DIABETIC RETINOPATHY CLASSIFICATION USING TRANSFER LEARNING AND CNN

By

Sushmoy Barua

ID: 46332

Batch: 27th

Session: 2017-2018

Supervised By

Sarnali Basak

Associate Professor

Department of CSE

Jahangirnagar University University

University in partial fulfillment of the requirements for the degree of

BSc Hons in CSE

IN

COMPUTER SCIENCE AND ENGINEERING

Department of Computer Science and Engineering

Jagannath University

Dhaka, Bangladesh

May, 2023

Candidate's Declaration

This is to certify that the work presented in this project entitled, "Diabetic Retinopathy Classification Using Transfer Learning And CNN", is the outcome of the research carried out by Sushmoy Barua nder the supervision of Sarnali Basak, Associate Professor, Computer Science and Engineering Department, Jahangirnagar University, Dhaka, Bangladesh.

It is also declared that neither this project nor any part thereof has been submitted anywhere else for the award of any degree, diploma, or other qualifications.

Signature of the Candidate

Sushmoy Barua
ID: 46332
Batch: 27th
Session: 2017-2018
Department of CSE
Jahangirnagar University

Supervised by

Sarnali Basak
Associate Professor
Department of Computer Science and Engineering
Jahangirnagar University

Dedication

To my beloved parents

Contents

List of Figures

List of Tables

List of Algorithms

Acknowledgement

First of all, I would like to express my eternal gratitude to the Almighty God for the special blessing leading to the writing of this report.

I am extremely thankful and want to pay gratitude to my supervisor Sarnali Basak, Associate Professor, Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh for his effective guidance, encouragement, suggestion, cooperation, and valuable advice and special supervision in the preparation of this paper. His willingness to give his time so generously has been very much appreciated.

Before I get into thick things I would like to add a few heartfelt words for the people who are part of this project in numerous ways. People who give unending support right to stage the project idea are convinced.

I also acknowledge with a deep sense of reverence, my gratitude towards my parents, and other numbers of my family and friends, who have always supported me morally.

Last but not least, I would like to thank everyone I learned something from throughout my life.

Abstract

Diabetic retinopathy is a serious eye disorder caused by diabetes. It is brought on by damage to the blood vessels in the retina, the eye's backlight-sensitive tissue. Hyperglycemia can lead to retinal blood vessel damage and subsequent fluid leakage or bleeding in diabetics. This may cause the retina to enlarge and scar, which may ultimately result in blindness or vision loss. Diagnosing diabetic retinopathy has been accomplished through the use of machine learning methods such as decision trees, support vector machines, and convolutional neural networks (CNNs). In this investigation, we utilized the CNN methodology. Many convolutional, pooling, and fully connected layers comprise the CNN architecture, which is used to learn and extract information from input images. During the training phase, these layers' weights are adjusted to minimize the discrepancy between the expected output and the ground-truth label. Once trained, the CNN can be used to categorize new images of the retina into different kinds of diabetic retinopathy. In this study, we used transfer learning models to classify diabetic retinopathy. We used VGG-19, ResNet-50, Inception V3, and CNN models. Each model produced remarkable results, scoring 94.89%, 93.55%, and 97.69%. But with a test accuracy of 93.59%, VGG-19 did better than the others.

Chapter 1

Introduction

1.1 Diabetic Retinal Disease

Retinopathy of diabetics is a condition that can strike diabetics. The retina, the light-sensitive lining in the rear of the eye, is gradually failing. Diabetes reduces the body's ability to consume and store sugar, or glucose. An overabundance of glucose in the blood, which can damage many body parts, including the eyes, is the defining feature of this condition. Diabetes gradually damages the microscopic blood vessels in the retina and other small blood vessels all throughout the body. When these microscopic blood vessels leak blood and other substances, diabetic retinopathy occurs. The enlargement of the retinal tissue causes blurriness or worsening of vision. Due to the retinal tissue swelling, vision becomes hazy or worsens. The condition typically impacts both eyes. The likelihood of developing diabetes retinopathy increases as a person's diabetes worsens. Diabetic retinopathy can cause blindness if left untreated.
The following are signs of diabetic retinopathy:

- Observing spots
- Vision distortion
- Experiencing a dark or vacant area within your visual field.
- Nighttime vision impairment

When diabetics have high blood sugar levels for an extended period of time, the eye's lens may become clogged with fluid, it is in charge of concentration. This alters the lens's curvature, altering how things look. When blood sugar levels are under control, the lens assumes its former shape, which enhances vision. Individuals with diabetes whose levels of glucose can be properly managed will see a slower progression of retinopathy caused by diabetes. More than one-third of Americans, based on an identical study, are not aware that the only way to assess whether a person's diabetes may cause blindness is through an extensive eye examination. For this purpose, the American Osteopathic Association (AOA) advises that every one of the diabetics undergo a comprehensive dilated eye examination at least once per year. If retinopathy caused by diabetes fails to be detected and treated promptly, it can cause serious vision loss. The treatment for this disease differs depending on its severity. Individuals with retinopathy due to diabetes may need laser treatment to stop hemorrhaging and stop additional blood vessels from leaking. Drugs may need to be injected into your eye by your optometrist in order to treat inflammation or stop the growth of new blood vessels. If you have diabetes, you can help prevent or slow the development of diabetic retinal degeneration by carrying out the actions that follow:

- using prescribed drugs
- Maintaining proper diet
- Controlling high blood pressure with regular exercise
- Abstinence from alcoholic beverages and smoking.

What causes diabetic retinopathy?

Microaneurysms (MA), which appear as small red circular spots on the back of the eye, are the earliest sign of DR because of the vulnerability of the blood vessel's lining. It is smaller than 125 meters in size and has distinct limits. In accordance with the study by Michael et al. [21], there are six different forms of MA that can be identified in conventional fluorescence imaging and AOSLO reflectance.

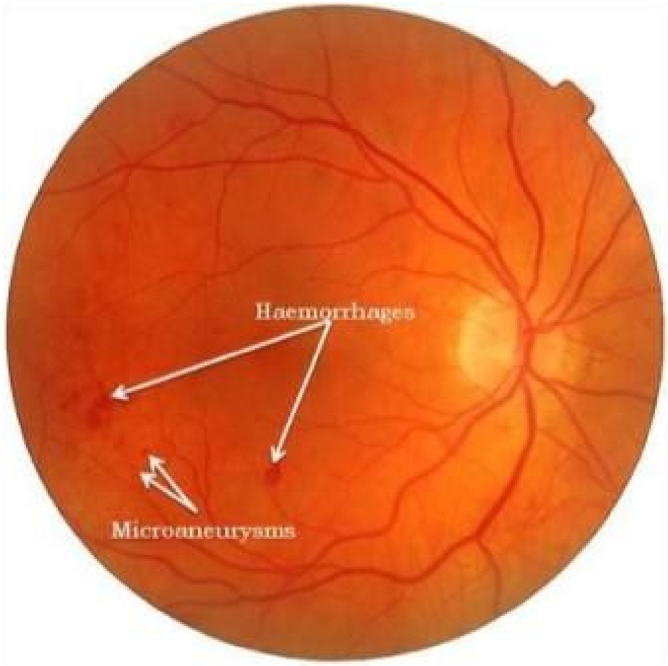

Figure 1.1: Microaneurysms

The lens of the eye displays an additional area when an injury is bigger than 125 m in length and has wavy edges. The two distinct types of HM are flame (superficial HM) and blot (deeper HM).

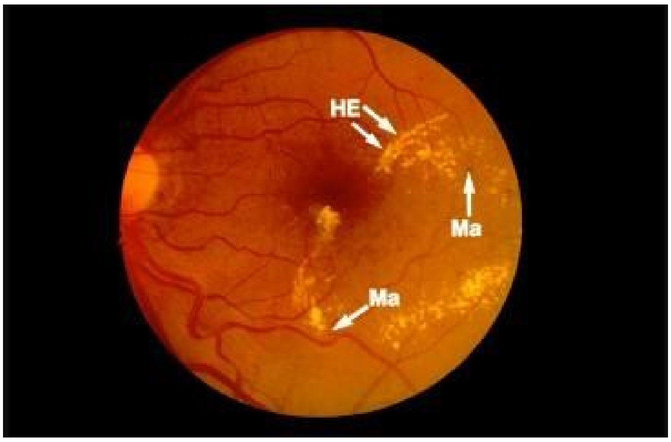

Figure 1.2: Hemorrhage bleeding in retinal tissue

The lens's harmed vessels for blood

- Fluid can leak into the macular, which is essential for clear center vision in the retina. Despite its small size, The region of the retina that allows humans to see hue and specific details is called the macular for short. The liquid makes the macular growth, causing blurred vision.
- New blood vessels may develop on the retina's layer in an effort to improve the flow of blood. These fragile, atypical blood vessels could rupture and leak blood into the retina, impairing vision.

Diabetic retinopathy comes in two different varieties:

1. In its very early stages, **non-proliferative diabetic retinopathy (NPDR)** has no or few signs. The retinal vessels that supply blood in NPDR are harmed. Small blood artery bulges called microaneurysms have the potential to let fluid into the retina. The leaking fluid might cause macular hyperplasia.
2. The disease's severe stage is known as **proliferative diabetes-related retinopathy (PDR)**. Currently, circulatory issues are depriving the lens of oxygen. As a consequence, the retina and vitreous, the gel-like fluid that fills the back with the eye, may produce new, vulnerable vessels for blood. Sight may become distorted if blood leaks into the retina from developing arteries. PDR can also result in glaucoma development and retinal detachment brought on by the growth of scar tissue. Glaucoma is an eye condition that causes gradual damage to the optic nerve. The fluid-draining region of the eye develops new blood vessels as a result of PDR. This causes an abrupt increase in eye pressure, which damages the optic nerve. It can result in significant vision loss and even eventual blindness if addressed.

A picture of a healthy and retinopathy-afflicted eye:

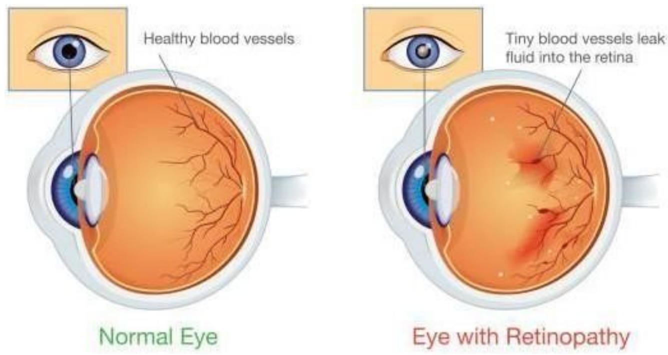

Figure 1.3: Damaged eye Vs healthy eye

What is the treatment for diabetic retinopathy?

The procedure of photocoagulation (laser therapy) may prevent the leakage of blood and other substances into the lens of the eye. In regions of the retina with aberrant blood vessels, a beam of light from a laser may be used to create microscopic burns in an effort to stop the blood from leaking.

Treatment for diabetic retinopathy is based on the illness's stage. A treatment's main goal is to slow or stop the progression of the illness.

Ongoing surveillance could be the only treatment for non-proliferative diabetes-related retinopathy in its early stages. You can limit the progression of the illness by adhering to your doctor's dietary and exercise recommendations and managing your glucose levels. Drugs can be injected directly into the eye to stop the growth of aberrant veins and potentially minimize the harmful effects of diabetic retinopathy. As the condition progresses, the aberrant blood vessels may leak blood and fluid into the retina, causing macular edema. To stop this spilling, laser therapy (photocoagulation) may be employed. To stem the leaks, a laser beam of light causes microscopic blisters in regions of the retina with aberrant arteries. Proliferative diabetic retinopathy, which results in extensive blood vessel production in the retina, can be treated with a pattern of dispersed laser therapy burns across the retina. The result is a shrinking and disappearance of aberrant blood vessels. Some side vision may be lost as a result of this technique to protect center vision.

1.2 Transfer learning

The ability to transfer knowledge between jobs is innate in humans. We apply the knowledge we gain while learning about one activity to solve related tasks in the same manner. We can transfer or use our knowledge across tasks more readily if they are closely related to one another. Simple illustrations include

1. Learning how to drive a car is equivalent to riding a motorcycle.
2. Learn to play jazz piano well if you play classical piano.
3. Study neural networks and possess mathematical and statistical skills

First off, it's critical to keep in mind that transferable learning, an extremely specialized kind of deep learning, is not a recent concept. A methodology based on transfer learning principles and the conventional way of creating and improving machine learning models is very unlike.

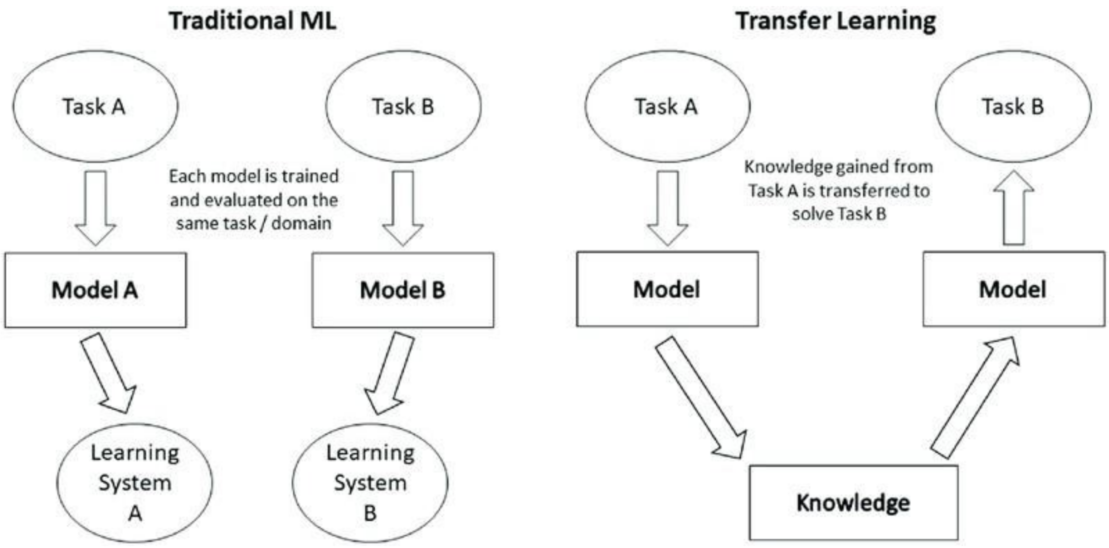

Figure 1.4: Comparing the two types of educational systems.

Traditional learning takes place in isolation and uses a small collection of tasks, datasets, and the training of unique, isolated models. Nothing that can be used with another model is kept up to date. With transfer learning, By drawing on previously learned information (features, weights, etc.) from older models, you may get around issues like a fresh assignment requiring fewer variables.

Some Uses for Transfer Learning
With predictive modeling, you can apply transfer learning to your own problems. The following are two typical methods:

1. Making a model is one possibility.
2. Another thing is to train a model.

Make a good approach: The challenge you choose must be tied in some manner to the concepts you learned when mapping input to output data, have a lot of data, and be relevant to predictive modeling.

1. **Construct the source model:**Making a successful model for this first project is the next stage. The model needs to outperform a naive model in order to show that some feature learning has occurred.
2. **Reusable design:** It is then possible to use the model fit on the source job as the basis for a relevant model creation for the next work. The model may be used in its whole or just a piece of it, depending on the modeling technique used.
3. **Change the design:** The framework could need to be modified or altered depending on the input-output pair data that are available for the relevant job.

Utilize Pre-trained Model Technique

1. **Deciding on the source model:** A source model that has already been trained is one of the frameworks that is available. Research organizations typically publish methods on difficult, huge datasets, and this might influence the choice of potential alternatives.
2. **Reusable prototype:** A model on the second job of interest can then be built using the model that was previously taught as a foundation. This can involve using all of the data or just a portion of it, depending on the modeling technique employed.
3. **Adjust the design:** The framework might need to be extended or adjusted based on the data input and output given for the pertinent task.

Extracting Features: To determine the most important components of the best description of your issue, another tactic is to use deep machine learning.This approach, which is commonly called "representation learning," typically produces performance that is noticeably superior than what would be achievable with a manually generated representation.

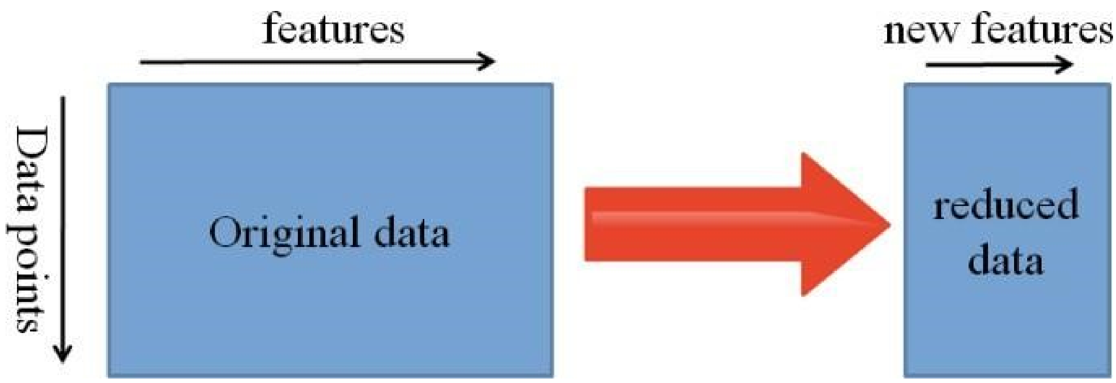

Figure 1.5: Feature Extraction.

1.3 Background and Motivation:

Early disease detection improves the efficacy of treatment in the medical field. Due to a shortage of insulin, diabetes is a disorder where blood glucose levels rise. It affects 425 million adults globally. Diabetes has an effect on the kidneys, heart, nerves, and heart. Globally, diabetes-related retinopathy causes 2.6% of blindness. In Bangladesh, there were 7.1 million instances of diabetes in 2015, with a prevalence of 7.4%, and 1.54 million people are at risk of developing diabetes-related blindness, based on studies from the international diabetes association (IDF). A multilayer neural network learning algorithm called deep learning (DL) has recently gained popularity. Artificial intelligence and human-computer interactions have significantly advanced as a result of DL's fresh perspective on machine learning [22]. Image recognition is now a useful tool because of deep learning's speedy advancements; it can avoid errors brought on by varying medical standards and aid in faster diagnosis. As a result, deep neural networks and algorithms for the recognition of images have been applied in multiple recent studies [17] to identify DR. The objective of this study is to build an automatic diagnosis of DR using the classification of images of the fundus.Our objective is to attain real-time end-to-end categorization of the patient's state from the fundus image by categorizing the images according to the degree of DR. It releases strain from the doctors' manual operation and experience to rely more on the automated and highly accurate diagnosis and treatment of DR. For this project, a range of picture preprocessing techniques are being used to extract several important attributes, which we will then categorize into the relevant categories. In order to find the DR in five data sets, we use CNN's architecture. The purpose of this study is to enhance ResNet-50's performance in identifying DR by applying preprocessing techniques and a revised structure.

Chapter 2

Literature Review

2.1 Identifying the stages of diabetic retinopathy by means of higher order spectra

DR was classified using component extraction-based categorization and DL [1]. To remove DR lesions such as blood vessels, exudates, and microaneurysms, multiple techniques were devised [2]. After exudates have been eliminated, DR [3] [4] [5] [6] [7] is graded. The size and number of microaneurysms were employed as features in a support vector machine to divide the DIARETDB1 dataset into positive and negative classes [8].

2.2 Reevaluating the original machine vision architecture

Furthermore, it has been noted that DL-based structures, such as CNNs, outperform feature extraction-based approaches [9]. The two primary methods of DL training for DR classification are learning entirely from the start and transferable learning. Additionally, it has been observed that DL-based systems, like CNNs, perform better than feature extraction-based techniques. Learning from scratch and transfer learning are the two main types of DL training for DR classification.

2.3 The creation and evaluation of a deep learning method to identify diabetic retinopathy in images of the retinal fundus

In order to train a convolutional neural network (CNN), 128,175 images of the fundus were used and the information set was then divided into two groups using it: photos in the first group had a severity rating of 0 or 1, pictures with levels 2, 3, and 4 were found in the second group. The sensitivity rating for [10] was 97.5% and 93.4% of the 9963 images in the EyePACS-1 dataset were unique, Using Messidor-2 and EyePACS-1, a specific condition can be detected. On the Messidor-2 dataset, the model attained a sensitivity of 96.1% and a specificity of 93.9%. The model, when tested with a cut point chosen for high specificity, produced results for the EyePACS-1 dataset with a sensitivity of 90.3% and a specificity of 98.1%.

2.4 Diabetic retinopathy and convolutional neural networks

Convolutional neural networks (CNNs) were used to build the initial model, which was trained on a dataset including more than 70,000 photos. Using the stochastic gradient descent approach [11], this model was able to classify DR into five groups with a specificity of 95%, accuracy of 75%, and sensitivity of 30%. The second model has a sensitivity of 96.8% and a specificity of 87% for DR auto-identification when it was trained from scratch using the MESSIDOR-2 dataset. Overall, it appears that both models are effective at detecting DR, with the second model outperforming the first model in terms of sensitivity and specificity.

2.5 Deep learning-based automated detection of retinopathy caused by diabetes

The first approach [12], which used a CNN trained entirely from scratch on the Kaggle dataset, had a sensitivity and specificity of 96.2% and 66.6%, respectively.Using a dataset of 71,896 fundus photos, the second strategy trained a CNN DR classifier and achieved sensitivity and specificity of 90.5% and 91.6%, respectively. Using a dataset of 75,137 fundus photos, the third model created a DL model with sensitivity and specificity scores of 94% and 98%. Overall, these models seem to perform well in recognizing DR or categorizing fundus photos for referable and non-referable classes. The third strategy obtained the highest sensitivity and specificity values.

2.6 Pre-trained convolutional neural networks are used for adjusting in the research comparing (DR) screening methods

The pre-trained Inception-V3 [13] and Exception models were improved to divide the Kaggle dataset into two categories. Saving the time and materials needed for deep learning was the aim. After balancing the dataset using data augmentation, the study obtained accuracy ratings of 87.12% on the Inception-V3 model and 74.49% on the Exception model. The study's findings suggest that improving pre-trained models can be a viable strategy for classifying fundus images and may result in time and resource savings over training a model from the start.

2.7 Deep convolutional neural networks for classifying images to identify retinopathy caused by diabetes

Wan et al. [14] applied transfer learning and hyperparameter tuning to multiple pre-trained models using the Kaggle dataset and compared the results. With hyperparameter tweaking during training, the VggNet-s model had the greatest accuracy rating (95.68%). The issue of insufficient training data for retinal vascular segmentation was also solved using transfer learning [15]. When examining 25,326 retinal pictures of people with diabetes from Thailand, an Inception-V4 model-based DR categorization outperformed human expert graders [16]. These studies show that several deep neural network techniques can effectively classify fundus images or solve segmentation issues, reaching excellent scores and, in some circumstances, outperforming expert graders.

2.8 Deep Neural Networks for the Detection of diabetic retinal degeneration

Based on abnormalities and required treatments [17], the 4476-piece training dataset was divided into four groups. The images were then divided into four 600 x 600 images and scaled down to 300 x 300 images before being fed into a number of Inception-V3 models that had already been trained, known as the Inception@4. The study discovered that the InceptionV4 model's accuracy results were superior to those of the VggNet and ResNet models. Then, a web-based DR classification system was created using the InceptionV4 model. Overall, the study shows how well pre-trained models like Inception-V3 perform when used to categorize images of the fundus in a system that is accessible via the internet. .tex

Chapter 3

Methodology

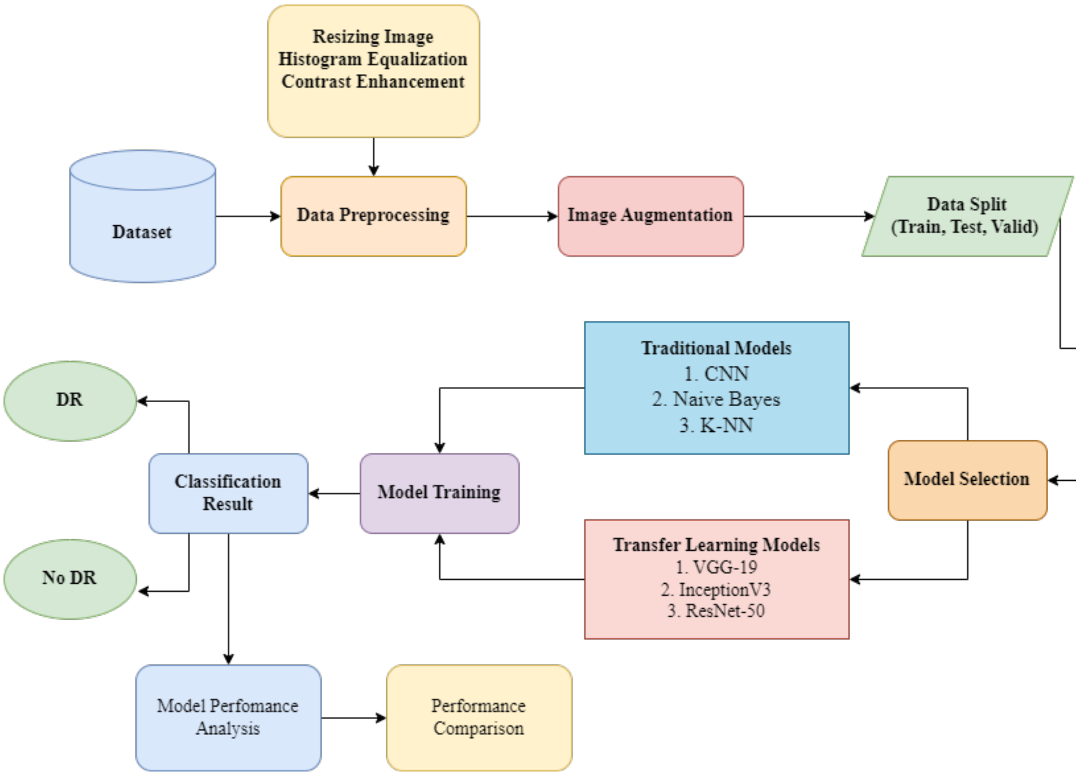

Figure 3.1: Project Workflow

3.1 Dataset Preparation and Collection

Almost 3000 fundus photos from the Kaggle dataset are rated in binary form for detecting DR and No DR. These were divided into 2600 photos for model training and 550 images for model testing. The photographs are available on the provided website.

3.2 Data Preprocessing

According to our research, hemorrhages, hard exudates, and cotton wool patches are all very easy to spot. However, yet, neither aneurysms nor aberrant blood vessel expansion were found in our data. The last two circumstances might be significant if we want to utilize our model to match human performance.

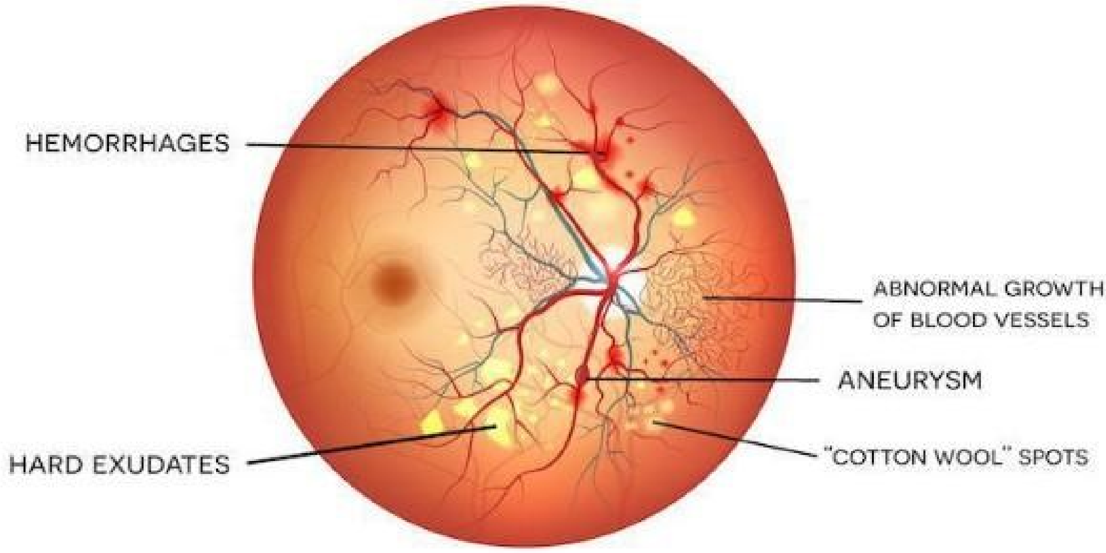

Figure 3.2: DR Detect features

In the preprocessing stage, preprocessing techniques are employed to solve the issue of variable lighting conditions in photographs. Accurately determining severity levels is challenging due to some of the dataset's photos being too dark or having an unbalanced color palette. Furthermore, some photos include blank spaces that lack any pertinent information. In order to prevent losing features in the useful sections of the image when it is magnified, it is advised to crop these areas.

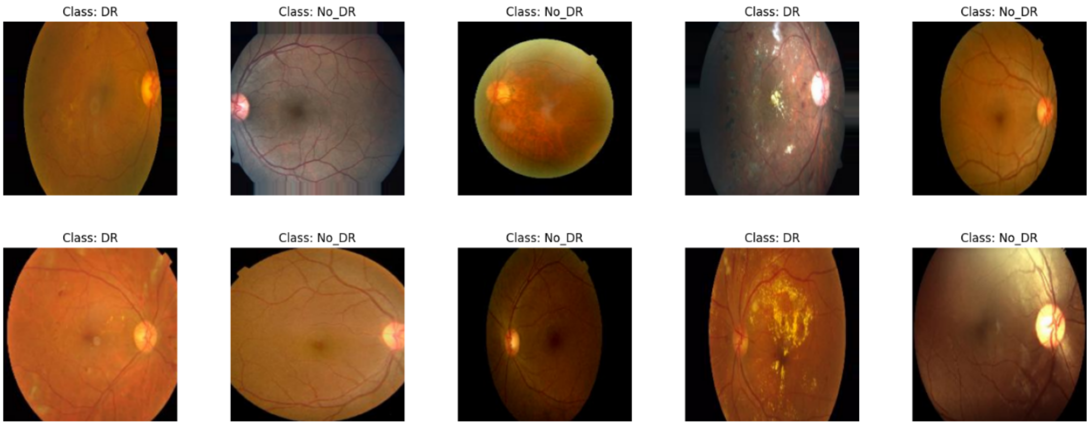

Figure 3.3: Original source pictures

To remove this color bleed, we transform this original image from BGR to RGB. Therefore, we can crop the images to remove any extraneous space. After being cropped, the images are converted to grayscale for help detect their stage and resize them.

1. **Histogram Equalization** Histogram equalization enhances the contrast of retinal images, crucial for improving diabetic retinopathy classification accuracy. By redistributing pixel intensity values uniformly, this technique makes features like microaneurysms and hemorrhages more visible. The process involves computing the image histogram, calculating and normalizing the cumulative distribution function (CDF), and mapping the original intensity values to new levels based on the normalized CDF. Enhanced images allow for better feature extraction, making critical indicators of diabetic retinopathy more pronounced. This preprocessing step improves the performance of deep learning models by providing clearer images for training, thereby aiding in early and accurate detection of the disease.

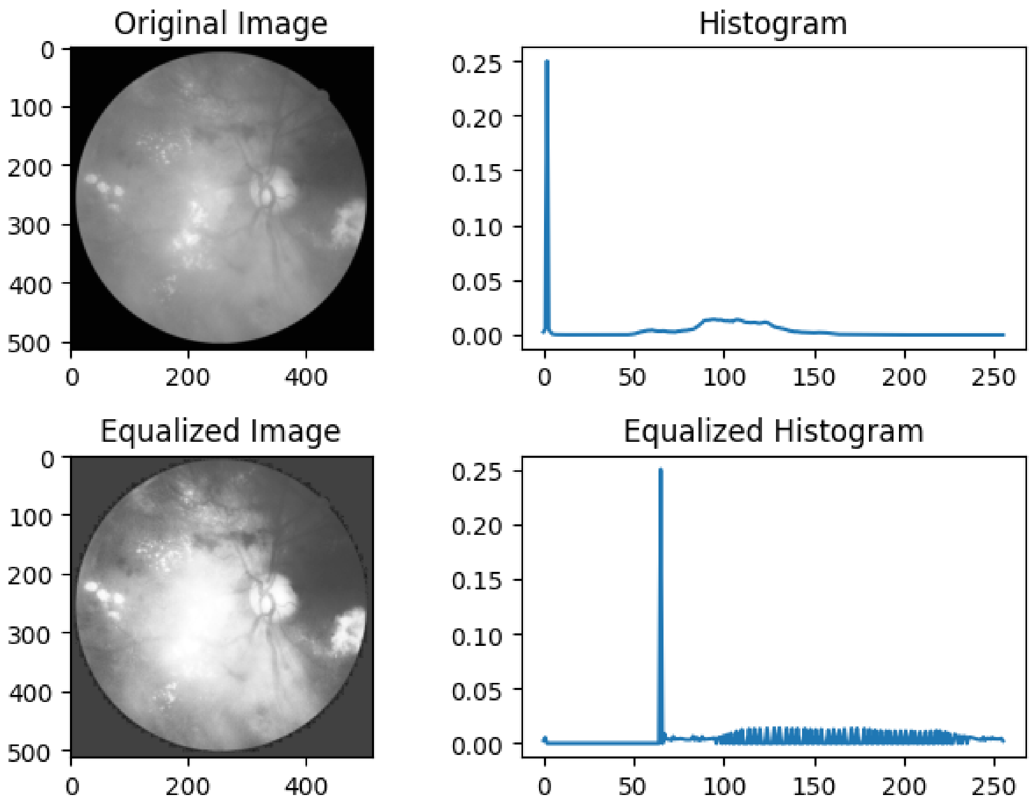

Figure 3.4: Histogram Equalization

2. **Clahe Contrast Enhancement** Contrast Limited Adaptive Histogram Equalization (CLAHE) is a technique used to enhance the contrast of retinal images, significantly aiding in the classification of diabetic retinopathy. In contrast to conventional histogram equalization, CLAHE performs histogram equalization to each tile separately while working on discrete areas of the image known as tiles. The key benefit of CLAHE is that it limits the contrast amplification to avoid noise enhancement, which is particularly useful for medical images. The process begins by dividing the image into non-overlapping tiles. For each tile, the histogram is computed and equalized, but with a limitation on contrast amplification set by a predefined clip limit.After that, borders between the different tiles are removed using bilinear interpolation, creating a consistently enhanced image. This technique makes subtle features like microaneurysms, hemorrhages, and exudates more prominent without amplifying noise. In the context of diabetic retinopathy classification, CLAHE enhances the visibility of critical features, improving the performance of feature extraction algorithms. This preprocessing step is crucial for training deep learning models, as it provides clearer and more detailed images, leading to more accurate and early detection of diabetic retinopathy.

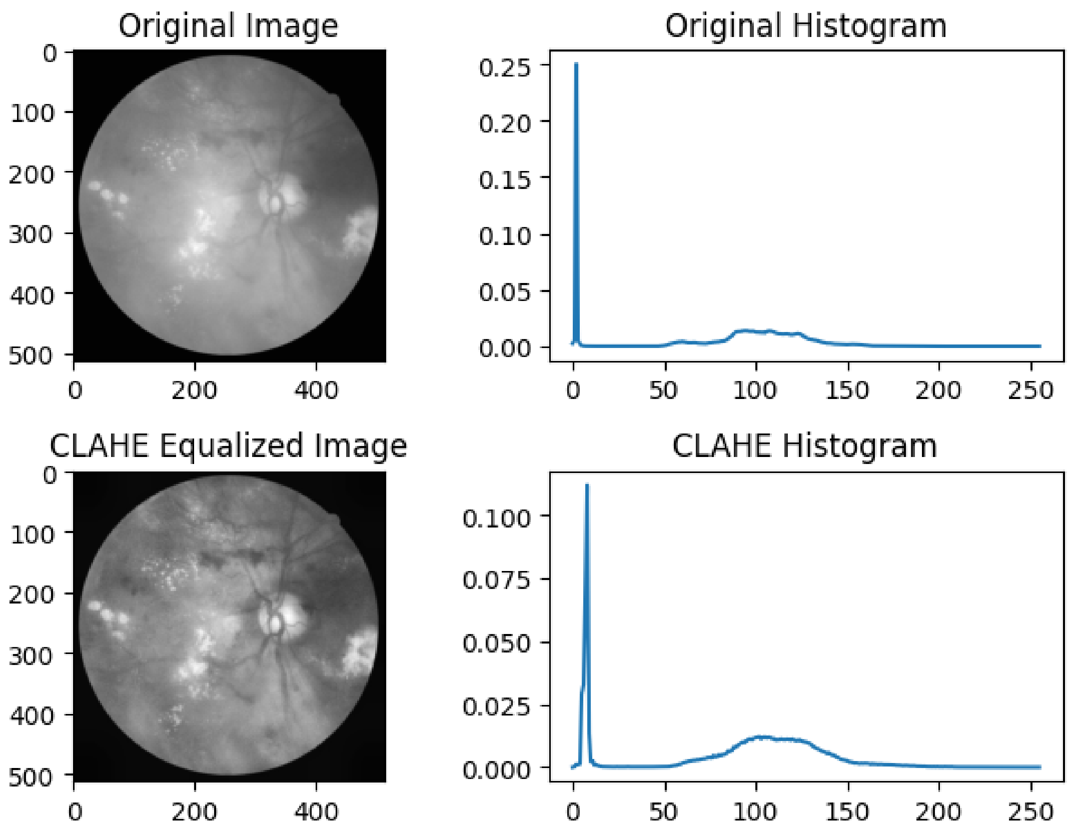

Figure 3.5: Clahe Contrast Enhancement

3. **Data Augmentation:** By artificially expanding the training dataset, data augmentation is a critical strategy used to improve the performance of diabetic retinopathy classification algorithms. With this technique, overfitting is decreased and deep learning models' capacity to generalize is enhanced by generating altered versions of preexisting images using a variety of modifications. Common data augmentation techniques include rotations, flips, translations, zooms, and intensity adjustments. For retinal images, these transformations simulate the variability and diversity seen in real-world clinical settings. For instance, rotating images can help the model recognize features from different angles, while intensity adjustments can mimic the variations in lighting conditions during image acquisition. By increasing the diversity of the training dataset, data augmentation helps models become more robust and capable of accurately identifying features such as microaneurysms, hemorrhages and exudates across different conditions. This enhanced robustness leads to better performance in real-world scenarios, ensuring more reliable and accurate detection of diabetic retinopathy. Overall, data augmentation is a vital preprocessing step in training deep learning models for diabetic retinopathy classification, as it effectively improves model accuracy and generalization by providing a richer and more varied training dataset.

3.3 Train Test Split

A training set and a testing set are separated from the 3,150 retinal images in the dataset. Particularly, 550 images are used for testing and 2,600 images are used for training. This indicates a roughly 80% train-test split, with 20% going toward testing. The diversity and size of the training dataset are effectively increased by applying data augmentation techniques, such as rotations, flips, translations, zooms, and intensity adjustments, to the 2,600 training images.

3.4 Proposed Model

3.4.1 Convolutional Neural Network (CNN)

An advanced kind of artificial neural network intended for processing and analyzing visual data is called a convolutional neural network (CNN). They have transformed computer vision by making it possible to automatically and effectively extract characteristics from images. This has resulted in major advancements in object detection, image classification, and other tasks utilizing spatial data. CNNs are characterized by

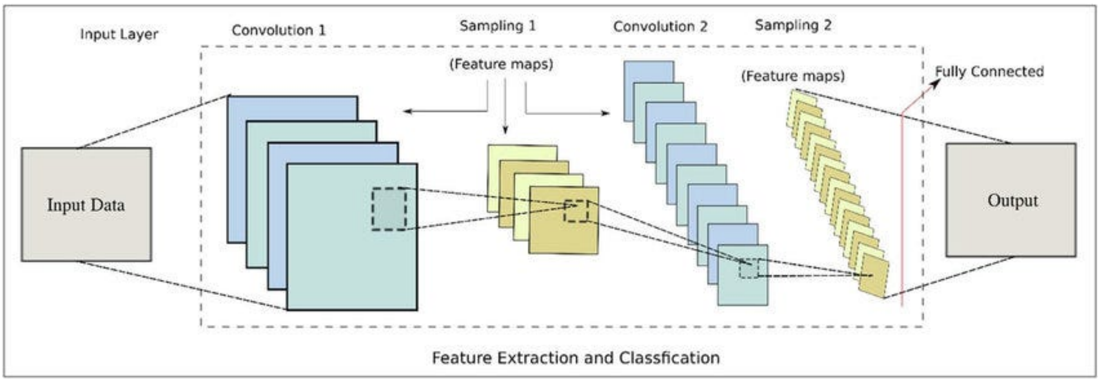

Figure 3.6: CNN architecture

their use of convolutional layers, which apply convolutional operations to input data. The primary components of a CNN include:

- **Convolutional Layers:** These layers use convolutional filters to identify different features in the input data, including textures, patterns, and edges. Every filter applies a dot product operation to the input image as it slides over it, creating feature maps that represent the input's spatial hierarchy.

- **Activation Functions:** To add non-linearity to the model, an activation function—typically the Rectified Linear Unit (ReLU)—is used after the convolutional process. The network is able to learn intricate patterns and representations because of its non-linearity.

- **Pooling Layers:** The most significant information is retained in feature maps despite their reduced spatial dimensions thanks to pooling methods like max pooling and average pooling. By lowering the number of parameters, this downsampling lessens overfitting and lowers the computational burden.

- **Fully Connected Layers:** The high-level features are flattened and then passed through one or more fully connected layers following a sequence of convolutional and pooling layers. These layers combine the learned features to carry out the final classification or regression tasks.

- **Output Layer:** The model's predictions are produced at the final output layer. This is usually a softmax layer that generates probabilities for each class in classification tasks.

Convolutional Neural Networks represent a powerful and versatile tool in the realm of machine learning, particularly for tasks involving visual data. Their ability to automatically learn and extract features makes them indispensable for modern computer vision applications. By leveraging CNNs, researchers and practitioners can build robust models that achieve impressive performance across a variety of challenging tasks.

3.4.2 Naive Bayes:

With an assumption of predictor independence, naive bayes is a probabilistic classification method based on Bayes' theorem. Because of its ease of use and efficiency, it is very helpful for classification tasks and is frequently employed in a variety of applications, including text categorization, sentiment analysis, and spam filtering.

Bayes' Theorem

The Bayes theorem, which is formally defined as follows and describes the likelihood of a class given the traits, is the foundation of the Naive Bayes model:

$$P(C_k|\mathbf{x}) = \frac{P(C_k) \cdot P(\mathbf{x}|C_k)}{P(\mathbf{x})}$$

Where:

- $P(C_k|\mathbf{x})$ is the posterior probability of class C_k given the features $\mathbf{x}$.
- $P(C_k)$ is the prior probability of class C_k.
- $P(\mathbf{x}|C_k)$ is the likelihood of features $\mathbf{x}$ given class C_k.
- $P(\mathbf{x})$ is the probability of the features.

Figure 3.7: Bayes' theorem

Types of Naive Bayes Models:

- **Gaussian Naive Bayes:**assumes that the distribution of characteristics is Gaussian, or normal. When the characteristics are regularly dispersed and continuous, it is frequently utilized.
- **Multinomial Naive Bayes:** Ideal for discrete features in text classification problems, like word counts or frequencies. It makes the assumption that a multinomial distribution governs the distribution of characteristics.
- **Bernoulli Naive Bayes:** utilized for boolean and binary characteristics. It is helpful for applications like document classification when features are binary indicators of word existence or absence because it assumes that features follow a Bernoulli distribution.

Large datasets and text classification issues are two areas in which the Naive Bayes model excels as a reliable and effective classifier. It is a useful tool in the machine learning toolkit and frequently yields competitive performance despite its simplicity and feature independence assumption.

3.4.3 k-nearest neighbors (KNN):

A straightforward yet effective non-parametric approach for regression and classification is K-Nearest Neighbors (KNN). Its foundation is the idea that predictions can be made by identifying the examples in the training data that are the most comparable. Because of its versatility and efficiency in a range of situations, such as classification and regression tasks, KNN is a widely used algorithm..

How KNN Works: The distance between the input data point and the training set's data points is compared by the KNN algorithm to determine how it operates. A data point is classified using either the average of the target values (for regression) or the majority class (for classification) of its k nearest neighbors.

- **Distance Metrics:** Various methods to calculate the distance between data points.
  - **Euclidean Distance:** Measures the straight-line distance between two points in a Euclidean space. It is defined as:

    $$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

    where $\mathbf{x}$ and $\mathbf{y}$ are two points in n-dimensional space, and x_i and y_i are their respective coordinates.
  - **Manhattan Distance:** Measures the sum of the absolute differences between coordinates of two points. It is defined as:

    $$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$
  - **Minkowski Distance:** A generalization of both Euclidean and Manhattan distances, defined as:

    $$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$$

    where p is a parameter that defines the type of distance.
- **Choosing K:** The number of neighbors K is a crucial hyperparameter in KNN.The choice of K affects the model's performance:
  - **Small K :** More sensitive to noise in the data.
  - **Large K:** More robust but may smooth out important patterns.
- **Classification:** KNN selects the class label that is most prevalent among the K nearest neighbors for classification tasks.
- **Regression :** When it comes to regression problems, KNN estimates the target value as the mean of the K nearest neighbors' target values.

The K-Nearest Neighbors (KNN) algorithm is a straightforward and effective method for classification and regression. Despite its simplicity, it can achieve strong performance in various tasks, especially when combined with appropriate distance metrics and careful selection of the K parameter. KNN is a valuable tool for practitioners due to its ease of implementation and interpretability.

3.4.4 Residual Network Architecture- ResNet

ResNet-50 is a 50-layer convolutional neural network that may be applied to computer vision tasks. A novel neural network called ResNet enables the training of very deep networks with more than 150 layers, solving the "vanishing gradient problem." ResNet was first mentioned in a computer vision research paper by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in 2015 titled "Deep Residual Learning for Image Recognition." ResNet uses the SKIP CONNECTION method to get over the gradient value reduction problem during backpropagation.

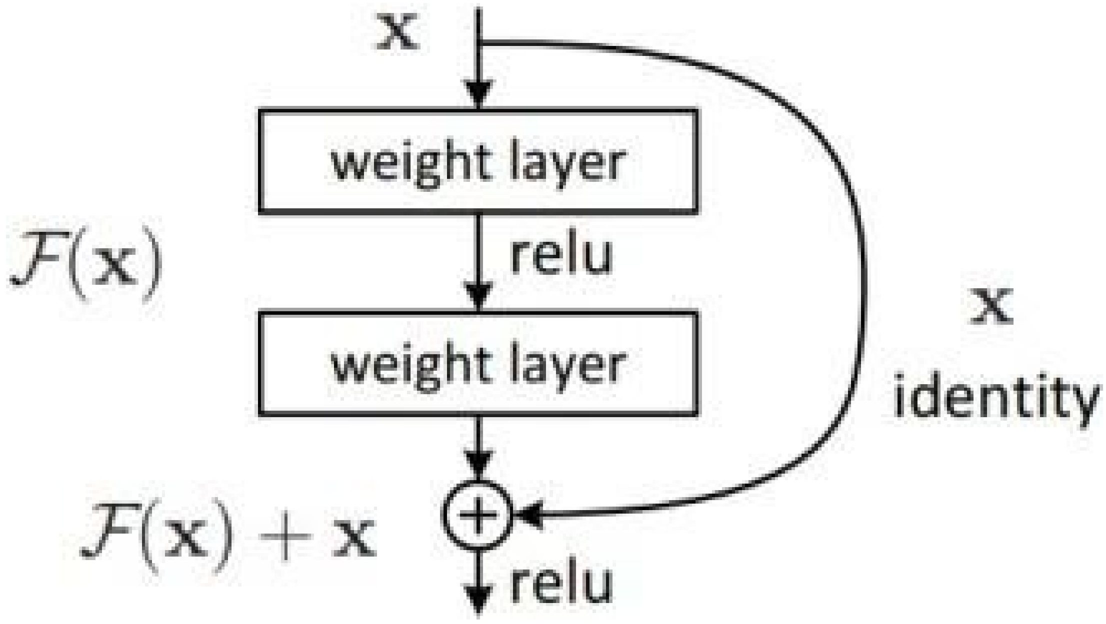

Figure 3.8: Skip Connections.

A direct link between some model layers that results in a different output is known as skip connections. The input is multiplied by the layer weights and a bias factor when a connection with a skip is available.

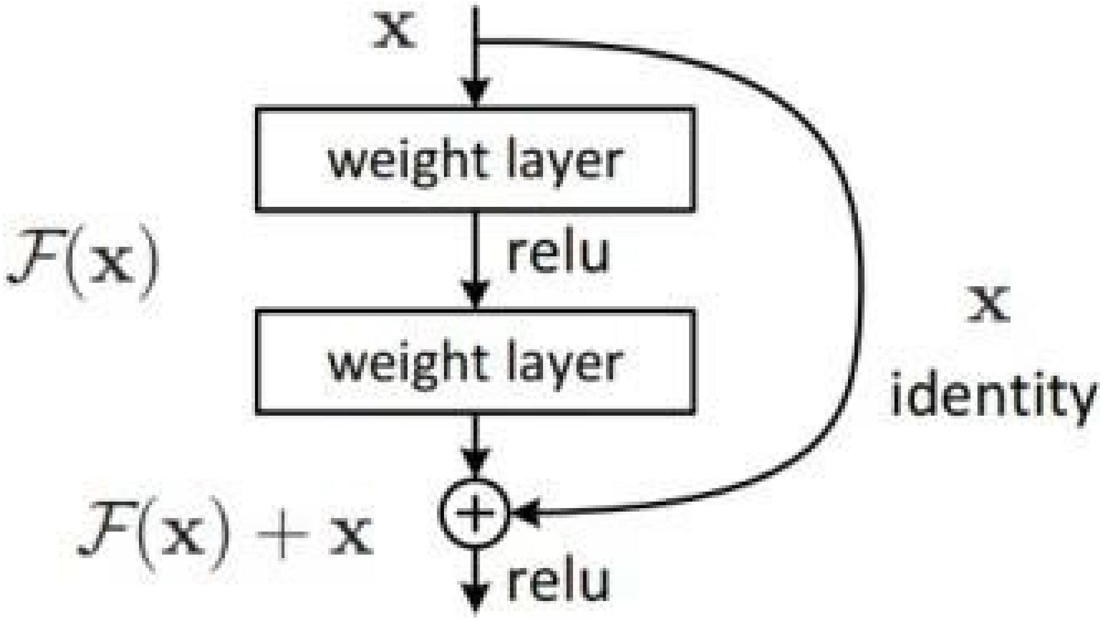

Figure 3.9: Skip Connections.

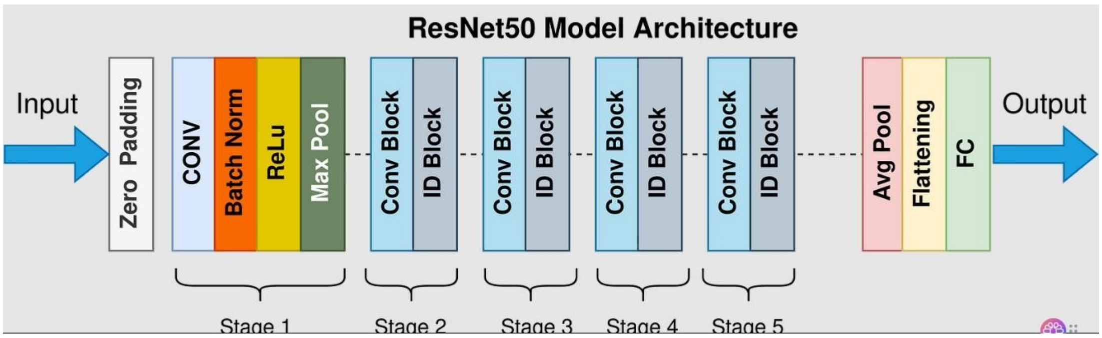

Figure 3.10: ResNet-50 Architecture.

3.4.5 VGG-16

Small convolution filters make up VGG networks, and VGG16 has 13 convolutional and 3 fully linked layers. The network uses a 3×3 receptive field for convolution and a 224×224 input picture.The Rectified Linear Unit (ReLU) activation function, which requires less training time and memory without sacrificing accuracy, is used by VGG in place of local response normalization. Furthermore, VGG retains a fixed convolution stride of 1 pixel to preserve spatial resolution following convolution.

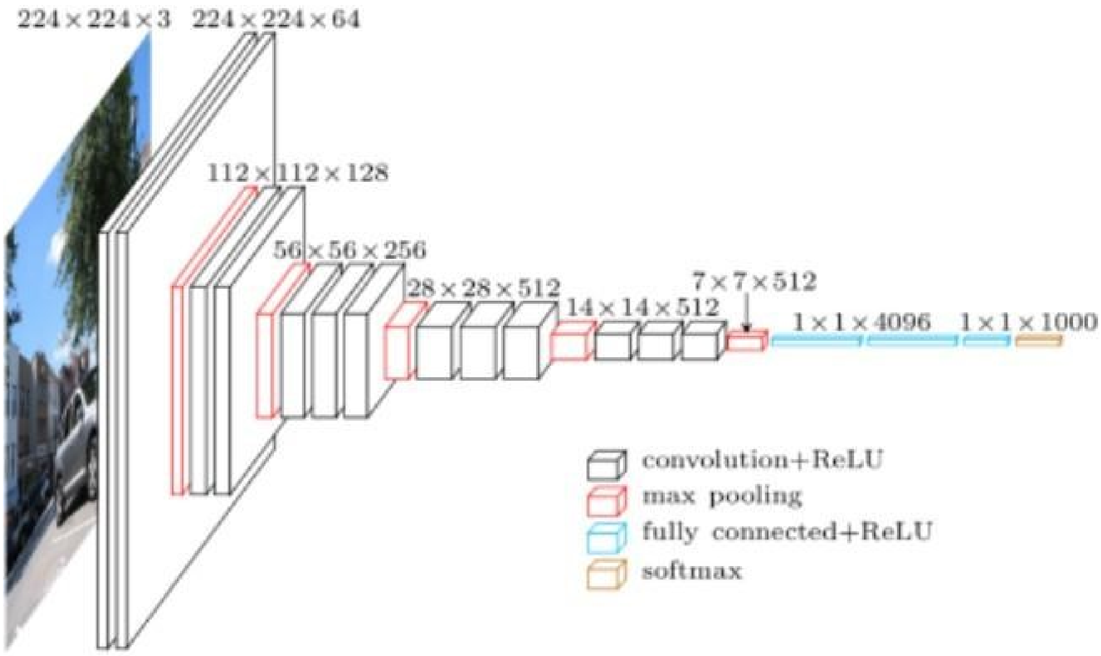

Figure 3.11: VGG-16 architecture.

3.4.6 InceptionV3

For high performance in image classification tasks, InceptionV3 is a deep convolutional neural network. Convolutions, average pooling, max pooling, concatenations, dropouts, and completely linked layers are among the intricate structures that make up its several inception modules. The network can record different degrees of spatial information and details thanks to these modules. The key architectural components of InceptionV3 include:

- **Factorized Convolutions:** InceptionV3 employs factorized convolutions to reduce the computational cost. For instance, a 3×3 convolution is decomposed into two 1×3 and 3×1 convolutions, which significantly reduces the number of parameters and thus the computational load.
- **Auxiliary Classifiers:** To address the vanishing gradient problem and improve the learning process, InceptionV3 incorporates auxiliary classifiers connected to intermediate layers. These auxiliary networks provide additional gradient signals and act as regularizes.
- **Efficient Grid Size Reduction:** InceptionV3 introduces efficient ways to reduce the grid size, such as using convolutions with stride instead of pooling layers,

which helps maintain computational efficiency while preserving the model's accuracy.

- **Label Smoothing:** This technique helps in regularizing the network by preventing the model from becoming overly confident about the training data labels, which can improve the generalization ability of the network.

- InceptionV3 accepts input images of size 299×299 pixels and utilizes the Rectified Linear Unit (ReLU) activation function throughout the network. This activation function accelerates the training process without sacrificing accuracy. The network uses batch normalization extensively across the layers, which helps in stabilizing and accelerating the training.

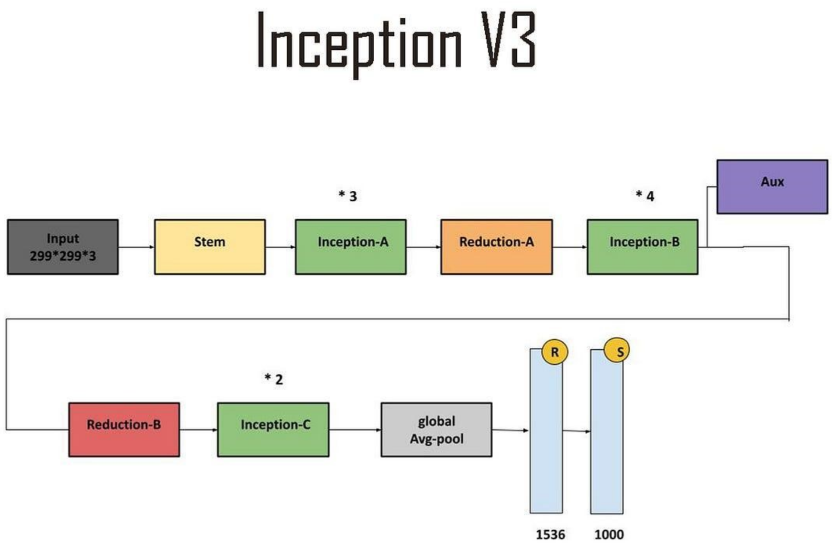

Figure 3.12: Inception V3 model architecture.

3.5 Model Training Parameter

Name	Value
Optimizer	Adam
Loss Function	Categorical cross-entropy
Metrics	Accuracy
Activation Function	Softmax
Batch Size	32

Figure 3.13: Model hyperparameter.

Chapter 4

Performance Analysis

4.1 Model Training with CNN:

In the model training phase using Convolutional Neural Networks (CNNs) for diabetic retinopathy classification, we focus on training the CNN to effectively learn from image data and generalize well to new, unseen data. The training process involves defining a suitable CNN architecture with multiple convolutional and pooling layers to capture essential features from images, followed by fully connected layers for classification. During training, we monitor performance by plotting training loss versus validation loss, as well as accuracy versus loss graphs. These visualizations are crucial for assessing the model's learning progress and identifying issues such as overfitting or underfitting. By analyzing these graphs, we can determine the best epoch—where the model achieves the highest accuracy with the lowest validation loss—ensuring a balance between learning and generalization. This comprehensive approach allows us to optimize the CNN model's performance for accurate and reliable diabetic retinopathy detection.

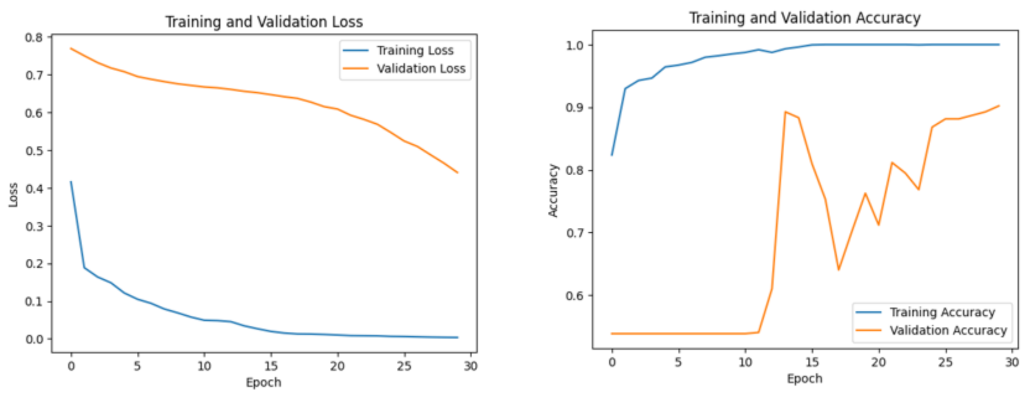

Figure 4.1: Training and Validation

4.2 Model Training with Naive Bayes:

In the model training phase using Naive Bayes for diabetic retinopathy classification, we leverage its probabilistic framework to classify images based on learned feature distributions. Naive Bayes models, including Gaussian, Multinomial, and Bernoulli variants, are chosen based on the nature of the feature data—continuous, discrete, or binary, respectively. During training, the model estimates the prior and likelihood probabilities from the training data to make predictions. Performance is evaluated by assessing metrics such as accuracy and confusion matrices on a validation set. These evaluations help in understanding the model's effectiveness and its ability to generalize to new data, ensuring accurate and reliable classification of diabetic retinopathy.

	precision	recall	f1-score	support
DR	0.85	0.87	0.86	113
No_DR	0.87	0.86	0.86	118
accuracy			0.86	231
macro avg	0.86	0.86	0.86	231
weighted avg	0.86	0.86	0.86	231

Figure 4.2: Classification Report for Naive Bayes

4.3 Model Training with KNN:

In the model training phase using K-Nearest Neighbors (KNN) for diabetic retinopathy classification, we apply this distance-based algorithm to classify images by comparing them to the nearest training samples. KNN works by evaluating the similarity between the input data point and the training data points using a chosen distance metric, such as Euclidean or Manhattan distance. During training, the model determines the optimal number of neighbors K and uses this to classify new data based on the majority class among the K nearest neighbors. Model performance is evaluated through metrics such as accuracy and confusion matrices, which are calculated on a validation set to ensure that the model is effectively learning and generalizing from the data for reliable classification results.

	precision	recall	f1-score	support
0	0.89	0.92	0.90	113
1	0.92	0.89	0.91	118
accuracy			0.90	231
macro avg	0.90	0.91	0.90	231
weighted avg	0.91	0.90	0.90	231

Figure 4.3: Classification Report for KNN

4.4 Model Training with VGG-19:

In the model training phase using VGG-19 for diabetic retinopathy classification, we will evaluate the performance by plotting the training loss versus validation loss, as well as the accuracy versus loss graph. These visualizations are crucial for understanding how well the model is learning and generalizing to new data. The training loss and validation loss graphs will help us identify if the model is overfitting or underfitting, by comparing the error on the training dataset against the error on the validation dataset over each epoch. Additionally, the accuracy versus loss graph will provide insights into the model's performance and convergence. By analyzing these graphs, we will be able to pinpoint the best epoch— the point where the model achieves the highest accuracy with the lowest validation loss, indicating the optimal balance between learning and generalization. This comprehensive evaluation ensures that the model is trained effectively for accurate and reliable detection of diabetic retinopathy

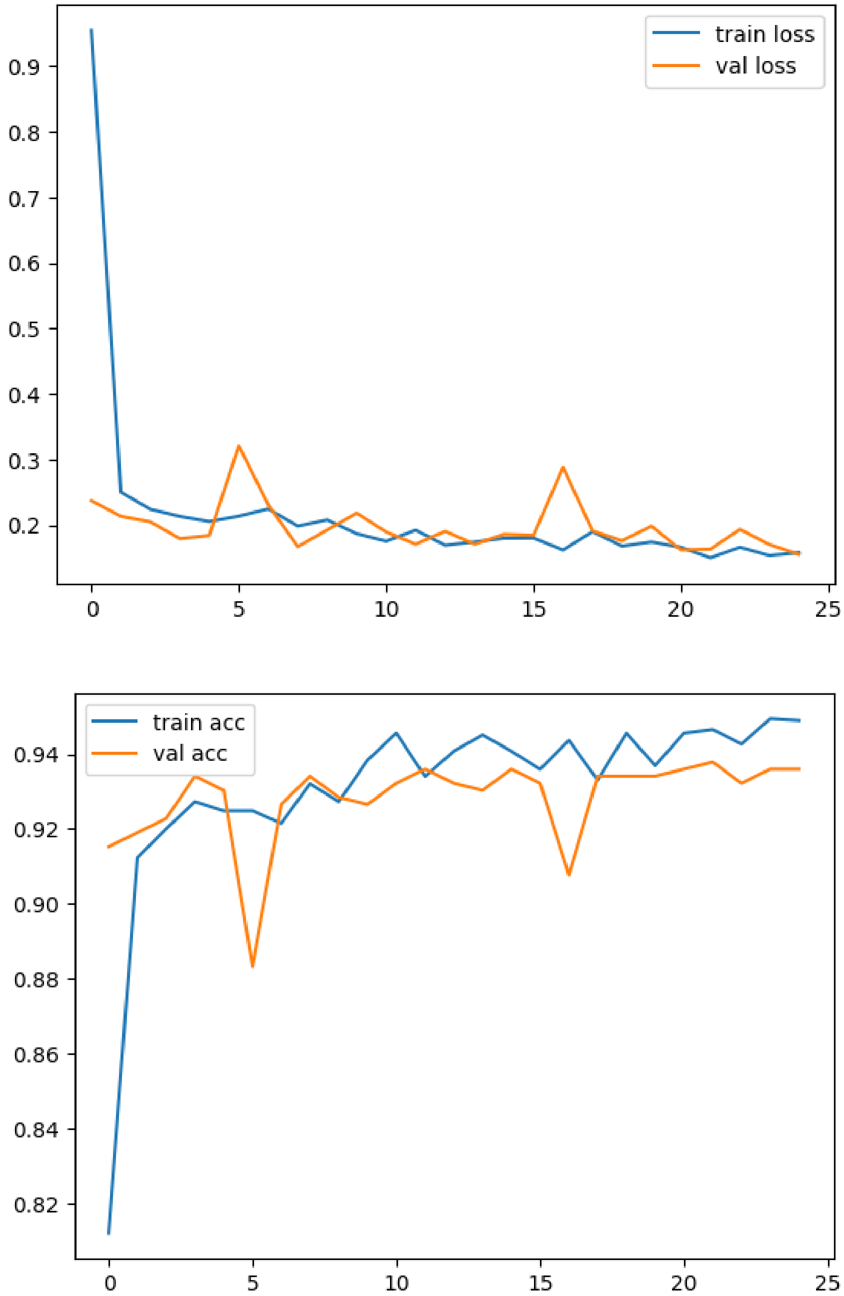

Figure 4.4: VGG-19 Model Performance (Accuracy vs loss).

Figure 4.5: VGG-19 Model Performance.

4.5 Model Training with InceptionV3

In the model training phase using InceptionV3 for diabetic retinopathy classification, we will evaluate the performance by plotting the training loss versus validation loss, along with the accuracy versus loss graph. These visualizations are essential for assessing the model's learning progress and its ability to generalize to unseen data. The training loss and validation loss graphs will reveal any signs of overfitting or underfitting by comparing the errors on the training and validation datasets across epochs. Additionally, the accuracy versus loss graph will provide a clear picture of the model's performance and convergence behavior. By scrutinizing these graphs, we can determine the best epoch— the point where the model achieves peak accuracy with minimal validation loss, signifying the optimal training stage. This thorough evaluation ensures that the InceptionV3 model is trained effectively, enhancing its capability for precise and dependable detection of diabetic retinopathy.

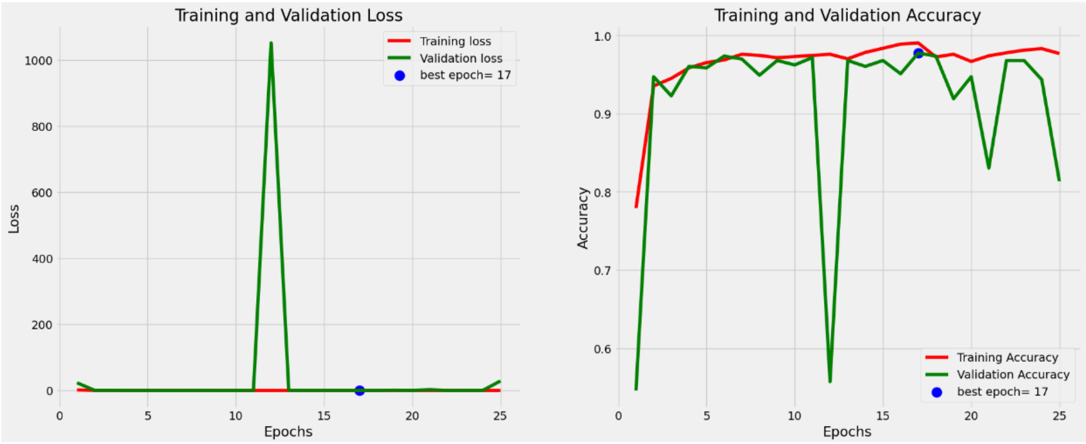

Figure 4.6: Training with Inception V3

4.6 Model training with ResNet-50

In the model training phase using ResNet-50 for diabetic retinopathy classification, we will assess the performance by plotting the training loss versus validation loss, as well as the accuracy versus loss graph. These visualizations are crucial for understanding the learning dynamics and the model's generalization ability. The training loss and validation loss graphs will help identify potential overfitting or underfitting by comparing the errors on the training and validation datasets over each epoch. Furthermore, the accuracy versus loss graph will provide insights into the model's performance and convergence trends. By analyzing these graphs, we can pinpoint the best epoch— the point where the model achieves the highest accuracy with the lowest validation loss, indicating the optimal balance between learning and generalization. This comprehensive evaluation ensures that the ResNet-50 model is trained effectively, leading to accurate and reliable detection of diabetic retinopathy.

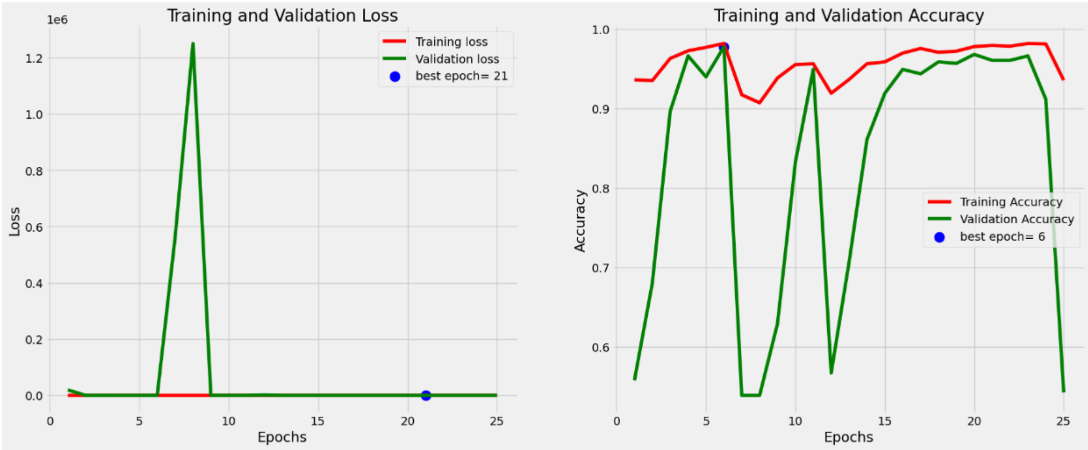

Figure 4.7: Training with ResNet-50

4.7 Comparative Study:

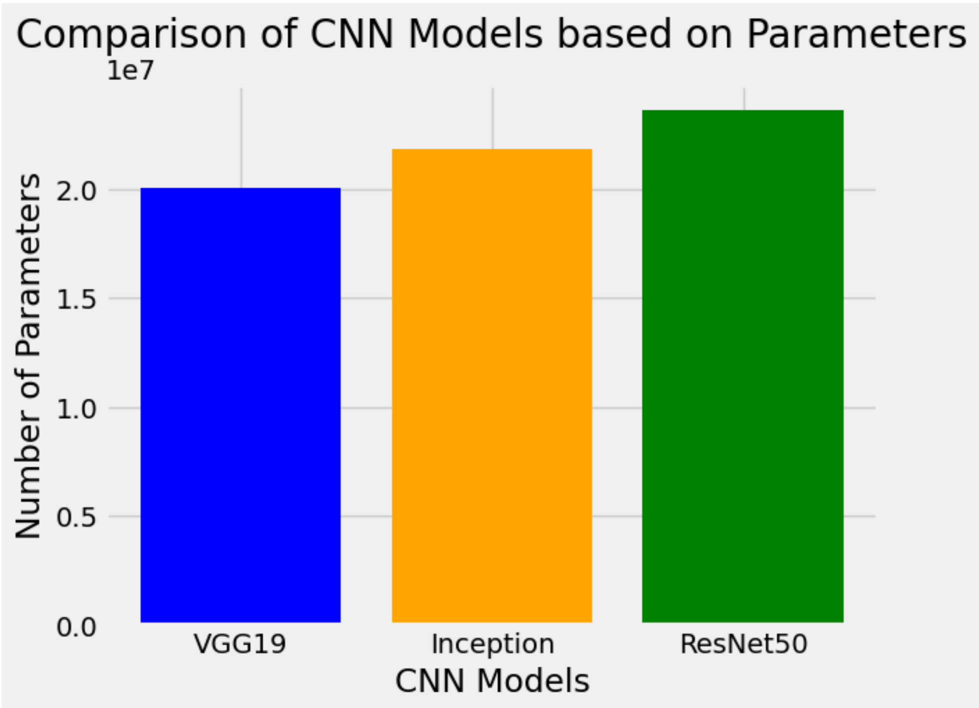

Figure 4.8: Comparison between the models

4.8 Result

4.8.1 Performance Comparison:

Model Name	Model Accuracy
CNN	90%
Naive Bayes	86%
KNN	90%
VGG-19	94.89%
Inception-V3	97.60%
ResNet-50	93.55%

Table 4.1: Performance of different models

summarizes the performance of various machine learning models used for diabetic retinopathy classification, presenting their accuracy rates. The table shows that the VGG-19 model achieves an accuracy of 94.89%, reflecting its strong performance due to its deep architecture and ability to learn complex features from images. The Inception-V3 model surpasses all others with an accuracy of 97.60%, indicating its superior capability in handling image classification tasks through its advanced inception modules. The ResNet-50 model also performs well with an accuracy of 93.55%, benefiting from its residual learning framework that helps in training deeper networks. Both CNN and KNN models achieve accuracies of 90%, demonstrating solid performance in image classification but with slightly less effectiveness compared to the more advanced architectures. Naive Bayes, with an accuracy of 86%, performs the least among the models listed, reflecting its limitations with image data that requires more complex feature extraction methods. This comparison highlights the strengths and limitations of each model, offering insights into their relative effectiveness for the task of diabetic retinopathy classification.

4.9 Evaluation Metric:

Four downstream tasks (cross-lingual sentiment analysis, named entity recognition, binary Text Classification, and multi-class sentiment analysis) are used to evaluate Bangla-BERT for Bangla language understanding. Furthermore, we compared Bangla-BERT to the multilingual version of BERT, taking into account other improved neural approaches like word2vecfor results for every job as a starting point.

1. **Accracy:** For machine learning classification models, accuracy is a performance parameter that is defined as the ratio of true positives and negatives to the total number of positive and negative observations. To put it another way, accuracy is the percentage of times we expect our machine learning model to accurately predict a result out of all the predictions it has made. It is defined mathematically as the ratio of all true positives (TP) to all true negatives (TN).

   $$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

   where TP is True Positives, TN is True Negatives, FP is False Positives, and FN is False Negatives.

2. **Precison:** The percentage of correctly predicted positive labels is estimated by the model precision score. Another name for precision is the predictive value of the positive. It displays the true positives to the total of the true positives and false positives (FP) mathematical ratio.

   $$\text{Precision} = \frac{TP}{TP + FP}$$

3. **Recall:** The model's recall score indicates how well it can distinguish true positives from forecasted positives. The ratio of genuine positives to the total of true positives and false negatives (FN) is shown mathematically.

   $$\text{Recall} = \frac{TP}{TP + FN}$$

4. **F1-Score:** The model score as a function of the recall and precision scores is indicated by the F1 score. An alternative to accuracy measures, the F1 score gives equal weight to both recall and precision when evaluating a machine learning model's accuracy performance. A harmonic mean of the precision and recall scores can be used to quantitatively express it.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

4.9.1 Confusion Matrix:

A confusion matrix can be used to evaluate how well a VGG-19 model works in the classification of diabetic retinopathy. This matrix takes into account the number of true positives, true negatives, false positives, and false negatives. It provides a concise summary of the model's accuracy and highlights areas that need development, making it crucial for improving the diagnosis of diabetic retinopathy.

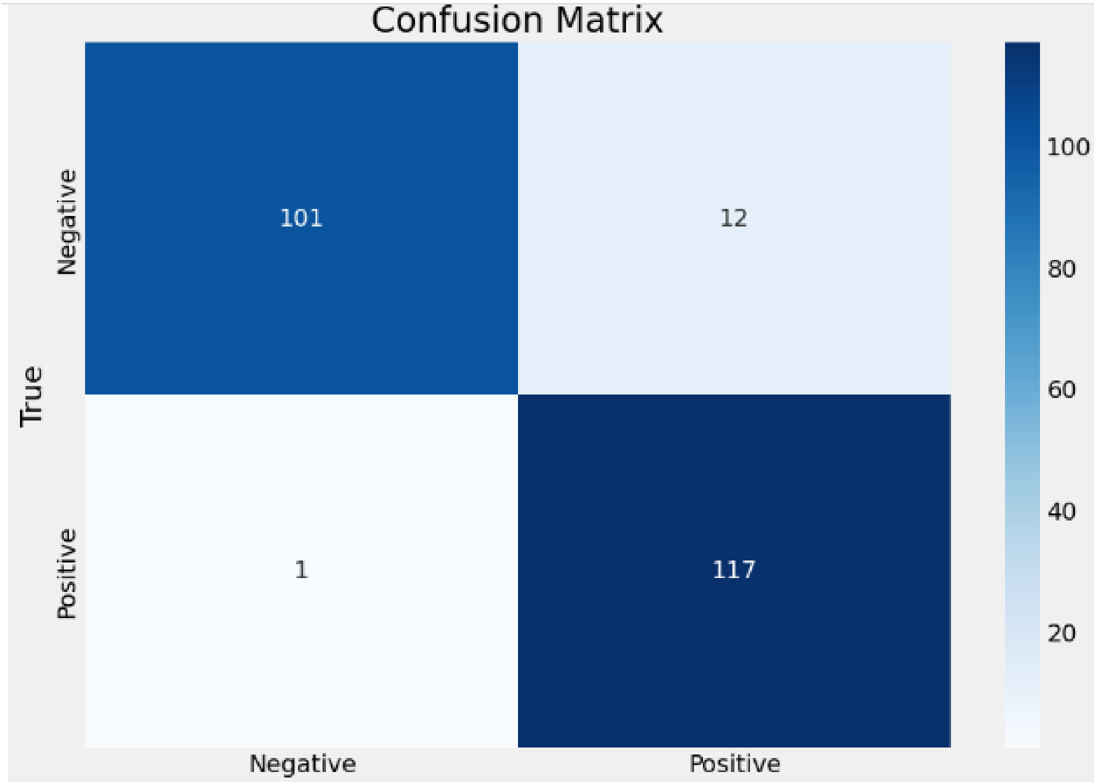

Figure 4.9: Confusion Matrix for VGG-19

4.9.2 Classification Visualization

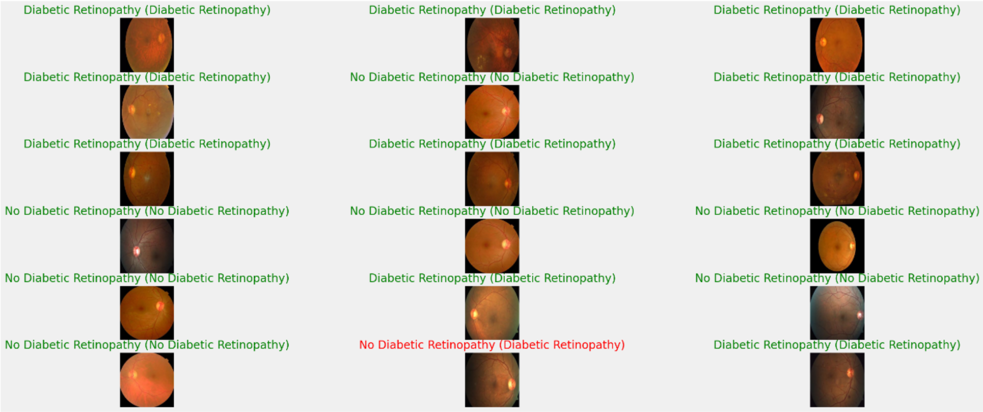

Figure 4.10: Result of classification

Chapter 5

Conclusions

5.1 Conclusions

In this study, transfer learning was employed to categorize DR into 2 classes using minimal training data. On the complete Kaggle DR dataset, the model fared better than alternative transfer learning techniques. With little training data, deep learning approaches may effectively classify DR and other medical picture classification issues. The performance of other trained deep convolutional models should be tested and compared.

5.2 Future Prospects of the Work

We'll apply the same model to a bigger dataset in contrast to the one we have now. We intend to apply the feature extraction part of pre-trained models to support vector machines and modify performance metrics like specificity and sensitivity, thereby boosting the trust of healthcare organizations in the usage of models in real time. We will evaluate the effectiveness of several photo pre-processing techniques, apply different transfer learning techniques, and use the pre-trained models to tackle challenging real-world image categorization problems. In the future scope of our novel work, integrating explainable AI techniques will be a crucial aspect of model development. Explainable AI aims to enhance the transparency and interpretability of machine learning models, particularly in complex medical applications like retinal image analysis. By incorporating explainable AI methods, such as attention mechanisms or gradient-based visualization techniques, we can gain insights into which regions of retinal images are most influential in the model's decision-making process.

References

[1] U. R. Acharya, C. M. Lim, E. Y. K. Ng *et al.*, "Computer-based detection of diabetes retinopathy stages using digital fundus images," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 2009.

[2] X. Chen, B. P. Nguyen *et al.*, "Automated brain tumor segmentation using kernel dictionary learning and superpixel-level features," in *Proceedings of the International Conference on Systems, Man, and Cybernetics (SMC 2016).* Budapest, Hungary: IEEE, 2016, pp. 2547–2552.

[3] ——, "An automatic framework for multilabel brain tumor segmentation based on kernel sparse representation," *Acta Polytechnica Hungarica*, vol. 14, no. 1, pp. 25–43, 2017.

[4] H.-T. Cheng, L. Koc *et al.*, "Wide & deep learning for recommender systems," in *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems (DLRS 2016).* ACM, 2016, pp. 7–10.

[5] Mayo Clinic, "Diabetic retinopathy," 2018, https://www.mayoclinic.org/diseasesconditions/diabetic-retinopathy/symptoms-causes/syc-20371611.

[6] Graham, "Kaggle diabetic retinopathy detection competition report," 2015, https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15801.

[7] Kaggle, "Diabetic retinopathy detection," 2015, https://www.kaggle.com/c/diabetic-retinopathy-detection/overview.

[8] J. Lachure, A. V. Deorankar *et al.*, "Diabetic retinopathy using morphological operations and machine learning," in *Souvenir of the 2015 IEEE International Advance Computing Conference (IACC 2015)*, 2015.

[9] K. Mader, "Kaggle diabetic retinopathy detection competition report," 2015, https://www.kaggle.com/kmader/inceptionv3-for-retinopathy-gpu-hr.

[10] B. P. Nguyen, Q. H. Nguyen, G.-N. Doan-Ngoc, T.-H. Nguyen-Vo, and S. Rahardja, “iprodna-capsnet: Identifying protein-dna binding residues using capsule neural networks,” *BMC Bioinformatics*, vol. 20, no. Suppl 23, pp. 1–12, 2019.

[11] MissingLink.ai, “Convolutional neural networks: architecture forging pathways to the future,” https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neuralnetworksarchitecture-forging-pathways-future.

[12] J. De Calleja, L. Tecuapetla, and M. A. Medina, “Lbp and machine learning for diabetic retinopathy detection,” in *Proceedings of 2014 Conference*, 2014, pp. 110–117.

[13] K. A. Anant, T. Ghorpade, and V. Jethani, “Diabetic retinopathy detection through image mining for type 2 diabetes,” in *2017 International Conference on Computer Communication and Informatics, ICCCI 2017*, 2017.

[14] B. P. Nguyen, H. N. Pham *et al.*, “Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records,” *Computer Methods and Programs in Biomedicine*, vol. 182, p. 105055, 2019.

[15] S. Sabour, N. Frosst, and G. E. Hinton, “Dynamic routing between capsules,” in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., 2017, pp. 3856–3866.

[16] N. Silberman, K. Ahlrich *et al.*, “Case for automated detection of diabetic retinopathy,” in *AAAI Spring Symposium - Technical Report*, 2010.

[17] A. Sopharak, B. Uyyanonvara, and S. Barman, “Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering,” *Sensors*, vol. 9, no. 3, pp. 2148–2161, 2009.