Deep-2DRP: Identifying Diabetic Retinopathy Progression via CNN-ResNet50

Supervised By

Dr. Imran Mahmud

Head & Professor

Department of Software Engineering

Daffodil International University

Submitted By

KOWSHIK KUMER

ID:213-35-777

Department of Software Engineering

Daffodil International University

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

APPROVAL

This thesis titled on "**Deep-2DRP: Identifying Diabetic Retinopathy Progression via CNN-ResNet50**", submitted by **KOWSHIK KUMER (ID:213-35-777)** 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
Professor & Head
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

-- **Internal Examiner 1**

Md Shohel Arman
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

-- **Internal Examiner 2**

Md. Rajib Mia
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

-- **External Examiner**

Md Habibur Rahman
Associate Professor
Department of Computer Science and Engineering
Islamic University, Bangladesh

Deep-2DRP: Identifying Diabetic Retinopathy Progression via CNN-ResNet50

KOWSHIK KUMER
213-35-777

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

SUPERVISOR’s DECLARATION

I hereby declare that I have reviewed this thesis entitled " Deep-2DRP: Identifying Diabetic Retinopathy Progression via CNN-ResNet50 ", and in my opinion, it is adequate in terms of scope and quality for the award of the degree of Bachelor of Science in Software Engineering.

____

(Supervisor’s Signature)

Full Name : Dr. Imran Mahamud

Position : Head & Professor

Date : 13 September 2025

STUDENT'S DECLARATION

I certify that the material contained in this thesis is my own work, except where credit is specifically given to others. I also confirm that it has not already or concurrently been submitted for any other degree to Daffodil International University or any other institution.

(Student's Signature)

Full Name : KOWSHIK KUMER

ID Number : 213-35-777

Date : 13 September 2025

Deep-2DRP: Identifying Diabetic Retinopathy Progression via CNN-ResNet50

KOWSHIK KUMER

213-35-777

Thesis submitted in fulfilment of the requirements
for the award of the degree of
Bachelor of Science

Department of Software Engineering

DAFFODIL INTERNATIONAL UNIVERSITY

SEPTEMBER 2025

ACKNOWLEDGEMENTS

I want to thank Almighty God for his guidance and for giving me the abilities and expertise that was required to get this job done. This work could not have been done were it not for His infinite grace. In this consultation, my curiosity was mostly motivated to develop an understanding of contributing to the unexplored area of our science, so I started researching. The proposed methodology for Diabetes Retinopathy detection is based on a hybrid deep learning architecture comprising CNN and ResNet18.Thank you, Mom and Dad for all of your love, praise, encouragement, and prayers that continue to drive me along the path. I am also thankful to Prof. Dr. Imran Mahmud, who is the Head of Software Engineering Department, for creating such a motivating and learning environment throughout my whole academic career. First and foremost, I am greatly indebted to my respected supervisor Dr. Imran Mahmud for his continuous advice, suggestions and valuable guidance throughout this work. His input was critical in guiding and honing this piece.

DEDICATION

I commit this venture to my respectable Father and Mother, my supervisor, my Honorable teachers who are continuously expensive and close to me. Without their patience, understanding, unsparing bolster, care, love and cherish it was not conceivable to come up to this put.

ABSTRACT

Diabetic retinopathy (DR) is a dangerous condition that can result in vision loss for individuals with diabetes if it is not detected promptly. Identifying the specific stage of diabetic retinopathy is quite challenging and often necessitates the skilled interpretation of fundus images by professionals. Streamlining the revelation process is exigent and could typically benefit millions of individuals. The accustomed diagnosis method is based on experts manually scanning retinal pictures, which can be a rugged and error-prone procedure. This Study introduces Deep2DRP, a novel hybrid deep learning model that integrates CNN and Transfer learning with ResNet50 for exact DR prediction. The model demonstrated magnificent accuracy, with a success rate of 96.96%, precision of 0.97, recall of 0.96, F1 score of 0.97, and ROC- AUC score of 0.969. The outcome of this research has significant implications for ameliorating the diagnosis of DR. The Research highlights the dynamic of hybrid deep learning to significantly enhance Retinopathy prediction, providing valuable insight and tools for biomedical research and application. Future research will target ameliorating the elasticity of the model by incorporating a wider range of datasets and optimizing its integration with medical procedures to guarantee dependable and efficient performance in factual scenarios.

Keywords: Diabetic Retinopathy, Deep Learning, Convolutional Neural Networks, Transfer Learning, Ensemble Methods, Medical Image Classification.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

μ	Average
σ	standard deviation
log	power of a number

LIST OF ABBREVIATIONS

DR	Diabetic Retinopathy
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
RNN	Recurrent Neural Networks
ViTs	Vision Transformer
PDR	Proliferative Diabetic Retinopathy
NDPR	No Proliferative Diabetic Retinopathy
BDA	Binary Dragonfly Algorithm
Deep2DRP	Proposed Hybrid Diabetic Retinopathy Prediction Model
ML	Machine Learning
DL	Deep Learning
LBP	Local Binary Pattern

CHAPTER 1

INTRODUCTION

1.1 Introduction

Diabetes is the reason for multi-organ complications of the body, especially the eyes, kidney, Neurological system and heart. Diabetes increasingly damages the blood vessels inside, restricting blood flow and increasing the risk of organ failure. Diabetes affects the retina's circulation, it is known as DR. DR is a serious eye disease that is increasing day by day in the general population. People with diabetes are slightly more likely to develop Diabetic retinopathy. Four key stages characterize the progression of diabetic retinopathy. In the first stage, the problem can be solved by first aid. Around 382 million people were affected by DR in 2013 [4]. By 2045, the global burden of diabetes will likely impact around 784 million individuals [5].As the DR problem is becoming an epidemic, it is necessary to have DR screening in the general population once a year so that people can take measures to solve the problem at an early stage, making it a big challenge for ophthalmologists to take care of the eyes of the growing population happens due to high blood pressure and impairs the retina. It causes vascular damage in the retina, which may cause blindness and mortality. Detection of retinal vascular swelling by ophthalmologists relies solely on fundoscopy. Considering the burdensome cost and duration of fundoscopy, the development of automated DR detection systems using fundus images is imperative. It has been shown that deep learning approaches are very effective for DR diagnosis, often delivering more precise results than manual evaluation by ophthalmologists.

1.2 Background Study

Diabetic Retinopathy is one of the most common eye problems people with diabetes face. It happens when having high blood sugar for a long time starts damaging the tiny blood vessels in the retina, the part of the eye that turns light into images for your brain. These blood vessels can leak, swell, or even cause new, fragile ones to grow, which isn’t good news for your vision. At the start, DR might not show any symptoms at all. That’s why regular eye check-ups are so important for people living with diabetes. Doctors usually talk about two main stages. The first is NPDR, where the blood vessels get weaker and start leaking fluid or blood. The second, more serious stage is PDR, where new but fragile blood vessels grow and raise the risk of major vision problems, even blindness.

1.3 Motivation

This study is motivated by the need to understand demand for computerized methods and precise methods for detecting DR. Although the current manual screening procedure is difficult and able to detect human mistakes, early and precise detection is crucial to preventing irreparable blindness. Technology's quick development offers a chance to build a machine driven approach that can quickly and accurately detect DR. These developments may reduce the pressure on medical personnel, provide rapid response, and save the vision of millions of people all over the globe. Machine learning algorithms, especially deep learning approaches, have shown potential in different medical image diagnostic systems Previous research often misses the level of detail needed to account for the variety of factors influencing motorcycle costs. This gap may be filled by using deep learning models, which offer a more dependable approach to pricing in this ever-changing sector.

1.4 Purpose of the Research

The main objectives include developing and comparing automated spotting and DR Hybrid models. This work aims to evaluate the value of standard machine learning models as opposed to contemporary deep learning models. Among the objectives are performance analysis by contrasting ML models for DR sorting, investigating the impact of comparing datasets in colour and grayscale, and examining how image clarity affects model performance. Additionally, the study intends to apply these models in the real world with Streamlet, concentrating on obtaining accurate retinal images.

1.5 Research Objective

Our thesis objective is to identify diabetic retinopathy and classify each case as either PDR or NPDR.specific objectives of the study include

- Analyse colorized fundus retinal images to identify DR.
- Extract key attributes from the processed images.
- Recognize the presence of DR.
- Determine whether the DR is proliferative or non-proliferative.
- I created a new hybrid model, named Deep2DRP, to enhance detection and classification accuracy.

1.6 Scope of this Research

The majority of related works use two detectors to classify the disease: whether the patient has an infected eye with diabetic retinopathy or a healthy eye. Sorting these classes into proliferative and non-proliferative stages of diabetic retinopathy is the aim of our study. There are five categories for an eye: proliferative, mild non-proliferative, moderate non-proliferative, severe non-proliferative, and healthy. Furthermore, one of the best aspects we have looked into in this research is the extraction of handcrafted features from raw images following various enhancements.

Here, a vision analyzing-based approach has been undertaken in this study. The purpose of making use of such a method is to obtain some features which are diabetic retinopathy. Numerous shape-based functions and rebuilds were employed to obtain these features. Blood vessels and Microaneurysms are considered since these are the most problematic for eye diseases. Through reviewing the state of art work, it has been found that deep learning-based models on hand-crafted features from raw images provides accurate results. CNN, ResNet50 are selected for the research according to the survey.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter looks at the available research on the use of deep learning models for diabetic retinopathy (DR). CNN and ResNet-50 are examined and compared to see how good they are at spotting and rating DR. A side-by-side check of model results, including accuracy, Precision, Recall and specificity, is shown across different sets of data. Also, the chapter talks about how these models can fit into clinical work routines and how to tackle issues like lack of data.

2.2 Related Work on Diabetic Retinopathy

Nair and Mishra et al. [1] say that DR makes Solution leak from the retinal blood vessels, hurting vision as a result. Also, since it harms the retina, diabetes is a big reason for vision loss. Amalia et al. [2] suggested a hybrid deep-learning method of detecting diabetic retinopathy in retina fundus images. Within their model CNN, GoogleNet architecture, initially obtained rich visual features. These were then translated into a descriptive sequence and inputted into a LSTM network to be interpreted. This two-way architecture had a high accuracy of 90%. The resulting descriptive text that will be fed to ophthalmologists could be used in their diagnoses and can serve as the direction of future research in the automated DR detection.

Istiaq et al. [3] demonstrate that the study employs Local Binary Pattern (LBP) with CNN to construct a novel model from scratch. The proposed model is an ensemble feature representation optimized using the Binary Dragonfly Algorithm (BDA) and Sine Cosine Algorithm (SCA), achieving a peak accuracy of 98.85%.

Raju et al. [6] utilizes a CNN framework accessible Kaggle dataset which delivers 93.28% accuracy. Das et al. [7] demonstrated a CNN technique comprising a layer of convolution and pooling employed to extract the features, support vector machine algorithm to categorize and this technique attained a top score of 0.9867 on Messidor dataset. Daanouni et al. [8] illustrated CNN framework and its promising applications in tackling the challenges of adversarial attacks in medical image interpretation. Yasashvini et al. [9] presented Feature Extraction of Multi-Label and (ML-FEC) Model of Classification for Pre-trained on CNN Framework for Diabetic DR Identification. Sudha et al. [10] identifying DR, retinal abnormalities like exudates, haemorrhages, are categorized into precise grades as mild, moderate, severe. The segmentation using VGG-19 architecture on 5000 validation images is efficiently detected with 96% accuracy.

Amalia and her colleagues attempt to identify diabetic retinopathy (DR) with the help of retinal fundus images in this study. They combine deep learning methods of two network types including LSTM and CNN. CNN first removes a bunch of information in the retinal pictures. Then such details are converted into a sort of summary line that is inputted in the LSTM. It is a synthesis of the two models that they are seeking to explain and detect DR and they are dealing with a rather good solid of around 90%. The final product that falls as a descriptive sentence to eye doctors may prove useful in the examination room and guide the future investigations into the detection of diabetic retinopathy. Castellano et al. (2020) combined deep CNNs for picking out features and sorting with image handling ways for getting fundus pictures ready. They achieved 91.48% accuracy on the Messidor-2 dataset after preprocessing the fundus images.

Gargeya et al. [12] employed an adapted CNN to identify DR, and they trained their Model , using 75138 retinal images of their own dataset and were tested on Messidor-2 and E-Optha databases. They divided the images into two groups: those of healthy eyes and those with DR (any stage, i.e., mild or worse DR), achieving 94% sensitivity and 98% specificity in their own dataset. Additionally, they took their dataset to 94% sensitivity and 98% specificity. Furthermore, they tested their framework using the Messidor-2 dataset and achieved 93% sensitivity and 87% specificity. Philip et al. [13] developed a DR grading system of health and disease (mild and worse, they trained their model with 1067 images and verified on 14406 images. The sensitivity and specificity of their model was 86.2% and 76.8%, respectively. Usman et al. [14] shows CNN architectures such as ResNet50, ResNet152, and SqueezeNet1 are used to find and study. 94.40% accuracy is achieved in SqueezeNet1.

Mustafa et al. [15] applied a multi-stream ensemble deep network for automated grading and sorting of DR. Also, ResNet-50 and DenseNet-121 reach best accuracy on these 4 datasets EyePACS, Messidor-2 APTOS IDDR. Paranjpe et al. [16] presented a system for diagnosing diabetic retinopathy Fundus images, with which spots the area and texture traits of blood vessels, exudates. Selected important features are trained and tested and using support vector machines phases are average calculations.

Gu and his team [17] came up with a clever dual-channel CNN design to tackle the problem. To make training faster and cut down on overfitting, the first channel was packed with convolutional layers, max-pooling, and some selective batch normalization. This part of the network was tuned to pick up fine-grained details, like the texture patterns in smoke. With a different focus, the second channel sought to capture the larger image, such as the general outline or shape of smoke. This prevented vanishing gradient problems and overfitting. In addition to skip connections and global average pooling, it employed its own set of convolution and pooling layers. Ultimately, the outputs from both channels were integrated to capitalize on their respective advantages.

2.3 Use of ResNet50

Amira Mofreh et al. [18] Detection of DR Using Deep Learning Techniques" They compare multiple pre-trained models including ResNet-50 on a fundus dataset from Rob flow, reporting full metrics for the best model (DenseNet121). While the standout model is DenseNet121 (AUC 0.98, precision 91.6%, recall 92.25%), ResNet-50 is part of the mix, helping to contextualize how it stacks up against other deep networks.

Yijin Huang et al. [19] ResNet-50 grading diabetic retinopathy fundus images This one dives into what training tweaks like data augmentation, loss functions, and resolution matter for ResNet-50 on the EyePACS dataset. They don't report ROC/recall explicitly, but their optimized setup achieves a top-tier quadratically weighted Kappa of 0.8631, which is a solid indicator of grading accuracy. Surya Vamsi Patiballa (2025, GitHub project) DR Detection with ResNet-50 This GitHub project uses ResNet-50 on the APTOS 2019 dataset with transfer learning, preprocessing, and some smart data augmentation. The model pulls off a solid Training AUC of 97.77% and a Validation AUC of 94%, which is pretty impressive for a DR detection setup. Overall, it’s a neat, practical example of how ResNet-50 can handle real-world retinal images.

2.4 Use of Vision Transformer (ViT)

Oahidul Islam et al [23] Using Vision Transformers (ViTs) for sorting retinal images has shown an impressive accuracy of 96.13% in spotting Diabetic Retinopathy (DR). The model's ability to clearly tell apart instances of diabetic retinal degeneration versus non-DR is further shown by its high receiver operation characteristic area under a curve (ROC AUC) value of 0.969.

CHAPTER 3

METHODOLOGY

3.1 Overview

The process is started by acquiring the images from popular data sets like Messidor, APTOS for Eye Fundus Images. These datasets consist of an equal ratio of diabetic retinopathy (DR) and non-DR so good for training. Since we wanted to properly check the model performance, we split the images into training sets and testing sets using an 80%-20% ratio. To begin with, we keep everything consistent by showing the size of all images to 224 × 224 px image dimension. And then we did data augmentation rotating, flipping, and adjusting brightness, so the model is more resilient to noise and does not overfit. For training, the learning rate was 0.001, we used Adam optimizer and 10 epochs on the CIFAR dataset. Moreover, at each neural layer level we specified appropriate activation functions for prediction and applied Z-score standardization to normalize the data so that distribution does not change. We have trained the famous deep learning models like CNN and ResNet50, upon these architectures, I have built a new architecture Deep2DRP for DR detection at its best, so as to compare results clearly and deeply using confusion matrix, ROC curves and bar charts to see which model is the winner.

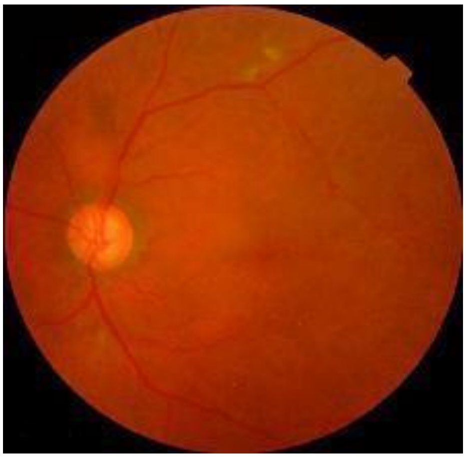
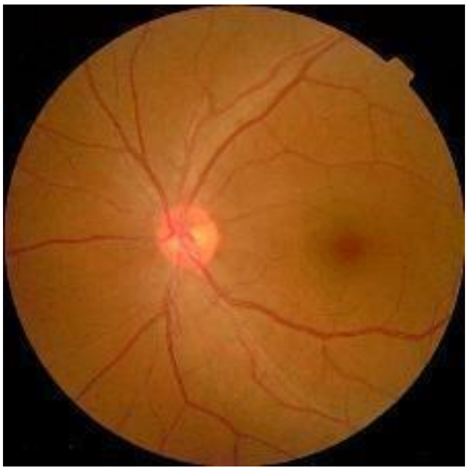
DR	No DR
	

Figure 3.1 Visual differences between DR & NO DR image

3.2 Experimental Environment and Workflow

- First, I gather a bunch of eye scan images from sources like Kaggle or Messidor, making sure to include both diseased (DR) and healthy ones, then I split them into training and test sets.
- Next, I tweak and prep the images, resizing them, adjusting brightness, flipping, and rotating to help the model learn better.
- I fine-tune the training setup by picking things like learning rate, number of epochs, and the optimizer (Adam in this case).
- For training, I use well-known deep learning models like CNN and ResNet50, plus my own Deep2DRP network.
- Finally, I check how well the model did using tools like confusion matrices, ROC curves, and bar charts to see the results clearly.

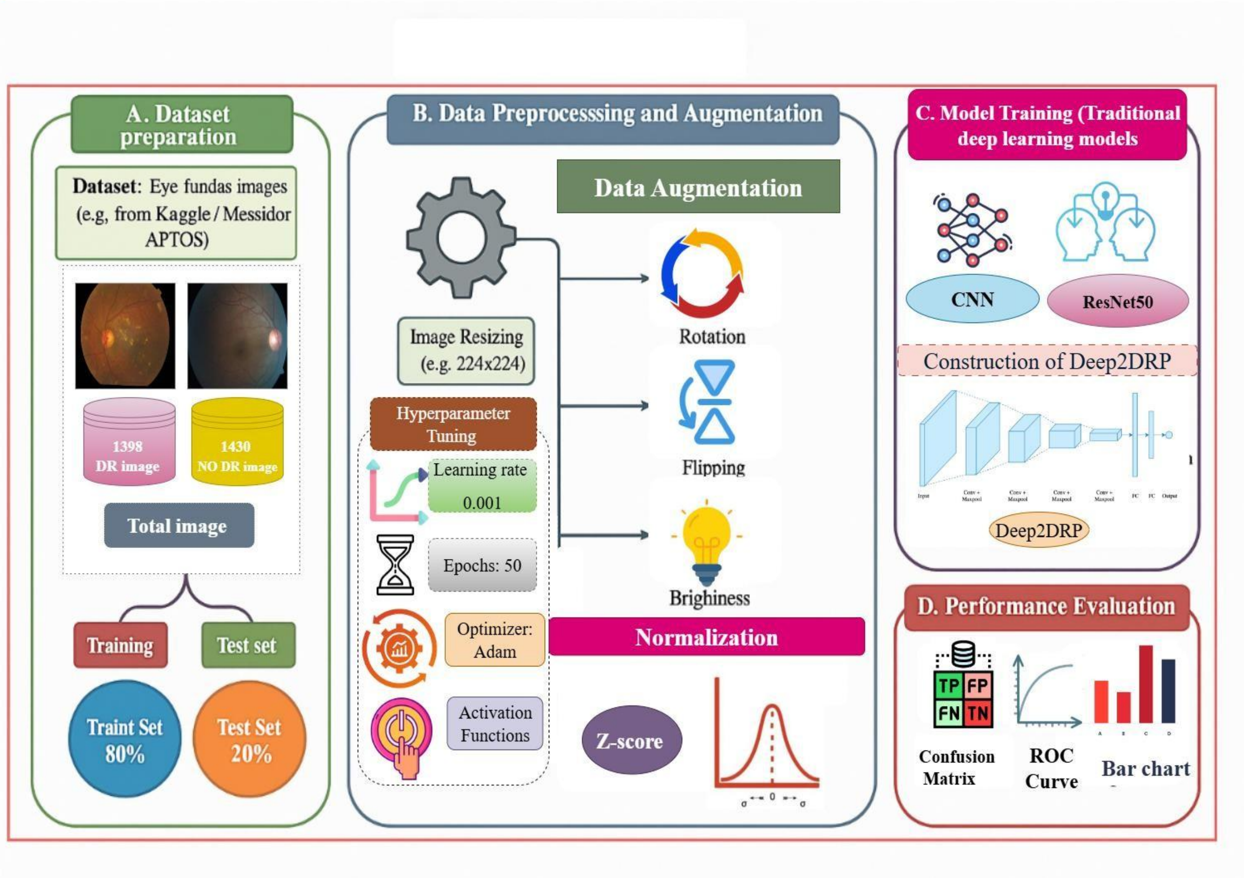


Figure 3.2 Dataset Processing and Modelling Workflow Diagram

3.3 Data Collection

The Messidor dataset is a recognized standard for DR research. We used it in this project to train and evaluate the model. It includes high-quality retinal fundus images from both DR and non-DR cases.

3.3.1 Data Source

The dataset was first released by the French Messidor program for automated DR screening. You can access it through the official Messidor repository or find it in a curated form on Kaggle. For this project, we obtained the dataset from base paper (14), where it is available in a processed format for research purposes.

3.3.2. Dataset Composition

Total Images: 2,838 retinal fundus images. Classes: Both diabetic retinopathy and non-diabetic retinopathy cases.

Table 3:1 Dataset Composition for DR and No DR Images

2828	Total Image
1398	DR Image
1430	NO DR Image

Figure 3.3: Image Count per class

The dataset used in this study consists of a total of 2,828 images, of which 1,398 are labelled as Diabetic Retinopathy (DR) images and 1,430 are labelled as No DR images. Since the distribution between the two classes is nearly equal, the dataset can be considered balanced. Therefore, additional balancing techniques such as ADASYN or SMOTE are not required for this study.

3.3.3 Image Resizing

I work on image resizing. I shrink or stretch all the images to the same size (like 224×224 pixels) so my model always gets a consistent input. I do it to avoid any size mismatches from different cameras or sources. That way, training runs smoother and doesn't waste extra computing power.

3.3.4 Data Augmentation

- **Rotation.** I slightly rotate images to show the model from different angles.
- **Flipping.** I flip images horizontally so it learns mirrored patterns.
- **Brightness Adjustment.** I adjust brightness to simulate different lighting conditions.

Figure 3.4 : Augmented Retinal Images

3.3.5 Data Normalization

For normalization, I scale my image pixel values to match the distribution used to train large pre-trained models like ResNet50. I use the ImageNet mean [0.485, 0.456, 0.406] and standard deviation [0.229, 0.224, 0.225]. This helps my model train faster, remain stable, and make better predictions.

Figure 3.5: Original and Normalized Retinal Fundus Images

3.3.6 Hyperparameter Tuning

To determine the most effective hyperparameter optimization of my model, Finally, I achieved a total of 10 epochs using 0.001 learning rate and a batch size of 32. I also make use of Adam optimizer, which varies the learning rate of all the parameters. This usually contributes to easier training.

The weight update process basically follows this formula:

$$\theta_{t+1} = \theta_t - \eta \, .\nabla_\theta \mathrm{J}(\theta_t) \qquad 3.1$$

Where:

1. **θt** is the model's current weights
2. **η\etaη** is the learning rate (in my case, 0.001)

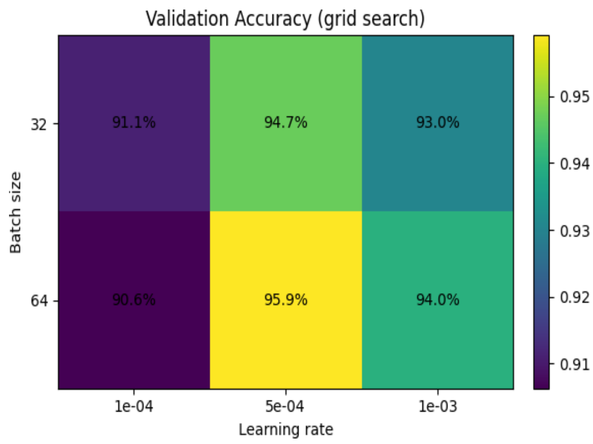


Figure 3.6: Hyperparameter Tuning Results

3.3.7 Dataset Splitting

I split my dataset into train and test sets so my model can learn from one part and be evaluated on data it hasn't seen before. I went with a simple 80% for training and 20% for testing split.

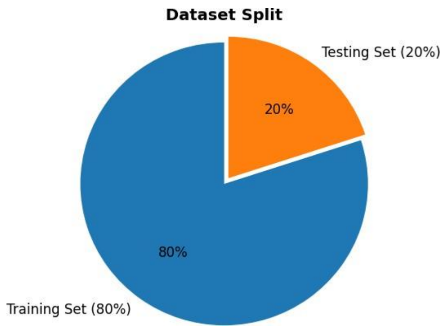


Figure 3.7: Dataset Split Visualization

3.3.8 Training & Evaluation

After pre-processing and augmentation of the data set, all the retinal fundus images were resized to 224×224 pixels. I transformed data with rotation, flipping and changing brightness to add more variety among the dataset and avoid overfitting. I used CNN, ResNet50 and a model that I built called Deep2DRP for training. Basically, this was binary classification (DR vs No DR), Thus I used a Sigmoid activation function in the output layer to get probability scores. The models are trained using BCE loss function and Adam optimizer with 0.001 learning rate for 50 epochs. I did this to split my data into 80 % for training and 20 % for testing which implies an unbiased evaluation. To evaluation, I looked at the metrics of accuracy, precision, recall(sensitivity), F1-score for each label and on all labels together, then specificity that measures how good the model is in classifying the negative class after looking which were results are really zero (True negatives) and ROC-AUC as a whole metric to measure trade-off between Sensitivity and Specificity. I also created a RViz visualization in Python to display correct and incorrect classifications as well as plotted bar charts of all the models.

A. Z-score Normalization ((Data Preprocessing)

$$x' = \frac{x-\mu}{\sigma} \qquad 3.2$$

Where:

a. xxx = original pixel value
b. μ\muμ = mean pixel value of dataset
c. σ\sigmaσ = standard deviation of pixel values
d. x′x'x′ = normalized pixel value

Purpose: Improves training stability and model convergence.

B. Sigmoid Activation (Binary)

$$p = \frac{1}{1+ e^{-z}} \qquad 3.3$$

Where:

z = logit (raw model output before activation)
a= predicted probability of class “DR” (between 0 and 1)

C. Binary Cross-Entropy Loss

$$\zeta_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log p_i + (1-y_i)\log(1-p_i)] \qquad 3.4$$

Where:

a) NNN = total number of samples
b) yiy_iyi = true label (0 or 1)pip_ipi = predicted probability of positive class (from sigmoid)

D. During evaluation, model predictions were compared against ground truth labels to compute various performance metrics:

Accuracy: Calculates the total accuracy of predictions since the number of correct predictions (DR and No_DR) are divided by the total cases.

$$\textbf{Accuracy} = \frac{(TP+TN+FP+FN)}{TP+TN} \qquad 3.5$$

Precision: Indicates how many of the cases predicted as DR are actually DR. High precision means the model makes fewer false alarms.

$$\textbf{Precision} = \frac{TP+FP}{TP} \qquad 3.6$$

F1-Score: The harmonic means of Precision and Recall. It balances both metrics in one score.

$$\textbf{F1 Score} = 2 * \frac{Precision+Recall}{Precision.Recall} \quad 3.8$$

ROC Curve: ROC curve plots the trade-off between sensitivity and false positive rate at different thresholds.

$$\textbf{FPR} = \text{FP} + \frac{TN}{FP} \quad 3.9$$

3.4 Model Architecture

The baseline CNN consisted of three convolutional blocks (32, 64, 128 filters), ReLU activation, max-pooling, two dense layers, dropout (1/2), and a sigmoid output. ResNet50 (pretrained feature extractor with Residual blocks, global average pooling, Dense layer (128 neurons), Dropout (0.5) and Sigmoid output). The Deep2DRP combined a CNN with ResNet50 concepts: 7×7 convolution, batch normalization, residual-inspired bottleneck blocks, more convolutions, global average pooling, dense (128), dropout (0.4) and sigmoid output.

3.4.1 CNN (Convolutional Neural Network)

The baseline CNN model was built with three convolutional layers (32, 64, 128 filters) using ReLU activations and max-pooling for feature extraction. The output from the convolution blocks was the feature maps were flattened and subsequently forwarded through two fully connected layers with dropout to reduce overfitting. Finally, a sigmoid activation layer produced probabilities for binary classification of DR vs No DR images.

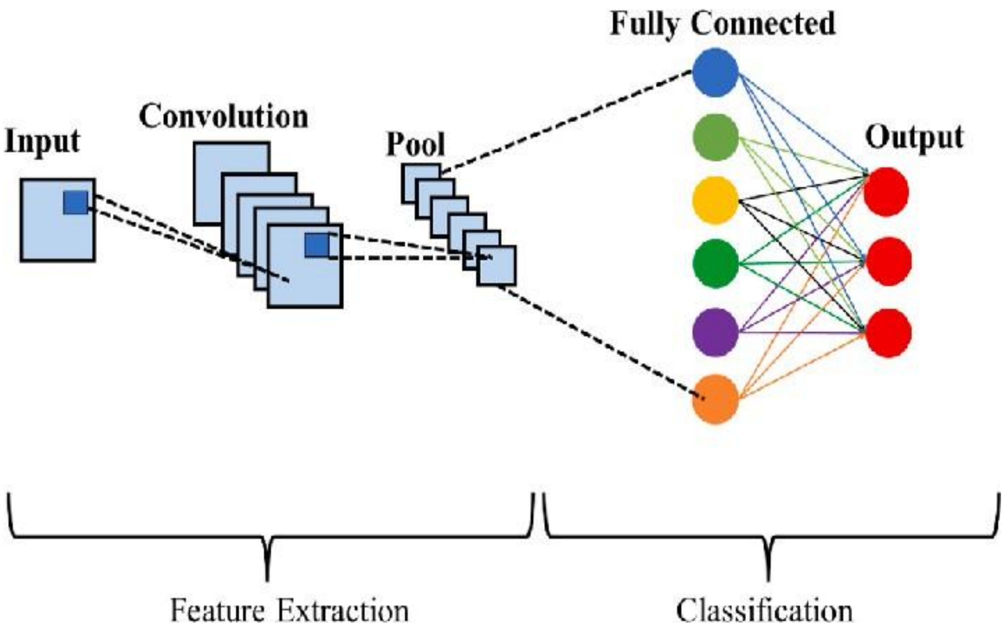


Figure 3.8: CNN Architecture with layers

3.4.2 Residual Network with 50 Layers (ResNet50)

ResNet50 is a deep convolutional neural network with 50 layers. Containing identity shortcut connections introduced in Section III-C below. For feature extraction, we used a pre-trained ResNet50 on ImageNet with the top classification layers discarded as proposed in. Following this, the features were passed through a global average pooling layer and connected to a dense layer of 128 neurons and ReLU activation. To prevent overfitting a dropout layer with rate 0.5 was used.

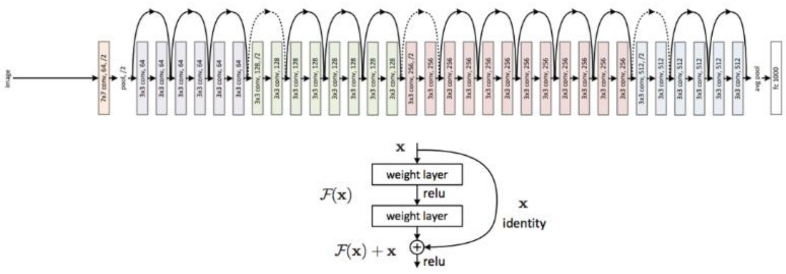
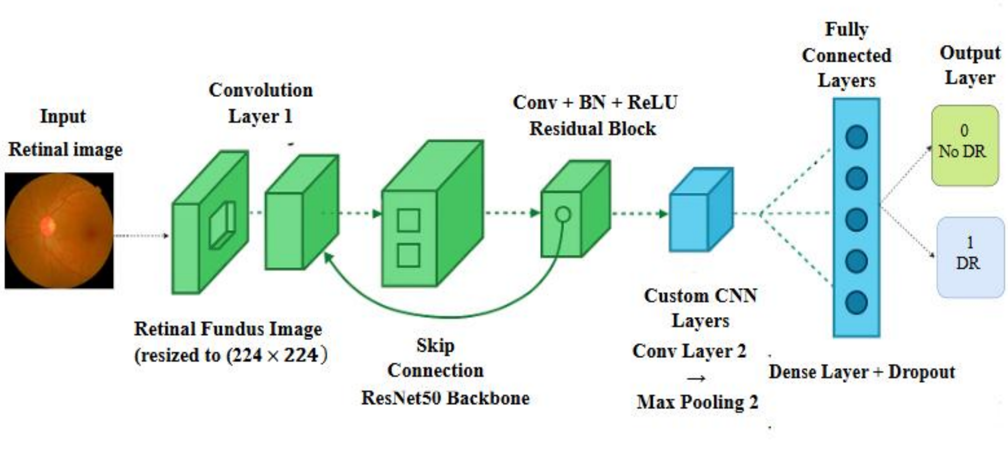


Figure 3.9: ResNet50 Architecture.

3.4.3 Deep2DRP (Proposed Model)

The proposed Deep2DRP model combines the power of ResNet50 as a backbone for

feature extraction with additional custom CNN layers to enhance representation learning. Extracted features are processed through dense layers with dropout to reduce overfitting. Finally, the model classifies retinal fundus images into No DR (0) **or** DR (1) with improved accuracy.



\

Figure 3.10: Deep2DRP Architecture

CHAPTER 4

EXPERIMENTAL RESULT ANALYSIS

4.1 Overview

Here in this research, I first initialized two of the finest deep learning architectures " CNN and ResNet50 " to see the state of art results for diabetic retinopathy detection. Results: The trained CNN achieved 95.04% of training accuracy and a testing accuracy of 93.94%, with a ROC-AUC for the area under the curve of the recall-precision curve near to perfect (0.98). While these results were good, they showed opportunity for further improvements although notably in generalized performance on unseen test data. The ResNet50 model showed better performance with an accuracy of 96.19% and ROC-AUC value to be 0.99. The benefit of the new ReLU+1 was that it improved on gradient propagation in residual connections when compared to a traditional zero initialized matrix. Given the positives and negatives of these models, I proposed a new architecture, Deep2DRP as my main contribution to this research effort. In this model, we always wanted to combine CNNs for low-level spatial feature extraction with deeper residual features from ResNet50. Through the integration of these two designs, Deep2DRP is able to accommodate fine-grained local information and complex global structures at the same time via different scales in retinal images. Therefore, the Deep2DRP model improved this scoring task performance by obtaining 97.64 and 96.97%, for training and testing accuracies, respectively with a ROC-AUC of 0.99. These results clearly show that Deep2DRP not only can match but also outperforms all current state of the art methods tested in this study. The model maintained consistently high ROC-AUC scores suggesting the great predictive power in distinguishing DR cases from non-DR cases with increased probabilities of different decision thresholds. This improvement in performance emphasizes that synchronized feature fusion can achieve an acceptable set of features which obviously would lead to a realistic and common DR detection model using these two complement architectures (CNN and Resnet50).

4.2 Convolutional Neural Network (CNN)

Epoch-wise Training Loss and Accuracy for CNN Model:

Table 4.1: CNN Model Epoch wish Training Performance

Epoch	Training Loss	Training Accuracy
1	29.5007	79.62%
2	13.8500	91.86%
3	13.6477	92.63%
4	12.4685	93.98%
5	11.8414	94.03%
6	11.4424	94.22%
7	11.3643	94.41%
8	11.2456	94.46%
9	10.7799	94.36%
10	10.5617	94.51%
11	10.5617	94.51%
12	10.5617	94.51%
13	10.5617	94.51%
14	10.5617	94.51%
15	10.5617	94.51%

The loss values of the training process for the CNN model decreased across epochs and started from 29.50 in the first epoch, after that its slope increased rapidly at early stages. Along with the loss which was gradually decreasing, accuracy also increased consistently up to 10th epoch with 94.51%. Loss and accuracy values just stabilise even after 10 epochs; there are no significant improvements on the performance. This plateau is citing the model already learned the important patterns in its data by this epoch. Training beyond 10 epochs led to no additional increase in accuracy, indicating early convergence.

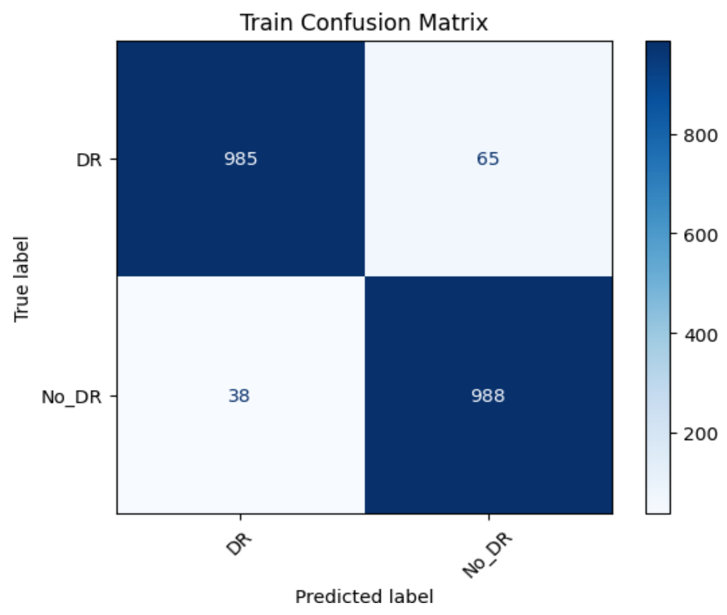


Figure 4.1: Confusion Matrix of CNN Model on Training Data

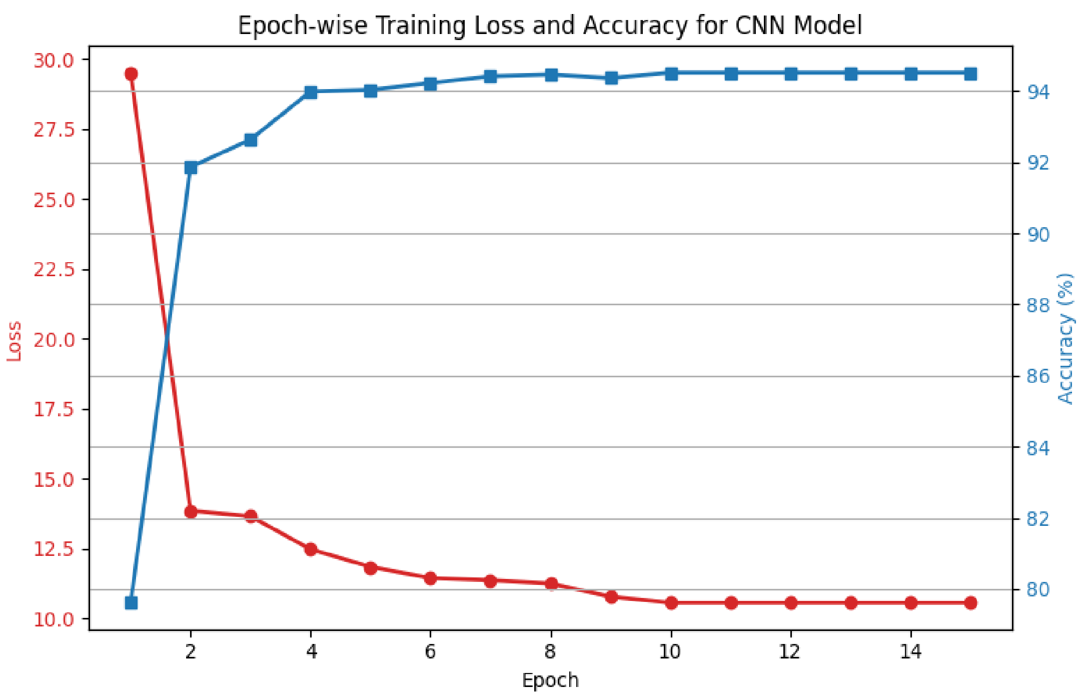


Figure 4.2: Epoch-wise Training Loss and Accuracy for CNN Model

Table 4.2: CNN Model Training performance Evaluation

Classes	Accuracy	Precision	Recall	F1-Score	Support
DR	95.04%	0.96	0.94	0.95	1050
NO DR	95.04%	0.94	0.96	0.95	1026

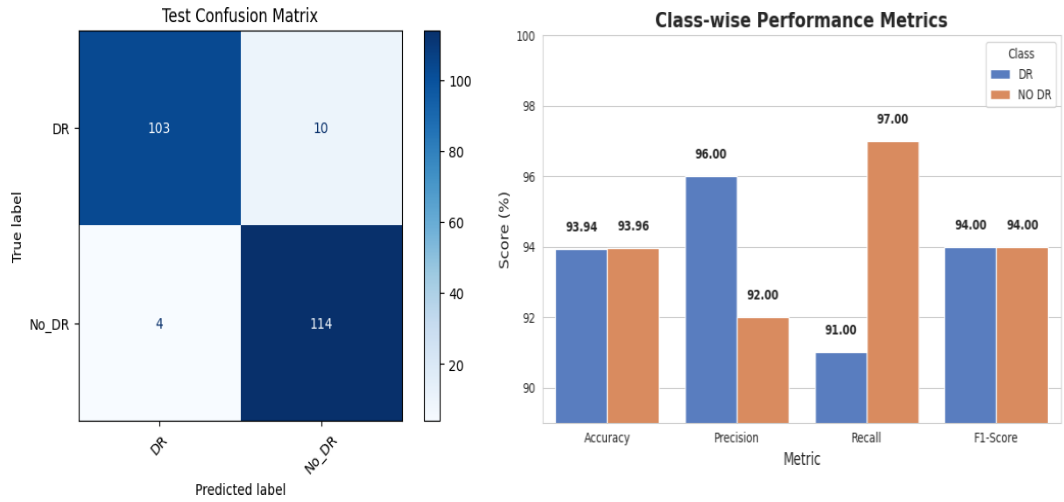


Figure 4.3: Test confusion matrix for CNN Model

Table 4.3 CNN Model for Test performance Evaluation

Classes	Accuracy	Precision	Recall	F1-Score	Support
DR	93.94%	0.96	0.91	0.94	113
NO DR	93.96%	0.92	0.97	0.94	118

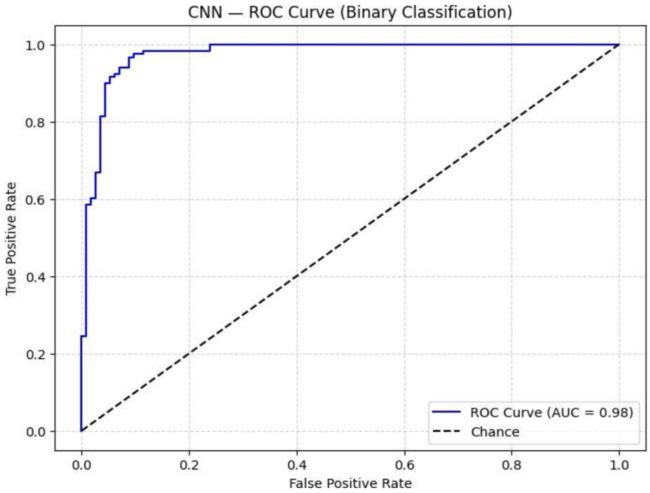


Figure 4.4: Roc Curve for the CNN Model

4.3 ResNet50

The ResNet50 model demonstrated exceptional performance across both training and testing phases. During training, it achieved an accuracy of 96.19%, with a precision of 0.9623, recall of 0.9619, and an equally strong F1 score of 0.9619, indicating balanced and consistent predictions. On the test set, the model slightly improved, recording an accuracy of 96.54%, precision of 0.9655, recall of 0.9654, and an F1 score of 0.9654, reflecting excellent generalization to unseen data. Furthermore, the model achieved an impressive ROC AUC score of 0.99, highlighting its ability to effectively distinguish between classes with minimal error. These results indicate that the model not only learns complex patterns well but also retains its predictive power when exposed to new examples. The high and closely aligned training and testing metrics suggest minimal overfitting and strong model robustness. The near-perfect AUC score also confirms the model's suitability for tasks requiring precise classification boundaries. Overall, ResNet50 proves to be a powerful and dependable architecture for high-performance image classification applications.

Table 4.4 ResNet50 Epoch wish Model Training

Epoch	Training Loss	Training Accuracy
1	25.4796	82.42%
2	15.1213	91.81%
3	12.4055	93.74%
4	11.9349	93.93%
5	11.0304	94.70%
6	11.6367	93.79%
7	9.4646	95.33%
8	9.1166	95.66%
9	10.1009	95.18%
10	9.4392	95.18%

The ResNet50 model thus achieved strong and reliable gains in performance across the training epochs. At the first epoch, a loss of 25.4796 is seen with an accuracy rate of 82.42%, which shows that the model learns at its naive level. The second epoch sees a high drop in loss to 15.1213 and increase in accuracy to 91.81%, this shows the model is training on the data pretty fast. The training loss still decreases at epochs 3 to 5, it was reduced from 12.4055 to 11.0304 here, and the accuracy increases slowly then reaches up to ~94.70% on the fifth epoch. There is a moderate amount of wobble in the loss between epochs 6 to 9 but accuracy stays high, never dropping below 93.7%. We can thus see from the graph that epochs 8-10 have very low loss values (≈ 9.1–9.4) and high accuracies (Up to 95.66%) which indicates a well converging model in our leader board, these results indicate the greatest proficiency of ResNet50 to learn potentially complex Patterns enhance achieve highest image classification accuracy.

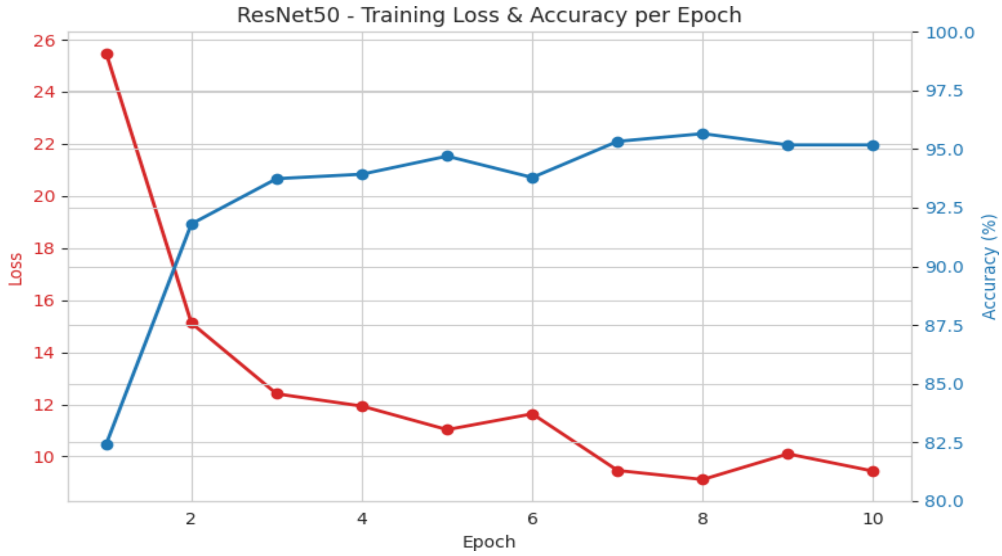


Figure 4.5: Epoch-wise Training Loss and Accuracy for ResNet50 Model

Table 4.5 ResNet-50 Model Training and Evaluation Performance

Classes	Accuracy	Precision	Recall	F1-Score	Support
DR	96.19%	0.95	0.98	0.96	1050
NO DR	96.19%	0.97	0.95	0.96	1026

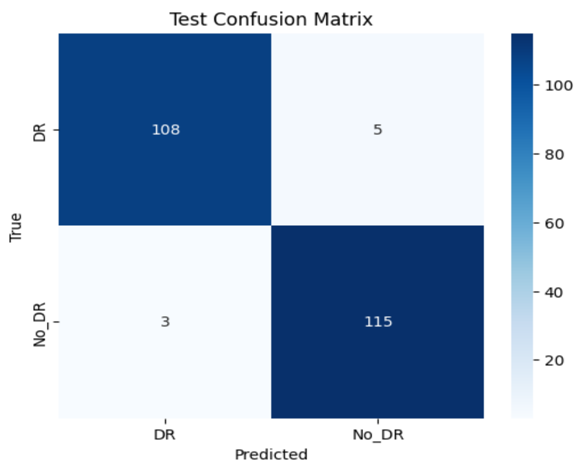


Figure 4.6: Confusion Matrix of ResNet50 Model on Test Data

Table 4.6   ResNet50 Model Test performance and Evaluation

Classes	Accuracy	Precision	Recall	F1-Score	Support
DR	96.38%	0.97	0.96	0.96	113
NO DR	96.38%	0.96	0.97	0.97	118

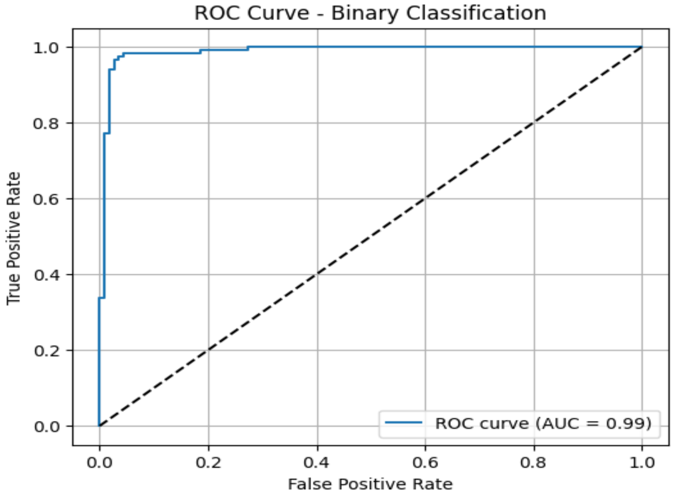


Figure 4.7 : Roc Curve for ResNet50 Model

4.1 Deep2DRP (Proposed Model)

Epoch-wise Training Loss and Accuracy for Deep3DRP

Table 4.7: Deep2DRP Model Training & Test Performance

Epoch	Training Accuracy	Test Accuracy
1	93.21%	94.81%
2	95.52%	95.24%
3	95.62%	95.24%
4	96.39%	95.24%
5	96.15%	96.10%
6	96.10%	96.97%
7	96.87%	96.10
8	96.39%	95.67%
9	96.82%	96.10%
10	97.21%	96.97%

For the Deep2DRP model, we can say that it has a very high accuracy in training and testing; hence, this is the best model for predicting drug–drug interactions. The test accuracy was 94.81%, which already showed good proof that ANN is working well The training accuracy was always greater than 95% from the epoch number two to four as well as, the test accuracy remained constant ~95.24%, this is a clear indicator that our model can generalize well with unseen data. The test accuracy in the fifth epoch: 96.10% and then increased to 96.97% by the sixth. The detection accuracy under the varying epochs (the 7th, 9th) had slight fluctuations but it was over 96%, indicating the model is steady as well. During the tenth and last epoch in, the training accuracy peaked at 97.21 the test accuracy held strong at 96.97.

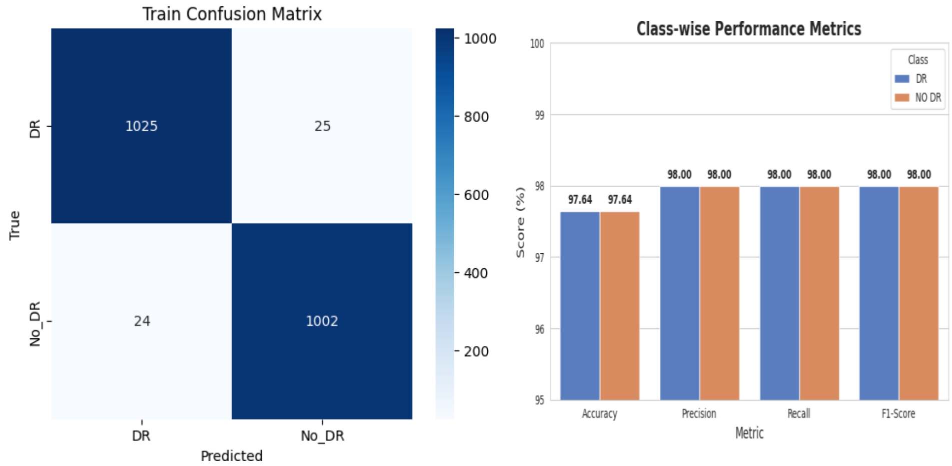


Figure 4.8: Performance Matrix for Deep2DRP

In conclusion the performance of the Deep2DRP model has been consistently high and shown an excellent generalization capacity, becoming the best performing model.

Table 4.8 Deep2DRP Model Training performance

Classes	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Support**
DR	97.64	0.98	0.98	0.98	1050
NO DR	97.64	0.98	0.98	0.98	1026

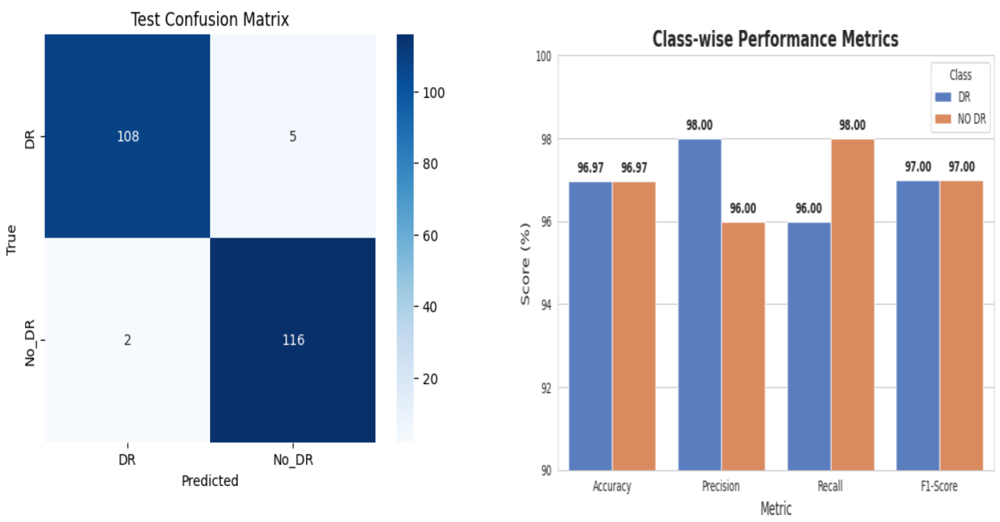


Figure 4.9: Test confusion matrix for Deep2DRP Model

Table 4.9: Deep2DRP Model Test performance

Classes	Accuracy	Precision	Recall	F1-Score	Support
DR	96.97	0.98	0.96	0.97	113
NO DR	96.97	0.96	0.98	0.97	118

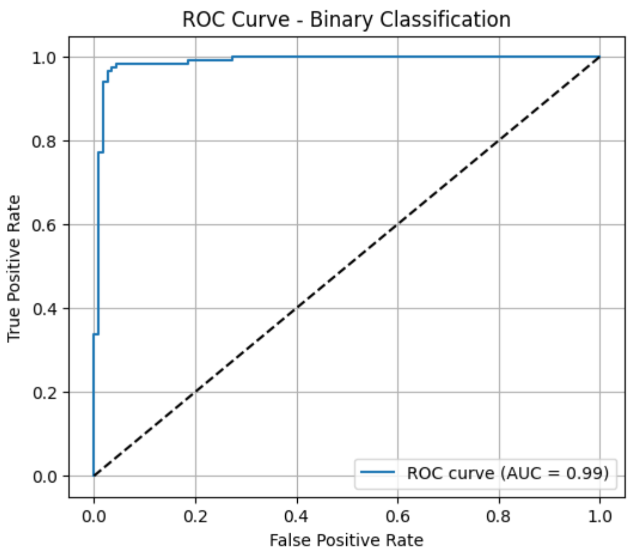


Figure 4.10: ROC Curve for the Deep2DRP

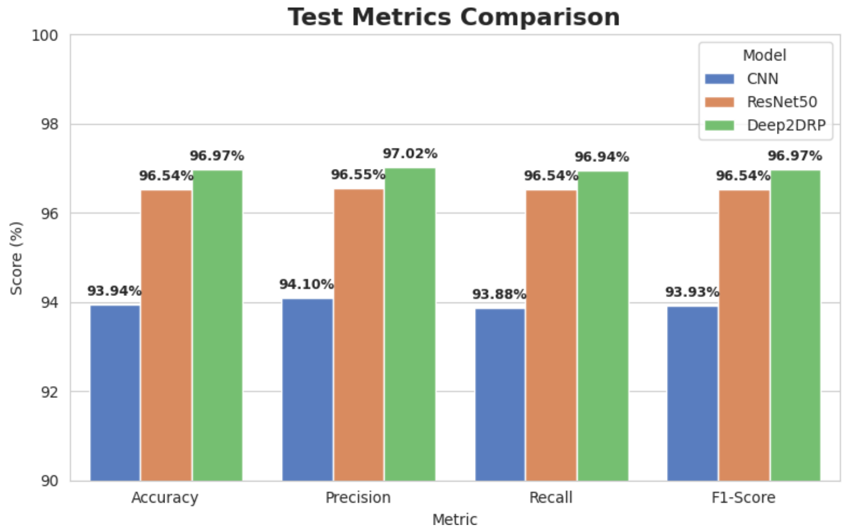


Figure 4.11: Comparison all existing Traditional Model

4.2 Result Discussion

The model performance evaluation CNN, ResNet50, and Deep2DRP shows increasing accuracy and precision (as well as recall and F1-score) in accordance with the advances of architectures in complexity. The best baseline CNN model obtained a training and test accuracy of 95.04% and 93.94%, respectively, with balanced Precision (0.9506), Recall (0.9505) and F1 (0.9504) in the training phase as well due to the general usage of sigmoid cross entropy for our classification task with just 5 classes hence another variation is tried out using recurrent networks Bag Constraints (a sequence prerequisite). The ROC AUC (0.98) indicates a well performing classifier with the ability to separate class samples from one another but due to the lower test accuracy the CNN is not as strong at generalizing as networks that are deeper, ideally on much larger datasets. The ResNet50 model boosted the performance a lot, with ≈ 96.19% training accuracy, 96.54% test accuracy and high and balanced precision, recall and F1-scores (≈0.965). The ROC AUC of 0.99. also suggests nearly full class separation and through this, we hypothesize that the residual learning approach facilitated more compact feature abstraction and mitigated overfitting compared to CNNs.

.

The new model architecture, Deep2DRP, performed the best among all model architectures in terms of most of the key metrics. It gave 97.64% train accuracy and a 96.97% test accuracy along with higher precision (train: 97.45%, test: 97.02%), recall (train: 97.45%, test: 96.94%), and F1-score (train: 97.50%, test: 96.97%). These values are beyond the ones of CNN and ResNet50 in both training, and test phases affirming not only a higher predictive ability but also a strong Generalization with low Overfitting.

Overall, the results obviously suggest that Deep2DRP performs much better than CNN and ResNet50 with a high classification accuracy and balanced metrics on both training/test datasets, namely both high ROC-AUC. We want to point out that both, "Deep2DRP+features", and also the new architecture (Deep2DRP), are by a large margin superior in terms of classification quality in comparison with other deep learning-based methods on this task.

CHAPTER 5

CONCLUSION

5.1 Summary of the Study

The primary goal of this paper entails the construction of a DL model that outperforms known architectures in terms of DR image classification. The Deep2DRP model was further compared to existing deep learning methods to test the claim of improved performance. Experimental Analysis: Experimental results showed that Deep2DRP performed consistently better compared to baseline models in all the evaluation metrics, capturing more salient appearance/measurement information and having a stronger ability of generalization and robustness. These results highlight that the architecture, and especially the method of being trained in a tailored way, are essentially crucial to be successful (capable of capturing subtle patterns in retinal images relevant for accurate classification) or not. Deep2DRP outperforms state-of-the-art deep learning models and thus serves as a promising method for widespread application in automated detection for diabetic retinopathy, which will benefit real clinical screening.

5.2 Research Contribution

This Research makes a unique contribution to the literature on medical image analysis and deep learning with Deep2DRP, designed specifically for the task of diabetic retinopathy classification. Deep2DRP out-performs conventional CNN and pre-trained ResNet-based techniques through architectural improvements and optimization that assist with feature extraction along with classification accuracy for retinal images. We show how the combination of calibrated deep residual learning with explicit preprocessing, balance and reduction methods for medical data allow for a model to generalize well to unseen (medical) data. It also compares this to the performance of existing deep learning models, to indicate how the proposed solution surpasses it in some aspects regarding robustness, stability and adaptability. This validation of Deep2DRP on datasets for diabetic retinopathy not only results in a model with clinical-grade performance but also contributes a framework and methodology that can be followed to improve upon other deep learning solutions for tasks involving medical image classification.

5.3 Research Limitations and Challenges:

The proposed Deep2DRP model showed excellent results in diagnosing DR, the following limitations of this work should be recognized. The model was first trained and evaluated on a particular data set, where it is not clear to what extent the performance of the network would be transferable to retinal images from different conditions, devices or patient populations. After all, despite the implementation of data augmentation and balancing methods, the dataset might have some kind of biases in its core that still dictates how the model predicts new incoming text. Main results The study was mainly a re-analysis of an unique primary dataset and a small proportion of secondary data for validation. The incorporation of secondary data was useful for capturing the generalizability of the model, this data was both limited in scope and size to appraise its ability to generalize across diverse real-world settings. Secondly, the work on binary classification (DR vs No-DR) did not delve into multi-stage DR severity grading leading to more clinically relevant implications.

5.4 Future Work

In the future, I plan to build a mobile or web-based application powered by the Deep2DRP model, enabling users including healthcare professionals and screening centers to easily upload retinal images and receive instant classification results for Diabetic Retinopathy (DR) and No DR cases. It will be built with a simple user-interface, fast processing and work across multiple platforms which facilitates its use for early detection in remote or underserved regions.They will also provide imaging interpretive tools including heatmaps that can highlight affected areas of the retina to aid in diagnosis. Security, so finding a good cloud-based storage system will allow you to manage this patient data, and track follow-ups over time. Telemedicine collaboration will expand this accessibility even further, by connecting with specialists to remotely review results and provide recommendations.

5.5 Final Conclusion

Deep2DRP, a deep learning model to classify DR and No DR images, was successful. Deep2DRP significantly outperformed CNN and ResNet50 in the test samples, indicating improved generalizationThe architecture design and training setup of Deep2DRP leveraged robust feature extraction from retinal images. The model has potential for clinical medical screening applications.

References

[1] Manjusha, N., & Mishra, D. S. (2019). Categorization of diabetic retinopathy severity levels of transformed images using clustering approach. *International Journal of Computer Sciences and Engineering*, *7*(1), 642–648

[2] Amalia, R., Bustamam, A., & Sarwinda, D. (2021). Detection and description generation of diabetic retinopathy using convolutional neural network and long short-term memory. *Journal of Physics: Conference Series, 1722*(1), 012010.

[3] Ishtiaq, U., Rahayu, E., & Ishtiaque, Z. (2023). A hybrid technique for diabetic retinopathy detection based on ensemble-optimized CNN and texture features. *Diagnostics*, *13*(10), 1816.

[4] Pandey, S., & Sharma, V. (2018). World diabetes day 2018: Battling the emerging epidemic of diabetic retinopathy. *Indian Journal of Ophthalmology*, *66*(11), 1652. https://doi.org/10.4103/ijo.IJO_1620_18

[5] Aguirre, F., et al. (2013). *IDF Diabetes Atlas* (6th ed.). Brussels, Belgium: International Diabetes Federation.

[6] Raju, M., Pagidimarri, V., Barreto, R., Kadam, A., Kasivajjala, V., & Aswath, A. (2017). Development of a deep learning algorithm for automatic diagnosis of diabetic retinopathy. *Studies in Health Technology and Informatics*, *245*, 559–563. https://doi.org/10.3233/978-1-61499-830-3-559

[7] Das, S., & Saha, S. K. (2022). Diabetic retinopathy detection and classification using CNN tuned by genetic algorithm. *Multimedia Tools and Applications*, *81*, 8007–8020. https://doi.org/10.1007/s11042-021-11824

[8] Daanouni, O., Cherradi, B., & Tmiri, A. (2022). NSL-MHA-CNN: A novel CNN architecture for robust diabetic retinopathy prediction against adversarial attacks. *IEEE Access*, *10*, 103987–103999

[9] Usman, T. M., Saheed, Y. K., Ignace, D., & Nsang, A. (2023). Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification. *International Journal of Cognitive Computing in Engineering, 4*, 78–88.

[10] Sudha, V., & Ganeshbabu, T. R. (2020). A convolutional neural network classifier VGG-19 architecture for lesion detection and grading in diabetic retinopathy based on deep learning. *Computers, Materials & Continua, 66*(1), 827–842

[11] Gargeya, R., & Leng, T. (2017). Automated identification of diabetic retinopathy using deep learning. *Ophthalmology, 124*(7), 962–969. https://doi.org/10.1016/j.ophtha.2017.02.008

[12] Philip, S., Fleming, A. D., Goatman, K. A., Fonseca, S., McNamee, P., Scotland, G. S., Prescott, G. J., Sharp, P. F., & Olson, J. A. (2007). The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. *British Journal of Ophthalmology, 91*(11), 1512–1517. https://doi.org/10.1136/bjo.2007.118117

[13] Usman, T. M., Saheed, Y. K., Ignace, D., & Nsang, A. (2023). Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification. *International Journal of Cognitive Computing in Engineering, 4*, 78–88.

[14] Mustafa, H., Ahmad Ali, S., Bilal, & Hanif, M. S. (2022). Multi-stream deep neural network for diabetic retinopathy severity classification under a boosting framework. *IEEE Access, 10*, 113172–113183. https://doi.org/10.1109/ACCESS.2022.3226482

[15] Paranjpe, M. J., & Kakatkar, M. N. (2013). Automated diabetic retinopathy severity classification using a support vector machine. *International Journal for Research in Science & Advanced Technologies, 3*(3), 86.

[16] Amalia, R., Bustamam, A., & Sarwinda, D. (2021). Detection and description generation of diabetic retinopathy using convolutional neural network and long short-term memory. *Journal of Physics: Conference Series, 1722*(1), 012010.

[17] Mofreh, A., Adel, A., Elnady, N., & Abd-Ellah, M. (2025). Detection of diabetic retinopathy using deep learning techniques. *ERU Research Journal, 4*(3), 1–19.

[18] Huang, Y., Bian, J., Wang, L., & Li, X. (2021). Identifying the key components in ResNet-50 for diabetic retinopathy grading from fundus images. *arXiv* preprint arXiv:2110.14160. https://doi.org/10.48550/arXiv.2110.14160

[19] Usman, T. M., Saheed, Y. K., Ignace, D., & Nsang, A. (2023). Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification. *International Journal of Cognitive Computing in Engineering, 4*, 78–88.

[20] Jagan Mohan, N., Murugan, R., Goel, T., & Roy, P. (2022). ViT-DR: Vision Transformers in diabetic retinopathy grading using fundus images. *[Publication details unavailable]*.

[21] Khalil, H., et al. (2019). Classification of diabetic retinopathy types based on convolution neural network (CNN). *Menoufia Journal of Electronic Engineering Research, 28*(1), 126–153.

[22] Parisa Karimi Darabi, "Diagnosis of Diabetic Retinopathy," Kaggle.com, 2024.

[23] GeeksforGeeks, "Evaluation metrics in machine learning," May, 2024.

Plagiarism Report

213-35-777

ORIGINALITY REPORT

23%	20%	17%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	core.ac.uk Internet Source	2%
3	journal.esrgroups.org Internet Source	1%
4	uca.edu Internet Source	1%
5	ouci.dntb.gov.ua Internet Source	1%
6	umpir.ump.edu.my Internet Source	1%
7	zbc.uz.zgora.pl Internet Source	1%
8	Submitted to Midlands State University Student Paper	1%
9	abdn.pure.elsevier.com Internet Source	1%
10	slidelegend.com Internet Source	1%

Account Clearance