National College of Ireland

Diabetic Retinopathy stages Classifications using Lesion features and CNN Models

MSc Research Project
Data Analytics

Sangeetha Pillay Anil Kumar
Student ID: 20232195

School of Computing
National College of Ireland

Supervisor: Prof. Qurrat Ul Ain

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Sangeetha Pillay Anil Kumar
Student ID:	20232195
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Prof. Qurrat Ul Ain
Submission Due Date:	19/09/2022
Project Title:	Diabetic Retinopathy stages Classifications using Lesion features and CNN Models
Word Count:	5085
Page Count:	20

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:	
Date:	18th September 2022

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):	

Diabetic Retinopathy stages Classifications using Lesion features and CNN Models

Sangeetha Pillay Anil Kumar
20232195

Abstract

It is generally known that diabetic retinopathy (DR) is one of the most prevalent global causes of vision loss in people between the ages of 25 to 74. Because (DR) may initially cause no symptoms or only very minor vision problems, predicting DR in the early stages is crucial. The ultimate aim of this study is to classify the stage of DR such as mild, moderate, proliferate and severe. This study makes use of two CNN models; the first will determine whether there is DR present or not, and the second will classify the DR picture according to its stage. The model made use of the lesion feature to aid in this classification. This was compared with a model that does not classify the stages of DR using the lesion feature (DR Stage Classification). In 10 epochs, the VGG-16 model of the DR stage classification model provided training accuracy of 56.77% and test or validation accuracy of 54.73%, whereas the VGG-19 model produced training accuracy of 60.28% and test or validation accuracy of 54.04%. And the Lesion Feature DR Stage Classification model outperformed both the VGG-16 and VGG-19 models, with training accuracy of 90.88% and test or validation accuracy of 52.03 % and training accuracy of 95.29% and test or validation accuracy of 48.65% in 250 epochs, respectively.

1 Introduction

1.1 Background

Diabetic Retinopathy is a common diabetic consequence seen diabetic patients with 12 years or more duration. It's caused by damage to the blood vessels of the retina, which is the light-sensitive tissue in the rear of the eye. If left undiagnosed and untreated, it can result in blindness.Hence it can be treated if it is diagnosed in the early stage. So when it comes to diagnosing DR in its early phases, clinical methods use fundus picture (Safi et al.; 2018). The ophthalmologist analyses the color fundus image and then assesses the patient's condition in the current clinical diagnosis. This detection is time-consuming and labor-intensive, resulting in increased inaccuracy. There is a chance that many patients with DR will not get timely diagnosis and treatment due to the high volume of DR patients who are growing daily and the lack of public healthcare resources in some regions, which could lead to untreated opportunities and eventually unrecoverable vision loss (Zago et al.; 2020). AIn order to provide prompt treatment, it was realized how important it is to automate the classification of DR phases.

1.2 Motivation

Diabetic Retinopathy (DR) is now one of the most common causes of vision loss. Currently 93 million people live with vision-threatening DR. The Centers for Disease Control and Prevention estimates that by 2050, 16.0 million people would have diabetic retinopathy and 3.4 million will have vision-threatening DR (Group*; 2004).
As was previously mentioned, the existence of lesion is the primary and visible symptom of DR. As a result, the primary goal of this study is to categorize the stages of DR and compare them to another model that incorporates lesion features. Figure 1 depicts the stages of DR in fundus image and Table 1 shows the type of lesions present in each DR stage.

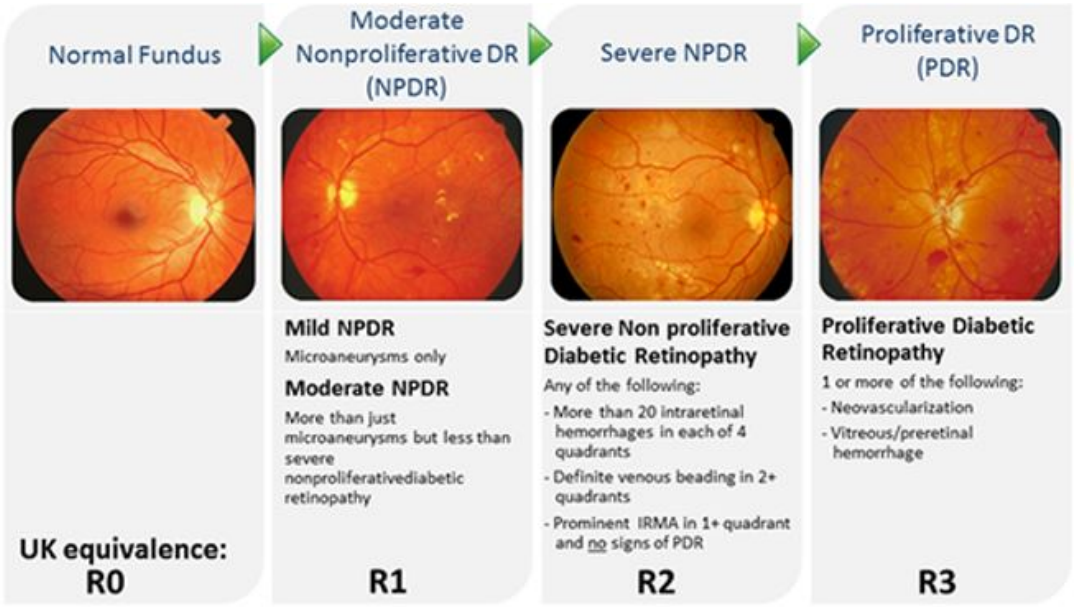

Figure 1: Stages of DR[1]

Table 1: Stages of DR and the presence of Lesions[2]

DR stages	**Type of Lesion**
No DR	No lesions
Mild DR	Microaneurysms
Moderate DR	More than just microaneurysms but less than severe DR
Severe DR	Any of the following: more than 20 intraretinal haemorrhage in each of 4 quadrants; definite venous beading in 2+ quadrants; Prominent intraretinal microvascular abnormalities in 1+ quadrant and no signs of proliferative DR
Proliferative DR	One or more of the following: neovascularization, preretinal haemorrhage

[1]https://s3.ap-south-1.amazonaws.com/boost-content-cdn/CustomPages/Images/36d3ec27.png
[2]https://www.mdpi.com/1424-8220/21/11/3704/htm

Although numerous studies and papers have presented automated approaches for the classification of stages of DR, this study differs from others by first detecting the existence of DR and, if the image contains DR, categorizing the stages of DR using the lesion features in it.

1.3 Research Question

1. How successfully can deep learning and image processing be utilized to detect the stage of DR with the presence of lesion?

2. How can Deep Learning technique be used to improve and solve the problems associated with manual classification of DR stages with the presence of lesion?

1.4 Research Objective

The goal of this study is to automate the stages of DR classification to reduce manual effort and give patients who are suffering from DR timely treatment. So as to lower the chance of visual loss.

1.5 Outline of Research Paper

There are five sections in this paper. The introduction is where the study starts. A review of studies on DR detection and classification is discussed in the second section. The research methodology and design criteria are covered in the next section, which is followed by the results and conclusion of the suggested model and study. My research questions are answered in the affirmative because this model's training duration is significantly shorter than that of earlier studies.

2 Related Work

2.1 Overview

A thorough evaluation of the literature is done in this section. It includes an overview of earlier research on the identification and classification of DR. In the corresponding papers, the methodologies are examined and discussed. This review is conducted in the automated techniques to detect and classify the stages of DR.

2.2 Detection of Diabetic Retinopathy

A robust diagnostic model was created using an ensemble-based model of learning that incorporated numerous well-known classification methods was discussed in Group* (2004). InfoGainEval and WrapperSubsetEval picked the top 5 and top 10 features of the Messidor dataset with accuracy of 70.7% and 75.1%. For medical use, this precision is insufficient. As a result, the ensemble model is inadequate for DR detection using fundus pictures.
A thorough analysis of works discussing DR, its features, root causes, difficulties and also discussing about ML models, comparisons, cutting-edge DL models, and future perspectives for DR detection was done in Odeh et al. (2021). In order to detect DR, this work exposed me to image processing using DL and ML algorithms similar in figuring

out the DR's stages. A comparison of ML and DL techniques was also conducted, and it became evident that ML techniques are significantly less scalable when dealing with high-dimensional data and take longer to develop than DL techniques. Hence I started focusing more on DL techniques in my research.
The issue of low quality fundus images and identify retinopathy symptoms was addressed in Nneji et al. (2022). The author employed WFDLN to process two channels of fundus pictures: contrast-limited adaptive histogram equalization (CLAHE) and contrast-enhanced canny edge detection (CECED) fundus images. In a weighted methodology, the outputs of both channels are combined, and the final recognition result is calculated using softmax classification. In contrast to the proposed model's overall sensitivity of 98.7%, specificity of 97.8%, and accuracy of 98.0% on the Kaggle dataset, the author's proposed model was tested on the Messidor dataset and reports sensitivity of 98.9%, specificity of 98.0%, and accuracy level of 98.5%.
The five severity levels of diabetic retinopathy are automatically recognized using a unique technique in Aljehane (2022) research, without the need for pre- or post-processing on retinal fundus pictures (DVFs). A new compressed layer, fine-tuning stages, and a semi-supervised multilayer deep learning method were used by the authors. This SLDR system's area under the ROC, specificity (SP), and sensitivity (SE) were assessed and contrasted with state-of-the-art methods (AUC). The lack of data pre- or postprocessing is, in my opinion, a drawback. To describe the photos, the author merely makes use of the gradient location-orientation histogram and the dense color scale-invariant feature transform (DColor-SIFT) (GLOH). Using a contrast enhancement method, I want to enhance.
Research on using deep learning and machine learning to diagnose eye diseases such glaucoma, ARMD, cataracts, and DR was assessed in Vyas and Khanduja (2021). This article's author looked at about 5 DR detection-related works. It described five different detection strategies in five studies. I kept researching the approaches mentioned in two of these five papers as my interest in them grew. To find DR, one must first pinpoint the lesion. I chose this subject for my research after reading this paper.
A CNN model to tackle the three most challenging problems in DR detection: classification, segmentation, and detection was employed in Wan et al. (2018). As a result, they combine transfer learning and hyper-parameter tweaking with ResNet, GoogleNet, VggNet, and AlexNet. Due to the amount of noise in the data and the scarcity of data sets, normalizing techniques such baseline normalization and denoising, as well as data augmentation, were employed to comprehensively preprocess the data. Although testing was conducted using the same test data, the results appear to be accurate.
In order to categorize DR in OCTA images and their related co-registered structural images, ensemble learning techniques in conjunction with deep learning are examined in Heisler et al. (2020). The superficial and deep vascular complexes' en face OCTA pictures were manually extracted before exporting, and system software was used to eliminate projection artifacts. The results of this hand extraction are unreliable and take a lot of time. The original20 and a variation known as Grad-Cam were the two class activation maps that the authors compared. This was utilized to identify the image regions that would be most helpful in classifying the image. For an ophthalmologist, making a diagnosis from OCT pictures is challenging and requires more experience. Therefore, I opt for the approach that involves using fundus photographs.
In order to identify DR and categorize its stages, authors of this (Houby; 2021) used CNN models using colored retinal fundus images. They applied the pre-trained VGG-16

CNN model for this work using a transfer learning (TL) strategy. The best-achieved accuracies for classifications into five classes, four classes, three classes, and two classes are, respectively, 73.7, 63.5, 80.5, and 86.5. This study and that are similar. However, the use of the lesion feature was included in this study.

2.3 Diabetic Retinopathy Detection with the presence of Lesion

A fully patch based CNN model for localizing retinal lesions was proposed by the authors of Zago et al. (2020). Two effective CNN models are included in this, which has improved prediction accuracy. The sensitivity for DR screening is 0.940-0.95 percent confidence interval (CI) and the ROC score is 0.912-. (0.940-0940). (0.921-.959). The author preprocessed the data as the initial stage of data processing by using an enhancement technique. This I used as a model input for my extraction of lesion features. The author divided each patch retrieved from a given fundus image into 2 categories for training the model: lesion or non-lesion using patch-based labeling. All lesion pixels and the same number of randomly selected non-lesion pixels are selected for each image in the training set, or dataset of just lesion images, in order to fit a five-layer CNN known as the selection model for a constrained number of epochs. This model converges quickly because the dataset with the preselected patches is balanced in terms of lesion and non-lesion pixels. The outcomes demonstrated that the VGG16 model performed well due to its extensive generalizability. The paper's main asset is the patch-based dataset. The advantage of this paper was that 28 images with pixel-level annotations were used for the model training, which reduced the burden. Which is also a drawback due to the poor classification seen in the initial CNN networks.

An automated approach for categorizing and localizing various research types in fundus images utilizing a small training set without deep feature extraction as a first step to generalizing these approaches to abnormal illness identification was proposed in Lam et al. (2018). VGG16, AlexNet, GoogLeNet, ResNet, and Inception-v3 were among the five CNN models used for the test. Utilizing a sliding window, the trained CNN was applied to the full scan to generate a multiclass probability distribution of the aforementioned pathologies over the image. The efficient parameter-to-performance ratio of GoogleLeNet-v1 led to its selection for the sliding window. With this outcome, lesion detection was highly accurate.

DenseNet121, Xception, ResNet50, and MobileNet were evaluated as models for automatic DR classification on fundus images in delaPava et al. (2021). This technique first pinpoints the primary ocular abnormalities associated with DR before diagnosing the condition. The model worked exceptionally well, and the outcomes were on par with those of cutting-edge methodologies.

2.4 Random Forest Variable Importance

Microaneurysms (MA) are difficult to identify automatically for three reasons. The low contrast between the lesion and the vast fluctuations in color, the tiny size of MA lesions and retinal backdrop, and the high prevalence of false positives in areas with comparable intensity values, such as noises, blood vessels, and non-homogeneous background, are the first three factors. Hence in Cao et al. (2017) Small 25×25 pixel patches taken from fundus images were used to test the MA's possibility to be detected.

To help doctors correctly diagnose retinal illnesses, the authors of the Chowdhury et al.

(2019) presented a Random Forest classifier. Age-related macular degradation and diabetic retinopathy are the two conditions that induce the abnormalities in the retina that are categorized in this research. This model, which had a 93.58% accuracy rate, assisted me in creating a lesion feature model.
Although popular machine learning techniques like KNN, SVM, and neural networks are beneficial for classification, they do not give information on the factors that best contribute to the predictive model. However, random forests (RF) do not necessitate that the predictor space be reduced before classification. The efficiency of RF variable importance metrics in finding the real predictor among a large number of candidate predictors is therefore examined in this research.

2.5 Conclusion

A thorough literature review was completed before the modelling this study.The literature review led to the following conclusions:

- The decision to use fundus image for this study was made after it was found that processing the image for the classification of DR was fast and easy.
- It was discovered that the first and most visible indication of DR is the existence of a lesion.
- In the detection of DR, the CNN models VGG-16 and VGG-19 performed better than other deep learning and machine learning model.
- Using variable importance in a random forest model was the simple method for extracting features from an image.

Using variable importance in a random forest model was the simple method for extracting features from an image.

3 Methodology

The classifying stages of DR can be done in a variety of ways. However, the use of fundus imaging, which is used in clinical approaches as indicated by Safi et al. (2018), was the one that intrigued me the most. As was already mentioned in the related work section, the characteristics of the lesion can be useful in identifying the stages of DR. Consequently, two models are being created for this study. The first model, termed "DR Stage Classification", uses two CNN models depicted in Figure 2. CNN 1 will classify whether or not an image has DR, and CNN 2 will classify the data as having one of the four stages of DR. This model is then compared with the second model named as "Lesion Feature DR classification Model". The purpose of the "Lesion Feature DR classification Model" is identical to classifying the DR stage. But there is a little difference in the method used. By improving the data used to train the model, this is accomplished. Lesion feature is used in the data. An image dataset with a labeled lesion image that specifies the type of lesion is used for this. Using the random forest feature importance variable, this lesion's features are extracted. The feature importance from the lesion is then used to improve the DR image dataset. The stages of DR in this model are then determined using the same classification as in the "DR Stage Classification" method.

The method is depicted in Figure 3 The model employs the KDD approach. Python was utilized to carry out this KDD technique.

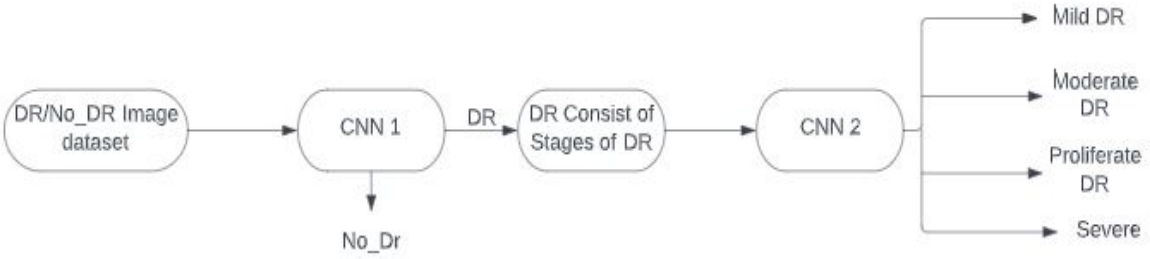

Figure 2: DR stage classification, Methodology

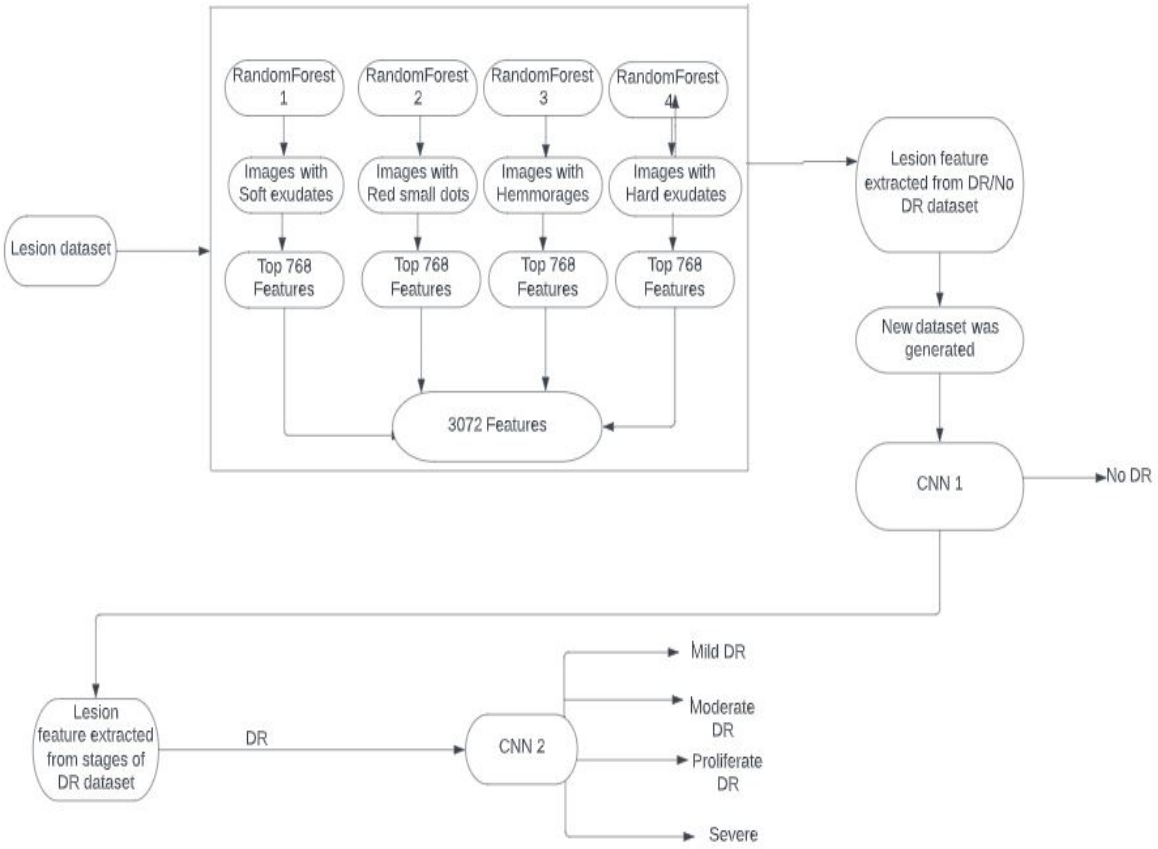

Figure 3: Lesion Feature DR stage classification, Methodology

3.1 Data Acquisition

Two labeled datasets were utilized to train and test the model. They obtained from Standard Diabetic Retinopathy Database (Zago et al. (2020)) and Kaggle, two open-source websites. Even though it was obtained from open source ethics declaration was produced and it was approved from the institution for the purpose of this study.

3.1.1 DIARETDB1 - Standard Diabetic Retinopathy Database

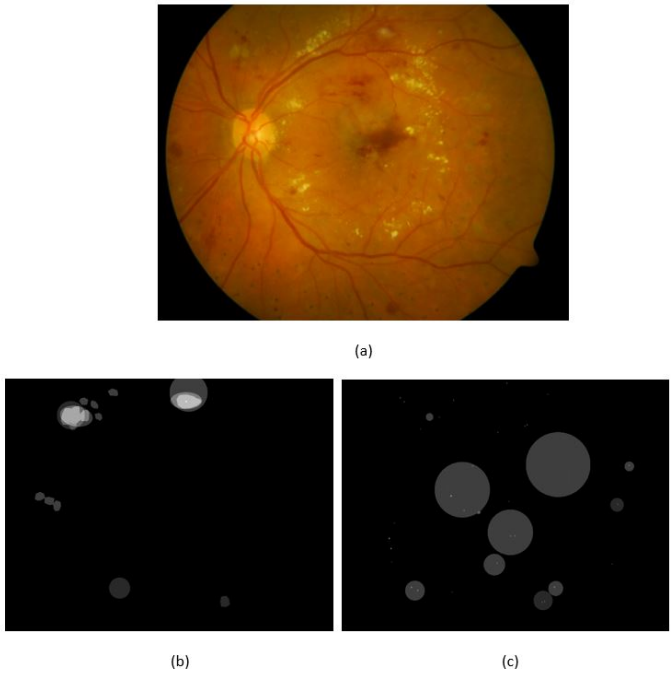

Figure 4: Standard Database for Diabetic Retinopathy. Calibration Level 1 (DIARETDB1) The database offers a ground truth image made up of the superposition of each expert's annotation for each type of analyzed lesion. (a) depicts fundus image, (b) depicts smalldots groundtruth image & (c) depicts the red dots groundtruth image (inspired from Zago et al. (2020))

This datasets consists of 89 color fundus images, 84 of which show lesions and 5 of which are normal and does not convey any indications of DR. Four experts reviewed the complete dataset, classifying the lesions into soft exudates, hard exudates, red small dots, and hemorrhages, and labeled them accordingly. The degree of agreement is indicated by the various gray levels in the ground truth photographs as depicted in Figure 4. This dataset was used to discover the lesion feature.

3.1.2 Diabetic Retinopathy 224×224 - Kaggle

The retina scan images that have been filtered using a Gaussian filter make up the dataset. According to the stages of DR, every image has been labeled. This dataset includes 3,662 photos in total, divided into 5 stages (classes): No DR, Mild, Moderate, Proliferate, and Severe. These images have already been scaled down to 224×224 pixels so they may be quickly incorporated into DL models. In order to train the classification of DR stages, this dataset was employed.

3.2 Data Pre-processing

The quality of the model can be impacted by image preprocessing, making it a crucial stage in modeling. While the model is still being trained, images were enhanced in real-time using the Keras ImageDataGenerator. These data preprocessing for the 2 models was carried out independently, as is briefly explained in this section.

3.2.1 DR Stage Classification

The dataset had already been reduced in size and put through a Gaussian filter. Binary classification is the first CNN model. As a result, the labelled dataset was categorized so that all of the images with other labels were classified as "DR" and the labels with No DR were classified as "No_DR" as shown in Figure 5. The dataset's impartiality or imbalance was verified because it is necessary for creating a model.It is evident from the Figure 6 that the images DR and non-DR photographs are nearly equal. Therefore, there is no class imbalance.The next stage was to divide the data into test and train data at a ratio of 2:3 because there was no class imbalance. As a result, the sklearn function *train_test_split*( ) was used.

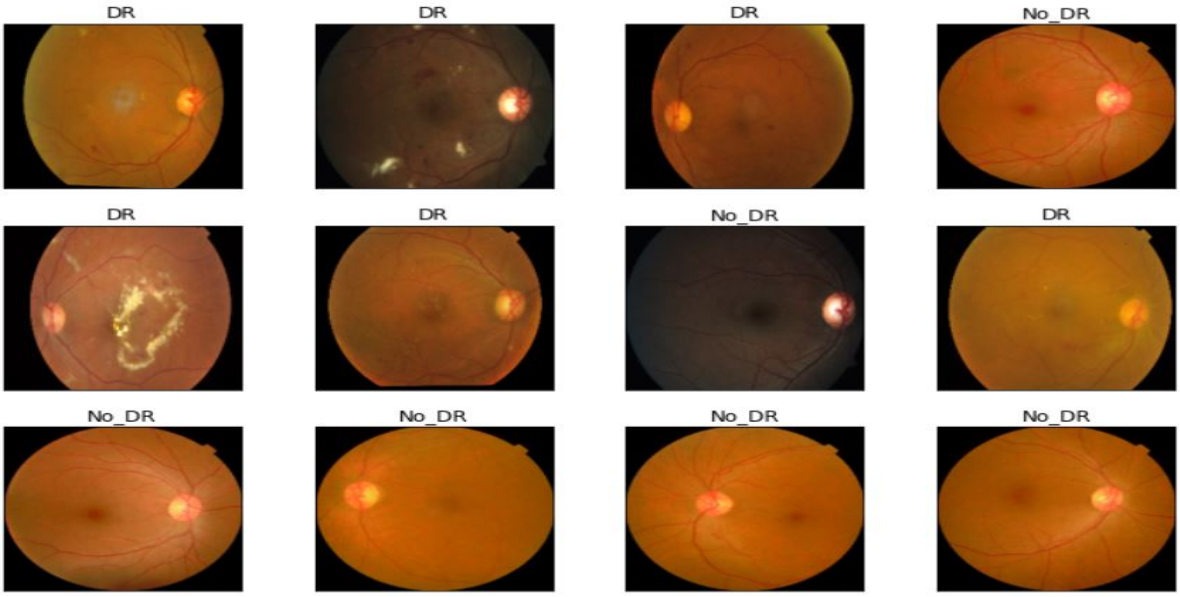

Figure 5: Images with DR and No DR

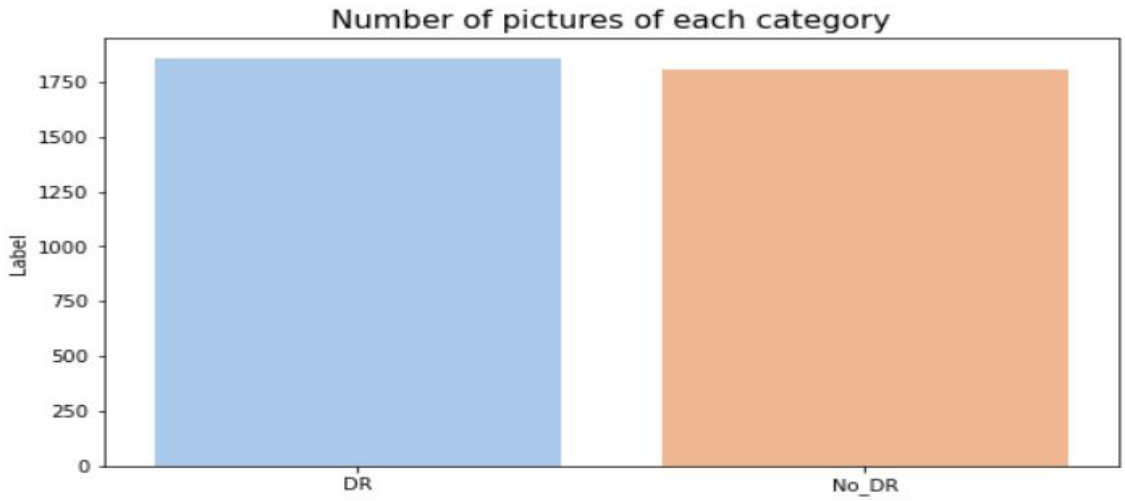

Figure 6: Class imbalance is absent

Now, only the stages of DR from the remaining four classes will be categorized in this second CNN model. The same dataset was used for this, but it was altered by removing photos with the label "No_DR" from the dataset as shown in Figure 7. Then the data was divided into test and train data at a ratio of 2 : 3. The same sklearn function *train_test_split( )* was used.

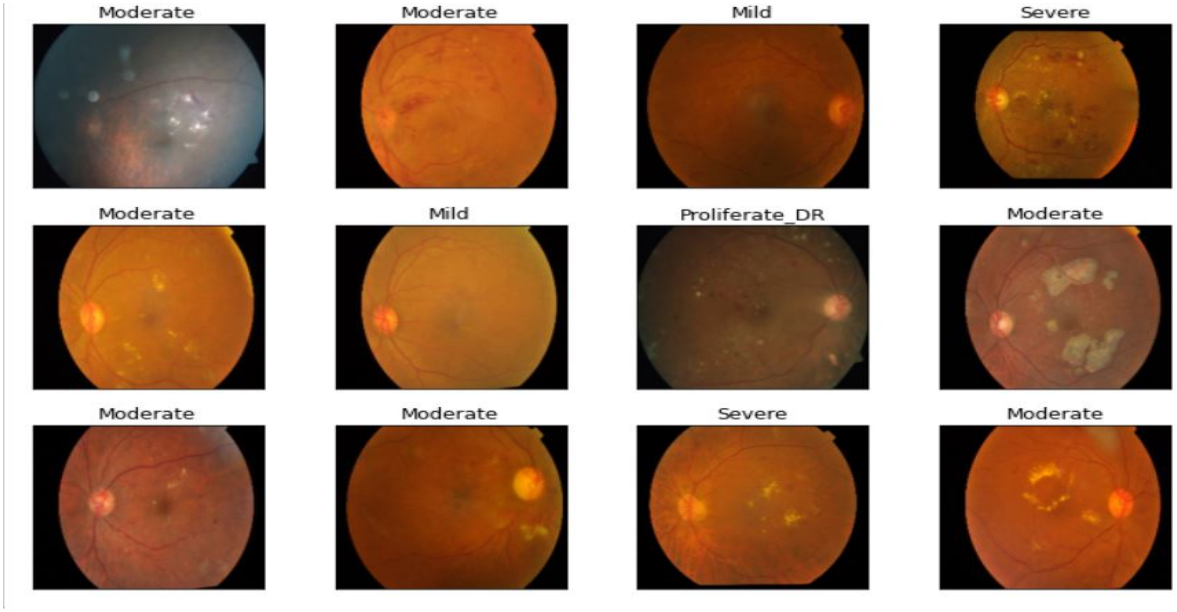

Figure 7: Images stages of DR

3.2.2 Lesion Feature DR classification Model

The literature study found that the presence of a lesion aids in the early diagnosis of DR phases. In order to identify the stages of DR, this model makes use of the lesion features. For modelling this model Diabetic Retinopathy 224 × 224 - Kaggle enhanced with the important features of lesions, for that DIARETDB1 - Standard Diabetic Retinopathy Database was utilised. A pandas dataframe was built for this. The dataset consists of four different types of lesions (multiple class), together with an image, where the value "1" denotes the existence of that lesion in the image and the value "0" denotes its absence Figure 8.

	softexudates	redsmalldots	hemorrhages	hardexudates	Names
0	0	1	1	1	F:\NCI_Documents\Final Thesis\images\ddb1_fund...
1	0	1	1	1	F:\NCI_Documents\Final Thesis\images\ddb1_fund...
2	1	1	1	1	F:\NCI_Documents\Final Thesis\images\ddb1_fund...
3	1	1	1	1	F:\NCI_Documents\Final Thesis\images\ddb1_fund...
4	1	1	1	1	F:\NCI_Documents\Final Thesis\images\ddb1_fund...

Figure 8: Presence of lesion in the images

Random forest was used to detect the existence of each lesion in the provided images in order to determine the important features, and by applying variable importance, the top 768 important features of each type of lesion were discovered (Archer and Kimes; 2008). A new list of 3072 features was formed after summing together all 768 features from each of the photos. As a result, a fresh image dataset was generated. This feature dataset was compared with the labeled "DR" or "No DR" and a new dataset was constructed as shown in Figure 9 in order to build the binary classification model to detect the presence of DR or no DR. Using the sklearn function *train_test_split*( ) the dataset was then divided into test and train data in a 2 : 8 ratio.

	image	label
0	F:\NCI_Documents\Final Thesis\feature\309.png	DR
1	F:\NCI_Documents\Final Thesis\feature\1314.png	DR
2	F:\NCI_Documents\Final Thesis\feature\3279.png	DR
3	F:\NCI_Documents\Final Thesis\feature\2758.png	No_DR
4	F:\NCI_Documents\Final Thesis\feature\1429.png	No_DR

Figure 9: Enhanced labeled dataset having labels as DR and No_DR

The dataset used for categorizing the stages of DR was improved using the previous method and is depicted in Figure 10. The dataset was then split into test and train data using the sklearn function *train_test_split*( ) in a 2:8 ratio.

	image	label
0	F:\NCI_Documents\Final Thesis\feature\726.png	Moderate
1	F:\NCI_Documents\Final Thesis\feature\1126.png	Moderate
2	F:\NCI_Documents\Final Thesis\feature\270.png	Mild
3	F:\NCI_Documents\Final Thesis\feature\1338.png	Moderate
4	F:\NCI_Documents\Final Thesis\feature\1454.png	Proliferate_DR

Figure 10: Enhanced labeled dataset having labels as the stages of DR

3.3 Modelling

The CNN deep learning models were used to create the model. Following a review of the papers, it was observed that the VGG-16 and VGG-19 models offer a high level of accuracy

in DR detection (Aljehane; 2022) (Safi et al.; 2018). As a result, both models were created, and their levels of accuracy were compared. Additionally, it was observed that lesions can be utilized to identify the DR (Zago et al.; 2020). As a result, a second model "Lesion Feature DR classification Model" with features of lesion was created utilizing enhanced data. The model building is elaborately described in this section.

3.3.1 DR Stage Classification

This model consists of two CNN models. Models VGG-16 and VGG-19 were employed. The image classification field considers VGG-16 to be an important turning point. where the number 16 in VGG-16 indicates that it has 16 deep layers. VGG-16 was shown to be the best configuration out of all the others on the ImageNet dataset. It was therefore selected for this study because the dataset used is comparable to an ImageNet dataset. Although the VGG-19 design is identical to the VGG-16 architecture, the VGG-19 network has three more extra convolutional hidden layers (Mascarenhas and Agarwal; 2021). The structure of VGG-16 and VGG-19 is depicted Figure 11 and Figure 12.

Convolution layers of VGG-16 & VGG-19:

- Filter size: 3×3
- Stride: 1
- Padding: Same

Max-pooling layers of VGG-16 & VGG-19:

- Filter size: 2×2
- Stride: 2

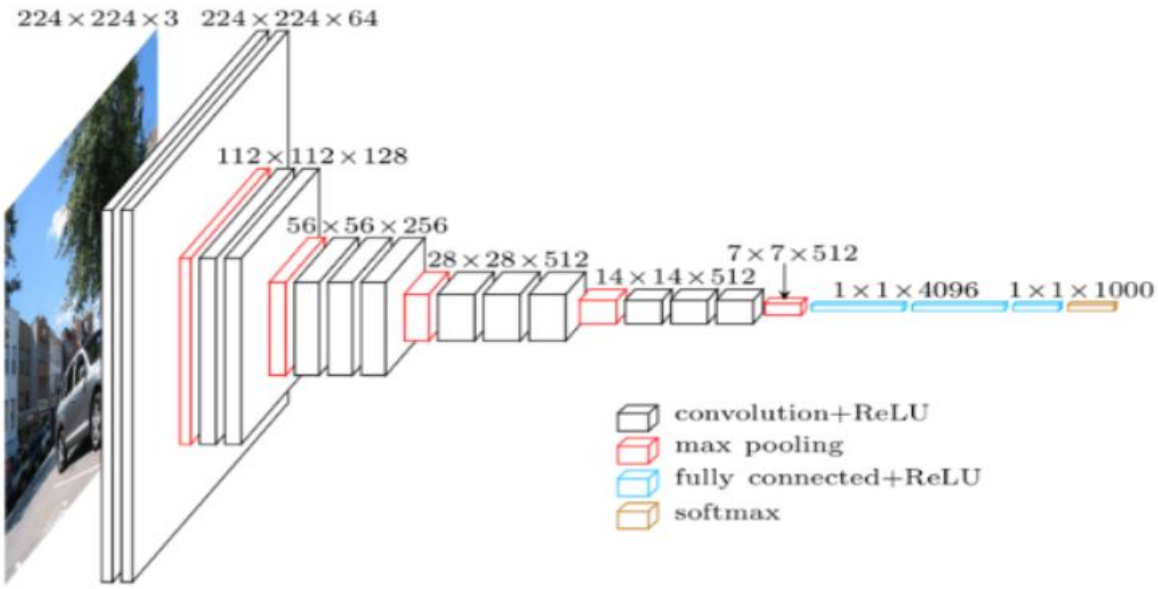

Figure 11: VGG-16 Architecture[3]

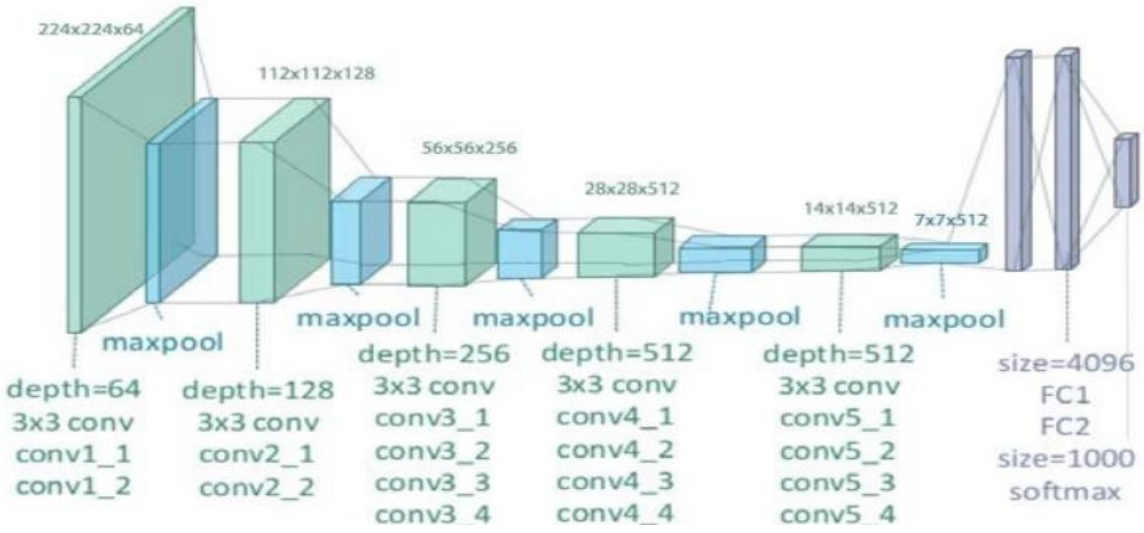

Figure 12: VGG-19 Architecture[4]

The first VGG-16 or VGG-19 are used to determine whether DR is present in the image or not. The next model will find it simpler to categorize the stages of DR because one class has been eliminated. The image will now be categorized into Mild, Moderate, Severe, or Proliferate stages of DR by the second model.

3.3.2 Lesion Feature DR classification Model

The "DR stage classification" only slightly differs from what is shown in the Figure 3.The lesion feature was added to the dataset to improve its ability to detect DR symptoms using the same model mentioned in the above section 3.3.1. In the section on data preprocessing, a brief mention of feature selection and dataset enhancement was made. For selecting the important lesion feature a random forest model was employed. It was extracted using variable importance from Random forest model (Chowdhury et al.; 2019). Given that there are four different types of lesions, including soft exudates, red small dots, hemorrhages, and hard exudates, four independent random forest models were built for the purpose of lesion detection. The random forest model with the highest accuracy was chosen from each one; they were all binary classification random forest models. The outcome is driven by variables of great importance, and the values of these variables have a big impact on the outcome values. This was done by creating a new list from each of the four sets of lesions using the $feature\_importances\_()$ function. After that, the dataset was used to generate the DR stage classification model was improved, and this dataset was used to train and test the DR stages. The same methodology used in the "DR stage classification" model used to classify the stages of DR.

3.4 Evaluation

Any DL model must include evaluation as a crucial phase. It enables us to compare the performances of various models and select the one with the best performance metrics.

[3]https://koushik1102.medium.com/transfer-learning-with-vgg16-and-vgg19-the-simpler-way-ad4eec1e2997

[4]https://koushik1102.medium.com/transfer-learning-with-vgg16-and-vgg19-the-simpler-way-ad4eec1e2997

Sensitivity and specificity are crucial for the diagnosis of diseases in the medical field because if a patient does not test positive, there is a possibility that their condition goes bad. As a result, the models are assessed and contrasted using performance criteria like accuracy. Graphs are produced by visualizing the results as well.

4 Design Specification

In this section, the design specifications of the DR Stage Classification and the Lesion Feature DR classification Model are separately discussed.

4.1 DR Stage Classification

The approach is divided into 2 phases. The Figure 13 depicts the process pipeline's architectural layout.

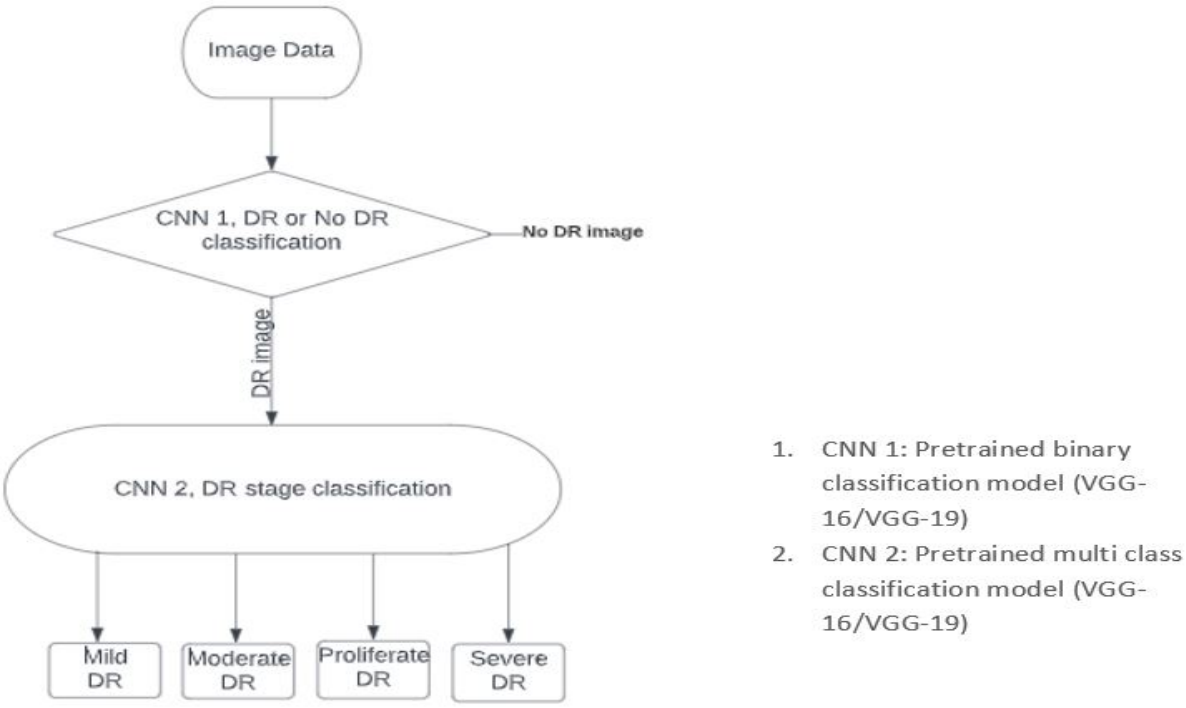

Figure 13: DR stage classification, model architecture

- **CNN 1**: This pre-trained VGG-16/VGG-19 model was trained using a dataset with labels that included DR and non-DR images. Whether DR is present in the image or not will be determined using this model.
- **CNN 2**: This pre-trained VGG-16/VGG-19 model was created using a labeled dataset that included the DR's stages as the labels. The stage of DR will be classified using this model.

4.2 Lesion Feature DR classification Model

The design is comparable to a DR Stage Classification model; however, the model learns the lesion characteristic in the prediction of DR and No DR as well as in the future stage of classification of DR as depicted in Figure 14. The design is identical to that of the DR Stage Classification model, but the model training has been modified with the dataset used. This also includes two CNN models.

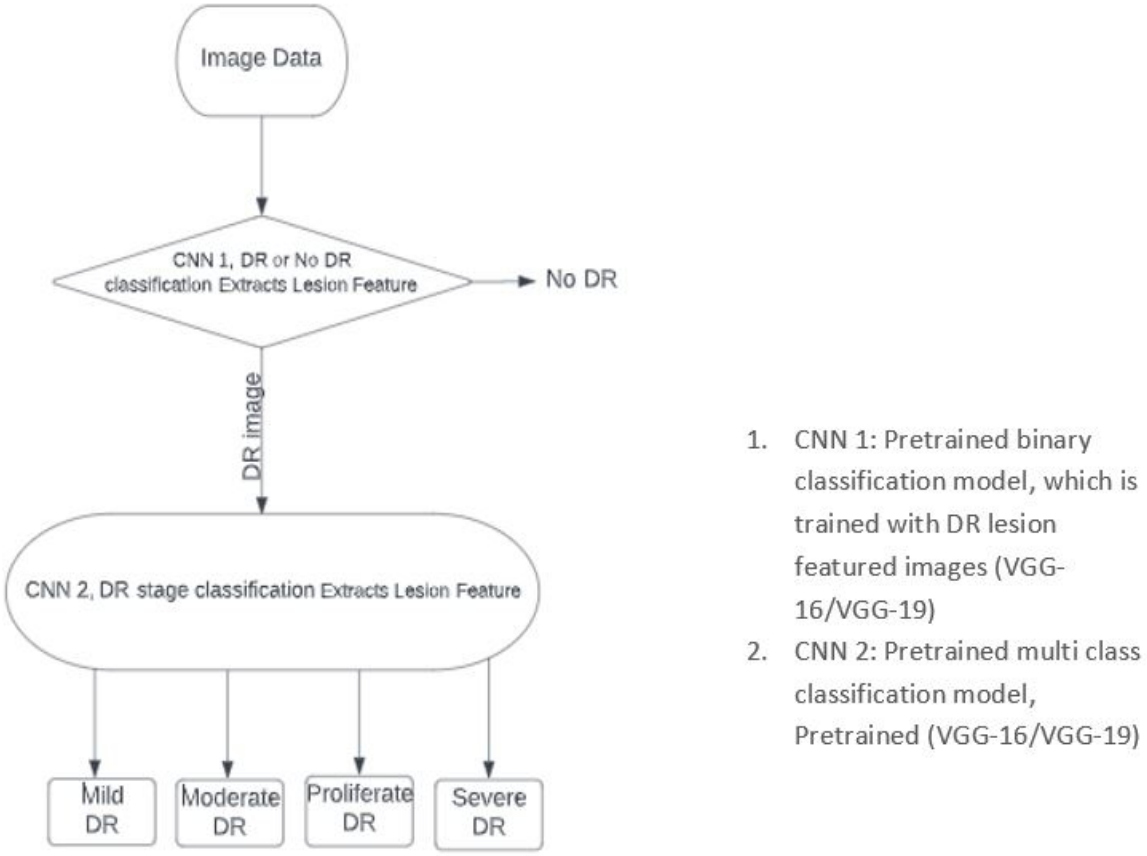

Figure 14: DR stage classification, model architecture

- **CNN 1**: This pre-trained VGG-16/VGG-19 model was developed using a labeled dataset including DR and non-DR as label. This dataset is enhanced by the lesion feature. This was accomplished by applying the random forest variable importance to extract the lesion feature. Ultimately, this pretrained model will classify the image, determining whether or not DR is present in the image.
- **CNN 2**: Using a labeled dataset that includes the DR's phases as the labels and was improved by the lesion feature, a pre-trained VGG-16/VGG-19 model was developed. In order to extract the lesion feature, the random forest variable importance was used. In the end, this pretrained model will categorize the image and place it in the DR stage.

5 Implementation

The model's implementation is briefly detailed in this section. Python was used to implement this research project. This was done independently for 2 models.

5.1 DR Stage Classification

All relevant libraries were imported as the initial step. The image dataset was stored in 5 distinct directories, one for every stage of DR. Python was used to load the image dataset, and a pandas dataframe was then created. They includes the image and corresponding label. Since the photos are kept in the directories of the stages DR, which is the label, the label's was extracted from the path by utilizing the *os.path.split*( ) function present in the *os* library. The next step was data preprocessing, which is briefly detailed in section 3.2.1.
The VGG-16 and VGG-19 CNN models were employed in this in the classification of image with DR or No DR (CNN 1) and also in classification of stages of DR (CNN 2). This is a pre-trained model. The creation of the image generators was the initial stage in designing this model. While the model is still being trained, the generators add to the images in real-time. This will increase the model's robustness and reduce the need for extra memory. Images for training, testing, and validation were generated.
TenserFlow Keras was employed in the construction of the CNN modeling. Using TensorFlow as its base, Keras is a neural network Application Programming Interface (API) for Python build on top of TensorFlow. Due to its user-friendly design, this makes it easier to implement complex neural networks. For that reason, this was used in the study. The model's parameters are shown in the Table 2.

Table 2: Architecture of Implemented CNN

Model	**Input Layer**	**Dense Layer 1**	**Dense Layer 2**	**Output Layer**
CNN 1	Shape=224x224; Weight=ImageNet; Pooling=Average	Shape=128; Activation=Relu	Shape=128; Activation=Relu	Shape=2; Activation=Softmax
CNN 2	Shape=224x224; Weight=ImageNet; Pooling=Average	Shape=128; Activation=Relu	Shape=128; Activation=Relu	Shape=4; Activation=Softmax

5.2 Lesion Feature DR Stage Classification

The initial phase of this model is the Lesion feature extraction. The required libraries were all imported. The labeled lesion dataset, which includes the existence of the type of lesion along with the image, was imported into Python as a pandas dataframe. The image dataset was then read as an array in the subsequent step.
The main improvement made to this model was to remove the lesion feature. Variable importance random forest was employed for this. The Table 3 shows the four different Random Forest models that were used.

The top 768 important features from each RandomForest model were then extracted using the $feature_importances_()$function, and a new list of the top 768 features from each of the four models was constructed. Then The CNN 1 model, which classify images

Table 3: Architecture of Implemented Random Forest model

Model	**Test size**	**Train size**	**Random state**
Random Forest 1: Soft exudates	0.2	0.8	5
Random Forest 2: Red small dots	0.2	0.8	1
Random Forest 3: Hemorrhages	0.2	0.8	14
Random Forest 4: Hard exudates	0.3	0.7	0

into DR and No DR, was trained using only the important features from the DR and No DR image dataset. The CNN 2 model, which classify the stages of DR, was then trained using the same technique using the important features that were extracted from the images of the DR stage.

6 Evaluation

Using the test and validation images, the models were evaluated. For the purpose of evaluating the models, accuracy, loss, and training time are calculated for each model. This is then compared in order to select the best model. According to the table, CNN 1 outperforms CNN 1 in the "DR stage classification" model, and CNN 2 outperforms CNN 2 in the "DR stage classification " as a result. CNN 1 and CNN 2 working together in the "Lesion Feature DR stage classification" model results in a better performance. This demonstrates that, although this was not evaluated in real time.

Table 4: Evaluation Matrix

Model	**Train Accuracy**	**Validation Accuracy**	**Training time (sec)**	**Epoch**
DR stage classification CNN 1: VGG-16	0.9575	0.9658	8928.95	10
DR stage classification CNN 1: VGG-19	0.9276	0.9452	9208.35	10
DR stage classification CNN 2: VGG-16	0.5677	0.5473	5392.03	10
DR stage classification CNN 2: VGG-19	0.6028	0.5405	5477.39	10
Lesion Feature DR stage classification CNN 1: VGG-16	0.9788	0.9212	3596.11	200
Lesion Feature DR stage classification CNN 1: VGG-19	0.9795	0.9007	4327.81	200
Lesion Feature DR stage classification CNN 2: VGG-16	0.9088	0.5203	2976.32	250
Lesion Feature DR stage classification CNN2: VGG-19	0.9529	0.4865	2367.61	250

The learning curve for the train was generated. The curve that results from the training dataset, known as a train learning curve, shows how well the model is learning

up new information. According to the Figure 15, the model's accuracy was improving significantly with each learning period. Additionally, with each epoch, the loss function was also decreasing.

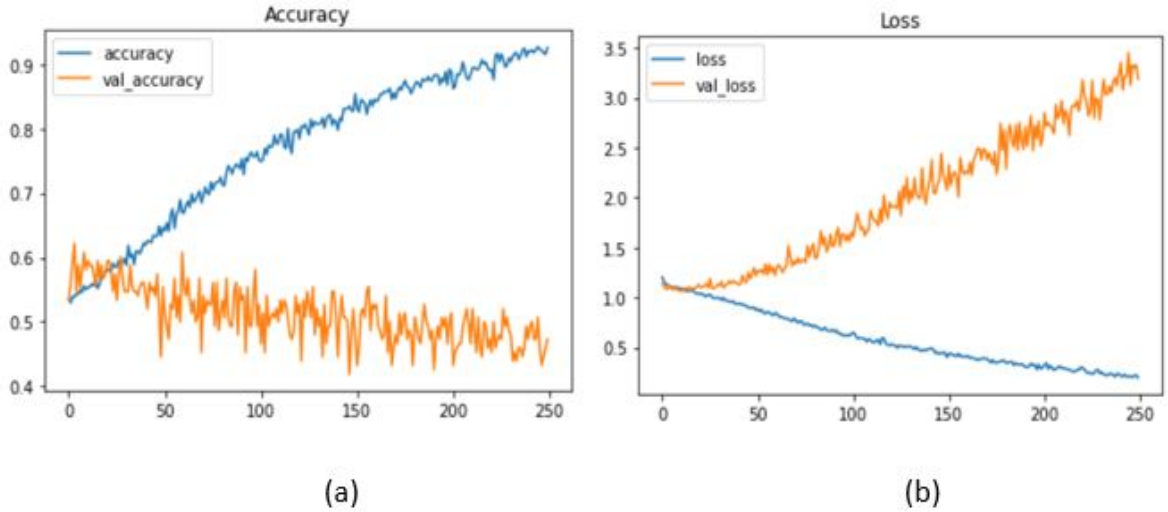

Figure 15: (a) Model Accuracy; (b) Model loss function

6.1 Discussion

Although the model's training accuracy was superior to its test/validation accuracy , this is unacceptable. This can be due to the image dataset that was used. Despite the fact that there was no imbalance in the image dataset as shown in figure, However, the image dataset lacked the required photos for every class to train the model. A machine learning Random forest model was used to extract the features. This model can be improved with another deep learning model, and the important features can be retrieved from it to be more helpful in DR stage detection. The sensitivity and specificity of the model should be evaluated before it is finalized because it is employed in the medical field.

7 Conclusion and Future Work

This study comes to the conclusion that extracting lesions from fundus images can help to identify the stage of diabetic retinopathy. Therefore, it provides a solution to the research question, "How effectively can deep learning and image processing be applied to detect the stage of DR with the presence of lesion?" with these results. The second research question, "How can Deep Learning technique be used to improve and solve the problems associated with manual classification of DR stages with the presence of lesion?" is addressed by the fact that the training time of the model is significantly shorter than that of the early studies. This accelerates the detection of DR.
The test and validation results are unacceptable, despite the fact that the training sets produced results that were acceptable. There is therefore a lot of room for improvement in the model in future work. The model's process flow appears to be proper, but the datasets that were utilized to develop the model are not very effective at identifying the stages of DR. Therefore, that might be used for future work. To extract the characteristics,

a machine learning Random forest model was employed. Another deep learning model can be used to enhance this one, and the important features can be extracted from it to increase the model's performance in DR stage detection. Additionally, as part of the ongoing research, it is also feasible to identify the lesion feature that aids the DR stage.

References

Aljehane, N. O. (2022). An intelligent moth flame optimization with inception network for diabetic retinopathy detection and grading, *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, pp. 370–373.

Archer, K. J. and Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures, *Computational Statistics & Data Analysis* **52**(4): 2249–2260.
URL: *https://www.sciencedirect.com/science/article/pii/S0167947307003076*

Cao, W., Shan, J., Czarnek, N. and Li, L. (2017). Microaneurysm detection in fundus images using small image patches and machine learning methods, *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 325–331.

Chowdhury, A. R., Chatterjee, T. and Banerjee, S. (2019). A random forest classifier-based approach in the detection of abnormalities in the retina, *Medical &; biological engineering &; computing* **57**(1): 193—203.
URL: *https://doi.org/10.1007/s11517-018-1878-0*

delaPava, M., Rios, H., Rodríguez, F., Perdomo Charry, O. and González, F. (2021). A deep learning model for classification of diabetic retinopathy in eye fundus images based on retinal lesion detection.

Group*, T. E. D. P. R. (2004). The Prevalence of Diabetic Retinopathy Among Adults in the United States, *Archives of Ophthalmology* **122**(4): 552–563.
URL: *https://doi.org/10.1001/archopht.122.4.552*

Heisler, M., Karst, S., Lo, J., Mammo, Z., Yu, T., Warner, S., Maberley, D., Beg, M. F., Navajas, E. V. and Sarunic, M. V. (2020). Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography, *Translational Vision Science & Technology* **9**(2): 20–20.
URL: *https://doi.org/10.1167/tvst.9.2.20*

Houby, E. M. F. E. (2021). Using transfer learning for diabetic retinopathy stage classification, *Applied Computing and Informatics* .

Lam, C., Yu, C., Huang, L. and Rubin, D. (2018). Retinal Lesion Detection With Deep Learning Using Image Patches, *Investigative Ophthalmology & Visual Science* **59**(1): 590–596.

Mascarenhas, S. and Agarwal, M. (2021). A comparison between vgg16, vgg19 and resnet50 architecture frameworks for image classification, *2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*, Vol. 1, pp. 96–99.

Nneji, G. U., Cai, J., Deng, J., Monday, H. N., Hossin, M. A. and Nahar, S. (2022). Identification of diabetic retinopathy using weighted fusion deep learning based on dual-channel fundus scans, *Diagnostics* **12**(2).
URL: *https://www.mdpi.com/2075-4418/12/2/540*

Odeh, I., Alkasassbeh, M. and Alauthman, M. (2021). Diabetic retinopathy detection using ensemble machine learning, pp. 173–178.

Safi, H., Safi, S., Hafezi-Moghadam, A. and Ahmadieh, H. (2018). Early detection of diabetic retinopathy, *Survey of Ophthalmology* **63**(5): 601–608.
URL: *https://www.sciencedirect.com/science/article/pii/S0039625717303235*

Vyas, A. H. and Khanduja, V. (2021). A survey on automated eye disease detection using computer vision based techniques, *2021 IEEE Pune Section International Conference (PuneCon)*, pp. 1–6.

Wan, S., Liang, Y. and Zhang, Y. (2018). Deep convolutional neural networks for diabetic retinopathy detection by image classification, *Computers & Electrical Engineering* **72**: 274–282.
URL: *https://www.sciencedirect.com/science/article/pii/S0045790618302556*

Zago, G. T., Andreão, R. V., Dorizzi, B. and Teatini Salles, E. O. (2020). Diabetic retinopathy detection using red lesion localization and convolutional neural networks, *Computers in Biology and Medicine* **116**: 103537.
URL: *https://www.sciencedirect.com/science/article/pii/S0010482519303968*