Classification of Diabetic Retinopathy using CNN

Tarlana Sahil[1], Pettem Arush[2]

Abstract— **Diabetic retinopathy (DR) is a common side effect of diabetes which is the major cause for the of the loss of vision and helpful to keep vision consistent by early detection and treatment .It is the most common cause of blindness for people with all over the world . Early detection and classification of DR is essential for timely intervention and effective management to prevent blindness. In this study, we propose a data-driven approach using machine learning to classify diabetic retinopathy.In this project we have tried various preprocessing techniques like gabour filter ,Gaussian blur ,Principal Component Analysis(PCA),Circular crop and Canny edge sharpening etc.we used multiple models for classification of diabetic retinopathy namely Efficient netB3 ,Resnet50 and VGG16 , opted for the best possible model which has good accuracy .For selection of the best model we choose F1 score as a comparision metrics .We mostly concentrated on data preprocessing in this for getting better results .**

Keywords: **Diabetic retinopathy, CNN, Gabor Filter, blindness**

Kaggle Project link: **kaggle**

*Overleaf report link:***overleaf**

I. INTRODUCTION

The leading cause for the loss of vision and blindness is diabetic retinopathy (DR). Damage caused to retina from diabetes can be simply stated as Diabetic Retinopathy. It is the main reason for blindness in most of the developing countries. In manual surgery, the fundus image of the retina must be interpreted and interpreted by an ophthalmologist, which is quite expensive. Diabetes is increasing worldwide. Anyone with type 1 or type 2 diabetes is susceptible to this disease.

As the condition worsens, you may begin to see lumps, black spots, blurred vision, or blurred vision in your vision. Diabetic retinopathy can be mainly classified into two types namely Non-proliferative diabetic retinopathy (NPDR)" and "Proliferative Diabetic Retinopathy (PDR)". Diabetic retinopathy has four stages which are progressively called as mild, moderate, severe, and proliferative retinopathy. A small ball in the retina is called a microaneurysm The growth of similar blood vessels is a sign of the early stages of the disease. Mean unrepeatable retinopathy is a disease in which the number and size of balloon-like swellings called microaneurysms increase. Because of this, the blood vessels that supply oxygen, oxygen, and nutrients to the retina stop working. During this stage of retinal degeneration, much of the vascular system becomes blocked, restricting blood flow to many parts of the retina.

The current DR screening process is time-consuming and complicated by the lack of appropriately trained ophthalmologists. Fluorescein angiography is performed, in which the pupil is dilated and a photo of the retina is taken using a special camera, and then the doctor examines the

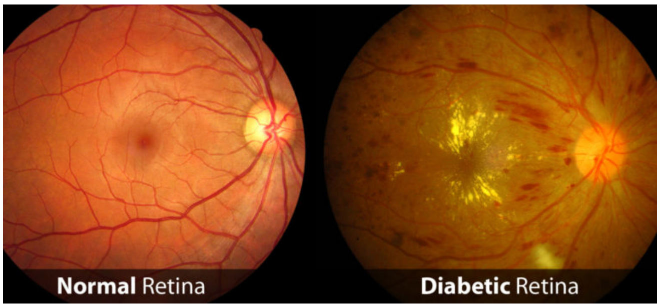

Fig. 1: Normal vs Diabetic retina

patient. It was estimated that in 2002, approximately 5 million people worldwide, or 5 percent of the world's blind people, suffered from diabetic retinopathy. Additionally, the lack of qualified doctors in rural areas where diabetes is common is also an important problem.The traditional ways for detection of Diabetic Retinopathy include's dilated eye examination,optical coherence tomography commonly called as (OCT) and analysing fundus images by specialists. even though these are effective, most of them are time-consuming and need specialized interpretation by trained professionals

A. Key Contributions

- CNN Model:Development of a CNN models like efficientnet b3 , resnet50 , vgg16 tailored for diabetic retinopathy classification.
- Accuracy: Compared the accuracy of all the models we developed and selected the best-performing modelfor DR detection.
- Impact in the medical field: This Project demonstrates the potential impact of Deep learning in medical image analysis.

II. PROBLEM STATEMENT

The main objective for our project is to make a ml model which is capable for classifying the fundus images into four stages which are progressively called as mild, moderate, severe, and proliferative retinopathy.

A. Objectives

- Dataset: Find a relevant dataset that is appropriate for training our model from a reliable source, such as publicly available repositories (e.g., Kaggle, ImageNet, roboflow).
- Data preprocessing: Remove any corrupted or low-quality images from the dataset.Resized the fundus images to a consistent input size required by the CNN model (e.g., 224x224 pixels)

- Compare the CNN models: Compare the accuracy of all the models and choose the best-performing model.
- Potential real-world applications: The CNN model can help in assisting medical professionals with early detection and monitoring of diabetic retinopathy, ultimately contributing to improved patient care and disease management.

III. RELATED WORK

Kajan et al. (2020) developed strategy of robotizing diabetic retinopathy (DR) determination utilizing pretrained profound neural systems (DNNs). They assessed VGG-16, ResNet-50, and Inception-v3, accomplishing an exactness extend of 94.1%to 95.1% for DR recognizable proof with their proposed approach.

Shankar et al. (2020) proposed an mechanized methodology based on profound learning to identify and classify fundus pictures of DR. Their strategy coordinates Inception-v3, VGG-19, and DenseNet-121, accomplishing an in general victory rate of 96.8% for DR discovery and classification.

Mateen et al. (2020) displayed a novel approach for recognizing exudates in fundus pictures related to DR, utilizing pretrained convolutional neural systems (CNNs). Their strategy illustrated tall precision in exudate location, with a affectability of 91.7% and a specificity of 99.5%, beating other modern profound learning models for exudate recognizable proof.

Abini M. A; S. Sridevi Sathya Priya (2023) With the pre-trained Convolutional Neural Networks (CNN) VGG-16 and MobileNet-V2, the accuracy rates for determining the extent of diabetic retinopathy from a picture were 90% and 92%, respectively.low accuracy rate when dealing with fungal pictures of retinas in good condition and unable to classify early stages of diabetic retinal disease.

Nandhini S; Sowbarnikkaa S; Mageshwari J (2023) Has the class imabalnace ,majority class has almost 50% of the total image and minority class has around 9% of the total images which resluts in the bad perfomance of the model in the minority classes.

IV. PROPOSED METHODOLOGY

Before proceeding with the data into the machine learning models, we performed various pre-processing techniques on the images of our dataset and verified whether the particular technique helped to increase model performance or not. We resized all the images to 224×224 pixels to ensure that they have the same dimensions. We are using the dataset that has already categorized the images into five folders based on their corresponding classes, ranging from 0 to 5. We also tried various augmentations on the training data using techniques such as shear, zoom, and horizontal flip to increase the size of the dataset and to make the model more robust to variations in the images.

For our model architecture, we opted to try three different pre-trained models namely ResNet 50, EfficientNetB3, and VGG16 and use all of them and select which performs well in terms of accuracy, F1 Score, etc. We have finetuned the model according to our images in the dataset and added extra layers considering better perfomance of the model . We have added a custom classification layer consisting of a GlobalA veragePooling2D layer followed by two fully connected layers with ReLU activation followed with two Dropout layers and a SoftMax activation layer at the end. We have freezed the weights of pre-trained layers of the models to utilize the information.

During model training, we have used a learning rate 0.001 with Adam optimizer .we tried both categorical cross-entropy loss and binary cross entropy function and decided to use the previous one as it is producing better results . In addition to accuracy, we have also evaluated F1, which is more appropriate for imbalanced datasets like ours. The model was trained for 20 epochs with an early stopping patience of 3,5 based on the model to prevent overfitting. Additionally, we employed a Model Checkpoint callback to save the weights of the model with the highest validation loss score.

After evaluation of the model was finished, we used the history during the training process and plotted the results using Matplolib to get a better understanding over the model and later we verified the model with the test dataset we have using the saved weights and displayed the results.

A. Dataset

We used publicly available datasets of ROBOFLOW DATASET UNIVERSE which is a part of APTOS 2019 dataset, this dataset we are using is augmented already and splitted the dataset into train ,test and validation with split ratio of 80,10,10 .Initially, we tried working with EyePACS dataset contains 35,126 retinal images of diabetic retinopathy which makes it difficult to handle with the data . Our dataset includes images from five different classes: no DR, mild , moderate , severe , and proliferative DR. The images were collected from primary care clinics and graded by trained human graders. The APTOS 2019 dataset contains 3,662 retinal images in original . The dataset includes images from five different classes, ranging from 0 (no DR) to 4 (proliferative DR). The images were collected from different sources and are of varying quality. upon data analysis we found the existence large number of noisy and Blur images

.

Our dataset consists of retinal fundus images classified into five classes ,the specifications of each class are discussed below:

1) No DR: This class represents images without any signs of diabetic retinopathy. These images typically indicate a healthy retina. This class consists of 3251 images.
2) Mild DR: Images in this class show mild signs of diabetic retinopathy, such as microaneurysms or small retinal hemorrhages. This class consists of 588.
3) Moderate DR: This class includes images with moderate diabetic retinopathy, characterized by a greater number of hemorrhages and signs of retinal damage. This class consists of 1547.
4) Proliferate DR: Images in this class exhibit severe diabetic retinopathy with significant retinal damage,

including the growth of abnormal blood vessels. This class consists of 484.

5) Severe DR: This class represents images with advanced and severe stages of diabetic retinopathy, often accompanied by extensive retinal damage and potential vision loss. This class consists of 315.

B. Data Preprocessing

We have tried various preprocessing techniques and verified the results after each trial and finally decided with a set of preprocessing techniques which we fell produce efficient results.The list of various techniques we tried consists of circular crop,conversion of gray scale images and incresing the sharpness using canny edge detection and addigng extra layers of edges to the orginal image ,gabour filter which was used in the many of the papers implementing diabetic retinopathy classification we also use Ben graham technique of gaussian blur kernel,adjusting the Brightness and Contrast of all the images using opencv libraries and finally Histogram equalization .we skipped using histogram equaiation as the results are not so good when tried incorporating it as a pre processing step.

Below we discussed some of the techniques that we used for preprocessing in detail:

1) Gabor Filter Technique:
The Gabor filter is a powerful tool used in image processing for various tasks such as texture analysis, edge detection, and feature extraction. It is defined mathematically as a complex sinusoidal plane wave modulated by a Gaussian function.

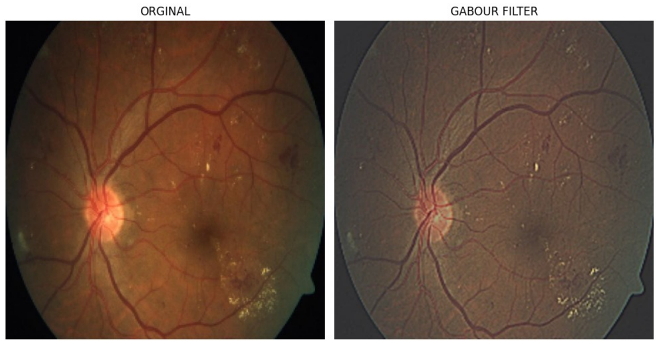

Fig. 2: Gabour Filter

Mathematical Formulation:
The 2D Gabor filter is given by the following equation:

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

where:

$$x' = x\cos(\theta) + y\sin(\theta)$$
$$y' = -x\sin(\theta) + y\cos(\theta)$$

Here,

x, y are the coordinates of the image.

λ is the sinusoidal factor's wavelength..

θ is how the normal is oriented in relation to the parallel stripe

ψ is the phase offset.

σ is the standard deviation of the Gaussian envelope.

γ is how the normal is oriented in relation to the parallel stripe

The Gabor filter operates on images by convolving them with this defined kernel.

2) Principal Component Analysis (PCA):
The most widely used method for reducing dimensionality in data is principal component analysis (PCA), which finds a lower dimensionality representation for the input data while maintaining information and variance in the data. Prior to feeding high-dimensional data—like images—to machine learning models, PCA is used to preprocess the data. PCA can be useful in converting the input retinal images into a lower-dimensional feature space for diabetic retinopathy classification using fundus images, possibly capturing the most pertinent information for the classification task.

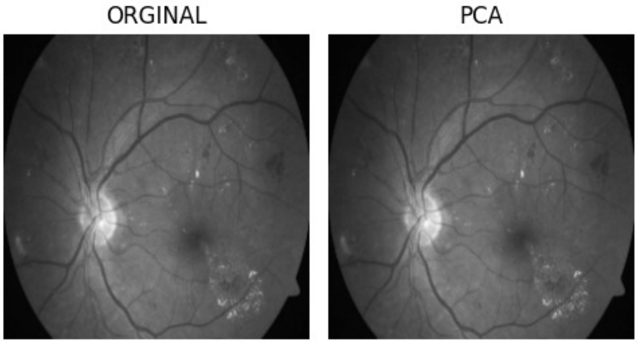

Fig. 3: PCA

Mathematical Formulation: Let A be the input data matrix of size $(n \times p)$, where N is the number of samples (retinal images) and p is the number of features (pixels). The covariance matrix Σ of X is computed as:

$$\Sigma = \frac{1}{N} A^T A \tag{1}$$

PCA finds the eigenvectors and eigenvalues of Σ by solving the equation:

$$\Sigma V = \lambda V \tag{2}$$

In this case, the eigenvalue λ corresponds to the eigen vector v, which is the eigenvector.The direction of the maximum variance in the data is represented by the eigen vector of the largest eigen value.Projecting X onto the subspace spanned by the top k principal components yields the transformed data Y:

$$Y = XP \tag{3}$$

where P is the matrix of size $(p \times k)$ containing the top k eigenvectors of Σ as columns. The transformed data Y has a reduced dimensionality of k, which can be significantly lower than the original dimensionality of p, facilitating more efficient processing and analysis.
The most pertinent and important information for the diagnosis of diabetic retinopathy can be retained with us while the dimensionality of the data is decreased to very less as possible by using PCA to preprocess retinal images. This can be helping in lowering the computational complexity and noise of the data , which may also useful in enhancing the performance of later machine learning models.

3) Canny Edge Detection:
The Canny edge detection algorithm is the most popular one for detecting edges in images. It involves several steps include Gaussian smoothening, gradient computation, non-maximum suppression, and edge tracking by hysteresis. Adding extra edges to an image using Canny edge detection can be achieved by manipulating the parameters of the algorithm. For example, by adjusting the threshold values used for edge detection, you can control the sensitivity of the

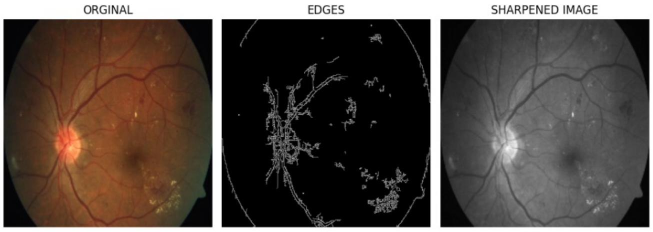

Fig. 4: Canny edge

algorithm and potentially detect additional edges in the image.
Mathematical Formulation:
Let I be the input image. The Canny edge detection algorithm computes the gradient magnitude M and orientation θ of each pixel in I. This can be expressed as:

$$M(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}$$

$$\theta(x,y) = \arctan\left(\frac{G_y(x,y)}{G_x(x,y)}\right)$$

where G_x and G_y are the derivatives of the image I in the x and y directions respectively.
The algorithm then applies non-maximum suppression to thin the edges:

$$M_{\text{thin}}(x,y) = \begin{cases} 0 & \text{if } M(x,y) \text{is not a local maximum} \\ M(x,y) & \text{otherwise} \end{cases}.$$

Finally, edge tracking by hysteresis is performed to detect continuous edges.

4) Automatic Circular Crop of Eye Retina Images:
In the analysis of eye retina images, a crucial preprocessing step involves the automatic circular cropping of these images. This process is essential for focusing on specific regions of interest, such as the optic disc or the fovea, which are pivotal for various diagnostic and research purposes in ophthalmology. Mathematical

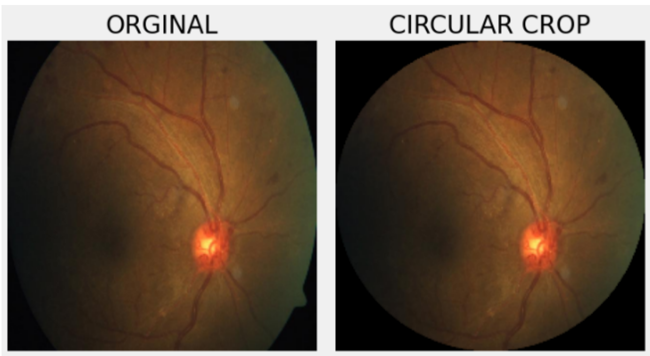

Fig. 5: Circular Crop

Formulation:
Consider an eye retina image I with dimensions $M \times N$. Let C_x and C_y represent the coordinates of the center

of the detected optic disc or fovea. The circular mask M_c can be defined mathematically as:

$$M_c(i,j) = \begin{cases} 1 & \text{if } (i-C_x)^2 + (j-C_y)^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases}$$

where R denotes the radius of the circular mask. Subsequently, the cropped image I_c can be obtained by element-wise multiplication of the input image I with the circular mask M_c:

$$I_c(i,j) = I(i,j) \times M_c(i,j)$$

This process ensures that only the pixels within the circular region of interest are retained, effectively achieving automatic circular cropping of eye retina images.

Algorithm Overview:

The automatic circular cropping algorithm for eye retina images can be summarized as follows:

1) Detect the location of the optic disc or the fovea region in the input image I.
2) Define a circular mask centered at the detected location to delineate the region for cropping.
3) Crop the image I using the circular mask to isolate the region of interest, typically containing the optic disc or the fovea.

The effectiveness of the algorithm largely depends on the accuracy of the optic disc or fovea stage, which in turn influences the precision of the subsequent circular cropping process.

Adjusting Parameters:

To add extra edges to the image, we can adjust the threshold values used in the Canny edge detection algorithm. Lowering the thresholds can increase the sensitivity of the algorithm and potentially detect weaker edges in the image.

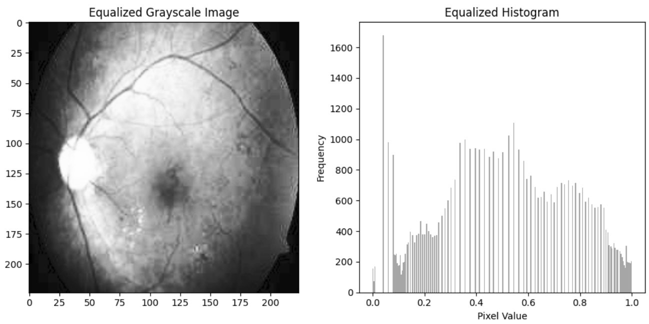

Fig. 6: Histogram equalization of a Greyscale image

C. Models

We selected to use three models for our diabetic Retinopathy classification. We used the pre-trained models from the TensorFlow and processed our dataset to feed to the model using the ‘flow from directory‘ method available in TensorFlow. The below is the list of models we have used, an

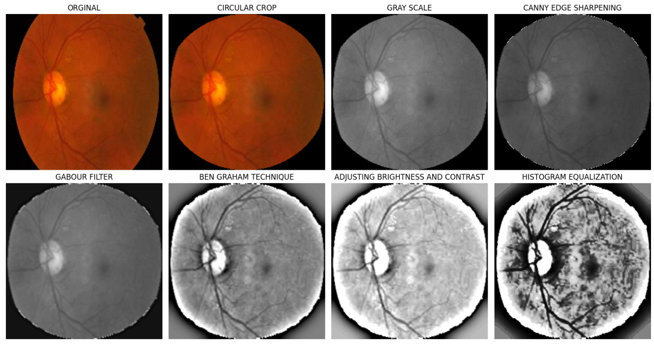

Fig. 7: preprocessing filters

explanation about their working, and the algorithms involved in them in detail:

1) ResNet-50: Microsoft Research unveil's the ResNet-50 model, which is a deep convolutional neural network(CNN) armature, in 2015. It is belongs to the family of the models named as ResNet, which uses residual connections to let the network skip one or more than one layers while it's being trained. With 50 layers and also pre-training on the ImageNet dataset, the ResNet- 50 model is a better possible choice for computer vision transfer literacy tasks. Residual connections, which enable's the network in learning the residual mappings rather than the asked underpinning chart-clunk directly, are the main invention of ResNet. This lessons the possibility of the evaporating/ exploding grade issue in veritably deep neural networks. Skip connections are used to apply the residual connections, where the input to the affair of several piled layers is added to a subcaste. Because of this, learning the residual mapping by the network is simpler than learning the original, unreferenced mapping.

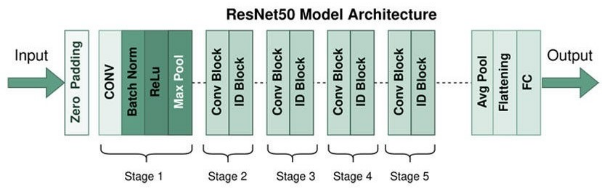

Fig. 8: Resnet50 architecture

2) EfficientNet-B3: EfficientNet- B3 The EfficientNet- B3 model is a part or a member of the EfficientNet models family, which aims in maximize performance with minimum calculations and good rameters. Grounded on a emulsion scaling fashion that uses a emulsion measure to inversely gauge the network's depth, range, and resolution, Google AI unveiled theEfficient- .Net models in 2019. Amid-range model in the family, EfficientNet- B3 strikes a balance between computational effectiveness and delicacy.

The depth, range, and input resolution of the network are gauged logically in the EfficientNet models through the use of the idea of model scaling. The delicacy and effectiveness trade- off of the model is the base for this scaling, which

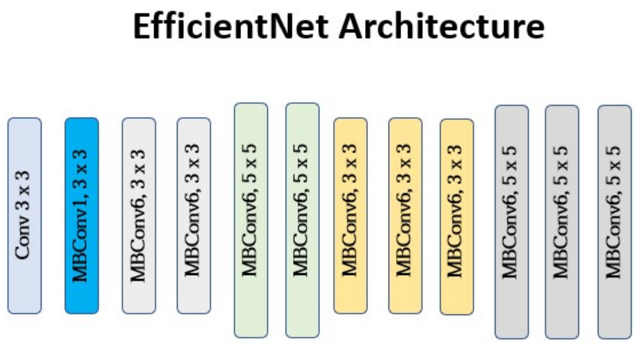

Fig. 9: EfficientNet-B3 architecture

TABLE I: Model Test Accuracy

Model	Test Accuracy
ResNet50	0.7472
VGG16	0.7055
EfficientNet B3	0.7611

is grounded on a emulsion measure. Neural armature hunt(NAS) is a system that looks at numerous combinations to identify the scaling factors.

3) VGG-16: The University of Oxford's Visual Geometry Group (VGG) unveiled the VGG-16 model, a deep CNN architecture, in 2014. One of the first models to successfully use a deep stack of convolutional layers, it has sixteen layers and shows the promise of deep networks for computer vision applications. In a variety of image classification applications, the VGG-16 model has been extensively utilized as a baseline for transfer learning. The VGG-16 architecture is based on

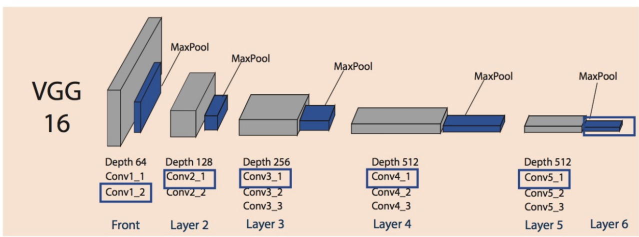

Fig. 10: VGG16 architecture

a simple pattern of max-pooling layers and stacking convolutional layers with small receptive fields (3x3). Thirteen convolutional layers make up the model, and then three fully linked layers. The convolutional layers are organized into five blocks, each of which has several convolutional layers with the same filter size. These blocks are as follows:

V. EXPERIMENTAL DETAILS

1) Initially, the model started to overfit after some epochs, so we implemented early stopping of the model by considering the parameter validation loss and patience level of 3 or 5 epochs based on the particular model we use. For VGG16, it is 3 epochs, and for the other two models, it is 5 epochs.
2) Regarding the Gabor filter, we tried the existing parameters mentioned in various papers but ended up with bad results in terms of model accuracy. Ultimately, we decided to do trial and error on all possible parameters, and we fixed the parameters to:

```
sigma = 1.99    # Standard deviation
lambda = 9.8    # Wavelength
gamma = 6.08    # Spatial aspect ratio
```

3) we also used gabour kernel in four possible directions [0,pi/4,pi,2,3pi/4] for enhancing the overall visibility of blood vessels and other important components in all directions
4) Regarding canny edge detection we tried various thresholds of lower and upper limit for detecting the blood vessels and decided to use a threshold range of 50,120,we detect the edges using canny edge detector and added three layers edges to the edges to the original model to increase the sharpness and visibility of blood vessels.

VI. RESULTS AND ANALYSIS

A. *Efficientnet b3:*

Efficientnet b3 gave Test Accuracy: 0.76%

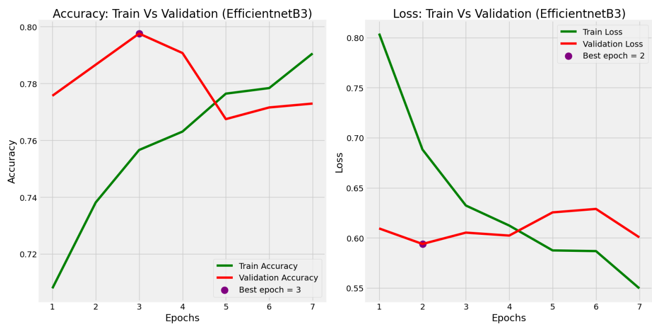

Fig. 11: Accuracy and Loss for Efficient B3

B. *VGG16:*

VGG 16 gave Test Accuracy: 0.70%

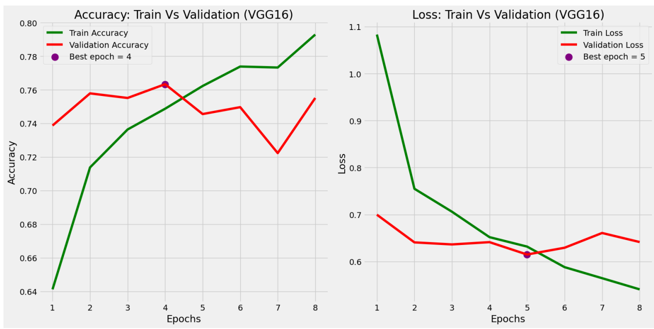

Fig. 12: Accuracy and Loss for VGG16

C. Resnet50:

Resnet50 gave Test Accuracy: 0.74%

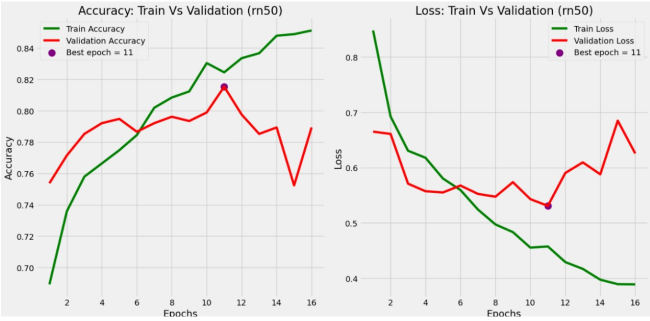

Fig. 13: Accuracy and Loss for Resnet50

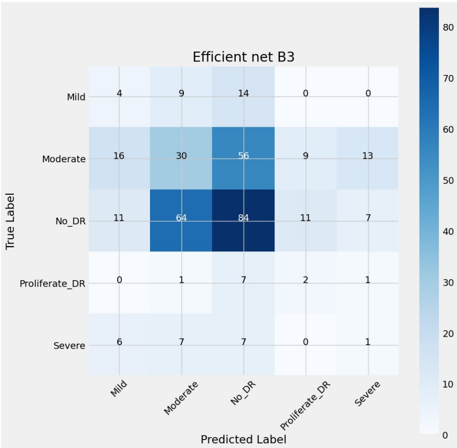

Fig. 14: Confusion Matrix for efficientnet b3 model

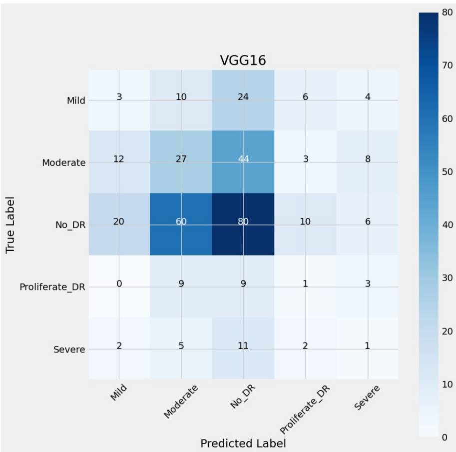

Fig. 15: Confusion Matrix for vgg16 model

VII. CONCLUSION

In conclusion, we decided to use the Efficient netB3 model for the Diabetic retinopathy classification as final one as it produces good accuracy when compared to others even though it has low F1 score which is mainly due to the class imbalance that we have in our dataset .we tried oversampling the images using the SMOTE but the obtained images are not so good.Also upon considering the augmentations due to varied augmentations and less number of data also may be reason for the low f1 score and recall.In further we want to try good ways for resampling the data and increase the size of the dataset ,and also make further modifications in the model for getting better results .Also we want to try other available preprocessing techniques for fundus images ,explore more in the field of medical imagning and thier diagnosis.

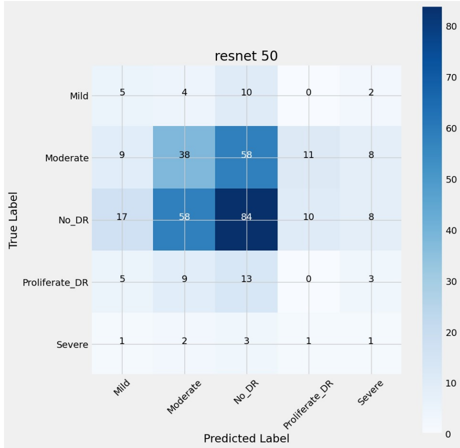

Fig. 16: Confusion Matrix for resnet 50 model

Model	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss
ResNet50	0.8597	0.3759	0.7893	0.6257
VGG16	0.7753	0.5655	0.7250	0.6586
EfficientNet B3	0.7844	0.5558	0.7729	0.6005

Fig. 17: overall perfomances of all models

VIII. References

- https://ieeexplore.ieee.org/document/10220715
- https://ieeexplore.ieee.org/document/10157960
- https://ieeexplore.ieee.org/document/10182642
- https://ieeexplore.ieee.org/document/10370255
- https://ieeexplore.ieee.org/document/10169453
- https://www.sciencedirect.com/science/article/pii/S235264832200