing down the progression towards vision impairment. However, this can prove to be difficult as the disease often shows few symptoms that can be neglected by the patient until it is too late to provide effective treatment.

2. Often, a day or two is generally required for a trained clinician to examine and evaluate digital color fundus photographs, which is a time-consuming manual process. By the time the clinicians submit their reviews, the delayed results can lead to miscommunication, lost follow-up, and delayed treatment.

3. The presence of lesions associated with the vascular abnormalities that are caused by Diabetic Retinopathy can be identified by the clinicians. This approach of screening while being effective can be highly resource-demanding. The expertise and equipment required are often lacking in areas, where the rate of diabetes in local populations is increasing and DR detection is most needed, as the number of individuals affected by diabetes continues to grow. The infrastructures needed to prevent blindness due to DR can be observed with growing insufficiency.

IV. OBJECTIVE

1. To detect lesions, i.e., exudates and other abnormalities in the fundus images using CNN.
2. To develop an automatic diabetic retinopathy detection system that analyses fundus images to classify them to assigned severity of the disease.
3. To test and validate the classified images which help to diagnose the retinopathic conditions.
4. To make a system that doesn't require any special skill set to operate.

V. METHODOLOGY

The entire setup is depicted in Figure 1. It includes the +20D lens, Raspberry Pi 4 Model B, hard drive, monitor, and keyboard. The images of the fundus taken using a smartphone camera are transferred to the raspberry pi using a custom-built application that uses Bluetooth for the completion of file transferring. The monitor is connected to the microprocessor using the micro-HDMI to HDMI cable. The transferred images are stored in the external hard drive which is connected to one of the USB ports. The images are

accessed using peripherals like a keyboard and mouse. The raspberry pi is powered using a 5V / 3A power adapter.

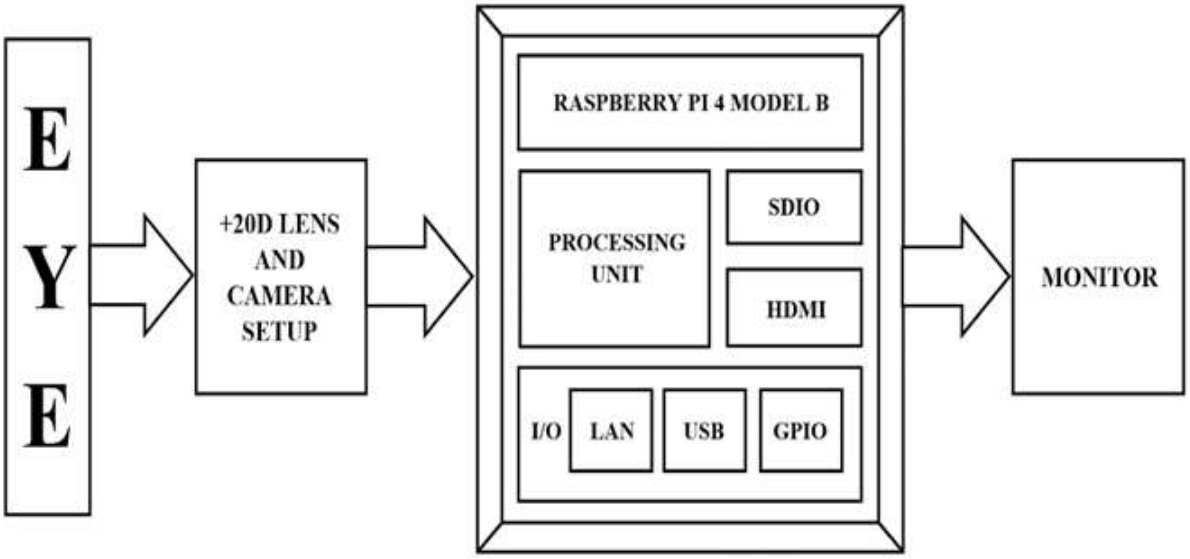


Fig 1: Block diagram of hardware setup

- Eye drops are used to widen (dilate) the pupils of the patient's eye to allow a better view inside the eyes.
- Tropicamide and Homatropine are generally used for dilating the pupil.
- Up to 30 minutes is allowed for the drops to show a clear effect on the pupil.
- The fundus of the patient's eye is viewed using a +20D lens. The lens is kept 50mm above the eyes. A +90D lens can also be used depending on the requirement.
- Using a smartphone camera, the image of the fundus is captured. The viewing is to be adjusted by changing the distance between the camera and the lens.
- The flash setup in the smartphone is used to illuminate the area.
- The captured image data is transferred to the Raspberry Pi through Bluetooth.
- For data transferring, a custom-built application is used which transfers images at full quality.
- These images are stored in a single directory.
- The mouse and the keyboard are connected to the USB slots in the Raspberry Pi.
- The Raspberry Pi is connected to the monitor through a Micro HDMI to HDMI cable. The processor is powered by a power adapter.
- The images will be saved in the external hard drive connected to the USB port and can be accessed at any point in time.

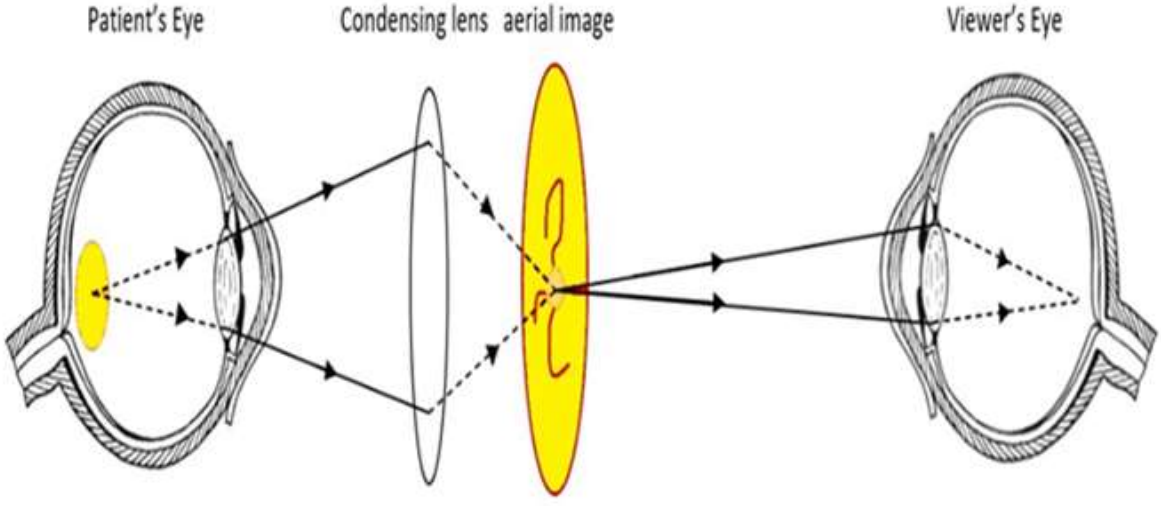


Fig 2: Schematic representation of viewing fundus of an eye.

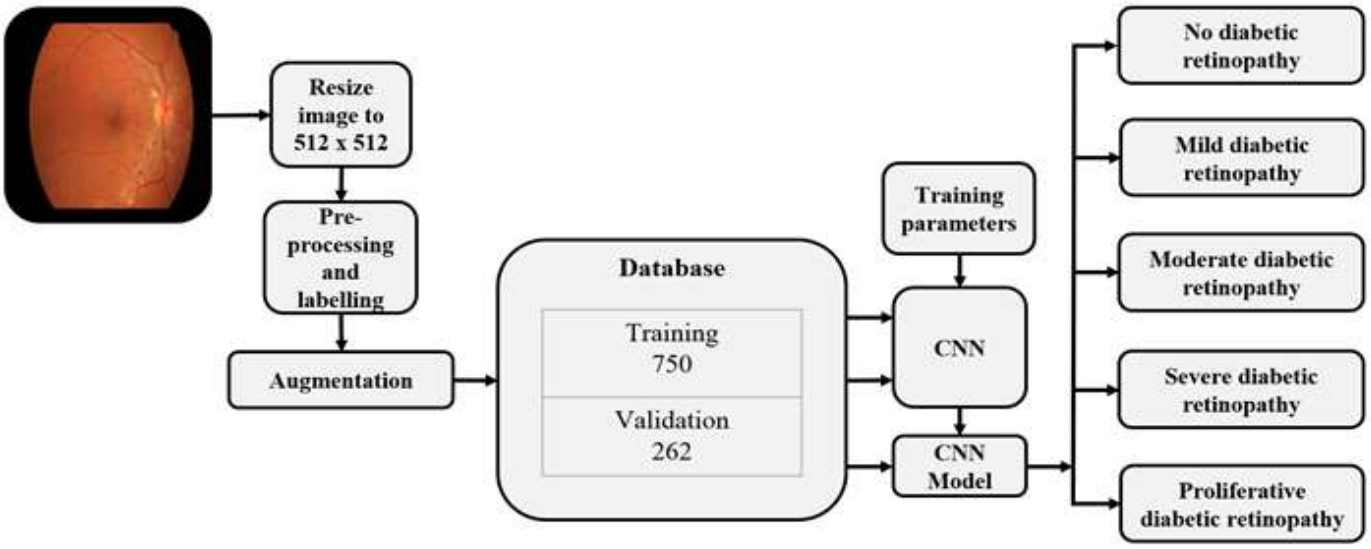


Fig 3: Flowchart of the system

The flow for implementation is:

- Start camera feed and take fundus photographs.
- Pass the data to the model for classification.
- Classify the image using the pre-stored images.

The entire process is done using CNN as shown in Figure 3. The CNN aims at differentiating fundus photographs into five different classes: No diabetic retinopathy, Mild diabetic retinopathy, Moderate diabetic retinopathy, Severe diabetic retinopathy, and proliferative diabetic retinopathy.

1. *IMPORTING LIBRARIES:*

   The libraries required are imported accordingly. They are:

   a. NumPy: For linear algebra
   b. Pandas: Data Pre-processing
   c. Matplotlib.pyplot: displaying and rendering plots or figures
   d. Imread, OS, globe: I/O

2. *LOAD DATASET:*

   1002 images are uploaded. These images are from the Kaggle website.

3. *EXAMINING THE DISTRIBUTION OF EYE AND SEVERITY:*

   In the histogram represented in figure 4(a), the rectangles on the left and right represent the left and right eye respectively with each valuing at 501. In figure 4(b), the rectangles represent the total number of cases in each class.

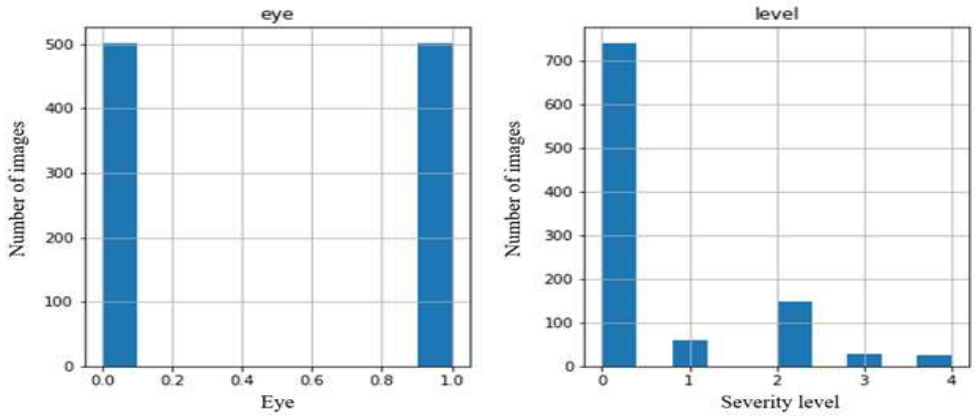


Fig 4**:** Histogram representation of (a) distribution of eye images, and (b) severity levels.

4. *SPLITTING OF DATASET INTO TRAINING AND VALIDATION SET:*

   Train_test_split is imported from sklearn.model_selection. This library is used to split the dataset into a random training subset and validation subset. Here 1002 images are split into 778 training subsets and 262 validation subsets.

5. *BALANCING THE DISTRIBUTION OF CLASSES IN TRAINING SET:*

The old training subset consisting of 778 images is evened out to 750 images.

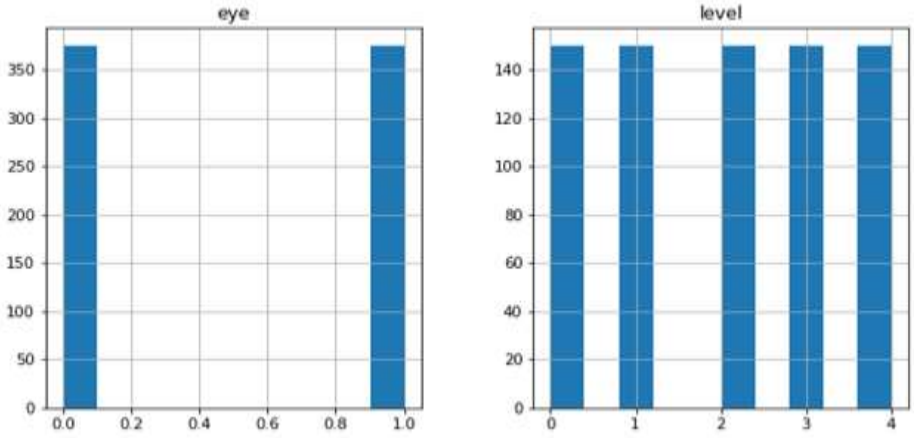


Fig 5: Balanced histogram representation of (a) distribution of eye images, and (b) severity levels.

6. *IMAGE PRE-PROCESSING:*

Here the libraries imported are required for pre-processing of input image data. The libraries imported are TensorFlow, Backend is imported from Keras, Preprocess_input is imported from Keras.applications.inception_v3 and NumPy.

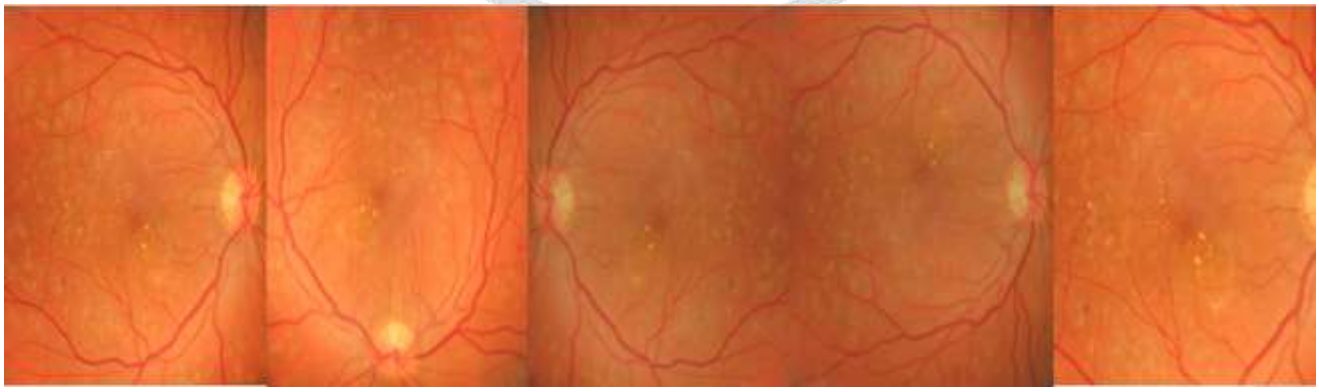
The TensorFlow library is used for machine learning. Keras is a library for neural networks used to train the model. Backend library helps to tie backend Operations directly to Keras. Inception V3 which is a pre-trained model is a type of convolution neural network used for Deep Learning. It has 48 layers of deep neural networks. It scales the pixels of the input image between -1 and 1.

The input image specifications of this library are (299,299,3) i.e., height, width, and channels. Preprocess_input is used to pre-process a batch of input images to a NumPy array or a tensor. The brightness, contrast, and saturation are set. Color mode is used to set to RGB. Preprocess_input function is applied which will encode these images into an array. The images undergo affine transformations by cropping them to a uniform size and rotating them. They are scaled and resized uniformly. To generalize the features of the fundus images to cover each case, the entire database is augmented and training of the CNN is done.

Augmentation is the process of artificially creating new training data from the existing dataset.

The process of augmentation is done by:

- Flipping the images in an up-down direction in the horizontal axis.
- Flipping the images in the left-right direction in the vertical axis.
- Counterclockwise rotation of images at its center point by 90°.
- Re-scaling of the images to 512x512.



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

Fig 6: Augmented image: (a) original image, (b) 90° rotation, (c) horizontal flip, (d) vertical flip, (e) Re-scaling

.

7. *VALIDATION DATASET:*

This data set has 262 images. These are not augmented.

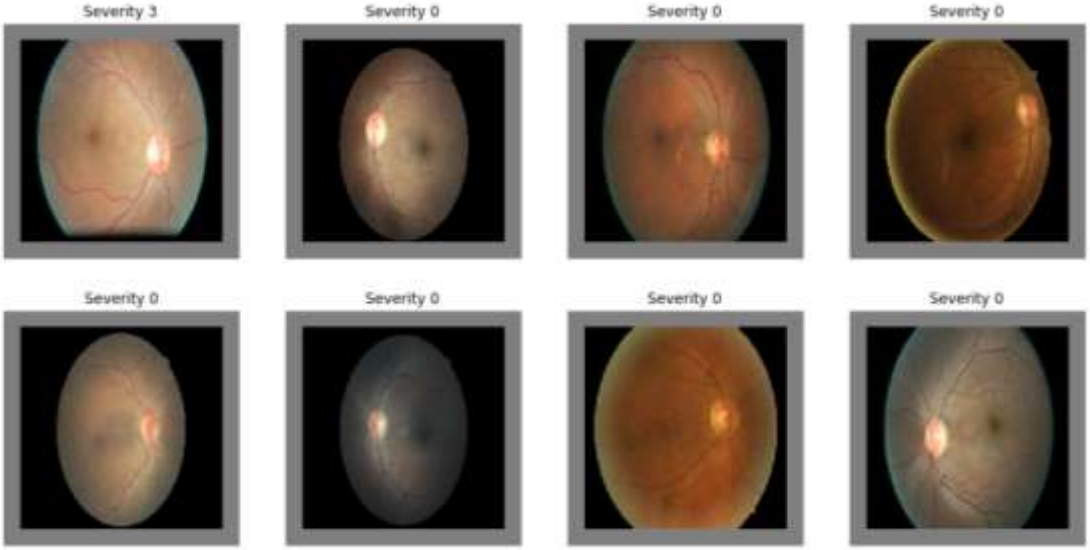


Fig 7: Sample images of validation dataset.

8. *TRAINING DATASET:*

This dataset has 750 images. These images are augmented. It helps in increasing the amount of data available by changing the orientation of the same image. This improves training on the network more efficiently to predict with higher accuracy.

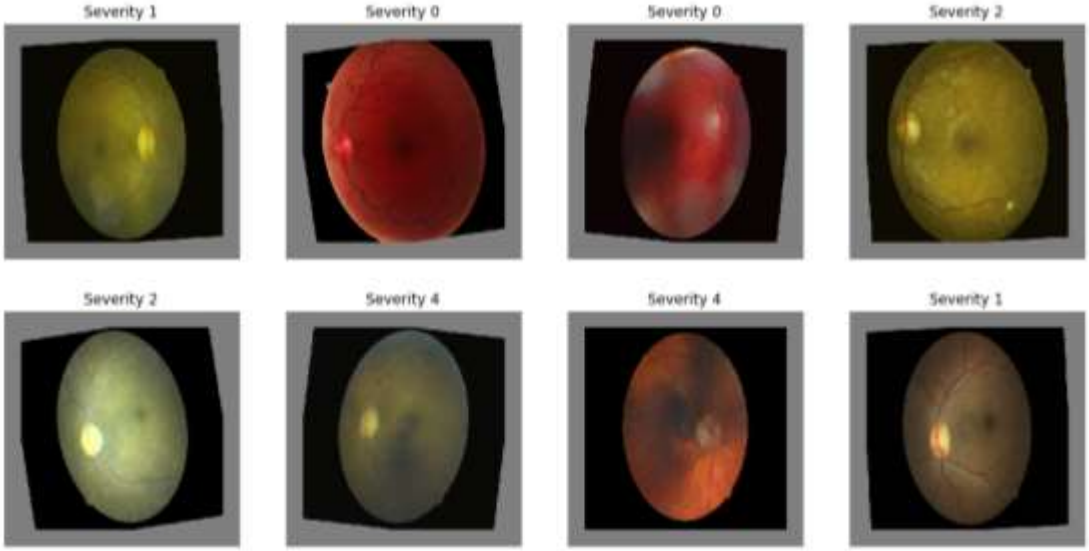


Fig 8: Sample images of the training dataset.

9. *ATTENTION MODEL:*

This model is constructed using pre-trained models as the base. The pre-trained models used are VGG16, InceptionV3, and InceptionResNetV2. A method called transfer learning is adapted to use only the weights and few CNN layers of these pre-trained models so that they can be applied to our image dataset. Transfer learning is a technique used to transfer the model parameters pre-trained from a commonly available dataset. The first layers of these models are replaced with several layers of normalization and a visual attention layer. The final layers are retained which perform the classification. The inception model works as a multilevel feature extractor that computes 1x1, 3x3, 5x5 convolutions in the module of the same network.

The attention mechanism is used to focus on those parts of the input data to create corresponding output based on that part of the input which was being focused. The region of focus is mainly on the weighted features or the main foreground rather than the blank regions or the background. Using attention allows configuring neural networks to give results of recognition or prediction by focusing on major areas of an image.

A sigmoid function is configured which functions as the activation function for convolution of output between 0.0 to 1.0. This activation function stores input values within a range of 0 and 1. So using this function gives an output for each pixel a value between 0 and 1. When these values are multiplied against the original convolution output, the areas near a 0 value are multiplied by a value of 0 which means that the output of these values is nearer to 0 and are not areas of focus. Areas

having value 1 and multiplied by values near 1 are outputs that are closer to original values which means that these are the areas of focus.

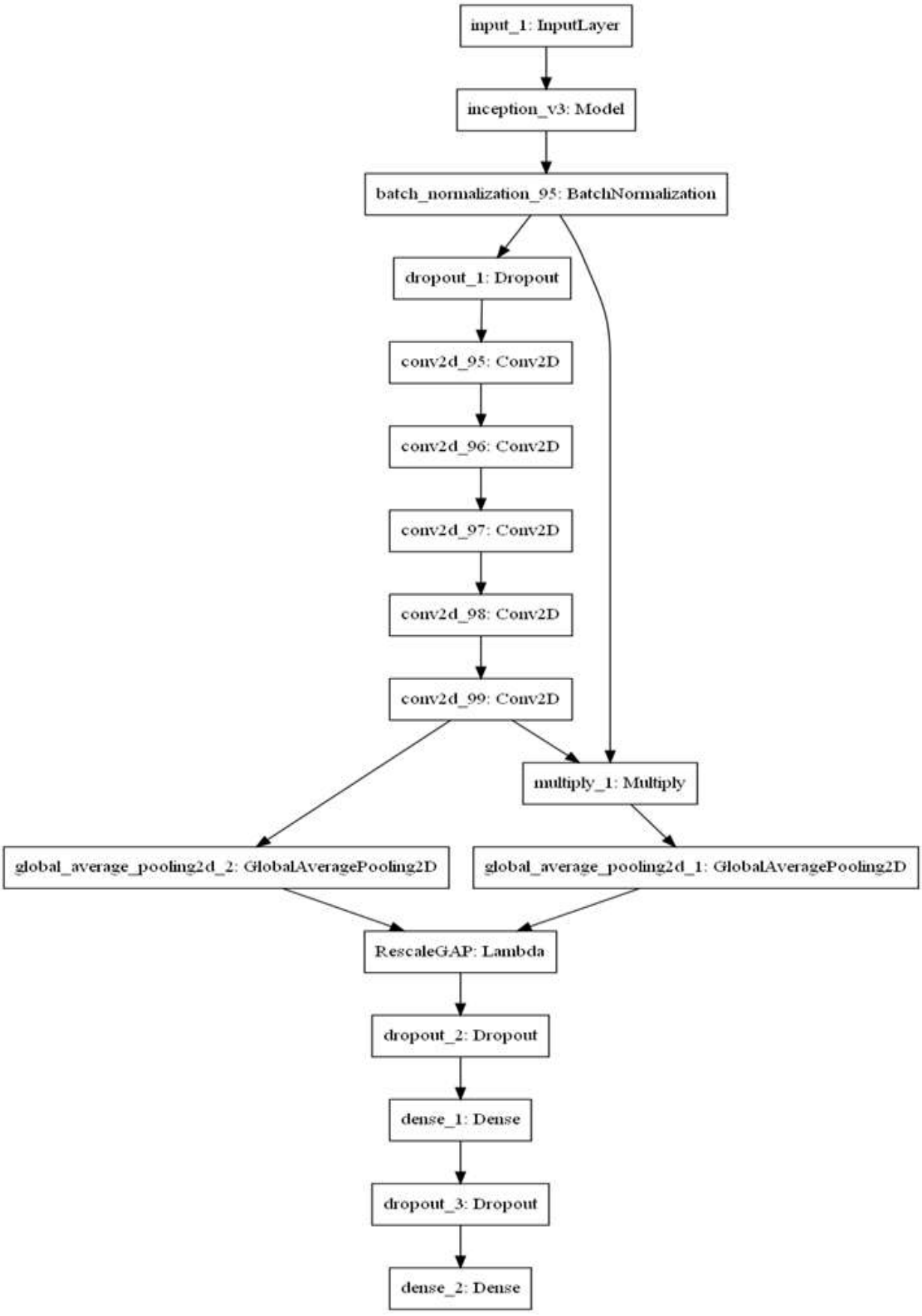
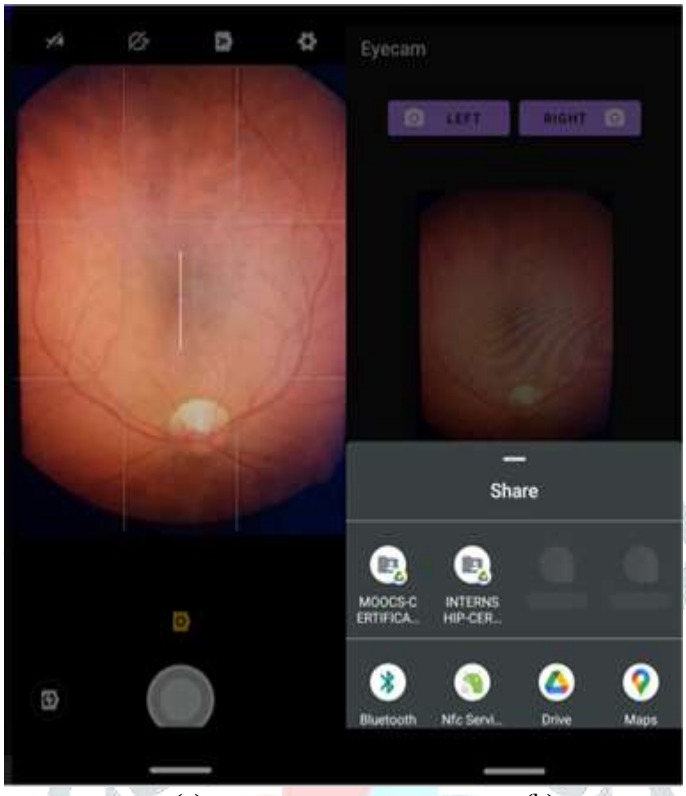


Fig 9: Flowchart representation of architecture layers.

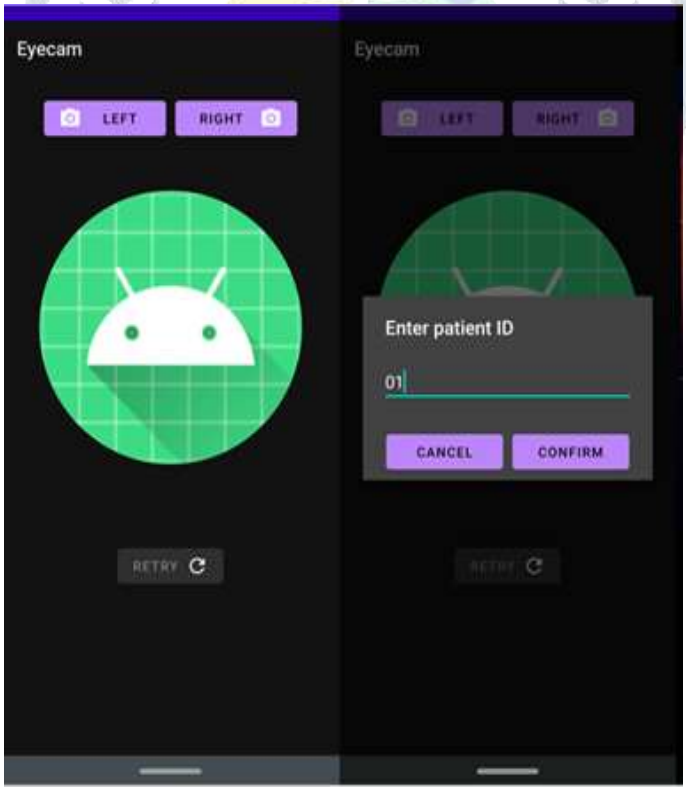
10. *SMARTPHONE APPLICATION - 'EYECAM':*

To capture images of the fundus and to transfer them to the microprocessor, a custom application 'Eyecam' is built. It is an android application that is capable of capturing images of the eyes using the smartphone camera, assigning names to the image files, and transferring them using the Bluetooth option within the application. The application is built using the Android Studio software. It is an integrated development environment provided by Google for Android. It is based on JetBrains' IntelliJ IDEA software. The application is built on the Android Studio V4.1. User permissions are required to access the camera and storage for the use of the application. The application has a size of 27.27 MB.



(a) (b)

Fig 10: 'Eyecam' application (a) camera view (b) sharing options



(c) (d)

Fig 11: (c) homepage (d)PatientID entry

VI. RESULTS AND DISCUSSION

The Inception V3 model gives accurate results up to 92% and can classify the images into their respective categories. Here some of the images shown below are classified into their respective types. Hence, if there are any abnormalities detected the model can identify and can report back by classifying the image as abnormal. This helps the doctors by saving their time in analyzing the image obtained with better accuracy.

Here, three images were captured using the 'Eyecam" application and were used as input to the trained Inception V3 model. Table 1 represented below gives the details and results of the classification process.

Image No.	CLASSIFICATION CATEGORIES					Actual Category	Predicted category	Predicted Percentage
	No DR	NPDR	Moderate NPDR	Severe NPDR	Proliferative NPDR			
	0	1	2	3	4			
01_right	96.67%	01.1%	00.2%	01.0%	01.03%	0	0	96.67%
02_left	02.37%	93.23%	03.0%	00.9%	00.5%	1	1	93.23%
03_left	04.89%	07.15%	86.19%	01.07%	00.7%	2	2	86.19%

Table 1: Results

Thus, the CNN model is accurate in depicting the degree of DR in the patient's eyes. Along with these values, attention mapping outputs are provided with color codes to further understand the extent of DR in the images.

Visual Attention maps are plotted for these eye images.

(a) (b) (c)

Fig 12: Eye Images of (a) 01_right (b) 02_left (c) 03_left

Here the green highlighted regions are the areas that are focused on for prediction.

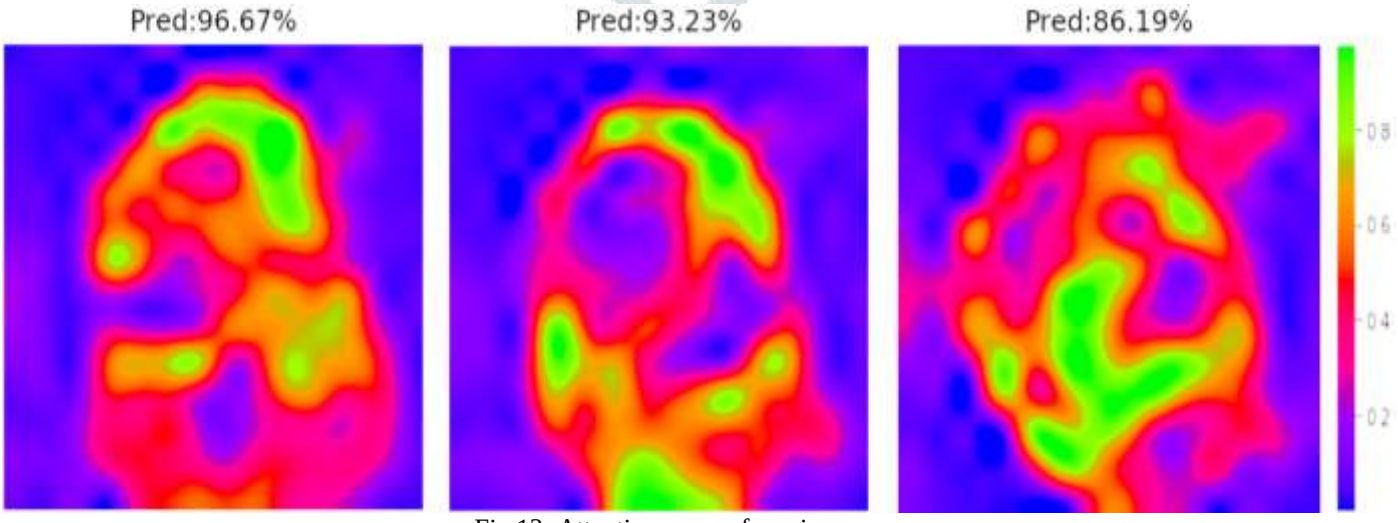


Fig 13: Attention maps of eye images

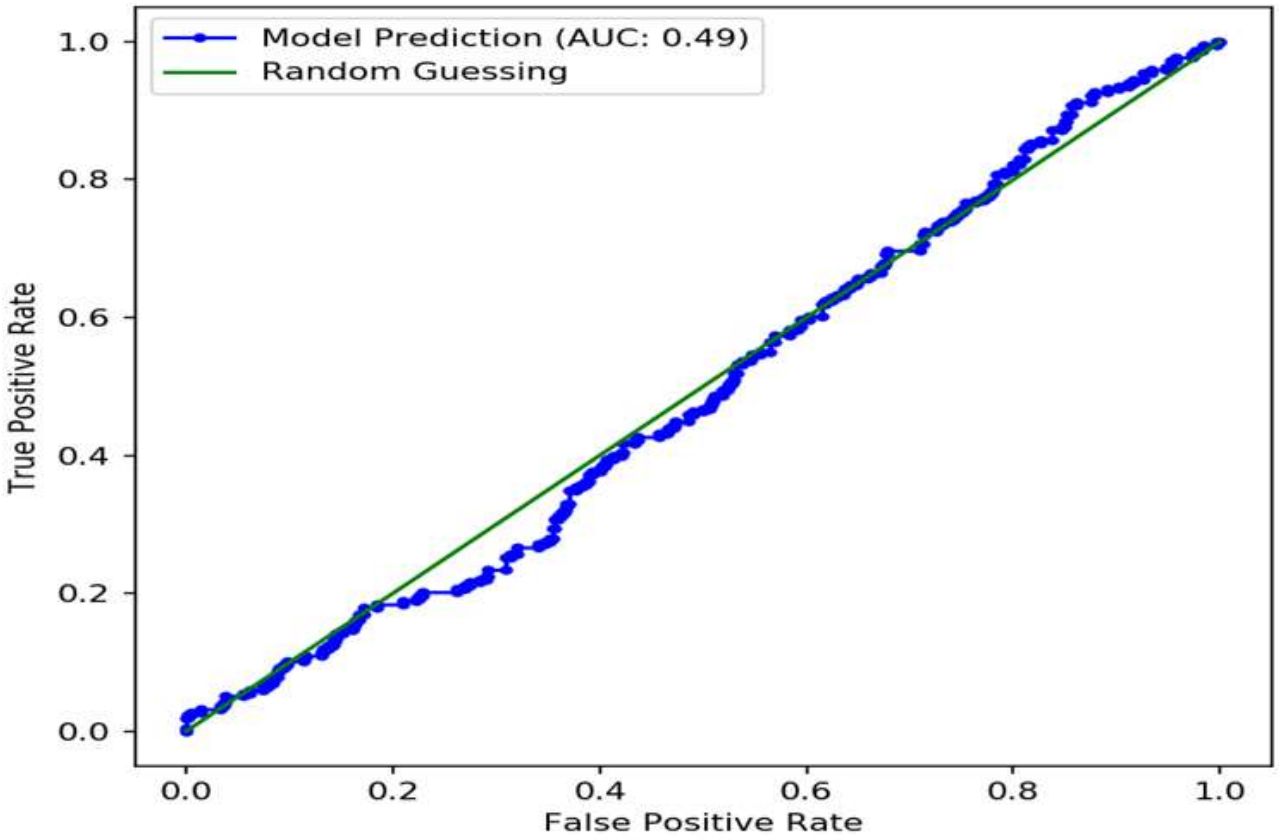


Fig 14: Model prediction graph.

The graph shows the true positive rate VS the false positive rate. This is crucial to determine whether the model has achieved notable accuracy. Sensitivity and specificity are important aspects of this prediction. True positive rate depicts sensitivity. False-positive rates depict the deviations from specificity.

VII. CONCLUSION

- In this work, a CNN system is created in order to get trained and to classify the fundus color photographs.
- Detects diabetic retinopathy through means of image classification based on CNN and using TensorFlow.
- Classify the images which helps the ophthalmologist in faster diagnosis.
- Prevents delay in results that otherwise would have led to a lost follow-up, miscommunication, and/or delayed treatment.
- With the ever-increasing number of individuals with diabetes, the need to prevent blindness due to DR can be reduced.
- Reduced manual processes that do not require a clinician to undergo training for examining and evaluating the digital fundus photographs of the retina. hence saving time.
- 'Eyecam' application used for image capturing and image file sharing helps save time.
- Thus, with extensions in the dataset, further classifications can be used by additional CNN within the cascade CNN.
- More target attributes such as age and gender can be further added to the classification.

REFERENCES

[1] Esfahani MT, Ghaderi M, Kafiyeh R. "Classification of diabetic and normal fundus images using new deep learning method," Leonardo Electron J Pract Technol 2018; 17(32):233–48, 2018.

[2] Gulshan V, et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," Am Med Assoc 2016;316:2402–10 2016.

[3] Shanthi T, Sabeenian RS. "Modified Alexnet architecture for classification of diabetic retinopathy images," Comput Electr Eng. 2019;76:56–64, 2019.

[4] Jiang H, Yang K, Gao M, Zhang D, Ma H, Qian W. "An interpretable ensemble deep learning model for diabetic retinopathy disease classification." In: 2019 41st Annual International Conference of the IEEE engineering in medicine and biology society (EMBC); 2019. p. 2045–8, 2019.

[5] Harangi B, Toth J, Baran A, Hajdu A. "Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features," In 2019 41st annual International Conference of the IEEE Engineering in Medicine and biology society (EMBC); p. 2699–702, 2019.

[6] Dulanji Lokuarachchi, Kasun Gunarathna, Lahiru Muthumal, Tharindu Gamage, "Automated Detection of Exudates in Retinal Images," IEEE, 2019.