Interpretability of Diabetic Retinopathy images for EfficientNet

Project Report

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering

By

Yashesh Patel

(22MCED13)

Guided By

Dr. Rupal A. Kapdi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2024

Interpretability of Diabetic Retinopathy images for EfficientNet

Submitted By

Patel Yashesh Ruchir

22mced13

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

27 May 2024

Interpretability of Diabetic Retinopathy images for EfficientNet

Major Project - II

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering (Data Science)

Submitted By

Patel Yashesh Ruchir

(22mced13)

Guided By

Dr. Rupal A. Kapdi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

27 May 2024

Certificate

This is to certify that the major project entitled **"Interpretability of Diabetic Retinopathy images for EfficientNet"** submitted by **Patel Yashesh Ruchir (22mced13)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering (Data Science) of Nirma University, Ahmedabad, is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this major project part-II, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Dr. Rupal A. Kapdi
Guide & Assistant Professor,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

23/5/2024
Dr. Swati Jain
Associate Professor,
Coordinator M.Tech - CSE (Data Science)
Institute of Technology,
Nirma University, Ahmedabad

Dr. Madhuri Bhavsar
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr Himanshu Soni
Director,
School of Technology,
Nirma University, Ahmedabad

Statement of Originality

I, **Patel Yashesh Ruchir**, Roll. No. **22mced13**, give undertaking that the Major Project entitled **"Interpretability of Diabetic Retinopathy images for Efficient-Net"** submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering (Data Science)** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Yashesh..

Signature of Student

Date:

Place:

Endorsed by

Dr. Rupal A. Kapdi

(Signature of Guide)

Acknowledgements

I wish to convey my heartfelt gratitude to my internal guide, **Dr. Rupal A. Kapdi**, who holds the position of Assistant Professor in the Computer Engineering Department at the Institute of Technology, Nirma University, Ahmedabad. I appreciate her valuable guidance, enthusiastic approach, and continuous encouragement during the entire duration of the Project and Thesis work. Her counsel, recommendations, and constructive feedback have significantly contributed to the development and refinement of the project.

I take great satisfaction in expressing my gratitude to **Dr. Madhuri Bhavsar**, who serves as the Head of the Computer Science and Engineering Department at the Institute of Technology, Nirma University, Ahmedabad. I appreciate her generous assistance and the provision of essential infrastructure, contributing to a conducive research environment.

I extend sincere appreciation to **Dr. Himanshu Soni**, the respected Director of the School of Technology at Nirma University, Ahmedabad, for the invaluable motivation he has provided during the entire duration of this endeavor.

I also extend my appreciation to the institution and all faculty members of the Computer Engineering Department at Nirma University, Ahmedabad, for their exceptional focus and valuable suggestions in support of the project work.

Patel Yashesh Ruchir

22mced13

Abstract

Diabetic Retinopathy (DR) emerges as a consequence of either type-1 or type-2 diabetes, and it is crucial to detect complications early to prevent visual issues such as retinal detachment, vitreous hemorrhage, and glaucoma. The interpretability of automated classifiers in medical diagnoses, like diabetic retinopathy, is of paramount importance. The primary challenge lies in extracting meaningful insights from these classifiers, given their inherent complexities. In recent years, considerable efforts have been devoted to transforming deep learning classifiers from opaque statistical black boxes with high confidence to models that are self-explanatory. A persisting concern revolves around the effective preprocessing of data before classification. Despite the proven efficacy of supervised machine learning schemes in application, challenges persist in dealing with data redundancy, feature selection, and human expert interference. Consequently, we propose a combinatorial deep learning approach for interpreting diabetic retinopathy (DR) detection. Our method integrates the Shapley Additive Explainability (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) techniques to analyze the output of deep learning models effectively. The outcomes of our experiments demonstrate that our proposed approach surpasses existing schemes in the accurate detection of DR.

Abbreviations

DR	Diabetic Retinopathy
SHAP	SHapley Additive exPlanations
LIME	Local Interpretable Model-agnostic Explanations
SVM	Support Vector Machine
PNN	Probabilistic Neural Network
KNN	K-Nearest Neighbor
CNN	Convolutional Neural Network
PSO	Particle Swarm Optimization
DRE	Diabetic Retinopathy Estimation
DRM	Diabetic Retinopathy Module
ReLU	Rectified Linear activation Unit
GLCM	Gray Level Cooccurrence Matrix
HR	Hazard Ratio
ESRD	End-Stage Renal Disease
HSI	Hue, Saturation, Intensity
CAM	Class Activation Maps
PDR	Proliferative Diabetic Retinopathy
NPDR	Non-Proliferative Diabetic Retinopathy
IRMA	Intra-Retinal Microvascular Abnormalities
DME	Diabetic Macular Edema
ETDRS	Early Treatment Diabetic Retinopathy Study
ANN	Artificial Neural Network
DNN	Deep (Feed-Forward) Neural Network
PSO	Particle Swarm Optimization

–

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 What is Diabetic Retinopathy?

Diabetic Retinopathy (DR) is a medical situation in which human having diabetes have impact on either eye or both of their eyes are affected with vision loss and blindness. It occurs because of the imbalance of blood sugar level in such persons and this damages the capillaries of the photosensitive skin tissue at the end of human eye ball known as retina. All people should note that not every diabetic person develops diabetic retinopathy only some unfortunate ones but still the risk remains high and also that some of the other mild symptoms are blurry vision, having hard time identifying colours and floating spots in vision.

Diabetes can often lead to malfunctioning of capillaries, because of the cells of the capillary becomes impermeable to insulin and this with time leads to depositing glucose which in return causes blockages, which are responsible for supporting retina with supply of fresh blood, water and other bodily fluids instead end up damaging retina with filling up the eye ball, or retina in this case, with blood and fluids [3] because of rupturing of small blood vessels present in retina because of developing of too much pressure in tiny area and finally the blood and fluids burst open through the capillaries into the eye ball. Now as a consequence such an eye develops optical characteristics known as lesions which are blood vessel area, cotton wool spots, hard exudates, haemorrhages and microaneurysms.

Blood vessel area as an eye lesion is rupturing of tiny capillaries in the conjunctiva and also in vitreous humor leading to contamination inside the eye ball. Cotton wool spots are cloudy white velvety smudges in our eye ball which is regarded as irregular diagnostic result of an ophthalmoscopic medical test and cotton wool spots can sometimes be yellowish-white or greyish-white in appearance and irrespective of their colour they are also known as soft exudates [4]. Hard Exudates is fluid discharge from the fluid which is formed when a capillary is recovering from the damage with time these leftover some of these fluids and its discharge become solid and hard. Hard exudates are made up of lipid and proteinaceous material such as fibrinogen and albumin that seep out from the damaged capillaries of retina and these exudates gets settled mainly in the exterior plexiform surface of the retina. On preforming Optical Coherence Tomography (OCT) we get to know that hard exudates are round lesions which are indeed responsible for our eye having hyperreflective foci (HRF) spots on the retina leading to develop dibaetic macular edema in our case.

Haemorrhages in eye means flooding of vitreous humor with blooding flowing out of the raptured capillaries of retina blocking the light rays falling on the retina making it hard for the eye to capture the view of the outside world which is medically termed as intraocular haemorrhage. Microaneurysms are minute space of lumps in capillaries present in eye which ideally indicates the exacerbating diabetic retinopathy.

1.2 Types of DR images used

The types of diabetic retinopathy used by me are based on what class they belong to going from class 1 to class 4 we have mild DR, moderate DR, severe DR and proliferative DR (PDR). All these classes fall under fundus images which in medical terminology is called opthalmoscopic images. Mild, moderate and severe DR, i.e. classes 1, 2 and 3, falls under non-proliferative DR (NPDR) whereas class 4 being proliferative DR (PDR).

Class 1: Mild DR is the very first stage of diabetic retinopathy in which the eyes of patient develops a very few to several number of microaneurysms but still small amount number to be consider as alarming which is why we cannot say that for sure the diabetes patient is suffering from diabetic retinopathy. In this stage person suffers from tiny bit of blurry vision due to minute lumps in macula the center most region of retina.

Class 2: Moderate DR is second stage of non-proliferative DR in which person suffers from partially dry eyes due to significant increase in microaneurysms and also now the eye have developed haemorrhages along with cotton wool spots. The eyes in such cases faces lack of blood supply because most of blood supply to the retina is cut off leading to blockages in eye.

Class 3: Severe DR is the final stage of non-proliferative DR if the case gets any worse then patient's eye enters the early stage of proliferative DR which is class 4 type of DR and either of the two class 3 or class 4 patient needs to be immediately treated with best eye surgery from an expert eye surgeon. Now the patient has started experiencing almost complete loss of vision that the vision is almost blurry with few spots clear but the patient has hard time to read the colour or text or even describe the object in front few meters.

Class 4: Proliferative DR is the final stage before permanent loss of eyesight because of most of retina is damaged so is the vitreous humor. Now since the retina tear is so high that the patient experiences fluctuating vision (moving blurry images) and also vitreous haemorrhage has been developed leading to the dark or empty spots in vision.

1.3 Model used for image classification

Efficient-Net B7 model to train it the creators of this model opted to utilize images sized at the resolution of 600 by 600 pixels on the ImageNet-1K dataset. Efficient-Net is a fully convolutional model that propounds an introduction of novel approach to scaling such that the model easily scales across all resolution, all depth and all width with absolute uniformity with the deployment of a straight-forward yet remarkably effective compound co-efficient [5]. Different approach from usual way to randomly managing these variables, the Efficient-Net scales them uniformly with the help of pre-determined set of fixed values of scaling coefficients.

When the image which is inputted is highly volumetric in such rare cases we must expand the CNN with ancillary layers that are deeper and having poise weights merged to the network so the responsive area is expanded and we need to put further courses so that our CNN model is grasp even better and reliable amount of exquisite patterns [6]. Efficient-Net also netted with better accuracy and efficiency then previous models like MobileNet and ResNet when put side to side with some other random scaling CNNs.

EfficientNet is in accordance with the onset of network that came into being because of neural architecture search with utilization of AutoML MNAS(Multi-Modal Neural Architecture Search) framework. This CNN is custom tuned for achieving highest accuracy possible that is in a case where somehow the CNN becomes resource hungry then the accuracy of the model gets affected [7]. The direct consequence of this is it takes too long to anticipate any results and also other bad outcome is unperceptive speculation time.

EfficientNet outshines all CNN models that came before it remarkably well on majority of the benchmarking datasets. One can also utilize EfficientNet to properly scale-up other CNN models [8]. This feature of EfficientNet makes it cutting edge, high end and top of the line CNN model with astonishing lesser variables and FLOPS (floating-point operations per second) for ImageNet along with other frequently utilized readily available datasets for performing transfer learning.

1.4 Why it is required to interpret such a model?

The core objective behind the interpretation of such a model is to discover the least oblivious manner to explain the working of decision making process of such a model so that it is easy to be grasp by layman the importance of interpretation of models. By doing this we ensure there is fairness in the deployment of the model, that no other model is overlooked and each one is tested and the best one of them all is the only one which is used practically in real life. Hence as a result of such practice we introduce accountability in case where model fails to predict correct diagnosis in emergencies that may occur in sudden time where we may need to rely on AI to make a quick decision. Also this practice adds much appreciated transparency to the day to day working as when so if the model is starting to fail to give right predictions then the community using it will be kept in light as the daily report at the end of the day will be published directly on the public blockchain. Finally all these provide human with enough credence to start utilizing the model in practical life to resolve real-life issues which impacts a lot of people and also businesses around the world.

The inquisition of data and its features elucidates the innate traits piloting model predictions. This scrutiny accredits researchers with amplified ken of crucial aspects sculpting the decision making framework of the model, hence in doing so sped up the furtherance in feature engineering and data pre-processing procedures.

Unveiling both dissimilitude along side with impediments has expedited with the help of exegesis, sanctioning the refinement towards both the fallaciousness as well as curtailments enclosed by design augury. Experimenters guess track down occasion when is their prototype might possibly waver about its reliability in prophesying in turn grapple toward inferring of peculiar statistics, as a consequence set in motion fathomless scrutiny including rarefaction needed for their framework.

Cavernous swotting facsimile intermittently ply simultaneously inscrutable structure, shrouding its intramural working as well as offering cognition onerous. Nevertheless, construing the particular framework which is going to surface a path to the seminal research based divulgence alongside with keen acumen towards sinuous occurrence, exposing obscure arrangement moreover the correlation quiescent in the information available.

Earnest apprehension obtained from model explanations be of use to rampart the pliability and generalization of the model. Researchers can spot occasions where the model may yield to over-fitting on the training dataset or grapple to infer to novel data distributions, thereby assisting intensified advancements in model efficacy.

1.5 Related Work

The researchers, Skylar and Fang et al. [9], introduced a segment-based learning method to identify diabetic retinopathy, aiming to enhance recognition of images depicting diabetic retinopathy and associated lesions. This improvement is achieved through the mutual learning of classifiers and features from the data. To assess the segment-level Diabetic Retinopathy Estimation (DRE), a Convolutional Neural Network (CNN) was trained. Integration of all segment levels of Diabetic Retinopathy Management (DRM) is implemented to enhance the categorization of diabetic retinopathy images. The researchers assert that their algorithm's performance surpassed that of all existing schemes discussed in their article. However, it is noted that the researchers only considered scaling in their data pre-processing, and the segmentation process employed has limitations with a threshold. The proposed model for detecting diabetic retinopathy is based on feature identification in optical coherence tomography images. Automatic segmentation of retinal layers in an optical coherence tomography picture relies on gradient information. The researchers claim their algorithm efficiently detects broken edges due to the presence of blood vessels and false edges caused by image noise. Despite these claims, the interference of human experts in feature selection was not reduced, potentially increasing the error margin in the experiment's output and posing a risk of over-fitting.

The researchers, Alexa and Scarpa et al. [10], proposed a model that employs CNN for feature extraction prior to classification, aiming to enhance efficacy and diminish human expert interference in diabetic retinopathy detection. The researchers input the CNN feature extraction output into five different machine learning algorithm classifiers. According to their findings, the combination of CNN feature extraction and the J48 classifier demonstrated superior performance compared to existing traditional approaches, achieving accuracies of 99.89% and 99.59% for binary and multiclass classification, respectively. Notably, the researchers did not address data preprocessing to eliminate unnecessary and noisy data from the dataset before feature extraction and classification, potentially elevating computation costs and impacting classification outcomes. Additionally, the researchers proposed modifications to enhance their algorithm's future performance.

The researchers, Wejdan L and Shalash et al. [11], proposed model is designed to categorize diabetic retinopathy fundus pictures based on disease severity, utilizing convolutional neural networks along with appropriate pooling, Rectified Linear Activation Unit (ReLU) layers, and SoftMax to achieve a high level of accuracy. The probability score for each prediction class is calculated by the final fully connected layer of the deep learning architecture, with the class achieving the maximum probability score chosen as the predicted class. According to their experiments, their proposed algorithm achieved an accuracy of 96.6%. However, it's crucial to note that the dataset used by the researchers was limited, making it challenging to extrapolate findings to a larger dataset. Consequently, the researchers recommended a comparison of their algorithm using a more extensive database of images for a more comprehensive evaluation.

The proposed approach explores machine learning for the detection of diabetic disease using thermography images of the eye, incorporating the impact of thermal variations in eye structure abnormalities as a diagnostic imaging modality beneficial for ophthalmologists in clinical diagnosis. Thermal images undergo preprocessing, and texture characteristics based on the Gray Level Cooccurrence Matrix (GLCM) are extracted from gray images, along with statistical features from RGB and HSI images. These features are then categorized using a classifier with various feature combinations. The researchers employed a Support Vector Machine (SVM) for classification, implementing a 5-fold cross-validation scheme to enhance the algorithm's performance. According to their experiments, the algorithm achieved an accuracy of 86.22%, outperforming all traditional schemes mentioned in their related works. While the researchers, Ramendra et al. [12], did not explicitly state limitations in their work, existing research has highlighted the challenges of SVM in selecting an appropriate kernel function for experiments. In an extension of their work, the researchers proposed a deep learning approach utilizing the Lesion-aware deep learning system (RetinalNET) to detect the severity of diabetic retinopathy. Longitudinal research involved capturing bilateral fundic images of 91 Chinese type two diabetes patients with biopsy-confirmed DN who were not in the ESRD stage at the time of renal biopsy. An open framework for deep learning was employed to assess diabetic retinopathy severity using the Early Treatment Diabetic Retinopathy Severity Scale. The hazard ratio (HR) for the impact of diabetic retinopathy severity on ESRD was estimated using Cox proportional hazard models. Despite achieving better outcomes, the researchers acknowledged the potential influence of human expert interference and limitations in feature extraction on the experiment's output.

The proposed system automatically distinguishes between exudates and non-exudates regions in retinal images. The researchers incorporated the Gabor filter to preprocess their grayscale images, enhancing the visibility of lesions. Mathematical morphology was employed to segment candidate lesions. A combination of statistical and geometric features was used for feature selection. The proposed method demonstrated an accuracy of 98.58%, surpassing all existing methods mentioned in their related works. However, the researchers acknowledged limitations in the automation of diabetic retinopathy (DR) process grading. Consequently, they suggested that further investigation into the automation of DR process gradings, particularly in terms of severity, would be an intriguing topic for future study. In another study, researchers introduced a deep learning interpretable classifier for diabetic retinopathy. Their approach effectively classifies retina images into various severity levels with high performance [13]. For result interpretation, the researchers assigned a score to each point in the hidden and input space, providing an interpretable output. They further proposed a pixel-wise score propagation model for each neuron, dividing the output score into two components. To interpret the results, an ophthalmologist was engaged to identify statistical regularities aiding in the diagnosis of eye disease. Lastly, the researchers acknowledged limitations associated with retinal images containing manually marked lesions.

To identify diabetic retinopathy (DR), the researchers, Sowmya and Kulkarni et al. [14], introduced weighted class activation maps (CAMs) to visualize the location of suspected lesions. To enhance the diversity of fundus images, eight image transformation methods were introduced, and manual preprocessing of images was conducted before data augmentation. The researchers assert that their proposed model exhibits robustness and superior performance compared to existing schemes discussed in their related work. However, they acknowledge limitations, such as the potential for errors in manually preprocessing data, particularly when dealing with extensive datasets, which could impact the classification output [15]. The researchers applied filters to extract retinal vessels, employed Fuzzy C methods for exudate identification, and utilized Convex Hull to locate and remove optical disks. For the classification of fundus pictures into Normal, Non-proliferative Diabetic Retinopathy (NPDR), or Proliferative Diabetic Retinopathy

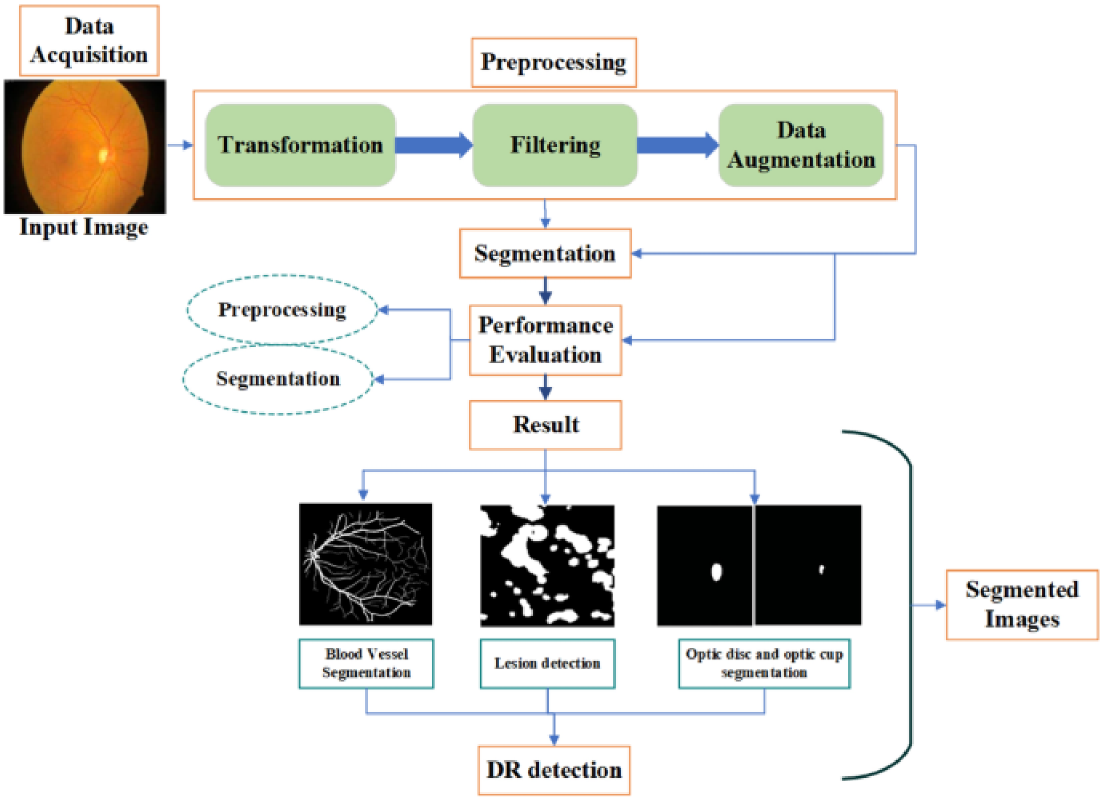

Figure 1.1: Deep learning for diabetic retinopathy [1]

(PDR), Support Vector Machines (SVM) were employed. Images underwent preprocessing before extracting features of retinal vessels, and the obtained images were trained using a two-stage support vector machine trainer. Their algorithm achieved an efficiency of 96.23%, surpassing existing systems discussed in related works, based on their experimental results. However, existing research points out the limitations of SVM for classification, particularly in dealing with large datasets and kernel selection.

The researchers, Praneeth et al. [16], proposed a CNN-based approach for diabetic retinopathy (DR) detection, addressing aspects like classification, detection, and segmentation. They incorporated transfer learning and hyper-parameter tuning to optimize their algorithm and improve its output. The Kaggle platform was utilized for model training, resulting in an accuracy of 95.68% from experimental results. Notably, the researchers did not consider data preprocessing, which could potentially enhance the algorithm's performance. A hierarchical Coarse-to-Fine Network (CF-DRNet) is suggested as an automated clinical tool for classifying five stages of DR severity grades using convolutional neural networks (CNNs). The CF-DRNet is hierarchical, significantly improving the classification performance of five-class DR grading. The Coarse Network classifies lesion characteristics into two categories: No DR and DR, with an attention gate module emphasizing critical lesion features and suppressing unnecessary background information. The Fine Network categorizes four stages of DR severity classes, including mild, moderate, severe non-proliferative DR (NPDR), and proliferative DR, derived from the Coarse Network (PDR). Five-fold cross-validation was employed to validate the experiment, demonstrating that the proposed CF-DRNet outperforms many state-of-the-art techniques using publicly accessible IDRiD and Kaggle fundus picture datasets. However, the researchers recommended developing a more complex Fine Network to further enhance classification performance, eliminating misunderstandings between the four DR severity grade levels. Additionally, future investigations will explore DR grading in conjunction with DR lesion identification to facilitate automated DR diagnosis. Despite several studies suggesting improvements in DR detection, the proposed system aims to introduce a novel feature extraction, data preprocessing, and learning methodology to address prior constraint-related issues identified in the literature.

Chapter 2

Literature Survey

2.1 Introduction

In the initial phases of diabetic retinopathy, the retina may exhibit microaneurysms, a result of the degeneration and loss of pericytes leading to the dilation of capillary walls. Rupture of the capillary or microaneurysm walls results in intraretinal hemorrhages. Non-proliferative diabetic retinopathy encompasses additional lesions such as soft and hard exudates, intraretinal microvascular abnormalities (IRMA), venous beading, and venous loops or reduplication. IRMAs manifest as large-caliber, tortuous vessels in ischemic areas and may indicate attempted vascular remodeling. The crucial distinction between non-proliferative and proliferative diabetic retinopathy lies in the presence of neovascularization, signifying the growth of new retinal vessels due to ischemia in pre-existing ones. Figure 2.1 displays various lesions observed in an illustrative fundus image of a retina.

At any point in the progression of diabetic retinopathy, diabetic macular edema (DME) may manifest as a significant endpoint, serving as the leading cause of blindness [14]. The presence of edema is characterized by abnormalities including exudates situated within one disc diameter of the foveal center, exudates within the macula, retinal thickening within one disc diameter of the foveal center, and microaneurysms or hemorrhages within one disc diameter of the foveal center.

In terms of the clinical grading protocols for diabetic retinopathy (DR), despite the widely accepted Early Treatment Diabetic Retinopathy Study (ETDRS) grading scheme

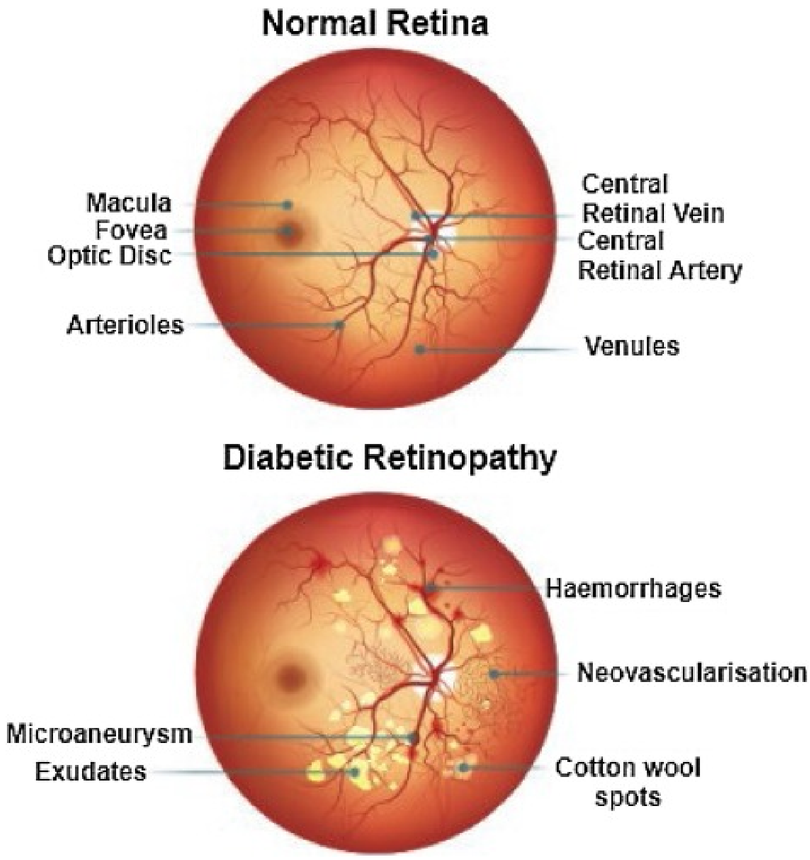

Figure 2.1: Lesions of Diabetic Retinopathy [2]

serving as the gold standard, its application in routine clinical settings has proven challenging and impractical [17]. In attempts to enhance patient screening and communication among caregivers, various alternative scales have been suggested. Despite the development of simplified diabetic retinopathy severity scales in numerous countries, the establishment of a unified international severity scale remains elusive. Addressing this, the Global Diabetic Retinopathy Project Group has put forth the International Clinical Diabetic Retinopathy Disease Severity Scale.

2.2 Techniques

Neural Networks :- Concerning the most basic form of a neural network, it pertains to an Artificial Neural Network (ANN) composed of three layers of neurons: one input layer, one hidden layer, and a final output layer. These networks, labeled as Shallow (Feed-Forward) Neural Networks, are characterized by having only one hidden layer [18]. In contrast, a Deep (Feed-Forward) Neural Network (DNN) incorporates more than two hidden layers. Each hidden and output layer comprises multiple artificial neurons, and every input node and hidden neuron node establishes connections with each neuron in the subsequent layer through connection links. It's worth noting that these networks accept a one-dimensional array as their input, making them incompatible for direct utilization with imaging data.

CNN :- Concerning Convolutional Neural Networks (CNN), which, in contrast to shallow neural networks, accept 2D arrays as input, their design draws inspiration from human vision, grounded in the fundamental mathematical operation of "convolution." A key distinction between CNNs and Deep Neural Networks (DNNs) lies in the fact that, for the latter, all neurons at a given layer contribute to computing the output of every neuron at the subsequent layer, a characteristic not shared by CNNs. Instead, CNNs employ filters or kernels to compute convolutions by sliding over a portion of the original image, generating a feature map [9]. In this approach, if the filter's size is x $\times$ x, only a window of x^2 pixels influences the computation of each unit in the next layer's feature map. This directly influences the receptive field, defined as the region in the input space affected by a particular CNN feature. Notably, the convolutional segment is often labeled "the feature extraction part" of the network, while the remainder is termed "the classification part." The former learns imaging features, subsequently condensed into a one-dimensional array and fed through the latter—essentially a Deep Neural Network—to classify the input image based on the generated features.

LIME :- Local Interpretable Model-agnostic Explanations (LIME) represents a method within machine learning that aims to furnish interpretable explanations for predictions made by intricate models. The primary objective of LIME is to enhance the interpretability of black-box models, especially when grappling with high-dimensional and intricate data.

The general workflow of LIME is outlined as follows:

1. Select an Instance: Identify a specific instance or data point for which you seek an explanation of the model's prediction.

2. Generate Perturbations: Introduce slight, random changes to the features of the chosen instance to generate a dataset of comparable instances. This dataset serves to

probe the model.

3. Predictions on Perturbed Data: Obtain predictions from the black-box model for the perturbed instances.

4. Fit an Interpretable Model: Train an interpretable model, often simpler (such as a linear model), on the perturbed instances and their corresponding model predictions. This simpler model is termed a surrogate model.

5. Interpretation: Utilize the interpretable model to comprehend how alterations in the features of perturbed instances correlate with changes in the model's predictions. This interpretation yields insights into the local behavior of the black-box model for the selected instance.

Notably, LIME is model-agnostic, implying its applicability to any machine learning model, irrespective of the underlying algorithm or architecture. It serves to bridge the gap between the intricacies and high-dimensionality of contemporary machine learning models and the necessity for human-understandable explanations.

LIME finds widespread use in diverse applications, including image classification, natural language processing, and other domains employing complex models like deep neural networks [10]. By producing locally faithful explanations, LIME contributes to the interpretability and trustworthiness of machine learning models, a critical aspect for applications in sensitive domains such as healthcare, finance, and criminal justice.

SHAP :- SHAP (SHapley Additive exPlanations) is a unified framework for interpreting the output of machine learning models. It is based on cooperative game theory and the concept of Shapley values, which were introduced by Lloyd Shapley in the context of distributing a payout among players in a cooperative game.

In the context of machine learning interpretability, SHAP values are used to attribute the contribution of each feature to the prediction of a particular instance. The main idea is to fairly distribute the "credit" for the model's output among its input features

[19]. SHAP values provide a way to explain the output of any machine learning model by decomposing the prediction into the contribution of each feature.

Here's a high-level overview of how SHAP works:

1. Define a Reference Model: A reference or baseline model is needed to compare against the actual model's predictions. This could be the average output of the model across the entire dataset.

2. Generate All Possible Feature Combinations: Consider all possible combinations of features and calculate the model's output for each combination.

3. Compute Shapley Values: For each feature and each instance, calculate its Shapley value, which represents the average contribution of that feature across all possible feature combinations.

4. Sum Shapley Values for Each Feature: Sum up the Shapley values for each feature to get the overall contribution of each feature to the model's prediction for a specific instance.

The resulting SHAP values provide a comprehensive and fair way to attribute the prediction to individual features, taking into account interactions between features.

SHAP values have been applied to a variety of machine learning models, including linear models, tree-based models (like decision trees and random forests), and complex models like deep neural networks. Keyang et al. [20],The interpretability provided by SHAP values is valuable for understanding why a model makes a particular prediction and gaining insights into the impact of different features on the model's output.

SHAP values are widely used in the field of explainable artificial intelligence (XAI) and have become a popular tool for model interpretability. There are libraries, such as the 'shap' library in Python, that provide efficient implementations for calculating SHAP values for various types of models.

The extensive analysis of the literature review presented in the tables reveal that the CNN based model is best choice for our research as there is a consistent emphasis on CNN across multiple studies, suggesting a robust and widely recognized deep learning technique used in the field of diagnosis of DR.

Table 2.1: Review/Survey Papers

Ref. No.	Authors	Year	Title	Methods surveyed	Dataset surveyed	Parameters used for Accuracy measure	Advancements	Drawbacks
[9]	Stolte et al.	2020	A survey on medical image analysis in diabetic retinopathy	Image registration, vessel segmentation, MA, EXU, HEM, NV, CWS, DRU Lesion segmentation and grading, DR grading, AMD Detection, MH grading and Drusen Detection	FIRE, STARE, DRIVE, HRF Image Database, DDR, DIARETDB1 v2.1, HEI-MED (DMED), E-Ophtha, DRiDB, VICAVR, Messidor, Duke Farsiu SD-OCT, and OCT Image Database	Support vector machine, k nearest neighbors, random forest, neural networks, deep learning and OD localization and segmentation	(1) recognizing features from images that were previously unknown or manually undetectable, (2) easier adjusting to diverse cohorts or imaging modalities, and (3) rapid and accurate learning that addresses user variability and specialist availability.	DR encompasses a wide range of possible lesions, defects, and defect locations that presents a unique challenge to effective diagnosis.
[10]	Tsiknakis et al.	2021	Deep learning for diabetic retinopathy detection and classification based on fundus images: A review	Custom CNNs, GANs and U-Nets	DiaretDB1, e-Ophtha, DRiDB, IDRiD, and MESSIDOR	Accuracy, Sensitivity, Specificity, F1 score or DICE and AUC	Such visualizations allow the clinicians to determine whether the model bases its prediction on relevant clinical features, which in the case of DR would include exudates, microaneurysms and haemorrhages.	Dark areas do not provide any meaningful information regarding any anatomical feature of the retina, when they are brightened up, and thus such images were discarded from both the training and the testing set.

Table 2.2: Review/Survey Papers

Ref. No.	Authors	Year	Title	Methods surveyed	Dataset surveyed	Parameters used for Accuracy measure	Advancements	Drawbacks
[11]	Alyoubi et al.	2020	Diabetic retinopathy detection through deep learning techniques: A review	CNN, auto encoder, sparse coding, FC, CONV, pooling layers	STARE, DRIVE, DDR, E-ophtha, DIARETdb1, HRF and Messidor	Accuracy, sensitivity and specificity	It is notable that the accuracy of the system which built their own CNN structure is higher than those using the existing structures.	The gap that needed to be covered is the existence of systems that could determine the five DR stages with high accuracy as well as detecting DR lesions. This point could be considered as the current challenge for researchers for further investigations.

Table 2.3: Review/Survey Papers

Ref. No.	Authors	Year	Title	Methods surveyed	Dataset surveyed	Parameters used for Accuracy measure	Advancements	Drawbacks
[14]	Mayya et al.	2021	Automated microaneurysms detection for early diagnosis of diabetic retinopathy: A Comprehensive review	Grey level transformation, Gaussian filter, local contrast enhancement, Resampling, CLAHE Hessian matrix, convolution with Guassian filter, ROC Green plane, Vessel Removal, colour normalization	DIARETdb ; e-Optha; RC-RGB-MA; DDR and AGAR300	Precision, recall, f-score, accuracy, sensitivity and specificity	Since the MA mask has mostly black pixels (millions) and only a few white pixels (hundreds), omitting white pixels still yields 99% accuracy.	While DL architectures excel in DR stage classification over traditional methods, their accuracy drops below 0.5 for larger patient groups.
[16]	Praneeth et al.	2023	Diabetic retinopathy detection using machine learning and deep learning techniques: A Review	CNN, ANN, PNN, DCNN, SVM, GBA, DT, RF, LR and MP	DRIVE, DIARETdb, MESSIDOR, STARE	Precision, recall, f-score, accuracy, sensitivity and specificity	Automated screening methods drastically reduce the time needed to make diagnoses, saving ophthalmologists time and money and enabling patients to begin treatment sooner.	The challenge for researchers lies in creating systems that accurately identify all five DR stages and detect DR lesions—a gap that requires further investigation.
[19]	Nadeem et al.	2022	Deep learning for diabetic retinopathy analysis: A review, research challenges, and future directions	CNN, ANN, PNN, DCNN, SVM, GBA, DT, RF, LR and MP	DRIVE, DIARETdb, MESSIDOR, STARE	Precision, recall, f-score, accuracy, sensitivity and specificity	A lot of survey, review and research papers related to topic is discussed just to put the number in terms of quantity of work done.	DR encompasses a wide range of possible lesions, defects, and defect locations that presents a unique challenge to effective diagnosis.

Table 2.4: Review/Survey Papers

Ref. No.	Authors	Year	Title	Methods surveyed	Dataset surveyed	Parameters used for Accuracy measure	Advancements	Drawbacks
[5]	Lakshminarayanan et al.	2021	Automated detection and diagnosis of diabetic retinopathy: A comprehensive survey	SMO, MLP, LMT, Lloyd's algo, GLCM and GLRLM	DRIVE, DIARETdb, e-Ophtha, EyePACS, OCTID, Rabbani	Accuracy, sensitivity, specificity, auc	Deep Learning algorithms are better than Machine Learning algorithms in terms of higher accuracy and reliability of DL models than ML.	ETDRS, the gold standard for DR progression grading, requires detailed evaluation and access to all 7 FOV fundus images, limiting its practicality.
[8]	Das et al.	2022	A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning	CNN, DCNN, AlexNet, VGG-19, GAN, GoogleNet, ResNet, DenseNet, IRCNN, LSTM, RBM, DRL	DRIVE, DIARETDB1, MESSIDOR, STARE, KENYA, HAPIEE, PAMDI and KFSH	F-score, accuracy, sensitivity and specificity	The MobileNet V1 has achieved a training accuracy of 98.90% and validation accuracy of 76.55%.	ML techniques face scalability challenges with high-dimensional data and entail longer analysis and model training times compared to DL techniques.

Table 2.5: Implemented Papers

Ref. No.	Authors	Year	Title	Methods used	Dataset used	Parameters used for Accuracy measure	Advancements	Drawbacks
[17]	Shaban et al.	2018	Automated staging of diabetic retinopathy using a 2d convolutional neural network	CNN	APTOS	Sensitivity, Specificity, cross-entropy loss, probability	Testing accuracy of 80.2% was achieved for training accuracy of 99.9% which is an improvement compared to the previous results.	The results are limited to small specific groups of people.
[18]	Winder et al.	2009	Algorithms for digital image processing in diabetic retinopathy	Quantitative Analysis, morphological analysis, High-pass filtering, Microaneurysms, Contrast enhancement and Watershed Transform	none	Morphological analysis, matched filtering, PCA, wavelets and edge detection	A lot of survey, review and research papers related to topic is discussed just to put the number in terms of quantity of work done.	None of the paper is practically implemented everyone is just discussing or just saying things without any facts from carrying out scientific experiments there is no quality of this paper except mindless reading is all you want to do.
[3]	Ishtiaq et al.	2023	A Hybrid Technique for Diabetic Retinopathy Detection Based on Ensemble-Optimized CNN and Texture Features	CNN, BDA, SFCN, SVM, SCA, GLCM, CLAHE, KNN, LBP	IDRiD, Eye-PACS, DI-ARETdb, MESSIDOR, STARE	Precision, f1-score, accuracy, sensitivity and specificity	Our proposed technique excelled in classification performance, achieving high accuracy (98.85%), sensitivity (98.85%), specificity (99.71%), precision (98.89%), and F1 Score (98.85%) compared to the mentioned techniques.	Deep learning models and architectures in DR analysis still face the issue of patch classification, as the anatomical location of the patch is not known.

Table 2.6: Implemented Papers

Ref. No.	Authors	Year	Title	Methods surveyed	Dataset surveyed	Parameters used for Accuracy measure	Advancements	Drawbacks
[4]	Mohanty et al.	2023	Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy	DenseNet 121, VGG16, XG-Bosst, LSTM, CNN, ReLu, Conv2D	APTOS 2019, DRIVE, STARE, EyePACS and Messidor	Precision, recall, f-score, accuracy, sensitivity and specificity	Our DenseNet 121 model outperformed existing methods on the same dataset with an impressive accuracy of 97.30%, surpassing all other architectures.	The change in learning rate affects the model in classifying the positive class images based on precision-recall and AUC.
[13]	Mushtaq et al.	2021	Detection of diabetic retinopathy using deep learning methodology	CNN, DenseNet-169, LBP, LTP, LESH, SVM, KNN, Regression, DT	Aptos 2019, DIARETdb and EyePACS	Precision, recall, f-score, accuracy, sensitivity and specificity	Our proposed model performed better than the regression model by achieving the accuracy of 90% however, 78% accuracy was yielded by the regression model.	Extensive preprocessing and augmentation due to diverse image conditions may risk losing vital features, underscoring the need for techniques that balance feature preservation and effective pre-processing, posing a potential drawback.

Table 2.7: Implemented Papers

Ref. No.	Authors	Year	Title	Methods used	Dataset used	Parameters used for Accuracy measure	Advancements	Drawbacks
[21]	Qummar et al.	2019	A deep learning ensemble approach for diabetic retinopathy detection	CNN, Resnet50, Inceptionv3, Xception, Dense121, Dense169	DRIVE and STARE	Precision, recall, f-score, accuracy, sensitivity and specificity	The results show that the proposed ensemble model performs better than other state-of-the-art methods and is also able to detect all the stages of DR.	The change in learning rate affects the model in classifying the positive class images based on precision-recall and AUC.
[15]	Gulshan et al.	2016	Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs	CNN, ANN, PNN, DCNN, SVM, GBA, DT, RF, LR and MP	EyePACS-1, MESSIDOR-2	Accuracy, sensitivity and specificity	The winning entry's performance was better than ophthalmologist grading the wrong eye of the same patient and exceeded the agreement observed between general physicians and ophthalmologists.	Only few cases are taken into considerations some types of DR gets ignored by the algorithm.
[22]	Dutta et al.	2018	Classification of diabetic retinopathy images by using deep learning models	CNN, DCNN, FNN, VGG-16, FCM, CONV, ReLu, BNN and DNN	FUNDUS	RMSE and accuracy	Both CNN and DNN models are effective for image tasks, but due to CPU training time constraints, DNN outperforms CNN in both training and validation accuracy.	Statistical values are reliable for predicting severity levels, but in the presence of noisy images, accuracy can suffer due to potential data quality issues.
[6]	Wong et al.	2016	Artificial intelligence with deep learning technology looks into diabetic retinopathy screening	CNN, ANN, PNN, DCNN, SVM, GBA, DT, RF, LR and MP	DRIVE, DIARETdb, MESSIDOR, STARE	Precision, recall, f-score, accuracy, sensitivity and specificity	A lot of survey, review and research papers related to topic is discussed just to put the number in terms of quantity of work done.	DR encompasses a wide range of possible lesions, defects, and defect locations that presents a unique challenge to effective diagnosis.

Table 2.8: Implemented Papers

Ref. No.	Authors	Year	Title	Methods used	Dataset used	Parameters used for Accuracy measure	Advancements	Drawbacks
[7]	Aldughayfiq et al.	2023	Explainable AI for Retinoblastoma Diagnosis: Interpreting Deep Learning Models with LIME and SHAP	LIME, SHAP, CNN, GMM	IDRiD, Eye-PACS, DIARETdb, MESSIDOR	PPV, NPV, precision, recall, F1-score, accuracy, sensitivity and specificity	This research showcases the effectiveness of deep learning in diagnosing retinoblastoma, emphasizing strong model performance and interpretability.	This study acknowledges limitations, including dataset bias and the need for real-world testing to assess the AI model's performance in clinical settings.
[20]	Alrajjou et al.	2022	An enhanced interpretable deep learning approach for diabetic retinopathy detection	LIME, SHAP, CNN, PSO, auto-encoder, GLCM, SVM	IDRiD, Eye-PACS, DIARETdb, MESSIDOR	Precision, recall, F1-score, accuracy, sensitivity and specificity	The enhanced CNN notably improves diabetic retinopathy detection with superior accuracy and enhanced interpretability through LIME and SHAP, promising more effective clinical diagnosis and treatment decisions.	Lack of data preprocessing to remove not required noisy data before feature extraction and classification has increased computational costs and impact classification outcomes.
[12]	Joseph et al.	2022	Explainable diabetes classification using hybrid Bayesian-optimized TabNet architecture	LIME, SHAP, SMBO, ELM, FDT, RNN	PIDD, ES-DRPD	Cross-Validation, precision, recall, F1-score, Cohen's Kappa, accuracy, sensitivity and specificity	The study found that adding Bayesian Optimization improved the TabNet model's diabetes classification, surpassing all benchmarked models with high accuracy values of 0.922 and 0.994.	The PIDD dataset used in this study is limited to female patients from a specific geographic region.

Chapter 3

Proposed Methodology

The architecture of the convolution neural network with rise in the number of the convolution layers has made it able to read and understand about the more detailed attributes of the image data provided. The sincere procedure of deployment of the convolutional neural network has set an example for focusing on every detail so to easily differentiate the hard exudates [11]. The very first layer has only one function to deeply study the very endpoints of the network which are situated at the way beneath the network as well as the very last convolution layer too should properly understand the attributes crucial to correctly identify the phases and lesions of diabetic retinopathy such as hard exudates such a lesion normally indicates that the patient has entered the stage 3 of DR which is severe NPDR. At the beginning of the convolutional neural network we have convolution blocks which along with activation maps are then able to perform batch normalization after every convolution layer in the network. With the growing number of feature maps we safely resort to one batch normalization per block and this is how we can freely save most resources possible.

Each and every maxpooling is operated alongside kernel sizes of 3x3 and 2x2 strides. Once we have traversed through the final convolutional block we notice that the network is reduced to only one dimension i.e. only representing a single attribute. The important thing is to prevent the graph from over-fitting we must give certain weights to each and every class depending upon the number of images in that particular class. The same way we operate upon dense layers doing the dropout of layers that are not required anymore this way we make sure that over-fitting does not occur [16]. This must be repeated over and over, once we have arrived at dense five node, this classification layer exploits a softmax activation function to speculate whether our classification is true or wrong. It is to be noticed that whenever we deploy leaky rectified linear unit as an actuation function we observe that it stops to depend upon specific nodes in the convolutional neural network. Simultaneously we utilize the ridge regression for regularization of weight and biases. Moreover also gaussian initialization is applied to decrease the initial training time.

The dataset which was utilized to experiment in this study was the APTOS 2019 blindness kaggle dataset which is publicly available on their friendly neighbourhood machine learning and data science community website https://www.kaggle.com/competitions/ aptos 2019-blindness-detection/data and this encompasses of 3662 train-images and 1982 test-images which makes a grand total of 5990 images occupying a sum of 10.22 GB of data. Moreover there is a private dataset within this competition which comprises of 13000 images with a storage requirement of 20 GB of disk space on your system. This is literally double of that is accessible to us freeloaders [21]. Now after reducing the resolution of images and then running the network on NVIDIA 3050Ti Laptop GPU we were able to train only certain part of dataset. Now the NVIDIA 3050Ti mobile GPU has 2560 cuda cores and is included with the NVIDIA CUDA Deep Neural Network library (cuDNN) for GPU learning. Utilization of this specific package lead to 1866 train-images, 1928 test-images, 9316 reduced-train images and 9306 reduced-train-cropped images which equals to a great grand total of sum of 22416 images overall out of which only 633 images were given to GPU once at a time.

The dataset was made up of images from the patients of assorted ethnicity, age groups and extremely diversified levels of lighting in the fundus photography. As a result of this variation the pixel intensity values within the images are infected and as a consequence inessential diversity is build up in the classification levels. Colour Normalisation is carried out on the images in response of unnecessary variation in the dataset [22]. Moreover the images have high resolution which occupied major portion of GPU memory. To answer this issue the images resolution is reduced to a resolution of 1024x680 pixels from original size of 3216x2136 pixels because this resolution in fact is able to hold on to the important features required for diagnosis of DR and further this reduced the dataset size that my GPU can handle.

The CNN has started its training on 1866 images in the beginning as it enabled for me to achieve a level where it is possible to avoid wasting time on training for obtaining quick classification. At the end of 15 epochs of training the network is further trained on a grand total of 22416 images for 6 epochs. Because of this the network suffered from heavy over-fitting as our dataset has 999 images of class 2 type which is moderate non-proliferative DR in contrast to class 1 having 370 images, class 3 having 193 images the least of all being the severe DR class and finally the last class has 295 images. Addressing this problem i have allocated real time class weights in the network [20]. Each batch that is loaded to performed back propagation on it for that batch the class weights is updated with a ratio depending upon how many images in that is of class moderate DR by doing this i try to reduce the over-fitting to this class whose images populated my dataset. The network is trained with the help of adam optimizer alongside with StepLR. A low learning rate of 0.001 further decreasing to 0.00025 as used for 15 epochs to make the weights steady enough. The authentic preprocessed images is only used for training the network only once.

Chapter 4

Discussion

4.1 Discussion

My work have proven that the level 5 issue for country level broadcasting of diagnosis of diabetic retinopathy is achievable from the adoption of CNN based approaches along with modern micro-services. The CNN that is utilized in this project has been adequate enough to grasp the necessary attributes required for differentiating the retinal images into their level of class. The most prominent thing of all it is able to identify most of the level 4 which proliferative diabetic retinopathy. On comparing with other researchers work where they have utilize huge datasets having high specificity, they have to pay a large price which is their model having lower sensitivity making their work unreliable.

The obvious advantage of implementing my model in your day to day life comes when one needs to associate levels(class type) to lakhs of new retinal scans, it can do this every minute making it much reliable which is fair. In medical practice daily those retinal scans termed as fundus images are actually being send to the clinicians so they can grade it to whatever ever class each eye belongs to but the problem lies in whenever a patient is taken in for in-person screening those fundus images appears to be not precisely graded which is irresponsible of the medical staff and may lead to severe eyesight loss of the patient. The extensively trained CNN comes in handy in those scenarios being able to make really fast diagnosis making the medical staff able to give immediate treatment for the patient happen in such short time notice.

Chapter 5

Conclusion

Diabetes is emerging as a rapidly growing health concern in recent times. Surveys indicate that individuals with diabetes have approximately a 30% likelihood of developing Diabetic Retinopathy (DR). DR progresses through various stages, ranging from mild to severe, ultimately leading to Proliferative Diabetic Retinopathy (PDR). Advanced stages of the disease can result in floaters, blurred vision, and, if left undetected in the early phases, may lead to blindness. Manual diagnosis of DR images demands highly skilled professionals, is time-intensive, and poses significant challenges. To address this, computer vision-based techniques for automated DR detection and classification have been proposed in the literature. This paper focuses on the classification of all DR stages, particularly emphasizing the early stages, a key limitation in existing models. The authors propose a CNN ensemble-based framework designed to detect and classify different stages of DR in color fundus images. The study utilizes the Kaggle dataset, the largest publicly available fundus image dataset, for model training and evaluation. The results demonstrate that the proposed ensemble model outperforms other state-of-the-art methods, successfully identifying all DR stages.

Bibliography

[1] R. Vij and S. Arora, “A systematic review on diabetic retinopathy detection using deep learning techniques,” *Archives of Computational Methods in Engineering*, vol. 30, no. 3, pp. 2211–2256, 2023.

[2] R. Biyani and B. Patre, “Algorithms for red lesion detection in diabetic retinopathy: A review,” *Biomedicine & Pharmacotherapy*, vol. 107, pp. 681–688, 2018.

[3] U. Ishtiaq, E. R. M. F. Abdullah, and Z. Ishtiaque, “A hybrid technique for diabetic retinopathy detection based on ensemble-optimized cnn and texture features,” *Diagnostics*, vol. 13, no. 10, p. 1816, 2023.

[4] C. Mohanty, S. Mahapatra, B. Acharya, F. Kokkoras, V. C. Gerogiannis, I. Karamitsos, and A. Kanavos, “Using deep learning architectures for detection and classification of diabetic retinopathy,” *Sensors*, vol. 23, no. 12, p. 5726, 2023.

[5] V. Lakshminarayanan, H. Kheradfallah, A. Sarkar, and J. Jothi Balaji, “Automated detection and diagnosis of diabetic retinopathy: A comprehensive survey,” *Journal of imaging*, vol. 7, no. 9, p. 165, 2021.

[6] T. Y. Wong and N. M. Bressler, “Artificial intelligence with deep learning technology looks into diabetic retinopathy screening,” *Jama*, vol. 316, no. 22, pp. 2366–2367, 2016.

[7] B. Aldughayfiq, F. Ashfaq, N. Jhanjhi, and M. Humayun, “Explainable ai for retinoblastoma diagnosis: Interpreting deep learning models with lime and shap,” *Diagnostics*, vol. 13, no. 11, p. 1932, 2023.

[8] D. Das, S. K. Biswas, and S. Bandyopadhyay, “A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning,” *Multimedia Tools and Applications*, vol. 81, no. 18, pp. 25613–25655, 2022.

[9] S. Stolte and R. Fang, "A survey on medical image analysis in diabetic retinopathy," *Medical image analysis*, vol. 64, p. 101742, 2020.

[10] N. Tsiknakis, D. Theodoropoulos, G. Manikis, E. Ktistakis, O. Boutsora, A. Berto, F. Scarpa, A. Scarpa, D. I. Fotiadis, and K. Marias, "Deep learning for diabetic retinopathy detection and classification based on fundus images: A review," *Computers in biology and medicine*, vol. 135, p. 104599, 2021.

[11] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Informatics in Medicine Unlocked*, vol. 20, p. 100377, 2020.

[12] L. P. Joseph, E. A. Joseph, and R. Prasad, "Explainable diabetes classification using hybrid bayesian-optimized tabnet architecture," *Computers in Biology and Medicine*, vol. 151, p. 106178, 2022.

[13] G. Mushtaq and F. Siddiqui, "Detection of diabetic retinopathy using deep learning methodology," in *IOP conference series: materials science and engineering*, vol. 1070, p. 012049, IOP Publishing, 2021.

[14] V. Mayya, S. Kamath, and U. Kulkarni, "Automated microaneurysms detection for early diagnosis of diabetic retinopathy: A comprehensive review," *Computer Methods and Programs in Biomedicine Update*, vol. 1, p. 100013, 2021.

[15] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *jama*, vol. 316, no. 22, pp. 2402–2410, 2016.

[16] P. Dp, "Diabetic retinopathy detection using machine learning and deep learning techniques: A review," *Industrial Engineering Journal*, vol. 52, pp. 11–22, 08 2023.

[17] M. Shaban, Z. Ogur, A. Shalaby, A. Mahmoud, M. Ghazal, H. Sandhu, H. Kaplan, and A. El-Baz, "Automated staging of diabetic retinopathy using a 2d convolutional neural network," in *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 354–358, IEEE, 2018.

[18] R. J. Winder, P. J. Morrow, I. N. McRitchie, J. Bailie, and P. M. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Computerized medical imaging and graphics*, vol. 33, no. 8, pp. 608–622, 2009.

[19] M. W. Nadeem, H. G. Goh, M. Hussain, S.-Y. Liew, I. Andonovic, and M. A. Khan, "Deep learning for diabetic retinopathy analysis: A review, research challenges, and future directions," *Sensors*, vol. 22, no. 18, p. 6780, 2022.

[20] S. Alrajjou, E. K. Boahen, C. Menga, and K. Cheng, "An enhanced interpretable deep learning approach for diabetic retinopathy detection," in *2022 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, pp. 127–135, IEEE, 2022.

[21] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. A. Khan, and W. Jadoon, "A deep learning ensemble approach for diabetic retinopathy detection," *Ieee Access*, vol. 7, pp. 150530–150539, 2019.

[22] S. Dutta, B. Manideep, S. M. Basha, R. D. Caytiles, and N. Iyengar, "Classification of diabetic retinopathy images by using deep learning models," *International Journal of Grid and Distributed Computing*, vol. 11, no. 1, pp. 89–106, 2018.