Deep Learning-Based Detection of Diabetic Retinopathy in Retinal Images

MSc Research Project
MSc in Data Analytics

Abhilash Chava
Student ID: X21178712

School of Computing
National College of Ireland

Supervisor: Vladimir Milosavljevic

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Abhilash Chava
Student ID:	X21178712
Programme:	MSc in Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Vladimir Milosavljevic
Submission Due Date:	14/08/2023
Project Title:	Deep Learning-Based Detection of Diabetic Retinopathy in Retinal Images
Word Count:	17,048
Page Count:	22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	*Abhilash Chava*
Date:	13th August 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Deep Learning-Based Detection of Diabetic Retinopathy in Retinal Images

Abhilash Chava
X21178712

Abstract

Diabetic retinopathy (DR) is an eye ailment that is progressive in nature and is caused by high blood glucose levels. It is the major cause of blindness among diabetics, particularly in less developed nations. The early identification of DR is essential for the preservation of vision; yet the present practise of having ophthalmologists review photographs of the retinal fundus is time-consuming and expensive. This is especially true in the early stages of the illness, when disease symptoms are less apparent in the photographs. The research suggests using deep learning methods, such as supervised, self-supervised, and Vision Transformer configurations, for classifying and detecting DR in photographs of the retinal fundus. Several models were evaluated and compared; based on the Kappa values, DenseNet was found to have the greatest performance, followed by MobileNet and VGG-19. Various models' Kappa scores are shown to be 77.0% for VGG-19, 77.6% for ResNet-50, 82.0% for DenseNet, 80.0% for MobileNet, and 73.2% for Vision transformer. Additionally, current retinal fundus datasets are investigated for DR detection, classification, and segmentation, and research gaps and potential topics for additional exploration are identified. The long-term objective is to devise a method that can diagnose DR in a speedy and accurate manner using only a smartphone.

Keywords: Diabetic retinopathy, Transfer learning, Staged Networks, Artificial Neural Networks

1 Introduction

Diabetic retinopathy is a vision loss problem that can have serious repercussions due to the link between diabetes and the disease. Diabetes hinders the body's natural capacity to take in and properly store sugar, which can have devastating effects on one's health. Damage to the eyes and other organs can result from having too much glucose in the blood. Damage to the retina and the tiny blood vessels that supply which is diabetic retinopathy, and it is the result of uncontrolled diabetes. The problem occurs when fluid leaks from these small blood vessels, causing the retinal tissue to swell and resulting in blurred or foggy vision. The key to preventing these complications and maintaining general health is keeping diabetes under control. Diabetic retinopathy, and the potential vision loss it might cause, can be mitigated with timely diagnosis and treatment (Nazir, T., et. al., 2021).

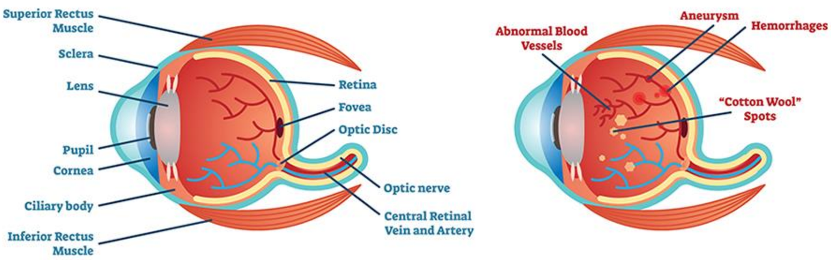

Figure 1: Difference between Normal Eye and Diabetic Eye (Nazir, T., et. al., 2021)

Diabetic retinopathy is a condition that can cause damage to one or both eyes and is more common in those who have had diabetes for a longer period of time. Fluid accumulation in the lens, caused by increased blood sugar levels, can cause blurred vision and distorted peripheral vision (Sarki, R., Michalska, S., Ahmed, K., Wang, H. and Zhang, Y., 2019). However, when blood sugar levels are kept in check, the lens can return to its normal shape, leading to better eyesight. If a patient's blood sugar levels are kept under control, the onset and development of diabetic retinopathy are both slowed. This emphasises the need of controlling diabetes to protect vision.

1.1 Motivation

Early detection of diabetic retinopathy is crucial for lowering the risk of vision loss from the disease. The illness might worsen if either the diagnosis or therapy is delayed. The use of deep neural networks (DNN) and other machine learning techniques has increased in the field of disease prediction, reducing the need for human analysts to aid in data interpretation. By allowing for early diagnosis, the risk of developing diabetic retinopathy can be reduced by using these state-of-the-art diagnostic tools. The purpose of this study is to save patients' eyesight and avoid difficulties by early detection utilising deep learning and data analysis rather than depending just on clinical examination. Reducing reliance on clinical analysis alone is one way to achieve this goal (Panwar, A., Semwal, G., Goel, S. and Gupta, S., 2022).

1.2 Business Objective

Even though state-of-the-art diagnostic tools are readily available, many people with diabetes still refuse or put off getting their eyes scanned. Since diabetes mellitus can cause blindness, a prompt diagnosis is crucial. For the goal of evaluating information related to diabetic retinopathy, diagnosing the illness, and projecting its evolution, deep learning and machine learning techniques have significant potential in the healthcare sector, especially in settings with limited resources. Rather than relying solely on clinical evaluation, data-driven analyses can be utilised to enhance the identification and treatment of diabetic retinopathy. In underdeveloped countries, these technologies might improve patient outcomes and aid in the battle against the dangers of diabetic retinopathy (Panwar, A., Semwal, G., Goel, S. and Gupta, S., 2022).

1.3 Research Questions

For the purpose of diagnosing diabetic retinopathy, this strategy relies on making use of sophisticated machine learning and deep learning techniques. Even if there has been some progress achieved, the development of vision transformers and a comprehensive investigation of important areas pertaining to this disease might lead to substantial advances in the fight against it. In this study, the following research questions will be addressed:

"*How can advanced deep learning techniques, including Vision Transformers, improve the accuracy and efficiency of detecting Diabetic Retinopathy in the retinal fundus images compared to traditional diagnostic methods?*"

In the subsequent chapter, a detailed discussion on the literature review conducted for various cases will be presented. This will be succeeded by the Methodology chapter, detailing the methods and processes implemented. The Implementation chapter will delve into the proposed framework and code execution. Lastly, the Results and Analysis chapter will focus on the examination of the outcomes.

2 Related Work

Diabetic retinopathy (DR) is a disorder that can cause vision loss of varying degrees of severity over time and can lead to irreversible blindness, and the chance of getting DR increases in those who have had diabetes for a longer period of time. Machine learning and deep learning algorithms have been developed to produce more accurate classifications since ophthalmologists have a hard and time-consuming time diagnosing DR. Researchers recommended a CNN architecture that has been pre-trained with Transfer Learning for the aim of identifying DR severity groups in retinal fundus pictures. Multiple computer vision and deep learning techniques were used into this design. By correctly classifying new pictures using feature vectors extracted from the DR images used as a test set, the suggested technique outperformed competing methods. Accurate classification of the various lesions seen in the diabetic fundus images was achieved using a combination of many Convolutional Neural Network (CNN) models and classifiers (Panwar, A., Semwal, G., Goel, S., & Gupta, 2022).

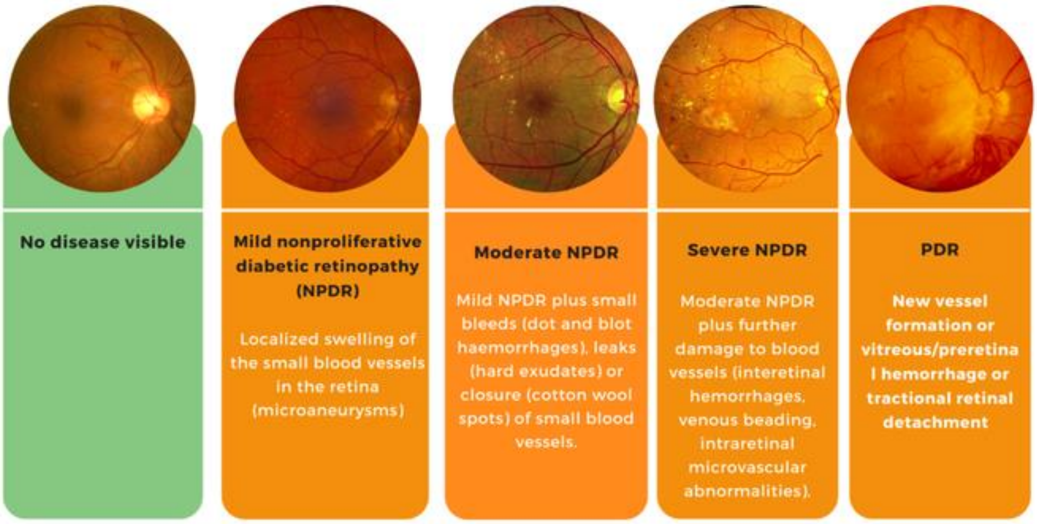

Figure 2: DR Taxonomy explanation (Farooq, M.S., Arooj, A., Alroobaea, R., Baqasah, A.M., Jabarulla, M.Y., Singh, D. and Sardar, R., 2022)

Farooq, M.S., Arooj, A., Alroobaea, R., Baqasah, A.M., Jabarulla, M.Y., Singh, D., and Sardar, R., 2022 study makes use of deep learning methods to improve the quality of research into the detection of diabetic retinopathy (DR). This study examines DR classification, grading, and staging by comparing and analysing several deep learning-based algorithms for identifying DR phases. Deep Neural Networks, Ensemble Convolutional Neural Networks, and Convolutional Neural Networks are just a few of the methods used for DR categorization. The researchers focus extensively on the advantages of deep learning systems for DR screening and prevention.

M. Ullah Imran, M. Arif Imran, and R. Noor Noor (2022) proposes a new method for detecting diabetic retinopathy by employing a hybrid neural network that increases entropy. The method avoids the challenges of distinguishing between the various stages of DR in medical imaging by employing discrete wavelet modifications to enhance picture visibility. The research group develops a powerful hybrid neural network to identify the signs of diabetic retinopathy in photographs.

In this research A.T. Nair, M.L. Anitha, and A. Kumar,2022 used deep neural networks to classify many models of diabetic retinopathy and then compare and contrast the results. Data collected from the retina is used to categorise retinal images according to DR severity, and techniques including fundus photography, OCT, and OCTAI are employed. Pre-trained designs like VGG16, EfficientNetB5, and ResNet50 are evaluated on the Kaggle (APTOS) dataset, with varying accuracy rates reported for each model.

In this study Masud (M), Alhamid (M.F.), and Zhang (Y), 2022, the researchers created a deep learning model to analyse images of diabetic retinopathy and categorise the severity of DR cases. High precision, a Cohen Kappa score, and weighted average accuracy, recall, and F1 score are only some of the ways in which this model excels above its competition of pre-trained models. There is hope that the proposed VGG-16 network design can accurately categorise DR in retinal images.

O'Shea and Nash (2015) recommend a deep learning model built on the VGG-16 network architecture for DR classification. The model shows promise for application in DR classification tasks on both the DRIVE and STARE datasets, with impressive results across the board in terms of accuracy, specificity, sensitivity, and precision. These studies add significantly to the current literature on the use of deep learning techniques for the diagnosis of diabetic retinopathy. They show that deep learning algorithms have great potential for detecting and classifying diabetic retinopathy (DR) at an early stage.

Let’s discuss in detail about the different individual areas for the DR detection,

2.1 Prediction of Diabetic Retinopathy using machine Learning

The prevention of blindness in diabetic individuals depends on their early detection of diabetic retinopathy. The study group set out to create a deep learning system that could foresee the likelihood of diabetic retinopathy within two years of diagnosis. Two variants of the deep learning system were developed, one of which took advantage of three-field colour fundus photographs and the other of which took advantage of only one-field colour fundus images. Teleretinal diabetic retinopathy screening data was used in both versions. The AUC for the three-field deep learning approach was 0.79 in internal validation, while the AUC for the one-field deep learning approach was just 0.70. When the deep learning system was combined with the risk variables, the AUC climbed to 0.81. Support vector machines had the best levels of accuracy and area under the ROC curve in the experiments. (Bora et al., 2021) and (Tsao et al., 2018) The researchers' method was successful and demonstrated the promise of combining machine learning with clinical criteria to reliably diagnose diabetic retinopathy.

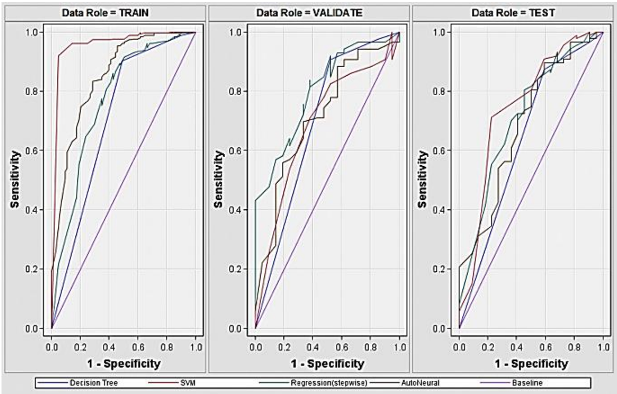

Figure 3: For the training, validation, and test data sets, the ROC curves of different machine learning algorithms (i.e., DT, LR, SVM, and ANN) (Tsao, H.Y., Chan, P.Y. and Su, E.C.Y., 2018).

It is unclear if the few models proposed up to this point in the prior study have been validated outside of the population for which they were originally designed. This has been questioned, therefore there is some uncertainty. By combining socioeconomic and clinical information, the researchers want to identify the features that are the most accurate and discriminatory in their interpretations (Alabdulwahhab, K.M., Sami, W., Mehmood, T., Meo, S.A., Alasbali, T.A., & Alwadani, F.A., 2021). This is because DR categorization utilising several machine learning methods on data from Saudis with diabetes is the focus of this research. Diabetic retinopathy is on the rise worldwide and can cause blindness, severe visual loss, and even death if left untreated. A more precise diagnosis of diabetic retinopathy may be possible with the use of signal processing and machine learning applications in biomedical imaging. The experiment involved taking pictures of the retina using a fundus camera from people who were either healthy or diabetic. The suggested algorithm's original contribution is the use of image preprocessing methods, morphological operations to detect statistical characteristics, and DWT for extracting the histogram-based feature. All of these are essential parts of the proposed approach. K-Nearest Neighbours, Support Vector Machine, and an Artificial Neural Network were used to categorise these characteristics (Mohanty, K.K., Barik, P.K., Barik, R.C., and Bhuyan, K.C., 2019). In order to make reliable DR predictions.

2.2 Computer Vision applications in Diabetic Retinopathy Detection

In their study, Kassani and colleagues (2019) suggest a novel approach to the extraction of characteristics for the diagnosis of diabetic retinopathy (DR). This method is based on a reworked version of the Xception architecture. In this architecture, many convolutional layers are combined to provide a unified multi-level feature. This data is then used to educate an ANN for DR classification purposes after it has been extracted. To compare the proposed technique to other popular architectures, employed Inception V3, MobileNet, ResNet50, and the initial version of the Xception architecture. The dataset for the 2019 Kaggle APTOS challenge was used for both experimentation and evaluation. Throughout their testing, the scientists saw that the Xception deep feature extractor outperformed the original design by a large margin. With an outstanding accuracy of 82.99%, sensitivity of 88.23%, and specificity of 87.01%, the proposed technique demonstrates the value of the modified Xception architecture in categorising DRs.

Meanwhile, Wang et al. (2020) aimed to improve detection of diabetic retinopathy by enhancing the ability to distinguish extremely tiny things. To get there, they used a feature pyramid network in tandem with an adjusted region proposal network. Data from the Messidor database and fundus images from the Shanghai Eye Hospital were used to validate the DR-grading model built using this updated R-FCN. Their research provided some grounds for cautious optimism. Data from the Shanghai Eye Hospital shows that the modified R-FCN achieved outstanding sensitivity and specificity scores of 99.39% and 99.33%, respectively. Furthermore, it showed a sensitivity of 92.59% when tested on the Messidor dataset and a specificity of 96.20%. In addition, when tested on the hospital data set, the most current version of the R-FCN lesion-detection model reached a 92.00% accuracy level. This demonstrates that the proposed R-FCN can accurately evaluate DR and identify lesions. These studies highlight the value of deep learning and convolutional neural networks (CNNs) for diabetic retinopathy diagnosis. The efficiency with which these techniques detect and categorise DR has increased. This was achieved by using transfer learning and modifying existing architectures. Considering the rising prevalence of diabetes and the critical need for early detection to avoid vision loss, these findings are timely and important. Incorporating feature pyramid networks and region proposal networks, as demonstrated in the modified R-FCN, and employing feature extraction techniques, such as multi-level feature extraction in the modified Xception architecture, contribute to increased accuracy and sensitivity when diagnosing diabetic retinopathy. The updated R-FCN includes these methods. Recent advancements in deep learning-based methodologies have provided ophthalmologists and other medical practitioners with helpful tools for more precise diagnosis and management of diabetic retinopathy. The studies by Kassani et al. 2019 and Wang et al. 2020 add significantly to the growing body of knowledge concerning the application of deep learning to the diagnosis of diabetic retinopathy. To combat the rising global incidence of diabetic retinopathy and to ensure that immediate interventions are taken to prevent persons with diabetes from incurring vision loss, accurate and efficient diagnostic models are required.

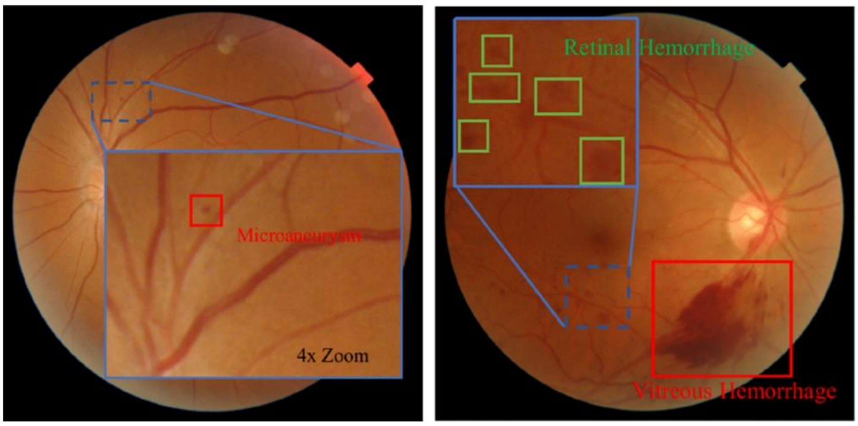

Figure 4: (a) A microaneurysm is shown in this image (b) Sample photograph showing haemorrhage in the retina and vitreous (Ramasamy, L.K., Padinjappurathu, S.G., Kadry, S., & Damaevius, 2021)

Screening for DR can be difficult, but an automated technique for early detection can preserve patients' eyesight and assist ophthalmologists save more. The results of this study paved the way for the creation of a diagnostic framework for DR. Gray-level features such as co-occurrence, run-length matrices, and Ridgelet Transform coefficients were used to extract and aggregate ophthalmoscopic data from retina images in the first step of the procedure. Sequential Minimal Optimisation (SMO) is a classification approach for diabetic retinopathy that uses retinal characteristics as the basis for classification. The researchers got a 98.87 percent sensitivity, 95.24 percent specificity, and 97.05 percent accuracy on the DIARETDB1 dataset, and a 90.9 percent sensitivity, 91.0 percent specificity, and 91.0 percent accuracy on the KAGGLE dataset (Ramasamy, L.K., Padinjappurathu, S.G., Kadry, S., & Damaevius, 2021). The study findings demonstrate the superiority and efficacy of the proposed strategy.

Testing and grading BDL algorithms is a formidable task. In order to determine whether or whether new tools provide more precise uncertainty estimates than older ones, and routinely examine the scalability and robustness of the methodologies. These evaluations are crucial for developers when deciding which BDL tools to use for their apps. One of the most widely disseminated criticisms of BDL methodology cites the following flaws in the UCI experiments: Both the medical and automotive industries are in desperate need of fresh standards since current practises have not been effective. In this article [15], the authors provide a novel BDL benchmark derived from a medical imaging application with a focus on diabetic retinopathy diagnosis. The researchers invented this standard. Due to the model's inherent ambiguity, it is often necessary to refer patients to specialists for further evaluation when the model's diagnosis is in doubt. Visual inputs (512 x 512 RGB retinal pictures) are used for this purpose. Using model uncertainty, automated diagnosis is used in the actual world, with approaches being ranked using criteria gathered from expert domains. Among the many tasks that fall under the scope of this kind of application are the identification of out-of-distribution and distribution changes. After that, they look at how well the refined BDL techniques are doing across a wide range of professions. The study's authors found that certain preexisting benchmark resolution techniques, like UCI, 'overfit' their uncertainty to the dataset. Therefore, the researchers found that when compared to more straightforward baselines, the tactics in question performed poorly (Filos, A., Farquhar, S., Gomez, A.N., Rudner, T.G., Kenton, Z., Smith, L., Alizadeh, M., De Kroon, A., 2022).

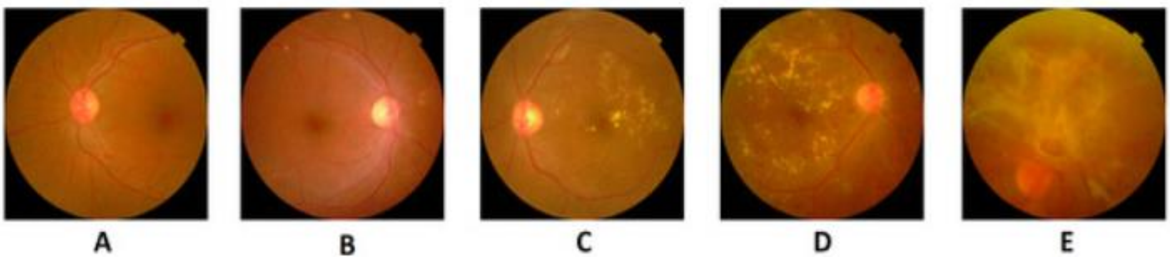

Figure 5: diabetic retinopathy in fundus pictures is divided into five stages: Either (A) no DR, (B) mild DR, (C) moderate DR, or (D) severe DR, (E) PDR DR (Ramasamy, L.K., Padinjappurathu, S.G., Kadry, S., & Damaevius, 2021)

2.3 Deep Learning Applications in the Efficient Prediction

Research has shown that automated diabetic retinopathy (DR) diagnosis using fundus photography can classify eye pathological signs with an accuracy of above 90%. Despite this, distinguishing between less severe, moderate, and severe cases remains a challenge. Sarki et al. (2019) used data from Kaggle and Messidor to train CNNs, which are a type of artificial neural network. Pre-trained on the ImageNet database, thirteen distinct CNN architectures had their performance enhanced by the inclusion of new data and an increase in the total quantity of training data. The ResNet50 model's accuracy for detecting No DR/Mild DR cases increased to 86% after being fine-tuned with the RMSProp Optimizer. Despite its empirical nature, the study addressed the limitations of CNN in detecting tiny lesions and allowed for early DR detection. Sarki et al. (2019) state that the model's robustness was demonstrated by its ability to accommodate varying degrees of image quality. Separately, Wan et al. (2018) used convolutional neural networks (CNNs) to automate DR identification by classifying, segmenting, and detecting features related with DR. The results indicated that CNNs with transfer learning produced a greater degree of accuracy (95%) after evaluating several different architectures such as AlexNet, VggNet, GoogleNet, and ResNet. The researchers opted to perform their training on Kaggle due to resource constraints (Wan et al., 2018).

The Eye WeS method for detecting DR and pinpointing lesion sites in fundus images was also introduced by Costa et al. (2019), with an emphasis on semi-supervised learning. The AUC performance was improved from 94.9% to 95.8% while maintaining just 5.0% of Inception V3's settings. The 2019 study by Costa et al. shows that when applied to several datasets, the same model can get an AUC of 97.1%, proving its utility. In conclusion, the introduction of CNNs and other deep learning algorithms has greatly improved the accuracy of automated DR detection. The challenges of detecting subtle symptoms have been addressed in these studies. Flexibility to a range of real-world scenarios and image quality levels has been demonstrated, and DR classification and lesion detection have also been successfully completed.

2.4 Findings

Thanks to the incorporation of deep learning methods, the detection of diabetic retinopathy (DR) has witnessed a significant improvement in both accuracy and sensitivity. Early DR diagnosis and classification is a challenging problem, but research has shown that convolutional neural networks (CNNs), deep neural networks (DNNs), and ensemble CNNs (ECNNs) have a great deal of potential. The results have been an increase in reliability, precision, and sensitivity. Pre-trained deep learning architectures have been shown to be useful for disease grading, and examples include ResNet50, EfficientNetB5, and VGG16. The possibility of developing DR within a specified time frame may also be predicted using a technique that combines clinical risk factors with deep learning algorithms. Collectively, our results demonstrate the value of deep learning in addressing challenges related to lesion detection, improving classification accuracy, and allowing for more precise and timely interventions to prevent vision loss in people with diabetes and diabetic retinopathy.

2.5 Research Niche

When applied to clinical risk indicators, deep learning algorithms provide a novel method for estimating the likelihood that diabetic retinopathy (DR) would worsen over a certain period of time. To avoid vision loss in diabetic patients, deep learning has been demonstrated to be effective in overcoming challenges related with lesion detection, enhancing classification accuracy, and enabling timely interventions. Notably, there is still the possibility of building a comprehensive framework, even when models like DenseNet are still not being used to their full potential. The purpose of this method is to find the most effective feature extraction strategy for DR detection by methodically analysing and comparing different transfer learning approaches. With the aim of boosting DR detection model precision, this method employs a data-driven, pragmatic mechanism for selecting the most effective transfer learning strategy. The unique technique has the potential to push the limits of current research, which bodes well for the development of targeted treatments for diabetic retinopathy.

3 Methodology

3.1 Proposed Methodology

Figure 6: CRISP DM methodology for DR detection (Source: ResearchGate)

The Cross Industry Standard Process for Data Mining, often known as CRISP-DM, is a process model that is used as the basis for a data science process. Let's talk about the actions that need to be taken in order to diagnose diabetic retinopathy (Azevedo, A., and Santos, M.F., 2008).

3.1.1 Business Understanding

High blood glucose levels due to inadequate insulin production define diabetes mellitus, which has increased from 108 million cases in 1980 to 422 million cases now. This disease negatively impacts several body systems, including the eyes, liver, kidneys, and joints. In 1980, 108 million people were diagnosed with diabetes. Regular eye exams, especially during pregnancy, are crucial for preventing diabetic retinopathy—a diabetes-related vision disorder caused by blocked retinal blood vessels due to high blood sugar. Defective vessels form, risking leakage, hemorrhage, retinal issues, and glaucoma. Timely detection and treatment are essential to halt irreversible vision loss. Timely diagnosis using resources like coloured fundus images is crucial. Because of the irreversible nature of later stages, Atwany, M.Z., Sahyoun, A.H., and Yaqub, M. 2022 stress the need of early diagnosis and therapy.

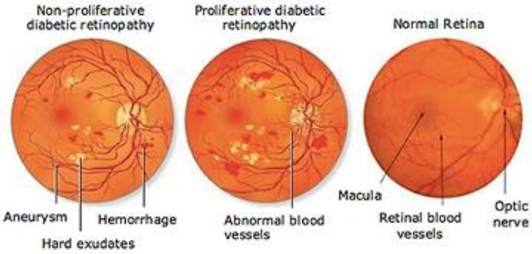

Figure 7: Different Stages of Diabetic retinopathy (Atwany, M.Z., Sahyoun, A.H., and Yaqub, M. 2022)

Damage to the retinal blood vessels from hyperglycemia makes this the greatest preventable cause of blindness globally. The illness can lead to blindness due to diabetic retinopathy, which progresses from "normal" to "mild," "moderate," "severe," and "proliferative" phases.

3.1.2 Data understanding

EyePACS (Source: https://www.chcf.org/[1]) provides researchers with access to a vast collection of retinal pictures captured using a variety of imaging techniques and shown in high resolution. There are always two valid points of view on every given issue. The left eye image for patient id 1 is named 1_left.jpeg, whereas the right eye image is named 2_right.jpeg. A clinician checked the data labelled by using a scale from 0 to 4 to determine if each image displayed symptoms of diabetic retinopathy.

0 No DR
1 Minor 2 Somewhat Strong

3 Very Critical Level 4 DR with Proliferation

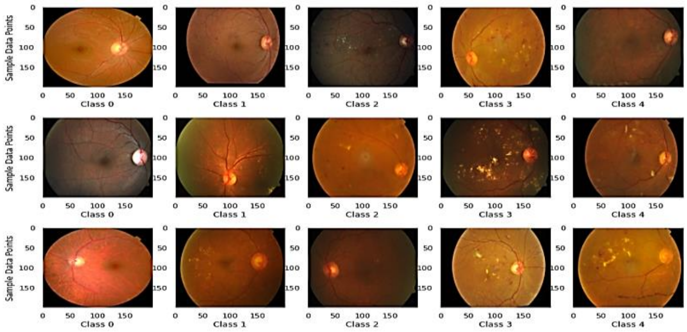

Figure 8: Random 15 images of the dataset depicting the different classes

Because of the wide range of camera brands and models used to capture these images, the left and right halves of certain photos may convey entirely different emotions.

[1] Source: https://www.chcf.org/

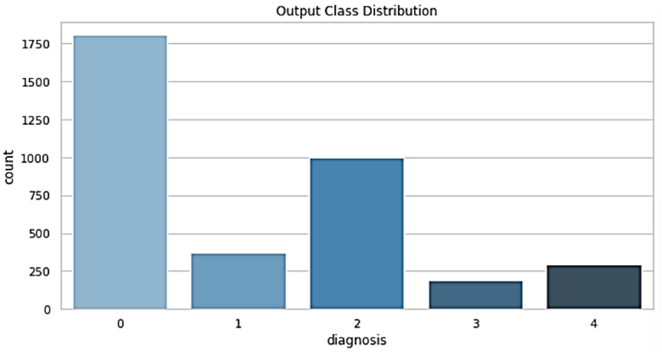

Figure 9: Count of different Classes present in the dataset

3.1.3 Data Preparation

The collection of data is a crucial first step in the process of creating a reliable model for the diagnosis of diabetic retinopathy. Here is a selection of well-known methods applied to the process of data preparation:

3.1.3.1 Data Cleaning

Data with unevenly distributed histogram values was removed, alongside the pre-existing duplicates. This step was essential to rectify the skewed nature of the dataset. To address noisy data points resulting from measurement errors, sensor artefacts, or other conditions, smoothing techniques like filtering were applied. Consequently, these interventions ensured the management of any noisy data points.

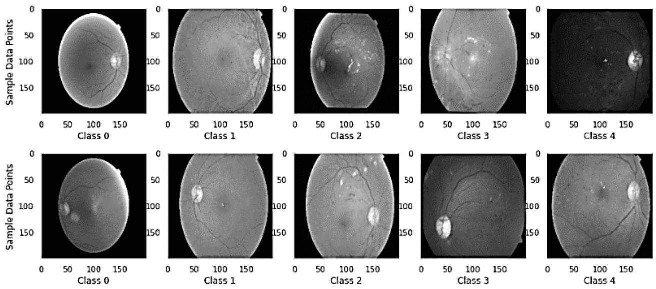

Figure 10: Gray Scaling to the original Dataset

3.1.3.2 Data Preprocessing

The retinal images are reduced during the resizing process until they were of the same resolution. This is essential for reaching a state of uniformity. By transforming the pixel intensity values to a uniform scale (such [0, 1] or [-1, 1]), and "normalised" an image by getting rid of the noise caused by tonal variations. Using filters or denoising methods (such median filtering and Gaussian blurring) on retinal images might improve image quality by removing noise. Denoising accomplished this goal.

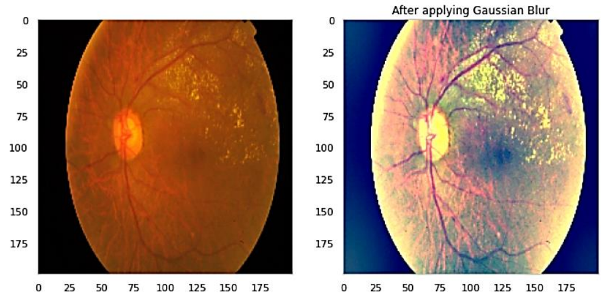

Figure 11: Gaussian blurring (Using Python)

$$f(i,j) = \frac{1}{2\pi\sigma^2} e^{-\frac{i^2+j^2}{2\sigma^2}}$$

3.1.3.3 Modelling and validation Data

From the cleaned-up dataset, training, validation, and test sets were established. A tiered approach was employed to guarantee a fair distribution of data among the different categories. This strategy was meticulously implemented, considering both the characteristics of the dataset and the specific demands of the diabetic retinopathy detection task. To fine-tune the parameters of the pre-processing pipeline, it was essential to continuously analyse how each stage influenced the performance of the model. The optimal analysis was then retained for the full model execution.

3.1.4 Modelling

3.1.4.1 Convolutional Neural Networks (CNN)

The layers of a Convolutional Neural Network (CNN) are coupled and are programmed to learn certain data attributes. Input, intermediate, and output layers are all part of this structure. The model relies heavily on key layers including convolution, ReLU activation, and pooling. Convolution adds filters to the input images, ReLU boosts training efficiency, and pooling downsamples the results for clarity. These actions are carried out in stages, with each one revealing a new set of traits. CNN's shared weights between layers allow for object translation tolerance. Class probabilities are then generated by a fully connected layer when feature extraction is complete. When it comes down to it, the classification layer is what determines what goes where. According to Gu et al. (2018), CNN shows excellent performance when it comes to the efficient learning and classification of visual data.

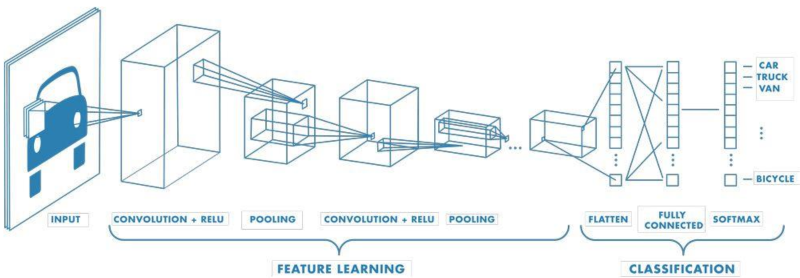

Figure 12: An illustration of a network that contains multiple convolutional layers. Each training image has a variety of filtering operations performed on it at a variety of resolutions, and the result of each convolved image is utilised as the input to the subsequent layer (O'Shea, K. and Nash, R., 2015).

3.1.4.2 Transfer learning

Transfer learning involves utilizing models that have been trained on extensive image datasets and then adapting them to detect diabetic retinopathy. This approach leverages the knowledge and representations from previously trained models for analogous tasks, like image categorization. Such a task might involve sorting photographs into various categories. Using this method eliminates the need to commence the training process from the beginning.

3.1.4.3 VGGNet

VGGNets are based on convolutional neural networks (CNN), and include many of the same key features into their own architecture. What follows is a schematic showing the simplest possible implementation of a convolutional neural network:

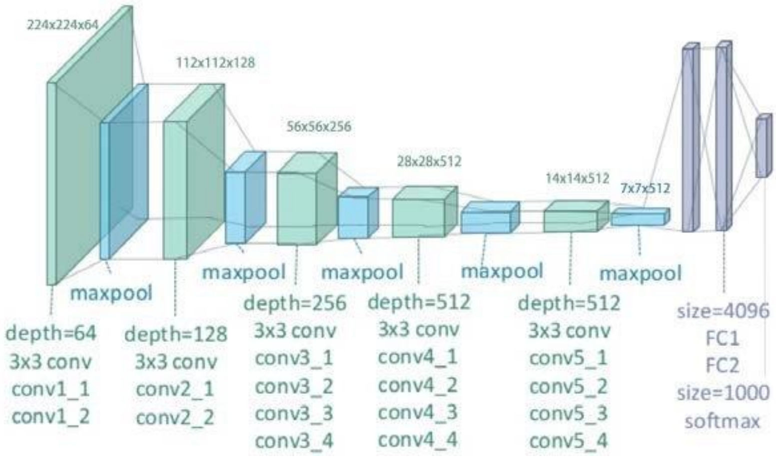

Figure 13: VGGNet (O'Shea, K. and Nash, R., 2015)

When requested to do its job, the VGG-based convNet receives a 224-by-224-pixel RGB picture as input. The preprocessing layer begins by subtracting the mean image values obtained from the whole ImageNet training set from an RGB image with pixel values ranging from 0 to 255. The VGGNet consists of three layers, or stages, with links between them. The first two tiers have 4096 channels each, whereas the third tier has just 1000 channels and gives each of the 1000 categories one channel.

3.1.4.4 DenseNet

DenseNet is based on a central notion, which is to construct dense blocks by combining feature maps from all previous layers. This allows each layer to benefit from the qualities of the ones behind it. Unlike traditional CNNs, it connects each dense block's convolutional layer, which allows it to build faster connections to lower layers. With the help of transition layers, networks may grow while maintaining a more compact structure.

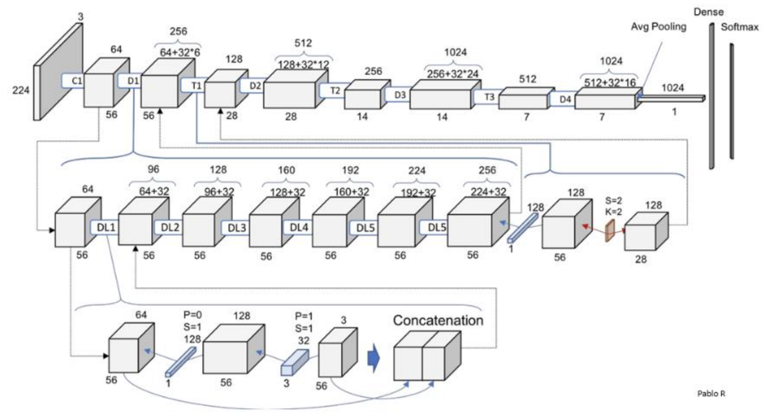

Figure 14: DenseNet (O'Shea, K. and Nash, R., 2015)

The DenseNet structure stands out for its ability to optimise feature use while minimising parameter count, producing excellent performance across a wide range of computer vision workloads. Notable applications include, among others, categorising photos, identifying objects, and segmenting images based on significance.

3.1.4.5 MobileNet

MobileNet is a versatile and efficient CNN framework. MobileNet use depthwise separable convolutions to generate lightweight models. This is accomplished by adding width and resolution multipliers as a means of negotiating between delay and accuracy. Pointwise and depth-wise convolutions are executed in depth-separable convolution layers. After a MobileNet reaches 28 layers, the width multiplier can be used to reduce the number of parameters. Images must be 224 pixels wide by 224 pixels tall by 3 pixels thick.

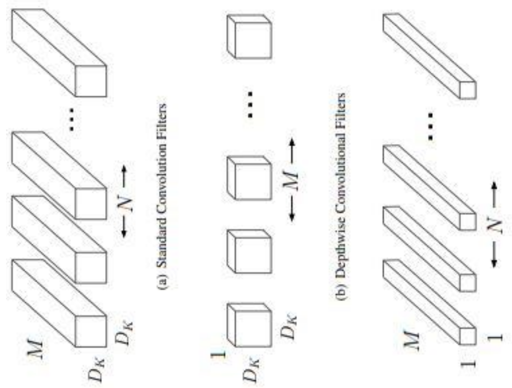

Figure 15: MobileNet Architecture (a) Normal Convolutional Filters (b) Depthwise Convolutional Filters (O'Shea, K. and Nash, R., 2015)

3.1.4.6 ResNet

ResNet, a Convolutional Neural Network (CNN) for computer vision tasks, addresses the vanishing gradient problem of deep networks. By introducing skip connections, ResNet overcomes gradient diminishment. It employs identity mappings, skipping layers and reusing activations, accelerating initial training. Subsequent retraining enlarges and empowers residual parts, expanding the feature space exploration in input images.

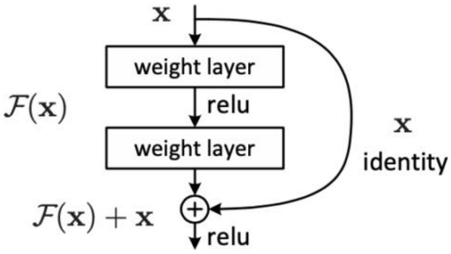

Figure 16: Residual Learning Block (O'Shea, K. and Nash, R., 2015)

3.1.4.7 ViT

Vision Transformer (ViT) competes with CNNs in computer vision tasks, offering nearly fourfold efficiency and comparable accuracy. ViT models challenge CNN dominance, showcasing impressive performance even with less processing resources. Derived from text-focused transformers, ViT predicts image class labels by representing images as patches, outperforming state-of-the-art CNNs with significantly fewer CPU resources.

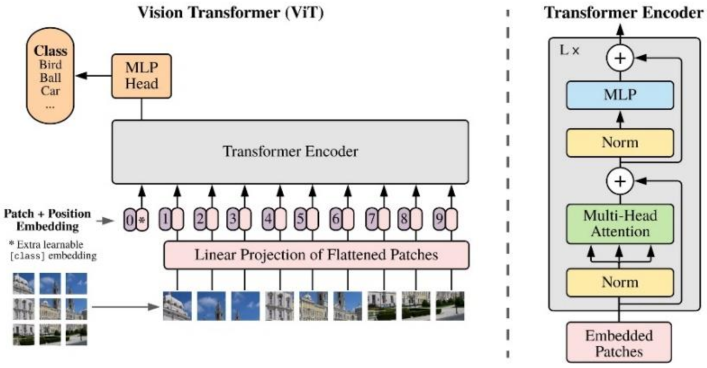

Figure 17: ViT Architecture (Source: Medium)

3.2 Evaluation

Models are selected by a procedure known as "greedy selection." The models' relative merits are used to determine which one will be put into use in each specific area. In this setting, use accuracy, specificity, and sensitivity as measures of the model's efficacy. Due to the skewed nature of the data in the healthcare sector, precision is not viewed as a crucial indicator; so, once the therapy is finished, it can also be selected. Despite the fact that sensitivity and specificity are both crucial indicators, they are often measured separately.

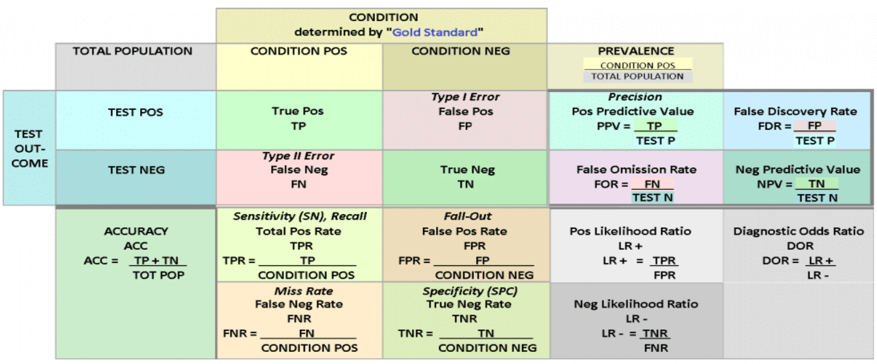

Figure 18: Confusion Matrix (Source: TowardsDataScience)

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

$$Speificity = \frac{True\ Negatives}{True\ Positives + False\ Positives}$$

TP Rate (Recall)	=	TP/(TP + FN)
FP Rate	=	FP/(FP + TN)
Specificity	=	1 – FP Rate
Precision	=	TP/(TP + FP)
Accuracy	=	(TP + TN)/(TP + TN + FP + FN)
F1 (F-Score)	=	2 * (Precision * Recall)/(Precision + Recall)

Kappa Score: For qualitative (categorical) items, inter-rater (and intra-rater) reliability may be calculated using Cohen's kappa coefficient. Due to its ability to account for the coincidental nature of an agreement, is widely regarded as a more reliable metric than a straightforward percentage of agreement. As with other indices of agreement, Cohen's kappa is fraught with interpretational difficulties and has sparked much debate. Disagreement between objects has been argued to be conceptually simpler by certain academics.

$$Kappa = \frac{P_0 - P_e}{1 - P_e}$$

Where,P_0 is the observed observations and P_e is the expected observations.

3.3 Deployment

The trained model is stored, and it is possible to deploy it in the form of an executable file.

4 Design Specifications

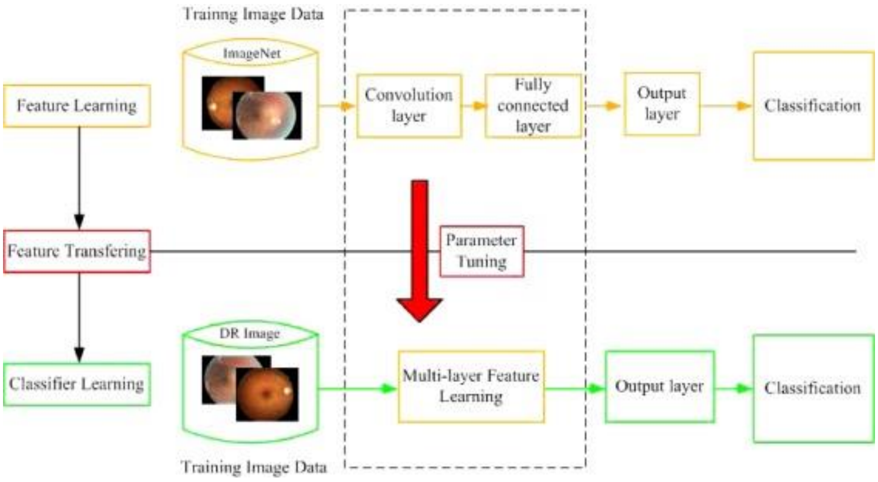

Figure 19: Modelling Pipeline for the DR detection

The above pipeline can be broken down as below,

Algorithm Flow:

1. **Data Collection:** Begin by gathering photos. This may be accomplished via autonomous channels, cooperation with medical research organizations, or the use of widely recognized databases. In this case, use generic datasets.

2. **Image Enhancement:** Next, improve the photographs for model training. Adjusting picture dimensions, adding minor modifications for improved model flexibility, scaling image values, and splitting the dataset for training and validation are all part of this process. For this, use two methods:

   a. **Proportional Split:** This ensures that the class distribution is uniform throughout both the Training and Validation sets.

   b. **Stochastic Selection:** Training and Validation images are chosen using a probabilistic technique.

3. **Feature Derivation:** Use a collection of well-known Transfer Learning architectures such as VGGNet, MobileNet, DenseNet, ResNet, and ViT. These are critical in generating the critical feature vectors that will be used as input for the final classification step.

4. Finally, use a neural network structure that interconnects layers to complete the final classification. This network has been painstakingly trained on the whole selected dataset.

5 Results and Analysis

This section explores further into assessing multiple model architectures in order to discover the most efficient one for quick detection and diagnosis of diabetic retinopathy across diverse systems. Aside from recognizing diabetic retinopathy, this system provides insights into major data areas that are critical for decision-making.

Table 1: Performance Comparison of different Architectures on retina image dataset

Model	Accuracy	Kappa score
ResNet50	0.75	0.77
VGG-19 (Nair, A.T., Anitha, M.L. and Kumar, A., 2022, March)	0.77	0.776
Dense net	**0.78**	**0.82**
Mobile net	0.77	0.80
VIT	0.72	0.73

According to the published performance measures in table 1, the kappa scores for models such as ResNet-50, VGG-19, DenseNet, MobileNet, and Vision Transformer are 77.0%, 77.6%, 82.0%, 80.0%, and 73.2%, respectively. Based on these numbers, DenseNet clearly outperforms the competition. The contrasting view of the various confusion matrices is shown below,

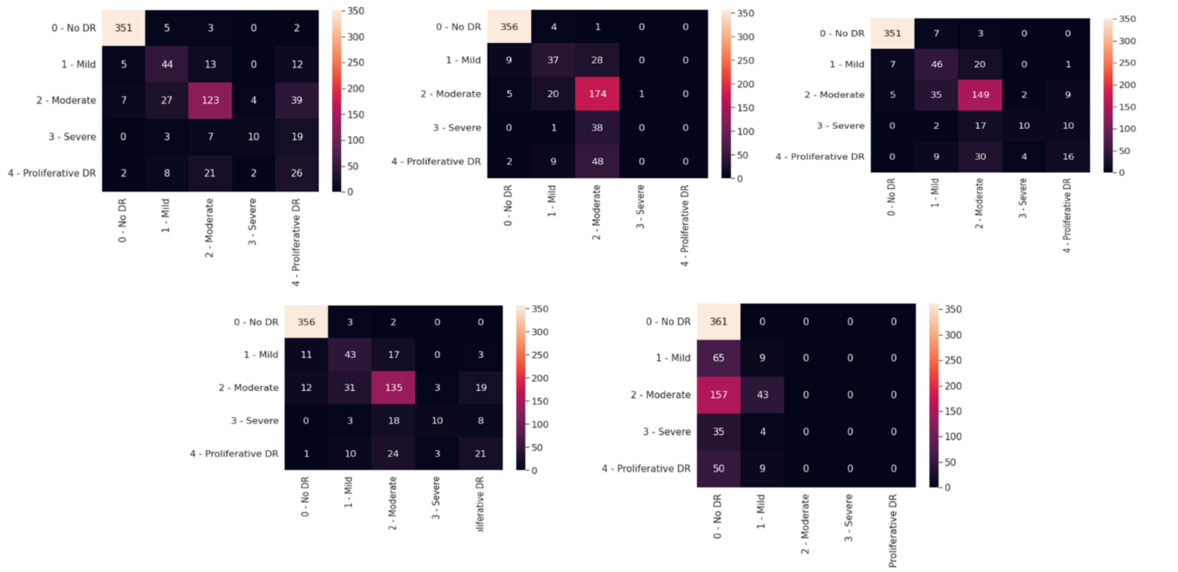

Figure 20: (a) ResNet50 (b) VGG19 (c) Densenet (d) Mobilenet (e) VIT model's Confusion matrix

The count is 9,146,948 when the trainable parameters for DenseNet are examined. When the DenseNet performance trajectory is plotted against the epoch variation in accuracy resulting from these parameters, it indicates a remarkable training accuracy of 92.56% by the 7th epoch. This tendency continues, culminating at 92.56% by the 10th epoch, as seen in the graph below. According to the study of Nair, A.T., Anitha, M.L., and Kumar, A. (2022, March), VGG-19 has an accuracy of 77.00% and a Kappa Score of 0.776. When compared, DenseNet's performance was better. The next section delves into the performance graphs connected with the DenseNet model.

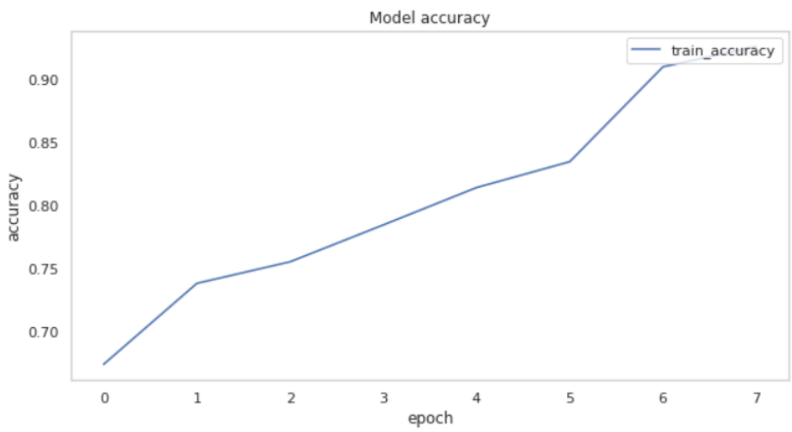

Figure 21: The fluctuation in accuracy relative to the epoch count for DenseNet.

As additional training samples are supplied, learning curves graphically illustrate the growth of training and validation losses. These curves indicate if including additional training data may improve model validation. When a model overfits, adding additional training data may improve its predictions on unfamiliar data. On the other hand, if a model underperforms, adding additional data may not help. The arcs in the loss graph seem to be well matched, with the validation loss slightly surpassing the training loss. The accuracy trajectory also shows a continuous increase during training, indicating that the neural network, especially with the DenseNet framework, outperforms alternative designs in diabetic retinopathy prediction. The next sections depict the changes in validation and training losses, as well as the accuracy metrics, across the epochs.

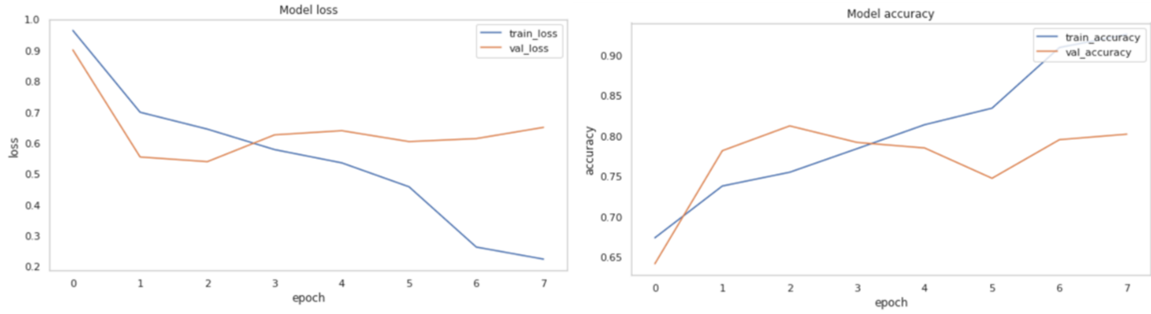

Figure 22: (a) Epochs Vs Loss (b) Epochs Vs Accuracy for DenseNet

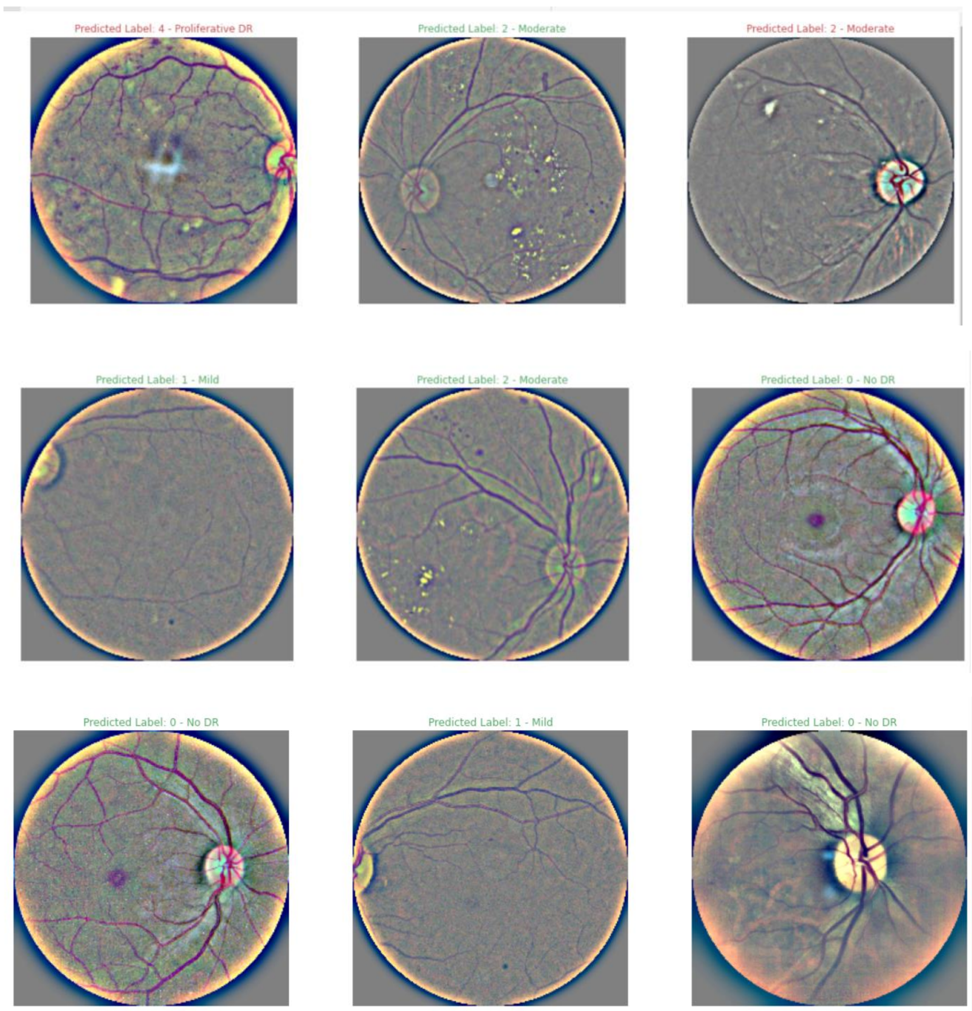

Figure 23: Test Features representation for DenseNet

The figure above shows that the model correctly anticipated traits, which are shown in green, whereas mispredictions are marked in red. Notably, DenseNet erroneously interpreted the characteristics of two photos.

5.1 Discussion

A system that makes use of Deep Learning Algorithms, such as CNN, ANN, and Transfer Learning, has a significant potential for application in medical settings all over the world. The research examined five different algorithms while concentrating on retinal pictures. DenseNet, MobileNet, and VGG-19 were all put through their paces, but the results showed that DenseNet had the edge. The results of the Kappa test for the different models ranged from 73.2% (Vision Transformer) to 82.0% (DenseNet), demonstrating the efficiency of the DenseNet model. The research used VGGNet as a point of reference and attempted to build a Greedy Selection ensemble model. This model, which was powered by DenseNet, emerged as the best performer, validating its resilience for the detection and classification of retinal images. Models of deep learning that are based on ensembles perform very well because of the varied views they include, which in turn reduces the risk of overfitting, improves error correction, and makes it easier to solve problems. They are capable of capturing ambiguity, providing stability, and improving generalization as a result of integrating capabilities that are complimentary. Even though they need a lot of processing power, ensembles are able to leverage collective intelligence to achieve higher performance across a wide range of challenging tasks. A thorough examination of several transfer learning models integrated with three-layer deep artificial neural networks was conducted. The Vision Transformer (ViT) model was also investigated. Notably, this approach provides insights not seen in any of the research referenced.

6 Conclusion

Systems that use Machine Learning (ML) and Artificial Intelligence (AI) outperform conventional manual approaches in terms of accuracy. This provides a substantial benefit for healthcare practitioners worldwide, since they may use Deep Learning methods such as Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Transfer Learning (TL) for both detection and classification tasks. The key dataset for this investigation was retinal pictures. This data was subjected to five distinct algorithm tests. DenseNet was the best performance, closely followed by MobileNet and VGG-19. The Kappa values for VGG-19, ResNet-50, DenseNet, MobileNet, and Vision Transformer were 77.0%, 77.6%, 82.0%, 80.0%, and 73.2%, respectively. These findings demonstrate DenseNet's higher performance. In compared to a prior study work, the suggested model yielded promising findings. However, computational difficulties were encountered throughout the procedure.

6.1 Future Work

Vision Transformers are likely to be effective for the Diabetic Retinopathy dataset, given the design of their architectures. Proper tuning for this DR dataset is essential and represents the next phase of the work. For further research, this study will be extended to other datasets to ensure the variability of the work is considered.

References

Abreu, A., Ferreira, D., Neto, C., Abelha, A. and Machado, J., 2021, February. Diagnosis of Diabetic Retinopathy Using Data Mining Classification Techniques. In International Conference on Advances in Digital Science (pp. 198-209). Springer, Cham.

Alabdulwahhab, K.M., Sami, W., Mehmood, T., Meo, S.A., Alasbali, T.A. and Alwadani, F.A., 2021. Automated detection of diabetic retinopathy using machine learning classifiers. European Review for Medical and Pharmacological Sciences, 25(2), pp.583-590.

Atwany, M.Z., Sahyoun, A.H. and Yaqub, M., 2022. Deep Learning Techniques For Diabetic Retinopathy Classification: A Survey. IEEE Access.

Bora, A., Balasubramanian, S., Babenko, B., Virmani, S., Venugopalan, S., Mitani, A., de Oliveira Marinho, G., Cuadros, J., Ruamviboonsuk, P., Corrado, G.S. and Peng, L., 2021. Predicting the risk of developing diabetic retinopathy using deep learning. The Lancet Digital Health, 3(1), pp.e10-e19.

Costa, P., Araújo, T., Aresta, G., Galdran, A., Mendonça, A.M., Smailagic, A. and Campilho, A., 2019, May. EyeWes: weakly supervised pre-trained convolutional neural networks for diabetic retinopathy detection. In 2019 16th international conference on machine vision applications (MVA) (pp. 1-6). IEEE.

Farooq, M.S., Arooj, A., Alroobaea, R., Baqasah, A.M., Jabarulla, M.Y., Singh, D. and Sardar, R., 2022. Untangling computer-aided diagnostic system for screening diabetic retinopathy based on deep learning techniques. Sensors, 22(5), p.1803.

Filos, A., Farquhar, S., Gomez, A.N., Rudner, T.G., Kenton, Z., Smith, L., Alizadeh, M., De Kroon, A. and Gal, Y., 2019. A systematic comparison of bayesian deep learning robustness in diabetic retinopathy tasks. arXiv preprint arXiv:1912.10481.

Imran, M., Ullah, A., Arif, M. and Noor, R., 2022. A unified technique for entropy enhancement based diabetic retinopathy detection using hybrid neural network. Computers in Biology and Medicine, p.105424.

Kassani, S.H., Kassani, P.H., Khazaeinezhad, R., Wesolowski, M.J., Schneider, K.A. and Deters, R., 2019, December. Diabetic retinopathy classification using a modified xception architecture. In 2019 IEEE international symposium on signal processing and information technology (ISSPIT) (pp. 1-6). IEEE.

Khan, S., Naseer, M., Hayat, M., Zamir, S.W., Khan, F.S. and Shah, M., 2021. Transformers in vision: A survey. ACM Computing Surveys (CSUR).

Masud, M., Alhamid, M.F. and Zhang, Y., 2022. A Convolutional Neural Network Model Using Weighted Loss Function to Detect Diabetic Retinopathy. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 18(1s), pp.1-16.

Mohanty, K.K., Barik, P.K., Barik, R.C. and Bhuyan, K.C., 2019, December. An efficient prediction of diabetic from retinopathy using machine learning and signal processing approach. In 2019 International Conference on Information Technology (ICIT) (pp. 103-108). IEEE.

Nair, A.T., Anitha, M.L. and Kumar, A., 2022, March. Disease Grading of Diabetic Retinopathy using Deep Learning Techniques. In 2022 6th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1019-1024). IEEE.

Panwar, A., Semwal, G., Goel, S. and Gupta, S., 2022. Stratification of the lesions in color fundus images of diabetic retinopathy patients using deep learning models and machine learning classifiers. In Edge Analytics (pp. 653-666). Springer, Singapore.

Qin, Z., Zhang, Z., Chen, X., Wang, C. and Peng, Y., 2018, October. Fd-mobilenet: Improved mobilenet with a fast downsampling strategy. In 2018 25th IEEE International Conference on Image Processing (ICIP) (pp. 1363-1367). IEEE.

Ramasamy, L.K., Padinjappurathu, S.G., Kadry, S. and Damaševičius, R., 2021. Detection of diabetic retinopathy using a fusion of textural and ridgelet features of retinal images and sequential minimal optimization classifier. PeerJ computer science, 7, p.e456.

Sarki, R., Michalska, S., Ahmed, K., Wang, H. and Zhang, Y., 2019. Convolutional neural networks for mild diabetic retinopathy detection: an experimental study. bioRxiv, p.763136.

Targ, S., Almeida, D. and Lyman, K., 2016. Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029.

Tsao, H.Y., Chan, P.Y. and Su, E.C.Y., 2018. Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms. BMC bioinformatics, 19(9), pp.111-121.

Wan, S., Liang, Y. and Zhang, Y., 2018. Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers & Electrical Engineering, 72, pp.274-282.

Wang, J., Luo, J., Liu, B., Feng, R., Lu, L. and Zou, H., 2020. Automated diabetic retinopathy grading and lesion detection based on the modified R-FCN object-detection algorithm. IET Computer Vision, 14(1), pp.1-8.

Wang, L., Guo, S., Huang, W. and Qiao, Y., 2015. Places205-vggnet models for scene recognition. arXiv preprint arXiv:1508.01667.

Zhu, Y. and Newsam, S., 2017, September. Densenet for dense flow. In 2017 IEEE international conference on image processing (ICIP) (pp. 790-794). IEEE.