A deep learning strategy using the CNN model DRDnet22 for the early detection of diabetic retinopathy from retinal images

Submitted by

Md. Sakib Ali Mazumder

191-35-2683

Department of Software Engineering

Daffodil International University

Supervised by

Dr. Imran Mahmud

Associate Professor & Head

Department of Software Engineering

Daffodil International University

This Thesis report has been submitted in fulfillment of the requirements for the Degree of Bachelor of Science in Software Engineering.

APPROVAL

This thesis titled on "**A deep learning strategy using the CNN model DRDnet22 for the early detection of diabetic retinopathy from retinal images**", submitted by **Md. Sakib Ali Mazumder (ID: 191-35-2683)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

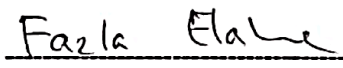
BOARD OF EXAMINERS

Chairman

Dr. Imran Mahmud
Head and Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

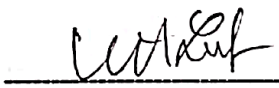
Internal Examiner 1

Afsana Begum
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University



Internal Examiner 2

Dr. Md. Fazle Elahe
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University



External Examiner

Mohammad Abu Yousuf, PhD
Professor
Institute of Information Technology
Jahangirnogor University

DECLARATION

I hereby declare that this thesis (**A deep learning strategy using the CNN model DRDnet22 for the early detection of diabetic retinopathy from retinal images**) has been done by me under the supervision of Dr. Imran Mahmud, Head & Associate Professor, Faculty of Science and Information Technology, Department of Software Engineering, Daffodil International University. It is additionally declared that neither this thesis nor any component has been submitted elsewhere for the award of any degree. All declarations are fully verified for completeness and the validity of their data element contents.

Submitted By:

Md. Sakib Ali Mazumder
ID: 191-35-2683
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Certified By:

Dr. Imran Mahmud
Head and Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

ACKNOWLEDGEMENT

First, to Almighty Allah, I express my heartiest thanks and gratefulness for His divine blessing in making it possible to complete the final year thesis successfully. I would like to express my gratitude to my family, for the constant support and faith they have shown me throughout my life.

I would want to convey my appreciation to my supervisor **Dr. Imran Mahmud**, Associate Professor & Head of the Department of Software Engineering for facilitating me with the mentorship I needed to complete this thesis. His scholarly and expert guidance, endless patience, proactive supervision, constructive feedback, and wholehearted support have made it possible to overcome obstacles and complete this thesis.

Also, I would like to thank all of my friends, both Alumni, and juniors, who have helped in some way. I would like to express my heartiest gratitude to all the faculty members and the staff of the Software Engineering department of Daffodil International University.

Finally, I would like to express my gratitude to my family, for the constant support and faith they have shown me throughout my life.

ABSTRACT

Diabetic Retinopathy, or DR, is a major problem with the eyes of diabetes patients. If the DR is found early, many people can avoid going blind and other complications. Several systems based on artificial intelligence have been proposed, and they are better at finding the DR than human analysis in terms of time efficiency and human dependencies. Manual screening using retinal fundus images, such as visual acuity testing, pupil dilation, and optical consistency tomography, requires highly skilled clinicians to find and evaluate the importance of many small details. This is a difficult, time-consuming, and error-prone task. Because of this, a computer-aided, automated process is a must. In this thesis, the APTOS 2019 dataset is used for training and testing. Which is made up of 3662 named pieces of data. DRDnet22 is a model that uses CNN to find and classify DR into 5 phases based on the severity level. Since each ConvNet gets different features, combining them with 1-D pooling and cross-pooling gives a better representation than just using the features from one ConvNet. For comparison, the traditional pre-trained model was also trained. Evaluation of performance indicates that the proposed model was more accurate with 81.6% than the traditional pre-trained models.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The diabetic consequence of diabetic retinopathy (DR) is caused by excessive blood sugar levels that damage the retina. Severe diabetic microvascular consequence diabetic retinopathy has a major influence on global healthcare systems. Despite being possibly avoidable and treated, DR is recognized in most nations as one of the top causes of blindness in the working-age population, with catastrophic personal and social implications. The sharp rise in DR patients numbers; is now considered a global epidemic. If DR is caught early, it is possible to keep many people from going blind. People with DR get small lesions in their eyes. The scale of severity is determined by the sort of lesions that form in the eyes.

1.1.1 Background Information Globally

About 425 million adults have diabetes in the world; by 2045 this will rise to 629 million. Almost all diabetic patients develop retinopathy within 5 to 15 years of the onset of diabetes. Diabetic retinopathy is the leading cause of vision loss in working-age adults (20-65 years). Approximately one in three people with diabetes have diabetic retinopathy and one in ten will develop a vision-threatening form of the disease.

1.1.2 Background Information in Bangladesh

Bangladesh has among the world's largest numbers of people with diabetes, with 35% of the population older than 35 years receiving a diagnosis of diabetes or prediabetes. In 2020, the country was home to around 1.85 million diabetic retinopathy patients, according to the International Agency for the Prevention of Blindness. 10% to 30% of diabetic patients in Bangladesh suffer from this progressive eye disease. With around 10 million diagnosed patients, Bangladesh remains at the forefront of the global diabetes epidemic, where one out of every ten adults is diabetic, according to 2017 data from the International Diabetes

Federation. From the study (Khatun, 2021) a harsh reality appeared that there is still insufficient awareness in the context of early detection and prevention of diabetic retinopathy and blindness due to that mass level people.

1.2 MOTIVATION OF THE RESEARCH

Knowing the fact that diabetic retinopathy can cause blindness it caught my attention that as the number of diabetic patient rates in Bangladesh is increasing day by day; Bangladeshi patients are in an alarming phase of which a lot of people are still not aware. The traditional manual diagnostic system takes a lot of time, patience, cost, and complex dependencies as well. I have come to understand that diabetic retinopathy may also be a drag that has a significant impact on patients in our society and our perception of diagnosing complications by researching diabetic retinopathy and related research and thesis papers. The need for automated DR detection and classifying severity levels is of paramount importance if we intend to avoid complications to improve the diagnosis system in our society, especially the less privileged persons or patients who can get benefited from this.

1.3 PROBLEM STATEMENT

The complications of diabetic patients are always at a gradually increasing rate. Diabetic Retinopathy among them. We know that diagnosis of these diseases is kind of problematic and complicated in our country. Early, many studies on detecting DR have already been done by many researchers. Some approaches of their study were quite impressive too. Some studies were done only with the preprocessed dataset as input. Which is not the most convenient way of solving with deep learning. Some studies obtained pretty high scores on classifying DR and Non-DR but they are lacking in classifying severity levels. Even though there are a lot of existing approaches, they are either not the best or are hard to use. So, we need an alternative that is both simple and stable.

1.4 RESEARCH QUESTIONS

- How efficient will be deep learning to detect diabetic retinopathy from retinal images as input?
- How much accuracy can be obtained with the proposed DRnet22 CNN model?
- Does DRnet22 perform better than pre-trained CNN models?

1.5 RESEARCH OBJECTIVE

- Deep learning is quite efficient to detect and classify diabetic retinopathy in terms of time complexity, and skilled clinicians' dependencies.
- The proposed CNN model can obtain a high accuracy score to classify the severity levels.
- There is still quite a bit lacking for the regular pre-trained model to acquire a high correction rate. DRDnet22 overcomes the lacking and performs better than the regular pre-trained models.

1.6 RESEARCH SCOPE

- Manual diagnosis using retinal fundus images, such as visual acuity testing, pupil dilation, and optical consistency tomography, requires highly skilled clinicians to locate and evaluate the importance of many small details. This is a difficult, time-consuming, and error-prone task.
- A Computer Aided Diagnostic (CAD) system will overcome the high complexity making this task easier for health clinicians and ophthalmologists in quick decision making
- By using this Deep learning strategy, professionals will spend less time solely looking at abnormal images and more time looking at every fundus image.
- An automated System will improve in general Diagnosis System of this traditional DR detection efficiency.

1.7 THESIS ORGANIZATION

In the first chapter, there is a section on the diabetic retinopathy detection system and how it is used. The background of the work, why the research was done, the problem statement, the research questions, and the research's main objective are also talked about. The other parts that are important to our research are:

In the next chapter, I will talk about the literature review. This is where we can observe what other researchers have done in the same field of detecting DR, what methods they used, where they fell short, and how my work compares to theirs. In chapter 3, I will describe how my workflow was done. In the methodology section, I will focus on

about Data collection phase, how to pre-process the data was done, and the rest of the workflow. In chapter four, a brief discussion about the results of the method was taken place. Following is the conclusion which is the final chapter is the last one. Here, I have talked about what work should be done in the future to make things better and more conclusive.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

In the literature review section, the researcher study and assess the work, research, conference papers, journals, books, articles, etc. that has already been done. By this, one can see what work has previously been accomplished on the subject matter, sum up the whole thing, and see what needs to be done. After assessing, they can improve things on their weaknesses and develop strategies to get around them so they can obtain an improved outcome.

2.2 RELATED WORK

In recent years, machine learning models have become very popular as a way to solve a wide range of problems (Das, et. al.,2017). For example, they can be used to classify images, process text, find faults in real-time, and help with health care. ML algorithms are often used to make predictions about diseases. A study (Hinton, G., 2018) evident that deep learning has a high amount of potential to transform traditional healthcare systems as well.

This section discusses standard DR recognition models from the literature. Numerous studies have already been done to find the best way to use a computational approach to find and classify DR from image data of the retina. Researcher moves toward Artificial intelligence and deep learning to identify retinopathy (Wong, et. al., 2016). In a study (Bodapati et. al., 2020), a Blended Multi-Modal Deep ConvNet was suggested as a way to predict how bad diabetic retinopathy will be. Experiments on the benchmark Kaggle APTOS 2019 competition dataset show that the proposed blended feature representations are better for training than the methods that are already being utilized. They notice that the best way to spot DR is to use cross-average pooling to combine features from Xception and VGG16. With the model the authors made, they can predict the severity level with an 81.7% accuracy and a 71.1% kappa statistic. In a study conducted by (Gangwar, et. al., 2021) a novel was proposed by the authors to detect DR. For the hybrid model, they used transfer learning on Inception-ResNet-v2 with pre-trained weights and added a custom

block of CNN layers on top of Inception-ResNet-v2. The proposed model was tested on the Messidor-1 diabetic retinopathy dataset and the APTOS 2019 blindness detection dataset (Kaggle dataset). The model did better than other results that were already out there. On the Messidor-1 dataset, the model achieved a test accuracy of 72.33% and on the APTOS dataset, the model got a higher accuracy. The LeNet convolutional network was used by (Perdomo et al., 2017) The CNN model does a great job of finding exudate in images of the eye fundus. Early results of this study were promising for finding exudate. A study (Zubair et al. 2021) proposed a model that uses VGG-NiN to classify the different phases of the DR using the fewest number of learnable parameters to speed up model training and convergence. They did this with the EyePACS dataset from Kaggle. Of the 88,702 photos in the dataset, 53,576 are not labeled, while 35126 are. Fundus Imaging Categorisation and Lesions Localization System was shown off by a study by (Alyoubi et al., 2021). They used two public datasets. The DRD dataset and the APTOS 2019 Kaggle dataset were both used. The DRD dataset has 13,673 images of the fundus that were taken with a 45º field of view (FOV). There are 6,266 normal images, DR images of 6,256, and 1,151 images that weren't labeled. For each category of DR lesion, 757 images have been labeled with a box (MA, soft EX, hard EX, and HM). The sizes of the 3,662 retinal images in the APTOS-2019 Kaggle Dataset are different. But the general public can only understand the ground truths of the training images. The dataset is put into five groups based on the DR phases. 1857 images of the DR category, and 1805 are normal. Their CNN512 got an accuracy score of 0.886 on the DDR dataset and 0.841 on the APTOS 2019 dataset. With a 0.216 mAP, the YOLOv3 model was able to successfully locate lesions on the DDR dataset. It was also used to find and identify the DR lesions. The authors (Ramya et al. 2022) came up with a new way to classify exudates from an image of the fundus. They called it the hybrid CNN-based binary local search optimizer–based particle swarm optimization algorithm. Image augmentation is used in this paper's proposed method to make the fundus image the right size without dropping any details. Recent research by (Bitto, Mahmud I.,2022) used the architectures of convolutional neural networks (CNNs) from the visual geometry group (VGG-16), ResNet-50, and Inception-v3 to differentiate between normal eyes, eyes with conjunctivitis, and eyes with cataracts. Inception-v3 is the best at recognizing eye diseases. ResNet-50 comes in second. Lastly, VGG-16 is the least

good at spotting eye diseases. Though there is a slight limitation of disease categorization as the used dataset did not focus on more categorization of eye diseases. This publication (Asiri, et. al.,2019) discusses cutting-edge deep learning-based diabetic retinopathy detection methods to empower researchers. (Prasad, et al., 2019) developed GUI-based (graphical user interface) CNNs. The Glaucoma datasets came from Medimg, and the Retinopathy datasets came from Kaggle. The rate of accuracy was set at 80%. Using methods like fine-tuning the parameters and cross-validation. (Hassan et al. 2021) proposed a novel cascaded decoupled convolutional network with two different modules that perform together to grade DR based on lesions and medical standards. The findings indicate that the proposed framework works better than other cutting-edge frameworks because it got a mean Dice score of 0.820 when it was used to divide up retinal lesions. (Upadhyay et al., 2022) The proposed model is a transition to the structure of the existing VGG-16, which is used to analyze the complex artifacts in the retinal region. The proposed deep neural network works well enough to get an accuracy of 89% for binary classification as DR and NO_DR.

2.3 CONTRIBUTIONS

Recent studies show that the diagnosis of Diabetic Retinopathy (DR) diseases has already moved to a higher level with a lot of different methods. There are many different kinds of algorithms, models, and methods were already proposed. For achieving a satisfactory outcome, they implemented a feature extraction method, augmentation, interpretation, and several other techniques. Even so, some of them depend a lot on complicated steps and methods for extracting features. Some studies also used other kinds of algorithms, which increased the amount of time needed to do the calculations and caused their methods to work slowly. Some work well with certain dependencies and some do have limitations. We also tried to identify multiclass grading retinopathy as part of our work.

CHAPTER 3

RESEARCH METHODOLOGY

The next method, shown in Figure 3.1, is consisting of five steps that depict the whole workflow of this research methodology. The steps of our method begin with collecting data, then the preprocessing phase of the data, extracting features using a Convolutional Neural Network, training the proposed DRDnet22 model, and figuring out and evaluating how accurate the model is by making predictions in the testing phase.

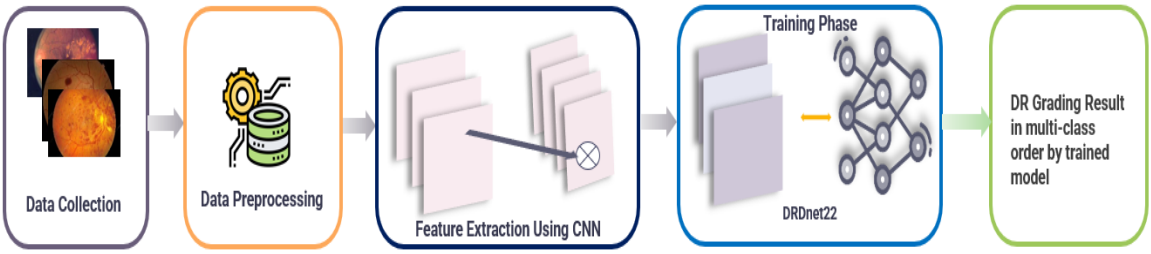


Figure 3.1: Workflow Phases

3.1 DATA COLLECTION

Our study was based on a dataset called Asia Pacific Tele-Ophthalmology Society Dataset, or APTOS-2019 Dataset, which was made available to the public by the Aravind Eye Hospital in India (APTOS 2019 Blindness Detection | Kaggle. (n.d.)). Messidor-1, Messidor-2, Aptos-2019, and EyePacks are some of the public datasets that can be used to grade the DR. The reason to think about the APTOS-2019 Dataset is that after this dataset was published, many researchers were interested in using it to build a solution and contribute to this area of research. The dataset consisting five stages to find out how bad the disease is: 0: "No DR," 1: "mild stage," 2: "moderate stage," 3: "severe stage," and 4: "proliferative diabetic retinopathy (PDR) stage." These stages help people in rural areas, where medical screening is hard to do, find out if they have this disease and take steps to

prevent it. Technicians from Aravind went to these rural areas to take pictures with high-resolution specialized fundus cameras, which are made up of a complicated microscope attached to a camera with a flash. They then relied on highly skilled doctors to examine the images and label them with a diagnosis.

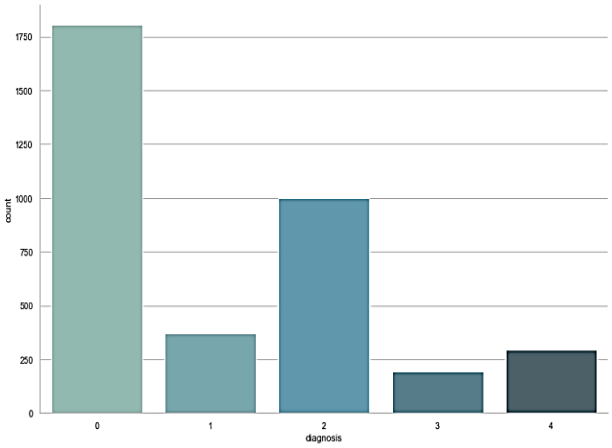


Figure 3.2: Dataset Multiclass labelling

There were 3662 training samples and 1928 test samples in the dataset as a whole. The images were made available in different sizes, such as 2416 × 1736, 819 × 614, and 3216 × 2136. Even though there were 1928 test samples, their label annotations were not available to the public and were kept hidden so that the participants' work could be judged in the end. This was because the dataset was for a competition. So, the test samples were not taken into account for this study. 3662 samples were used to classify the DR into multiple classes. This was done so that the severity of the DR could be graded in five stages. In Table 3.1, we can see the curated dataset along with the correct classes, numbers, and distribution. But because the images came from different places, they were all different sizes and shapes. Most data are in the Non-DR class, as we can see from the data demonstration.

Table 3.1**:** Dataset specifications

DR Stages	Number of images	Training Images (75%)	Testing Images (15%)
No DR	1805	1353	452
Mild	370	277	93
Moderate	999	750	249
Severe	295	221	74
PDR	193	145	48
Total	**3662**	**2746**	**916**

3.2DATA PREPROCESSING

The data from the source was not in a standard format. Inconsistencies like these could affect how well models work as a whole. If it is not dealt with, the possibility of poor convergence can become a serious problem during training. Resulting in failure to detect and classify the target features that we are looking for. Next shown Figure 3.3 depicts the sample dataset.

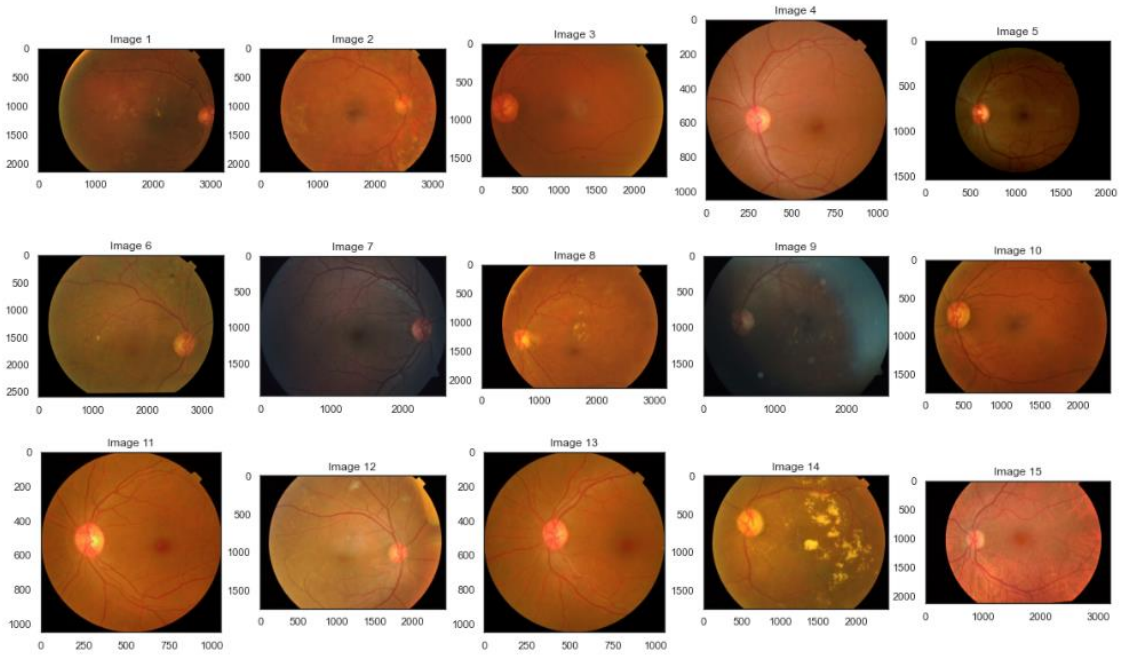


Figure 3.3: Sample Dataset

3.2.1 Data Resizing

As we can see from the sample dataset that was shown. There are multiple sizes of data. Which will make it hard for the model to learn and remember the way. So, all of the images' Height, Width, and Channels were resized to 224 × 224 × 3 to make them all the same size.

3.2.2 Auto Cropping to further improve

As there are multiple shapes of images; some are fully squared some are rectangular shapes. An auto cropping method should have to introduce to reduce the redundancy. Cropping the uninformative area should be done to recognize better by the models.

3.2.3 Color vision circle cropping

As we can see, images contain different kinds of lighting. Some are darkened, and some are overexposed. To reduce the complication, we could try to see the picture better by making it grayscale. Green Out of the three-color channels in the image (Red, Green, and Blue), the green channel shows the most contrast between the blood vessels, exudates, and hemorrhages. Unlike the red and blue channels, the green channel is neither too dark nor too bright. So, we took only the green channel out for analysis and sorting. Then, just the green channel is taken out, and contrast is increased to tell microaneurysms from veins. The collected retinal images have different amounts of light. This is resolved by removing the approximate background image, which would be made by median filtering and then adding the mean of the original image. It keeps the original image's amount of light. The following equation shows this,

$$I_{ie} = I - I_{bg} + u \qquad \text{…Equation (3.1)}$$

Where, I_{ie} is the illumination of the equalized image, I_{bg} represents the background estimated image, which is made by using a 51×51 median filter, and u is the mean average of the RGB channels.

3.2.4 Applying Gaussian Weights

In filtering phase of images, each pixel is affected by a different filter. This is a process that can be done in parallel. Regarding other neighboring pixels, the filtering process and how it is applied are both independent. An image is frequently composed of vast numbers of pixels, making it an excellent contender for parallelization. To smooth the circle-cropped image, apply a gaussian smoothening filter with size 3 and sigmaX 10. Therefore, preprocessing improves in adjusting the inconsistent lighting and enhancing the image's contrast.

3.3 DATA NORMALIZATION AND AUGMENTATION

Changing a picture's pixel intensity levels to fall within a predetermined range is the process of "normalizing" the image. The default range is frequently [0, 1] or [-1, 1]. By dividing the image by 255, we normalized the image. The pixel range will be changed to 0, 1. If it is not addressed, the possibility of insufficient convergence may become problematic throughout training. To give all images equal pixel values in the same range, this study rescaled all xi pixels in each image to a value of 1.0 / 255 using the min-max scaling equation.

$$xi = \frac{xi - \min(xi)}{\max(xi) - \min(xi)} \qquad \text{....Equation(3.2)}$$

Table 3.2 displays the values for each augmentation that increased the number of samples for this research while preserving key characteristics. The augmentation occurred automatically and just for the training data, preventing undesired predefined results or an inflated bias. This is highly relevant to notice.

Table 3.2**:** Data augmentation settings

Augmentations	**Value**
Horizontal flip	True
Vertical flip	True
Rotation	360
Shear	0.1
Zoom	0.1
Height shift	0.1
Width shift	0.1
Fill mode	Constant

3.4 PROPOSED FRAMEWORK

For identifying and classifying DR a method is proposed to perform the required task. The two major components of the proposed method are transfer learning and localization and categorization. Figure 3.4 depicts all functionality in its entirety. Two main quality dimensions to the work that is being shown. The first is "dataset preparation," and the second is a better DRDnet22 network that has been trained to classify eye diseases. For feature calculation, we used Custom DRDnet22 with DenseNet-121 as its base network. The image sample and location of the impacted region in the input image are accepted by the DenseNet-121 features extractor of the Custom framework.

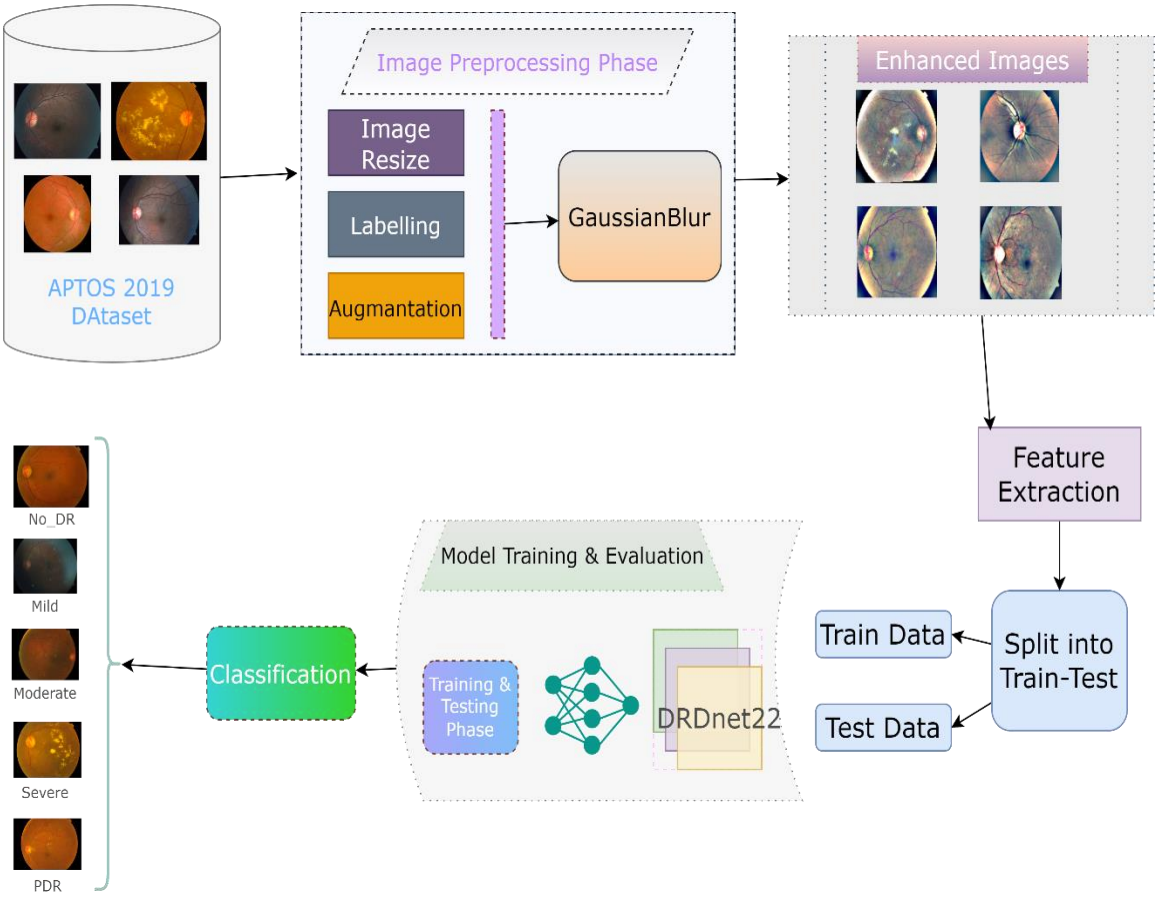


Figure 3.4: Flow diagram of proposed technique

3.4.1 Proposed Model Description

The database is the most important part of DenseNet-121. As shown in Figure 3, N × N × M0 shows the features maps (FMs) of the n-1 layer. We have tweaked the model, but we still think of it as the base model. M0 shows the total number of channels, while N shows the size of the FMs. A non-linear transformation H(.) is made up of several different techniques, such as a batch normalization layer (BN), a Rectified Linear Unit (ReLU) activation function, and a 1 × 1 convolution layer (ConvL), which are all utilized reduce the total number of channels. Again, ConvL was used to change the order of important points. Long dashed directions show the dense linkages, which are used to connect the n-1 layer to the n-layer and do concatenation using the output of the H-transform (.). In the

end, the n+1 layer makes N N (M0 + 2M).

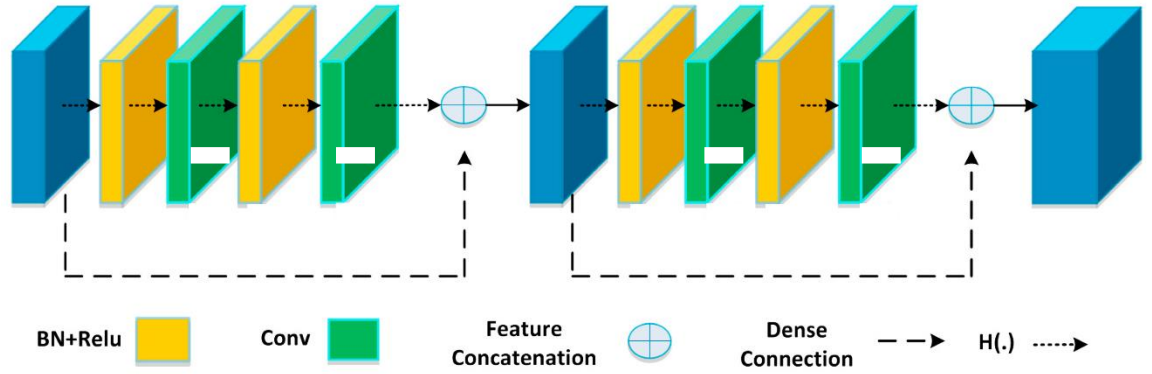


Figure 3.5: Densenet-121 Architecture

The feature representations have a significant impact on any machine learning model's performance, and this is also true for models used for DR recognition. Initially, several pooling strategies were conducted to combine the deep features obtained from several pre-trained models. Better explanatory and discriminative representation of the retinal images is provided by the final combined deep features. The classification models receive these combined features for DR identification or severity recognition. Figure 3.6 depicts how the feature extraction was done with the help of a convolutional neural network.

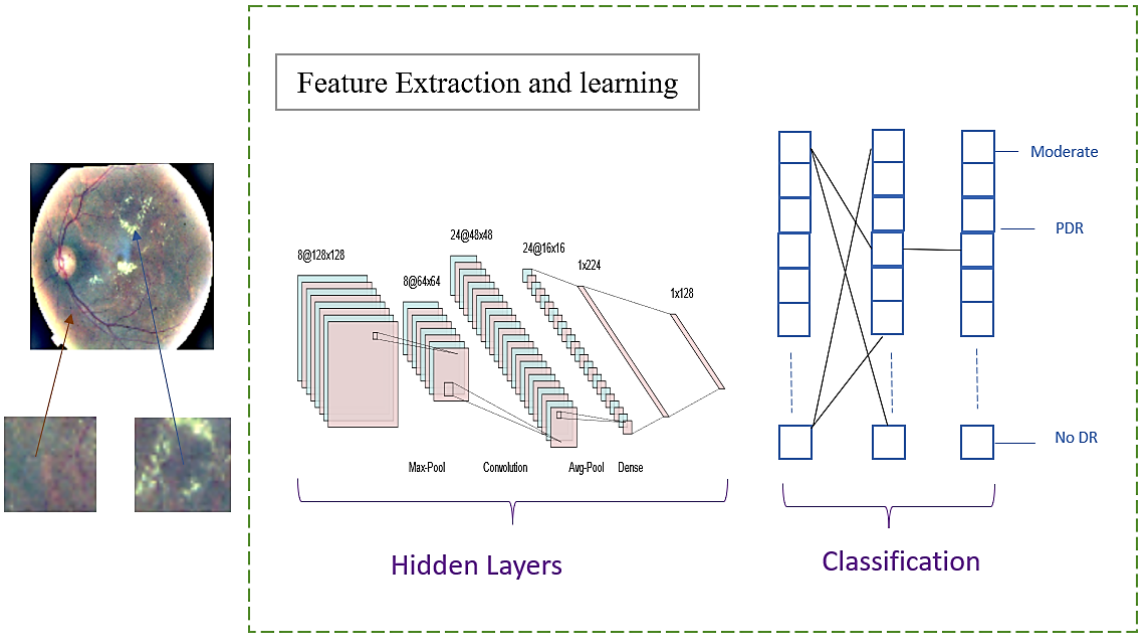


Figure 3.6: Feature extraction

3.5 SETTING THE HYPER-PARAMETERS

Before the actual training process, the hyperparameters of the DRDnet22 are set and the loss function is chosen. In Table 3.3, you can see the values of the hyper-parameters that were used during the training process. The batch size for this study is 20, which was based on the size of the data sources and the current specifications of the machine to avoid running out of memory during the training and testing phases. Adam, an optimizing compiler with a Learning Rate (LR) of 0.0005, was chosen for a training process that would be quick and effective. Adam was quick to learn and change, so even with a low LR value, the training time was cut by a lot. After only 50 epochs, there was strong convergence.

Table 3.3: Hyper-parameter configuration

Hyper-Parameter	**Value**
Batch Size	20
Epochs	50
Optimizer	Adam
LR	0.0005

3.6 LOSS FUNCTION

The success of a DL model depends not only on how accurate it is but also on how often it makes mistakes. For this study, the Categorical Crossentropy Loss (CCEloss) is in the following equation. was used because it minimized the difference between the classes that were expected and the classes that were found while measuring loss rates. It was also the best choice because, unlike its binary counterpart, it only gives the probability for the relevant classes when there are more than two classes. This made it a good selection for the work.

$$CCE_{loss} = -\sum_{c=1}^{m} Y_{o,c}\, log(P_{o,c}) \quad \text{… (Equation3.3)}$$

3.7 PRE-TRAINED MODEL DESCRIPTION

3.7.1 Visual Geometry Group – 16

A great 16-layer deep vision neural network model is the visual geometry group (VGG16). A 2D image is sent into the network as input (224, 224, 3). The first two layers have 64 channels and 33 filter sizes, and the padding on the first two levels is the same. After that, there are two convolutional layers with 256 filter sizes, then a stride (2, 2) max pool layer (3, 3). The next layer is a stride (2, 2) max-pooling layer, which is a duplicate of the layer above it. Figure 3.6 shows two convolution layers with three filter sizes and a total of 256 filters. Along with a max pool layer, there are two different kinds of three convolution layers. A two-layer convolution stack is then created using the transformed image. Instead of the 1111 filters used by AlexNet and 77% by ZF-Net, 33 filters are used in these convolution and max-pooling layers.

Figure 3.7: VGG16 working procedure

3.7.2 Inception V3

The Inception-v3 was made as a Googlenet module to help with image processing and finding things in images. It is the latest version of Google's Inception CNN. The goal of Inception-v3 was to improve the service and keep the number of parameters under control. It has below 25 million parameters, while AlexNet had 60 million. Figure 5 shows the process flowchart for how Inception-v3 works.

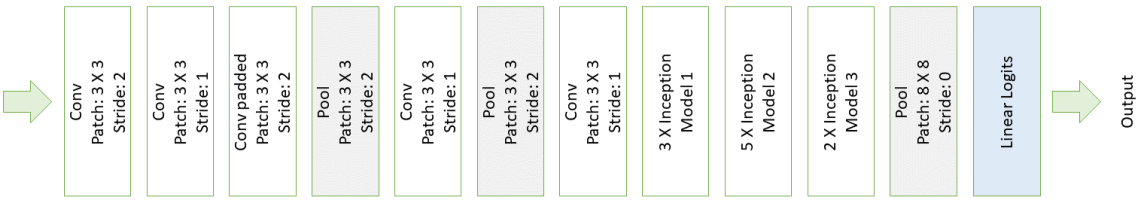


Figure 3.8: InceptionV3 working procedure

3.7.3 Residual Networks 50

A 50-layer CNN is ResNet50, which is a residual network. ResNet50 is a residual network with 48 layers, one for the max pool and one for the average pool. Figure 3.9 shows that every one of the five stages of the ResNet-50 framework has its conv and identity block. Every inversion block has three conv layers, and each identification block has three conv layers. In the ResNet-50, there are about 23,000,000 parameters that can be trained.

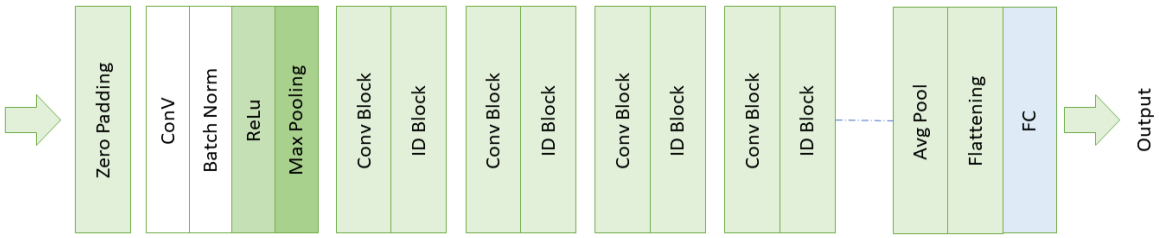


Figure 3.9: ResNet50 working procedure

3.7.4 EfficientNetB5

EfficientNetB5 nearly consists of 30 million parameters. The models from the efficients family performed better in ImageNet. Also, this is quite lightweight compared to other models.

Figure 3.10: EfficientNetB5 Work Procedure

3.7.5 Xception

Xception is a 71-layer deep convolutional neural network. The Xception model with transfer learning provides a more stable and dependable classifying model that can be applied to histopathological images with different levels of magnification (Chollet, F. 2017). There are 22.8 million parameters in Xception.

3.7.6 DenseNet-201

Transfer learning is a good way to get a good result from a small amount of data in a classification problem. And finally, hyper tuning can be used to improve the results of the model. In this model, InputLayer comes first, followed by a functional layer, a dropout, and then the Dense Layer.

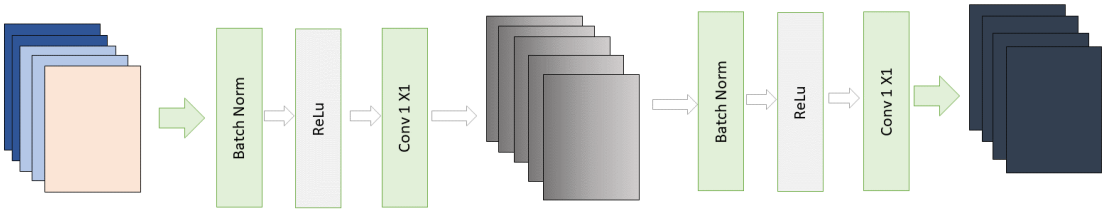


Figure 3.11: DenseNet Work Procedure

3.7.7 MobileNetV2

Mobile platforms are taken into account when making convolutional neural networks like MobileNetV2. It is based on a concave residual structure, and at the bottleneck, residual links connect the levels. In the secondary expansion layer, small and light depth-wise convolutions are used as an outlet of non-linearity to separator features. The MobileNetV2

architecture is made up of 19 bottleneck layers after the first fully convolutional layer with 32 filters.

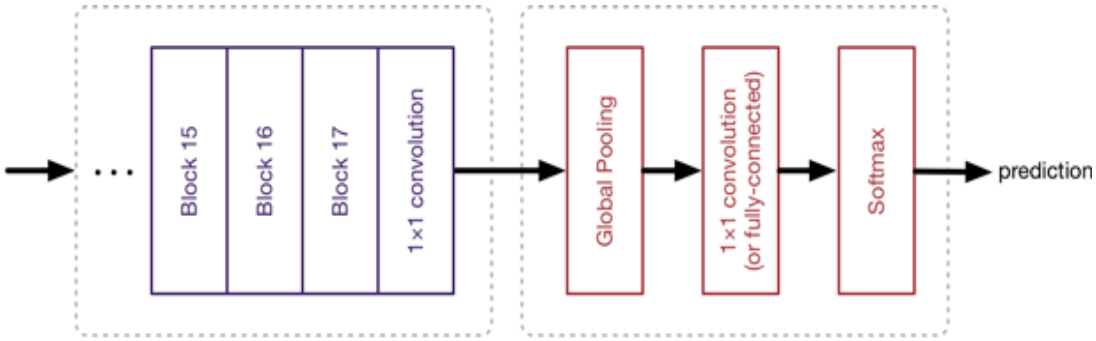


Figure 3.12: MobileNetV2 Work Process

3.7.8 NasNetMobile

The convolutional neural network used to train NASNet-Mobile is trained on more than a million images from the ImageNet database. The network can sort images into 1000 different types of objects. (Tsang, 2021)There are 769 layers in NASNetMobile as a whole. MobileNet models have fewer layers.

CHAPTER 4

RESULT AND DISCUSSION

4.1 INTRODUCTION

In this chapter, we will explore the proposed model performance evaluation in comparison with Keras's pre-trained models. For this purpose, the Model accuracy and loss graph, Confusion Matrix, Classification record of Precision, Recall, and F1 scores will be demonstrated.

4.2 EVALUATION METRICS

The traditional way to evaluate the performance level of vision-based models in medical imaging is to look at the True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) from the validation and test datasets. Then, the following equations are used to figure out the overall accuracy, precision, recall, and f1-score of the suggested model.

$$accuracy = \frac{TP + TN}{TP+TN+FP+FN} \quad \text{…(Equation 4.1)}$$

$$precision = \frac{TP}{TP+FP} \quad \text{…(Equation 4.2)}$$

$$recall = \frac{TP}{TP+TN} \quad \text{…(Equation 4.3)}$$

$$f_1\text{-}score = \frac{2 * precision * recall}{precision + recall} \quad \text{…(Equation 4.4)}$$

4.3 ANALYSING AND COMPARISON OF PERFORMANCE

After implementing the convolutional neural network models on our dataset, we have gotten the model's accuracy and the model's loss plotting graph. Training accuracy and validation accuracy are compared for gaining a graph of all the model's accuracy and the model's loss.

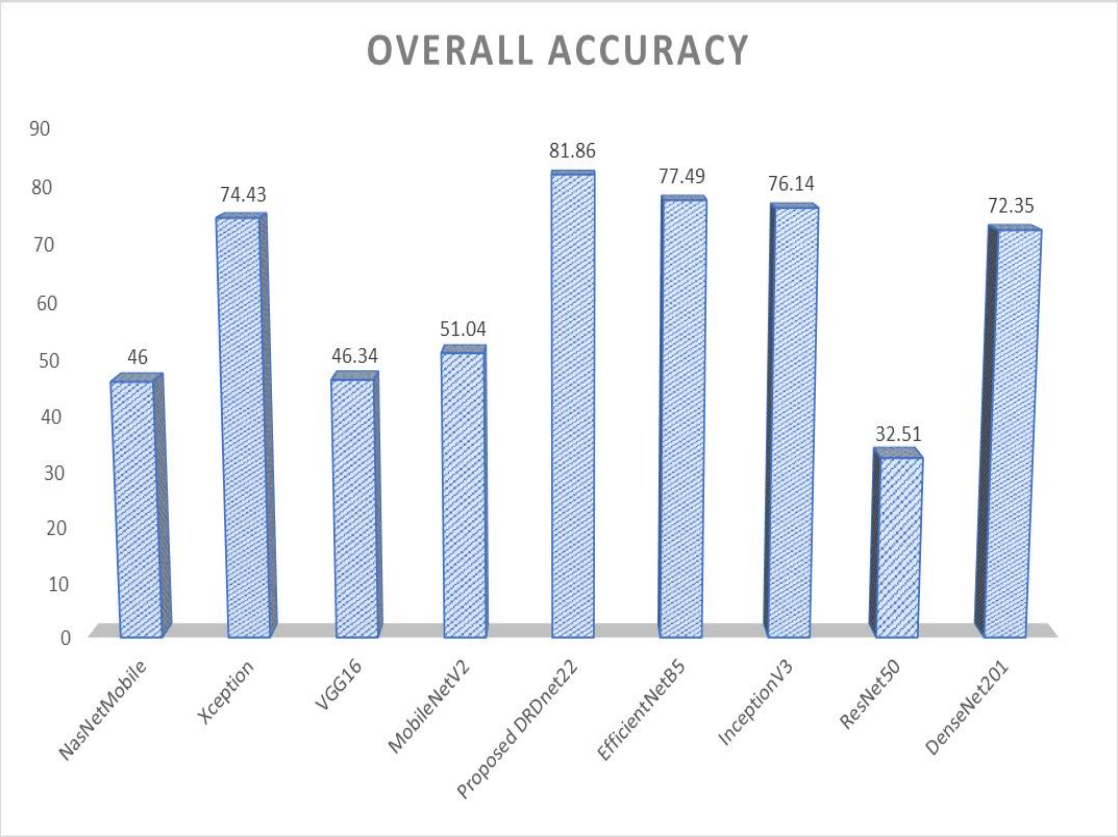


Figure 4.1: Overall accuracy comparison between proposed model and pretrained model

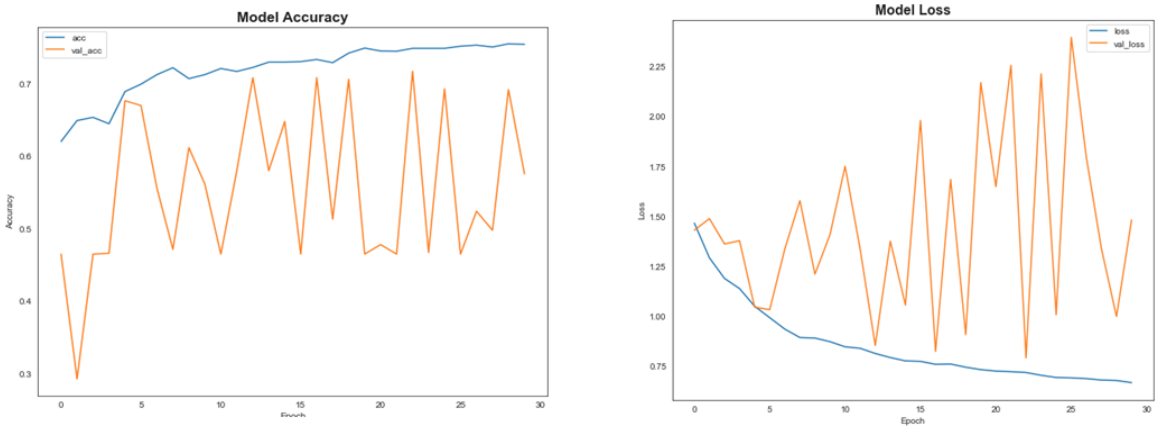


Figure 4.1: Model Accuracy and loss of DenseNet-201

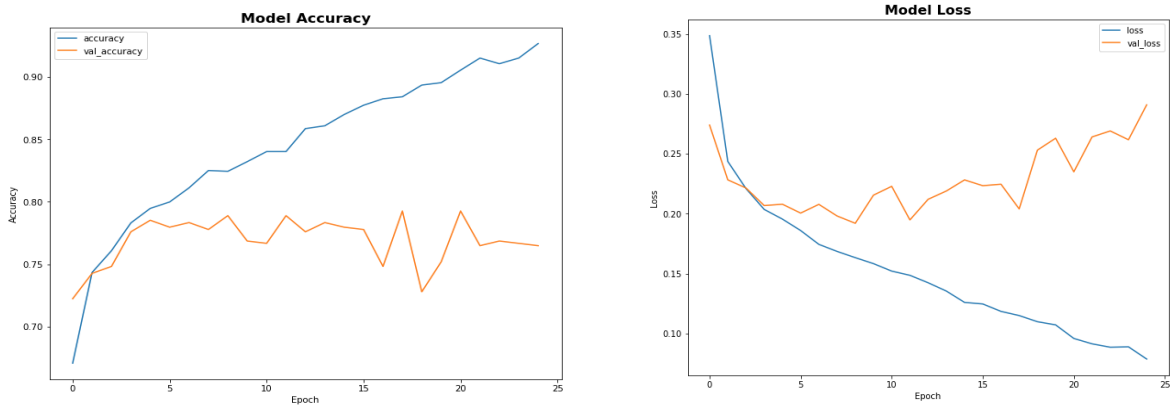


Figure 4.2: Model Accuracy and loss of InceptionV3

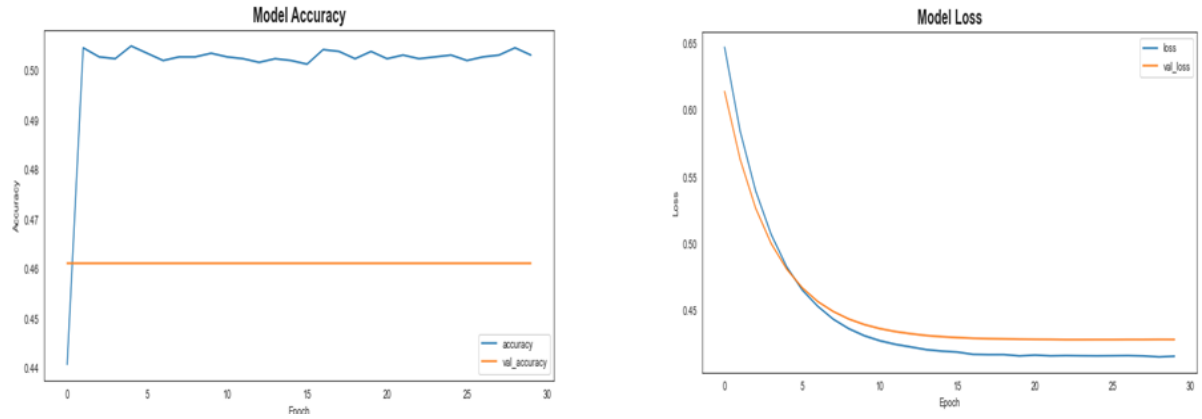


Figure 4.3: Model Accuracy and loss of VGG16

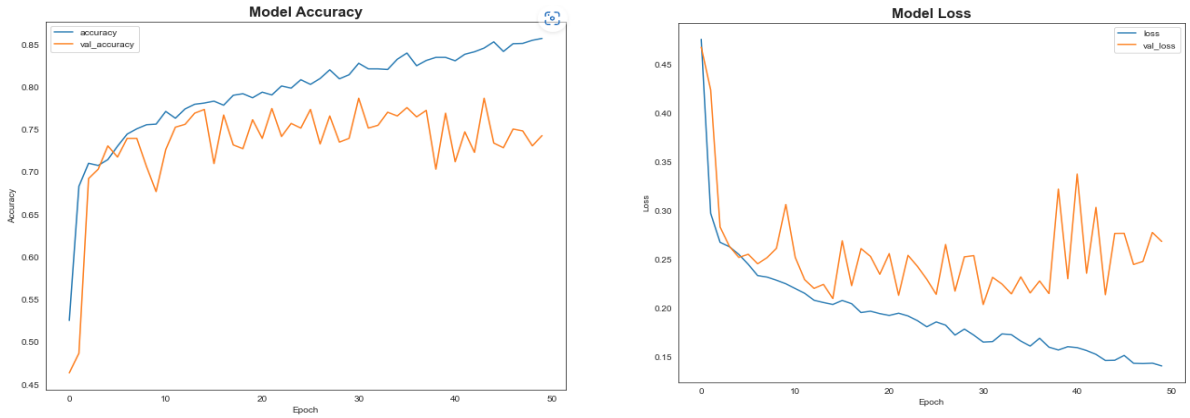


Figure 4.4: Model Accuracy and loss of EfficientNetB5

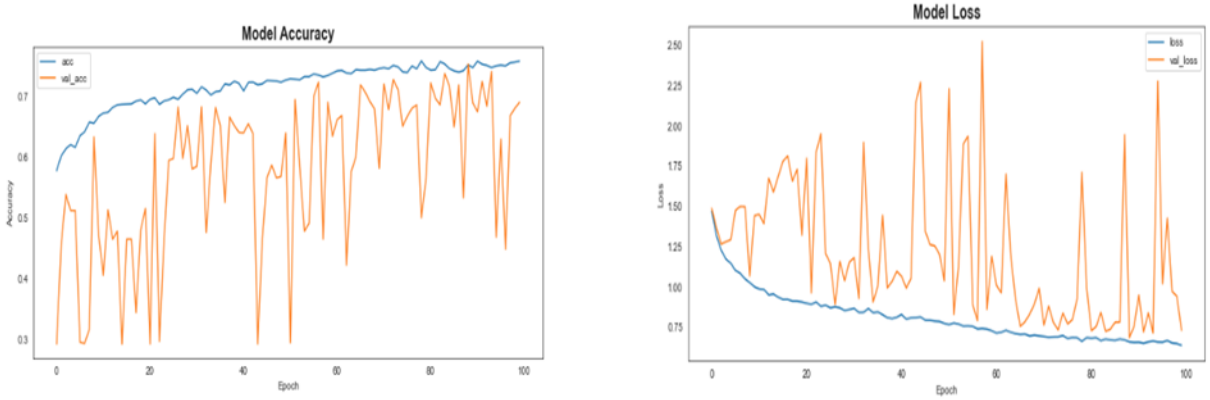


Figure 4.5: Model Accuracy and loss of Xception

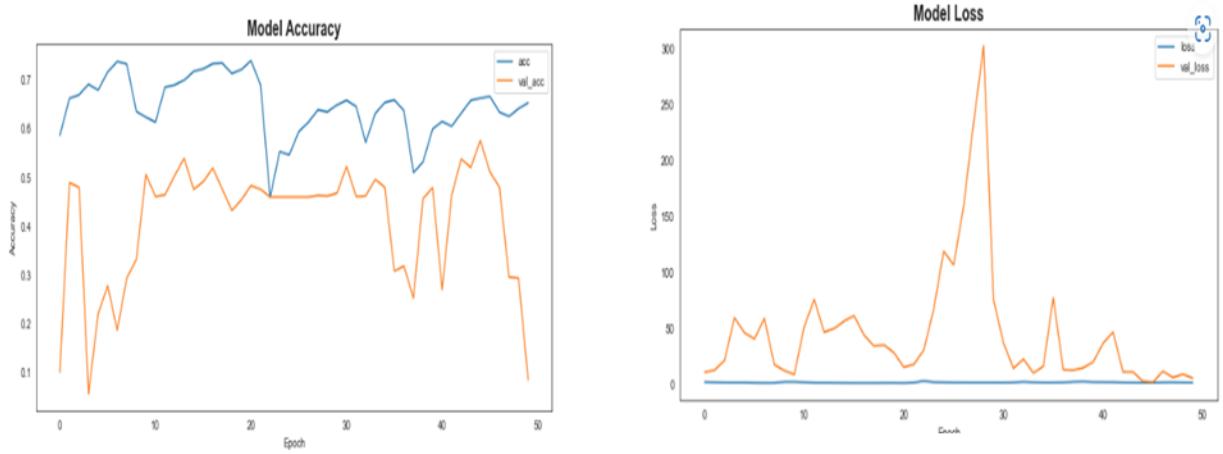


Figure 4.6: Model Accuracy and loss of MobileNetV2

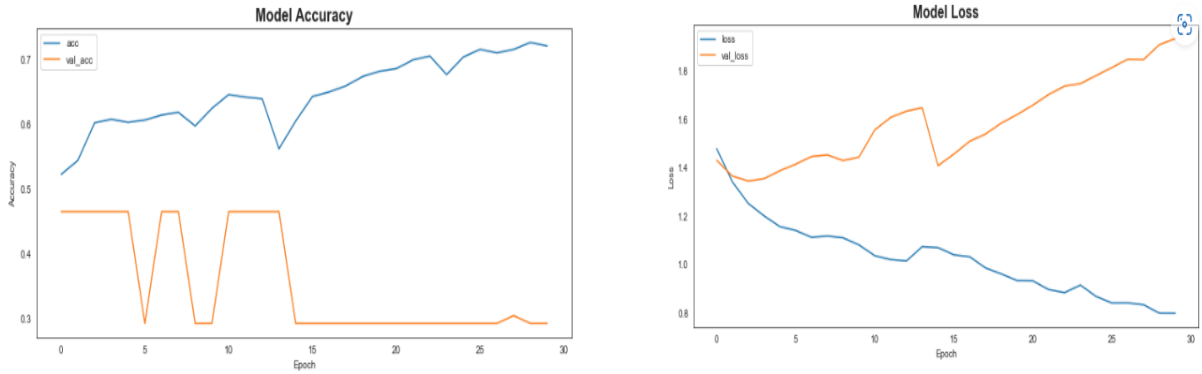


Figure 4.7: Model Accuracy and loss of NasnetMobile

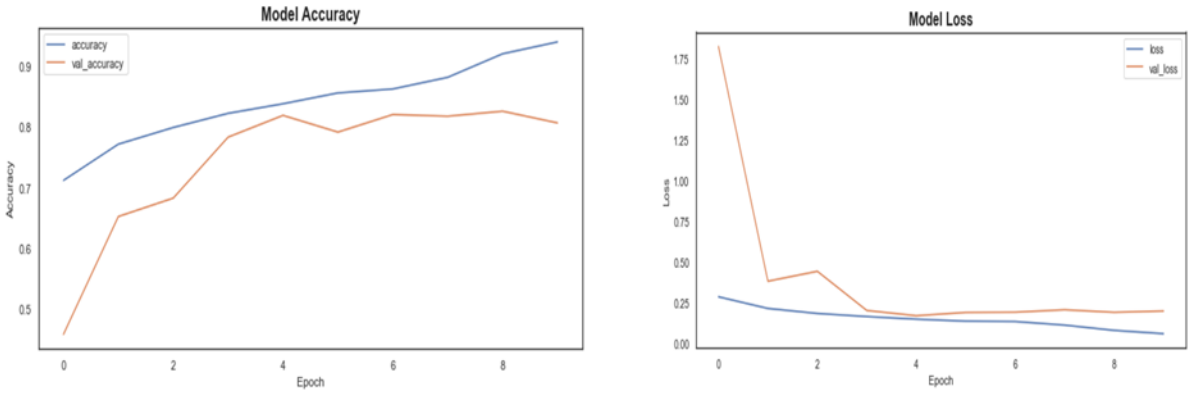


Figure 4.8: Model Accuracy and loss of Proposed DRDnet22

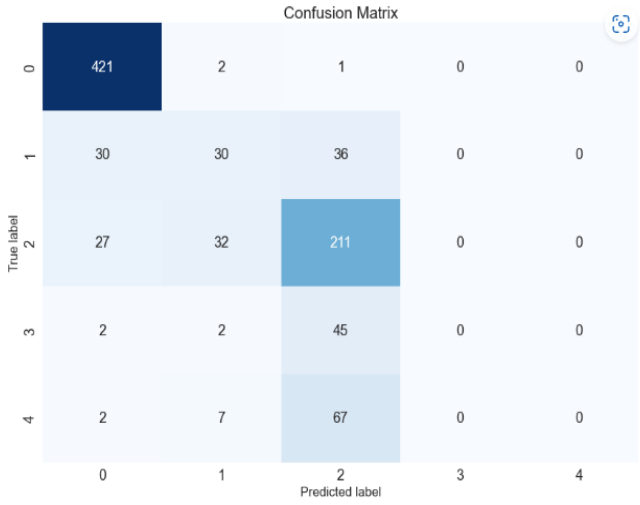


Figure 4.9: Confusion Matrix of DenseNet-201

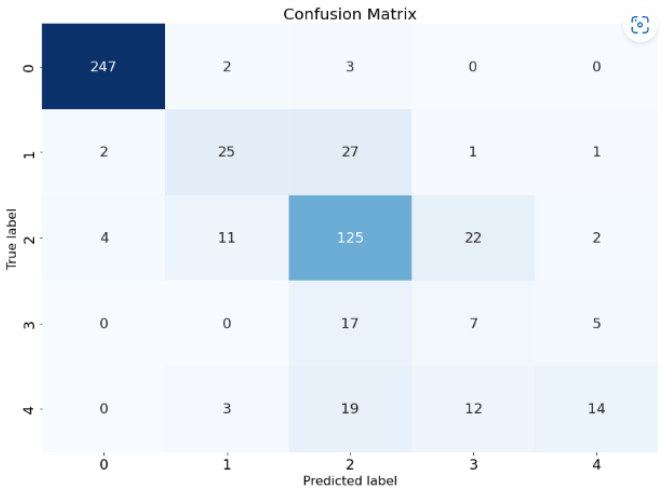


Figure 4.10: Confusion Matrix of InceptionV3

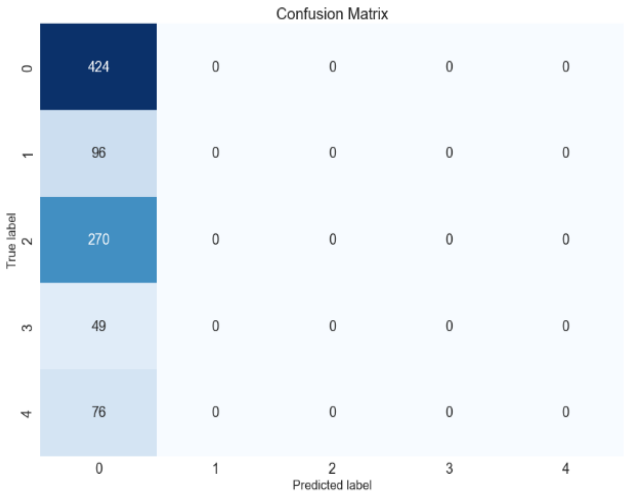


Figure 4.11: Confusion Matrix of VGG16

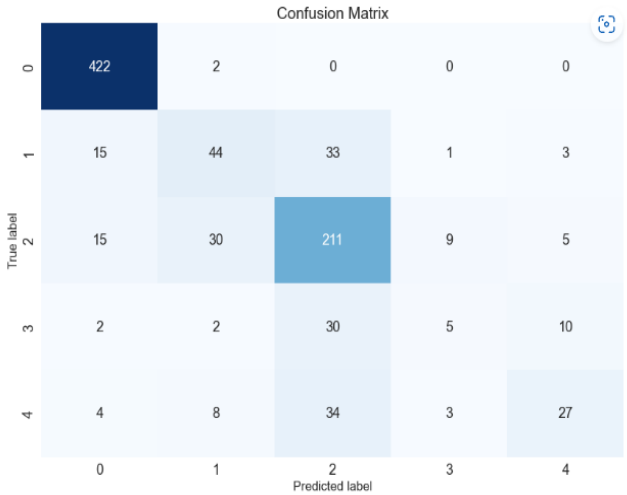


Figure 4.12: Confusion Matrix of EfficientNetB5

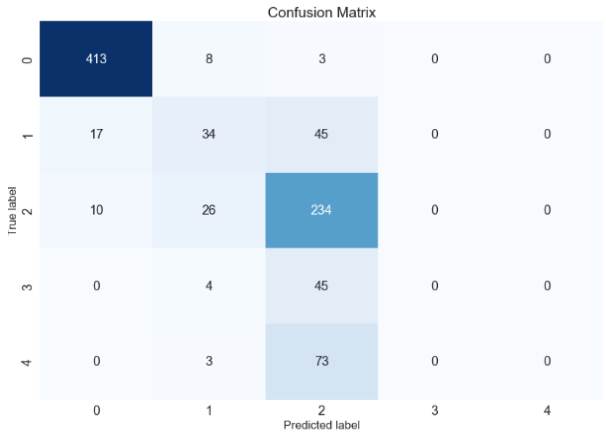


Figure 4.13: Confusion Matrix of Xception

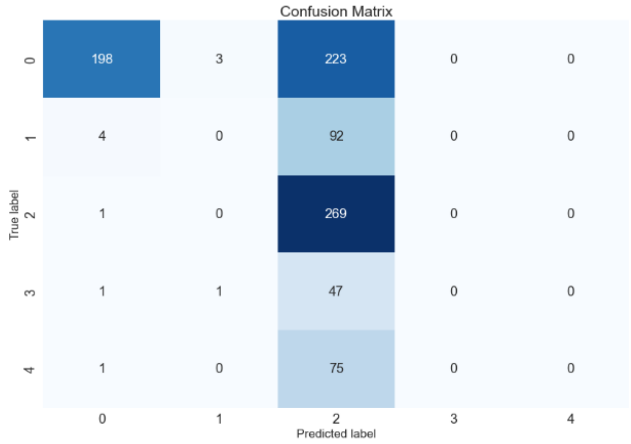


Figure 4.14: Confusion Matrix of MobileNetV2

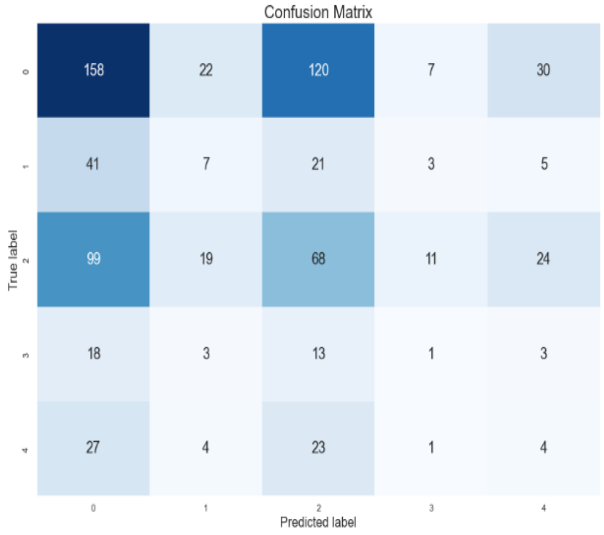


Figure 4.15: Confusion Matrix of ResNet50

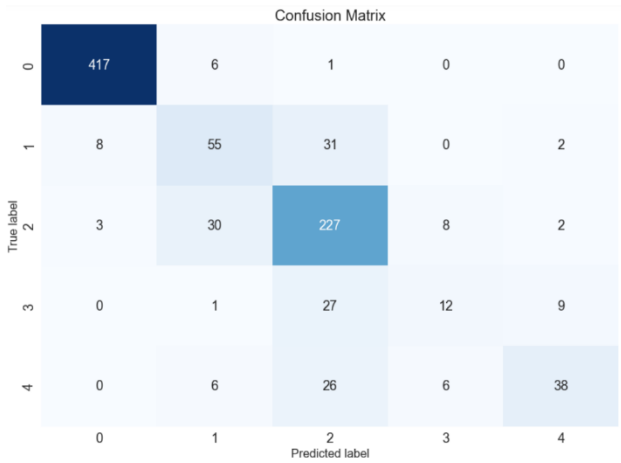


Figure 4.16: Confusion Matrix of Proposed DRDnet22

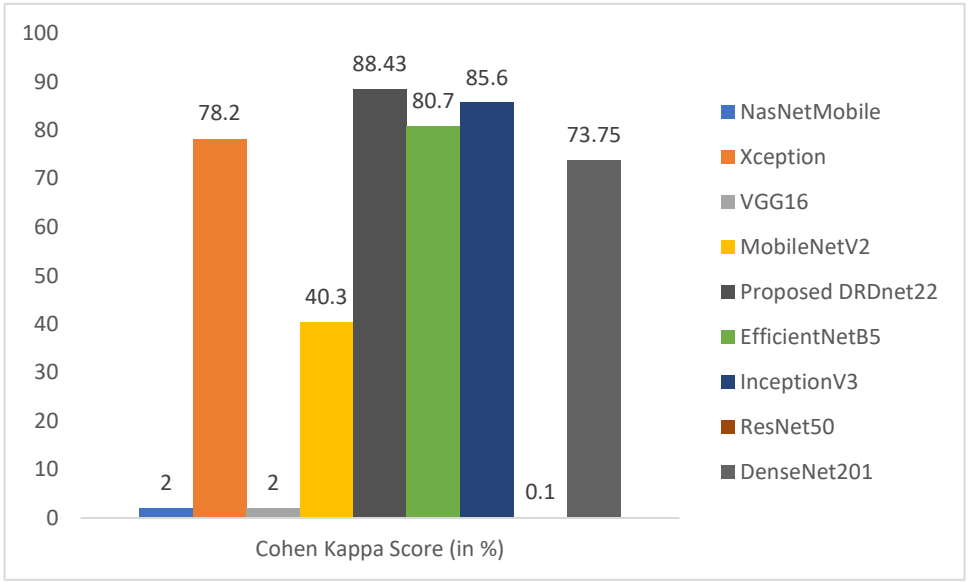


Figure 4.17: Comparison Cohen kappa score between all models

As observing the model loss and accuracy graph. We can denote that propose model is the stable among all with least amount of lost. Confusion metrics also denotes the same. From that comparison we can get a conclusive idea in terms of Cohen kappa score which is a significant way to measure model performance. Here, the proposed model is performing pretty well compared to regular pretrained model. Following there are Precision, Recall, and F1 score comparing among all of the models.

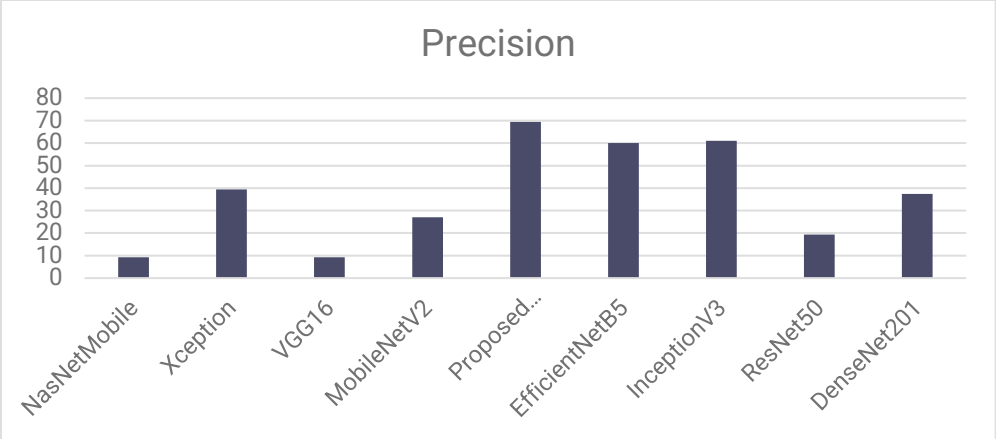


Figure 4.18: Comparison Precision score between all models

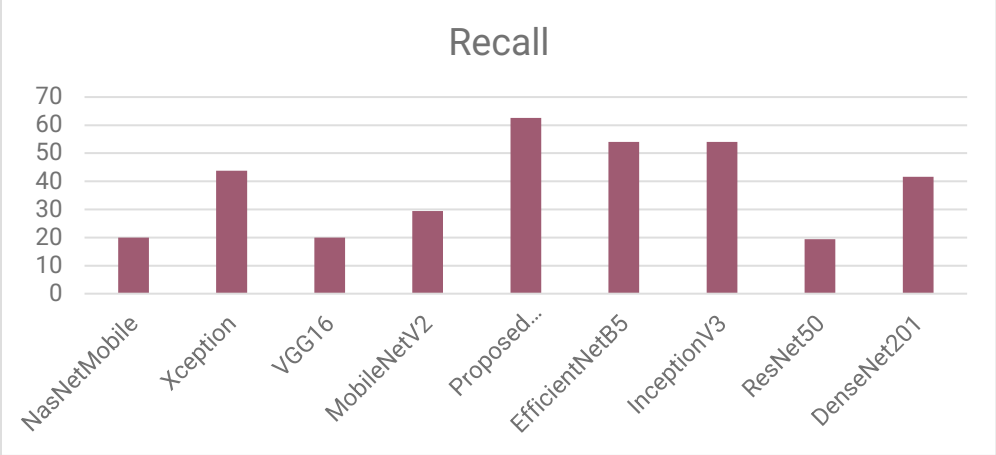


Figure 4.19: Comparison Recall score between all models

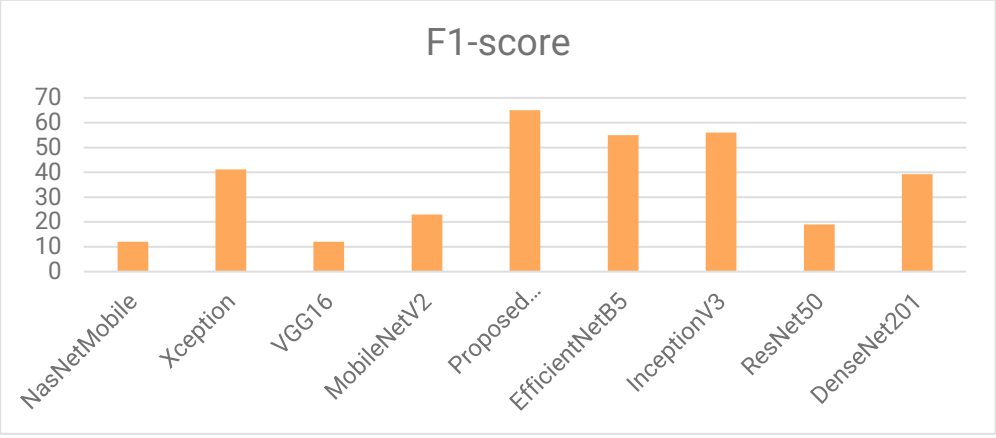


Figure 4.20: Comparison F1 score between all models

Chapter 5

CONCLUSION AND RECOMMENDATION

5.1 FINDINGS AND CONTRIBUTION

The purpose of this study is to find out that the proposed method classifies DR grades better than regular pre-trained model or not.

The dataset consisting 3662 labelled data were graded more accurately with an accuracy of 81.86% by the proposed DRDnet22. The main purpose was to classify as DR grading from input which can contribute to the traditional healthcare diagnostic system for betterment and ease of detection. Also, the proposed model is simple and lightweight as well which added a better advantage of time efficiency in detecting.

We evaluated the classification report to find Precision, Recall, F1 score. Also, we compared Cohen kappa score. Then we demonstrated the Confusion matrix. In this study, All the performance evaluation metrics identifies proposed model as the highest accuracy achiever.

5.2 LIMITATION

The study was based on a single dataset only. Thus, resulting in slight less improvement in test accuracy. As we know if the model tends to perform better if trained with lot of data.

5.3 RECOMMNENDATION OF FUTURE WORKS

This study focused only on detection of diabetic retinopathy and their classification. Deep learning can contribute more for the betterment of society if multiple disease detection can be done. We will work with more available and trustworthy dataset in identifying multiple diseases with the help of Deep learning approaches.

REFRENCES

Ahmad, J. (2022). Diabetic Retinopathy Trends in a Digital Screening Program of Eight and Half Years. Journal of Diabetes Research Reviews & Reports. SRC/JDRR-155. J Diabet Res Rev Rep.

Alyoubi, W. L., Abulkhair, M. F., & Shalash, W. M. (2021). Diabetic retinopathy fundus image classification and lesions localization system using deep learning. Sensors, 21(11), 3704.

APTOS 2019 Blindness Detection | Kaggle. (n.d.). https://www.kaggle.com/competitions/aptos2019-blindness-detection

Asiri, N., Hussain, M., Al Adel, F., & Alzaidi, N. (2019). Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey. Artificial intelligence in medicine, 99, 101701.

Bitto, A. K., & Mahmud, I. (2022). Multi categorical of common eye disease detect using convolutional neural network: a transfer learning approach. Bulletin of Electrical Engineering and Informatics, 11(4), 2378-2387.

Bodapati, J. D., Naralasetti, V., Shareef, S. N., Hakak, S., Bilal, M., Maddikunta, P. K. R., & Jo, O. (2020). Blended multi-modal deep convnet features for diabetic retinopathy severity prediction. Electronics, 9(6), 914.

Cavan, D., Makaroff, L., da Rocha Fernandes, J., Sylvanowicz, M., Ackland, P., Conlon, J., ... & Barratt, J. (2017). The diabetic retinopathy barometer study: global perspectives on access to and experiences of diabetic retinopathy screening and treatment. Diabetes research and clinical practice, 129, 16-24.

Cheloni, R., Gandolfi, S. A., Signorelli, C., & Odone, A. (2019). Global prevalence of diabetic retinopathy: protocol for a systematic review and meta-analysis. BMJ open, 9(3), e022188.

Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., ... & Gao, W. (2021). Pre-trained image processing transformer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 12299-12310).

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).

Das, K., & Behera, R. N. (2017). A survey on machine learning: concept, algorithms and applications. International Journal of Innovative Research in Computer and Communication Engineering, 5(2), 1301-1309.

Gangwar, A. K., & Ravi, V. (2021). Diabetic retinopathy detection using transfer learning and deep learning. In Evolution in Computational Intelligence (pp. 679-689). Springer, Singapore.

Hassan, B., Qin, S., Hassan, T., Akram, M. U., Ahmed, R., & Werghi, N. (2021). CDC-Net: Cascaded decoupled convolutional network for lesion-assisted detection and grading of retinopathy using optical coherence tomography (OCT) scans. Biomedical Signal Processing and Control, 70, 103030.

Hinton, G. (2018). Deep learning—a technology with the potential to transform health care. Jama, 320(11), 1101-1102.

Hossain, T., Jahan, N., Mazumder, M. S. A., Islam, R., Shamrat, F. J. M., & Khater, A. (2022, October). COVID-19 Detection Through Deep Learning Algorithms using

Chest X-ray Images. In 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1324-1330). IEEE.

Khan, Z., Khan, F. G., Khan, A., Rehman, Z. U., Shah, S., Qummar, S., ... & Pack, S. (2021). Diabetic retinopathy detection using VGG-NIN a deep learning architecture. IEEE Access, 9, 61408-61416.

Khatun, M. R. (2021, October 1). *Awareness of Diabetes and Diabetic Retinopathy among Diabetic Patients, Bangladesh*. https://www.jetir.org/view?paper=JETIR2110528

Lim, W. X., Chen, Z., & Ahmed, A. (2022). The adoption of deep learning interpretability techniques on diabetic retinopathy analysis: a review. Medical & Biological Engineering & Computing, 1-10.

Mazumder, M. S. A., Hossain, T., Shamrat, F. J. M., Jahan, N., Tasnim, Z., & Khater, A. (2022, October). Deep Learning Approaches for Diabetic Retinopathy Detection by Image Classification. In 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1504-1510). IEEE.

Nadeem, M. W., Goh, H. G., Hussain, M., Liew, S. Y., Andonovic, I., & Khan, M. A. (2022). Deep learning for diabetic retinopathy analysis: a review, research challenges, and future directions. Sensors, 22(18), 6780.

Nahiduzzaman, M., Islam, M. R., Islam, S. R., Goni, M. O. F., Anower, M. S., & Kwak, K. S. (2021). Hybrid CNN-SVD based prominent feature extraction and selection for grading diabetic retinopathy using extreme learning machine algorithm. IEEE Access, 9, 152261-152274.

Perdomo, O., Arevalo, J., & González, F. A. (2017, January). Convolutional network to detect exudates in eye fundus images of diabetic subjects. In 12th International Symposium on Medical Information Processing and Analysis (Vol. 10160, pp. 235-240). SPIE.

Prasad, K., Sajith, P. S., Neema, M., Madhu, L., & Priya, P. N. (2019, October). Multiple eye disease detection using Deep Neural Network. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) (pp. 2148-2153). IEEE.

Ramya, J., Rajakumar, M. P., & Maheswari, B. U. (2022). Deep CNN with Hybrid Binary Local Search and Particle Swarm Optimizer for Exudates Classification from Fundus Images. Journal of Digital Imaging, 35(1), 56-67.

Tsang, S. (2021, December 10). Review: NASNet — Neural Architecture Search Network (Image Classification). Medium. https://sh-tsang.medium.com/review-nasnet-neural-architecture-search-network-image-classification-23139ea0425d

Upadhyay, P. K., Batra, S., & Dhama, S. (2022). A Deep Neural Networks-Based Cost-Effective Framework for Diabetic Retinopathy Detection. In Data Science for Effective Healthcare Systems (pp. 199-211). Chapman and Hall/CRC.

Wong, T. Y., & Bressler, N. M. (2016). Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. Jama, 316(22), 2366-2367.

Zabeen, B., Khaled, M. Z., Husain, L., Aktar, A., Huda, K., Kamal, Y. A., ... & Azad, K. (2021). Risk factors associated with retinopathy in young people with type 1 diabetes in Bangladesh. Endocrinology, Diabetes & Metabolism, 4(2), e00197.