Transfer Learning for Detection of Diabetic Retinopathy Disease

Research Project
MSc Data Analytics

Alekhya Bhupati
Student ID: x18132634

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Alekhya Bhupati
Student ID:	x18132634
Programme:	MSc Data Analytics
Year:	2020
Module:	Research Project
Supervisor:	Dr. Catherine Mulwa
Submission Due Date:	23/04/2020
Project Title:	Transfer Learning for Detection of Diabetic Retinopathy Disease
Word Count:	7225
Page Count:	25

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.

I agree to an electronic copy of my thesis being made publicly available on TRAP the National College of Ireland's Institutional Repository for consultation.

Signature:	
Date:	23rd April 2020

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):	

Transfer Learning for Detection of Diabetic Retinopathy Disease

Alekhya Bhupati
x18132634

Abstract

Applying deep learning on medical data is a very challenging and crucial task. Transfer learning can reduce the cost of training to a great extent by using pre-trained deep convolution neural networks. Diabetic retinopathy is the major cause of blindness and it is increasing world-wide at an alarming rate. In this work, we proposes to apply the transfer learning methods for detection of diabetic retinopathy disease and its different stages. We have experimented various deep learning models such as VGG19, ResNet50 and DenseNet201 in order to determine the best classification model for DR detection. The large dataset for diabetic retinopathy consists of imbalance dataset. So this experiment has been performed for both balanced and imbalanced dataset. The results of the models has been analyzed using various metrics such as precision, recall, f1-score and accuracy.

1 Introduction

Diabetic patients are estimated to be 415 million over the world and it is estimated that out 10 adults one of them is diabetic from Gargeya and Leng (2017). Over the last decade, a lot of people have been diagnosed with diabetes, and with diabetes variety of eye diseases can occur, one of which is diabetic retinopathy (DR). The small blood vessels present in the human eye are damaged during the disease of Diabetic retinopathy and as a result reversible and sometimes irreversible permanent blindness in patients can be occured. There are fine vessels in the eyes that are affected by diabetic retinopathy and almost 45% of the diabetic patients suffer from it. During diabetic retinopathy, eyes get abnormalities such as microaneurysms, exudates (hard and soft exudates), hemorrhages, development of cotton wool spots in Paranjpe and Kakatkar (2014). DR is a progressive disease having different stages and patients suffer from blindness at the final stage. This research concentrates to use transfer learning approaches to detect the diabetic retinopathy disease also this work compares the different types of transfer learning algorithms using several metrics.

1.1 Motivation and Project Background

Usually, patients suffering from diabetic retinopathy are unaware of their disease so its detection before the time is very important Gargeya and Leng (2017). Diabetic retinopathy detection is very difficult because it is a time-consuming process that ultimately results in delayed treatment of the disease. With early diagnosis, it is estimated that 90% of patients can be cured of diabetic retinopathy Ishtiaq et al. (2019). Diabetic retinopathy can be cured manually or with aid of the automated system. In manually detecting the DR, ophthalmologists need to be expert and don't need any technical assistance for it Ishtiaq et al. (2019). Other limitations of the manual system are that these are more time consuming whereas proven to be inefficient when the large dataset is presented to them Paranjpe and Kakatkar (2014). More resources are also required if DR is being done manually *Neural Network Technique for Diabetic Retinopathy Detection* (2019). The automated systems, on the other hand, can detect very small indications of DR in patients, as the patient's retina is visible to the doctors. They happen to eliminate the need for manual labor for detection. Paranjpe and Kakatkar (2014) have provided an extended review of diabetic retinopathy, its stages, development of automated systems, the process of segmentation and classification used for the detection. The automated detection of DR, as proposed by the authors of this study, is shown in the form of block diagram below in Figure 1.

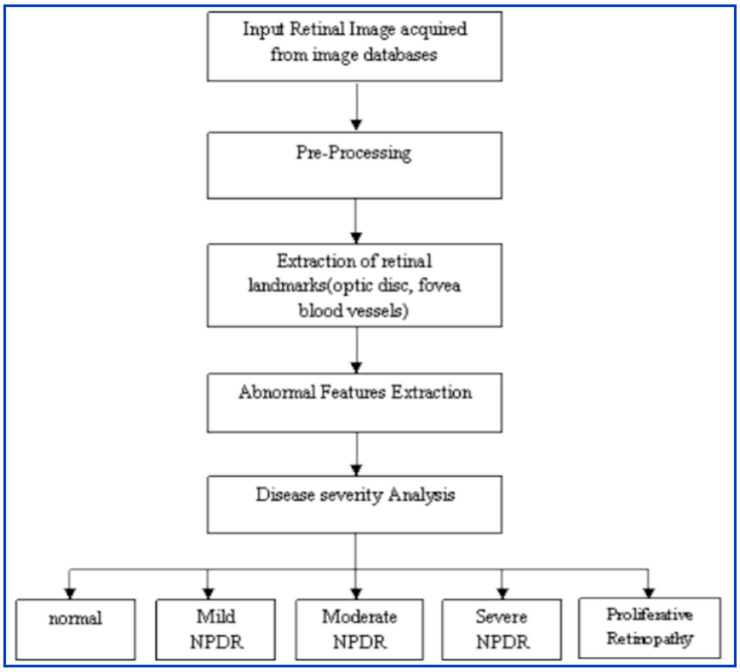

Figure 1: Extended Review of Diabetic Retinopathy

The system takes retinal images for processing, the next abnormal features are extracted, and upon the analysis of extracted features classification is performed. In the last stage, the images are classified in normal (without DR), mild and moderate (less severity), severe Nonproliferative Retinopathy (NPDR) and Proliferative Retinopathy (PDR). The accuracy of the system can be improved if information regarding the features chosen such as microaneurysms, exudates (hard and soft exudates), hemorrhages, development of cotton wool spots is presented to the system with texture. The performance of automated systems is measured with the help of specificity, sensitivity and accuracy Paranjpe and Kakatkar (2014). A similar analogy of the automated detection of DR is done by Kumaran and Patil (n.d.) who have categorized the DR in 4 categories including; initial or mild stage, moderate, severe and Proliferative stage which is the final stage. Patel et al. (2016) have also performed a detailed review on automatic detection of diabetic retinopathy. According to this, the DR can be classified into 3 stages based on the severity of the disease including mild, moderate-severe (NPDR) and PDR. They also have categorized the lesions that form as a result of DR based on their color, size, shape, edge, and classes. A thorough review of existing image processing techniques of DR, blood vessel and optic disc extraction techniques, lesion detection and feature extraction, techniques and existing tools and datasets available to test the proposed techniques have also been presented.

1.2 Artificial Intelligence and Detection of Diabetic Retinopathy

Using artificial intelligence techniques allows in early detection of DR, thus providing two advantages; less probability of human error, the minimum workload for an ophthalmologist, and a more efficient way of finding the lesions in the retina in less period. According to Ishtiaq et al. (2019), artificial intelligence methods tend to solve the detection of DR either with machine learning techniques or deep learning.

1.3 Machine Learning and its Associated Algorithms and Techniques for the Detection of Diabetic Retinopathy

Machine learning is an approach in which machine learns with the help of some algorithms and perform different tasks such as classification that is required for classifying the retinal images in case of presence and absence of diabetic retinopathy. A generic machine learning approach adopted by the researchers for the detection of diabetic retinopathy has been shown by Ishtiaq et al. (2019) diagrammatically as: First of all, a detection

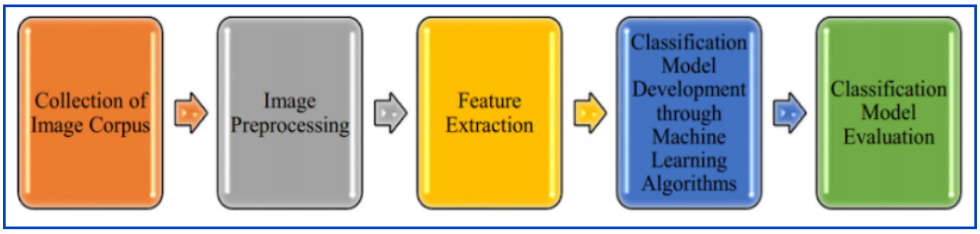

Figure 2: Methodology of Diabetic Retinopathy

model for diabetic retinopathy has constructed that constitutes images that are labeled. These images will serve as a training set in the algorithm and should belong to different categories of diabetic retinopathy. In these images, many unwanted features need to be removed, and for this purpose image processing on these images is applied. After that feature extraction is applied on the processed images so that discriminative features can be extracted and MVF (Master feature vector) is obtained. For the detection model, an algorithm of ML is developed and the MVF obtained serves as an input to this ML algorithm. This detection model learns with the help of different classification rules and for performance evaluation, test data consisting of unlabelled images are presented to the algorithm. Finally, the accuracy of the model is check with various measures and results are established Ishtiaq et al. (2019). Patel et al. (2016) have performed a thorough review on implementation strategies for detection of DR. As for machine learning techniques are concerned ANNs, Random forest, multilayer feed-forward neural network

(NN), SVMs, and Fisher discriminant analysis (LFDA) have been used by the researchers for classification purpose.

Neural networks are a branch of machine learning algorithms that acts on the principle of the human nervous system. and make use of 3 layers containing several nodes in each layer and neurons can be represented with a single node. ANNs tend to bring out some patterns or classification patterns that can be found in the presented data to it. Neural networks that belong to feedforward networks have been widely used by the researchers to detect diabetic retinopathy. Xu et al. (2017) have used convolutional neural networks for this purpose and have mentioned that numerous other ML approaches such as support vector machine and K nearest neighbors have been used in literature for this purpose.

1.4 Research Question

RQ: *"How can the transfer learning enhance/improve detection of the different stages in diabetic retinopathy disease ?"*

Sub RQ: *"How we can improvise the performance of models (ResNet50, VGG19 and DenseNet201) over imbalanced dataset ?"*

1.5 Research Objectives

Table 1: Research Objectives of Transfer Learning for Detection of Diabetic Retinopathy Disease

Objective	**Description**	**Metrics**
Obj. 1	A Critical Review of Diabetic Retinopathy Detec-tion and Identified Gaps (2014-2019)	-
Obj. 2	Exploratory Data Analysis to get insight about the feature for Diabetic Retinopathy Detection	-
Obj. 3	Implementation, Evaluation and Results of ResNet50	Precision, Recall, F1-Score
Obj. 4	Implementation, Evaluation and Results of VGG19	Precision, Recall, F1-Score
Obj. 5	Implementation, Evaluation and Results of DenseNet201	Precision, Recall, F1-Score
Obj. 6	Comparison of Developed Models	-

The outlined of the document is as follows: Section 2 investigates the related works of different techniques for detection of Diabetic Retinopathy disease. Section 3 represents Proposed design and methodology of the work. Section 4 Compares and evaluates the results of multiple algorithms.

2 A Critical Review of Diabetic Retinopathy Detection and Identified Gaps (2014-2019)

2.1 Introduction

This section investigates the literature review for detection of diabetic retinopathy. The critical review can be divided into multiple subsections where we explore the different technologies used for DR detection. Subsection 2.2 reviews the artificial intelligence and machine learning algorithms Then we investigates the various neural network techniques describe in subsection 2.3 . Subsection 2.4 discusses about various Deep learning approaches such as CNN and RNN methods and then in the subsection 2.5 we discusses the other techniques for DR detection.

2.2 A Review of Literature on Detection of Diabetic Retinopathy Using Artificial Intelligence and Machine Learning Techniques

An extended literature review has been performed by Ishtiaq et al. (2019), in which authors have studied the literature of detecting diabetic retinopathy with help of artificial intelligence methods, using machine learning and deep learning methodologies. Below is presented a detailed literature review that has been done in the field of detecting diabetic retinopathy using artificial intelligence techniques such as machine learning and deep learning methodologies. A combination of machine and deep learning techniques have also been employed by the researchers.

Ishtiaq et al. (2019) have classified the artificial intelligence approach used for the detection of diabetic retinopathy, found in the literature, as ML approaches and deep learning approaches. Further used algorithms are classified among these two approaches. Bellemo et al. (2019) have recently developed an artificial intelligence-based model using an adapted VGGNet architecture and residual neural network architecture. The proposed method classifies the images into different sets of images based on the severity of DR in the patients. Using the performance measures authors have established the results.

Number of ML approaches have been employed by the researchers for the finding the diabetic retinopathy and Ishtiaq et al. (2019) have enlisted all these algorithms that have been used to date. For classification Support vector machine (SVM) has been used as a part of ML algorithms. Researchers have used SVMs along with other methods to classify exudates, hard exudates, microaneurysms detection, and non-proliferative DB. Another ML algorithm presented by this study is the Random forest for classification and least used by the researchers to classify the hemorrhage detection in the images of the retina. k-Nearest Neighbor (kNN) algorithm that belongs to ML algorithms has also been enlisted by the study for classification and it has been shown that researchers have used the KNN algorithm for microaneurysms detection. Other machine learning approaches mentioned in this study include Local Linear Discrimination Analysis (LLDA) and Naïve Bayes (NB) that are probably-based algorithms used for classification. Similarly, Adaptive Boosting, decision trees, Self-adaptive Resource Allocation Network classifier and unsupervised classifiers, Ensemble classifier based classification algorithms have widely been used to detect DR.

Chetoui et al. (2018) have proposed a novel technique that extracts the texture features from the images of eyes in order to detect diabetic Retinopathy and employed Support vector machine for classification. The image classification is based on either the presence or absence of DR in the images. The proposed method uses Local Ternary Pattern (LTP), Local Binary Patterns (LBP) and Local Energy-based Shape Histogram (LESH) for extraction of texture features. A histogram of these extracted features is made as an input to the SVM. The proposed method is also tested on a real-time database of retinal images and performance evaluation has been done with accuracy, sensitivity and specificity measures. The area under the curve and average accuracy have been measured for performance evaluation.

Subhashini et al. (n.d.) have used the graphical user interface for the detection of diabetic retinopathy along with machine learning techniques. Authors of the study hold the view that image processing techniques can be paired with machine learning approaches and segmentation of the retinal images can be performed. The GUI of the proposed method makes it very viable for ophthalmologists to use the system easily and efficiently thus reducing their time of diagnosis. To remove the noise from the images, Gaussian filters have been used. Images of the user interest are taken as an input into the GUI based system, and the model of the proposed method performs Gaussian blurring on these images. After the removal of noise, k-means clustering is applied to find a region of interest. After that feature extraction and classification is performed with the help of machine learning algorithms. The results of the study have also been established by providing the images to the system and GUI of the software shows the percentage of the DR in a patient and recommends that instant appointment to an ophthalmologist is recommended. A convolutional neural has also been employed inside the system to make it more reliable.

Ogunyemi and Kermah (2015) authors have scrutinized various machine learning approaches and performed feature subset selection using the Lasso. Authors have used ensembles for combining the different classifiers and the ensemble classifier learns with the help of decision trees. The novelty of this approach is that the proposed approach uses the real-time dataset of patients with public health records. Authors have used plenty of variables that might affect the results of patients such as age, HB, dependency on insulin, etc. the results have been established with help of performance measures such as accuracy, specificity, Area under curve and sensitivity.

2.3 Neural Networks for Detection of Diabetic Retinopathy

Xu et al. (2017) have proposed an automatic method for diabetic retinopathy detection which classifies retinal images into normal and diabetic retinopathy images. According to authors, other classifiers used by the researchers such as SVM and KNN algorithms have failed to identify all symptoms of diabetic retinopathy in the images. The authors of this study have used convolutional neural networks (CNNs) for the classification of normal images and images with diabetic retinopathy. CNN's can learn the feature hierarchy that is required for classification purposes. Numerous multi-layer architectures of CNNs have been used and based on a series of experiments on real-time data of retina. The result of the proposed method has been established by conducting experiments and comparing them with d Gradient boosting machines. The performance measure of 5 different combinations of methods including convolutional neural network with data augmentation have been measured and shown that CNNs show 94.5% accuracy as compared to other methods.

According to Ishtiaq et al. (2019) artificial neural networks (ANNs) that have been used in the literature for detection of diabetic retinopathy include Probabilistic neural networks, Scaled Conjugate Gradient Back Propagation Network (SCG-BPN), Hopfield Neural Network (HNN), and Feedforward Backpropagation Neural Network (FFBPNN). The study also mentions several other ANNs and a combination of ANNs with other machine learning techniques that have been employed by the researchers for detecting lesions in retinal images.

According to Gargeya and Leng (2017), the existing algorithms developed for detection of diabetic retinopathy are limited because they have been tested on a singular small dataset, and when applied on real-time scenarios it limits the accuracy of detection. The feature extraction of other algorithms is manual-based, which limits its detection as well. To deal with the aforementioned issues, the authors of this study have proposed a completely automated algorithm that is based on neural networks. The developed algorithm can process colored retinal images and make a classification among retinopathy and non-retinopathy images. To automate the characterization of images, customized deep convolutional neural networks have been used. To test the proposed method, authors have used a large data set of 75137 DR images as well as validated on the public datasets as well. Results of the study show that feature-based deep learning methods detect diabetic retinopathy at very early stages.

Lim et al. (2014) have also used convolutional neural networks for the detection of diabetic retinopathy, because CNNs provide the best performance as far as the classification is concerned. The proposed approach of the authors identifies the regions containing the lesions. To provide input to the convolutional neural networks, these identified regions are transformed into tiles. To train CNNS, back-propagation has been used. To validate the proposed approach, authors have done experiments by taking two data sets and provided lesion-based classification as well as image-level classification. Authors have established the fact that this technique performs much better than SVMs and random forest methods.

Doshi et al. (2016) have proposed a novel model for diagnosing the DR that is based on using convolutional neural networks. The proposed model automatically learns those features of lesions and microaneurysms that are pivotal and there is no need for manual extraction. The input layer of CNN takes the images and the proposed CNN architecture takes 5 sets of convolution combinations. The parameters defined for CNN architecture include convolutional, pooling, dropout, hidden and feature pooling layers. The proposed model not only classifies the images but also shows what images have been misclassified. For evaluation, the quadratic kappa metric has been used by the authors based on 3 different CNN based models.

Lam et al. (2018) have used convolutional neural networks for automated detection of diabetic retinopathy. For specific extraction of features of the image, the convolutional network will take an image, and propose an architecture for it that will provide the best binary classification results. The model is then trained and to achieve the highest accuracy, both data preprocessing and augmentation methods are used so that early stages of DR can be detected. The proposed techniques even work very well if the sample size for training the data is kept small. The authors of the study have utilized two CNN architectures for training and testing data and using several other techniques, they have succeeded to find an optimal solution. For experimentation purposes, two real-time data sets have been used and the primary focus of the authors is to find the early stages of DR. A deep learning feature has also been used by the authors of the study, the transfer learning approach. The authors have established the fact that their proposed technique finds the microscopic level features in the retinal images.

Kumaran and Patil (n.d.) holds the view techniques used for detection DR, other than machine learning approaches tend to consume a lot of time, lack ability of processing images of less quality doesn't cater large database and noise issues, etc. the authors of this study have explained the usage of artificial neural networks for feature extraction of retinal images particularly retinal nerve fibers. According to the authors of the study, using the ANNs provides the best results in terms of finding lesions, removal of noise, the process of detection, localization, and segmentation of nerve fibers, extraction of abnormal areas in the images, DR categorization and performance evaluations. To implement the ANNs for DR detection, the first step is to select neural network structure, make neural networks learn with help of performing different calculations, and then evaluate its performance.

Neural Network Technique for Diabetic Retinopathy Detection (2019) have also proposed a technique for detecting DR that consists of different phases. First processing on data is performed, then segmentation is done on the data while in the next phase the extraction of blood vessels from the exudates and microaneurysms is done because they have a similar density in the images. Now for classification purpose, ANNs are used. Unsupervised learning is used where data learns from its own, and when there is a very percentage of error is observed in the data, it can be said that the system has learned to this point. The authors of the study have done experiments to validate their proposed method and compared the results with SVMs. Authors have proved that for NN shows higher accuracy, sensitivity, and specificity as compared to SVMs.

2.4 Review of Diabetic Retinopathy Detection Using Deep Learning Approaches

Deep learning techniques have also been used by researchers. Abràmoff et al. (2016) states that although they have used machine learning for the detection of DR previously, but improvement in the performance of the algorithm was marginal. In their recent study, the authors have used a deep learning approach with convolutional neural networks and proposed a technique that provides the highest performance. Using CNNs resulted in finding the novel associations among the images that are presented to the algorithm. The disease dataset provided to the system in this study is categorized into three categories including having no DR, vision-threatening DR and Macular edema. The severity of the disease has been given a scale, ranging from 0 at the lowest and 5 at highest. In the end, the proposed method classifies the images into negative (having no or very less severity DR), having threatening and very threatening DR and error images having low quality. Lam et al. (2018) have also employed a deep learning approach with the help of

CNN based architecture for detecting a very small and early stage of diabetic retinopathy.

Rakhlin (2018) has also worked on the detection of diabetic retinopathy using Deep Learning along with convolutional neural networks. Two data sets have been chosen by the author and utilized computer vision techniques for classification (using labeled images of having DR and no DR), segmentation (presence of any single object in form of the lesion) and detection which is hardest because small details need to be detected regarding DR. the classification model is trained with help of convolutional neural networks and it is combined with deep layer structure. The classification produced by the network in binary and learning of it is based on feature extraction. The retinal images are first made uninformed then presented to the quality assessment module in which sensitivity is measured. Augmented images that are generated randomly are fed into the DR model, and for accuracy two images of the eyes are combined. Authors have established the fact that their proposed approach performs better on a larger dataset.

Zago et al. (2020) have proposed a novel approach utilizing deep learning called deep network patch-based approach. Using deep CNN patches of an image are classified with and without lesions. A probabilistic possibility of the presence of lesion is produced with this approach. During the first stage usually, DR patients have microaneurysms and hemorrhages and algorithm tends to localize them at first. Authors term them as red lesions. Selecting the input sample is a difficult task and the performance of the classifier is increased with the help of developing a two-stage process. When the size of the data set increases, authors have used subsampling of it. The generic approach adopted by the authors of the study is that preprocessing of the images is done, and then the selection of lesions is done with the help of thresholding techniques. Then extraction of features is done and images are classified with having lesions or not having lesions. Results have obtained by running the algorithm on several datasets and show that the proposed model outperforms in terms of Se ranging.

2.5 Review of Other Techniques for Detection of Diabetic Retinopathy Disease

Other techniques have also been used for the finding of DR by the researchers. For instance, A thorough review of the detection of DR has been performed by Amin et al. (2016). Authors of this study have conducted a review on a database of retinal images that are available publically, then the performance measures used by the researchers. This study shows that for purpose of performance evaluation of an algorithm two parameters i.e. mean square error is used, along with while peak signal to noise ratio, whose higher value shows better performance of an algorithm. For the level of correctness of an algorithm sensitivity and specificity are used. Additionally, true positive (TP), true negative (TN), false positive (FP) and false negative (FN) are also used.

As to screen the disease of diabetic retinopathy at early stage, Optical Coherence Tomography (OCT) imaging, and spatial domain optical coherence tomography (SD-OCT) have been used by the ophthalmologists. As far as automated screening of DR is concerned, this study has categorized the algorithms and techniques used by the researchers based on abnormalities that result in diabetic retinopathy. Authors have mentioned the use of different techniques such as artificial neural networks, Bayesian outline work, SVMs, fuzzy logic reasoning, K-Means Clustering method, K-nearest Classification method and intelligent classifier Fuzzy SVM, the case-based reasoning (CBR) system and decision support system (DSS), Gaussian mixture model (GMM) based classifier, hybrid classifier, H-maxima transformation and Multilevel Thresholding, Fuzzy K-Median and length filter (FKMED), Dynamic thresholding, Local Binary Pattern (LBP). These techniques belong to an automated screening of diabetic DR and don't necessarily fall under the category of Artificial intelligence-based solutions.

Following table 2 shows the different machine learning and deep learning approaches found in the literature that have been used by the researchers for detection of diabetic retinopathy disease.

Table 2 : Comparative Analysis of Different Methods for Diabetic Retinopathy Detection

Authors	**Classification of Studies**	**Adapted Models**	**Experiment Outcomes**
Nguyen et al. (2020)	Convolutional layer, Pooling layer, Dropout layer, Flatten layer and Dense layer.	CNN,VGG-19 and VGG-16	82% Accuracy, 0.0904 AUC, 82% Specificity and 80% Sensitivity
Prabhjot Kaur (2019)	Optical Disk segmentation and Blood Vassal Extraction	Neural Network and Canny Edge detection algorithm	Better performance of Accuracy, specificity and sensitivity was noted as Compared to SVM classifier.
Doshi et al. (2016)	Convolutional layer, Pooling layer, dropout layer, Hidden layer and feature pooling layer.	CNN and quadratic Kappa metric	Effective performance for accuracy, specificity and sensitivity along with misclassification of images were evaluated through the proposed model.
Rakhlin (2018)	Classification, Localization or segmentation and detection.	CNN and VVG architecture	0.92 AUC, 98% sensitivity along with 71% specificity for Messidor-2 was observed in the proposed model.

2.6 Conclusion

The reviews are clear evidence that the in the literature review comparative analysis between the different Deep neural networks has not been performed by using transfer learning methods. Previous work finds out the results in terms of sensitivity and specificity we demonstrate this work by comparing the results between the balanced and imbalanced dataset using precision, recall, f1-score and accuracy as the metrics.

3 Methodology

3.1 Introduction

We have used modified KDD (Knowledge Discovery in Databases) methodology for the detection of diabetic retinopathy disease.The purpose of this project is to identify the existing issues and implement the transfer learning methods for detection of diabetic retinopathy disease. The implemented methodology can be carried out in following stages in order for diabetic retinopathy detection shown in Fig 3. These steps are further elaborated in subsections.

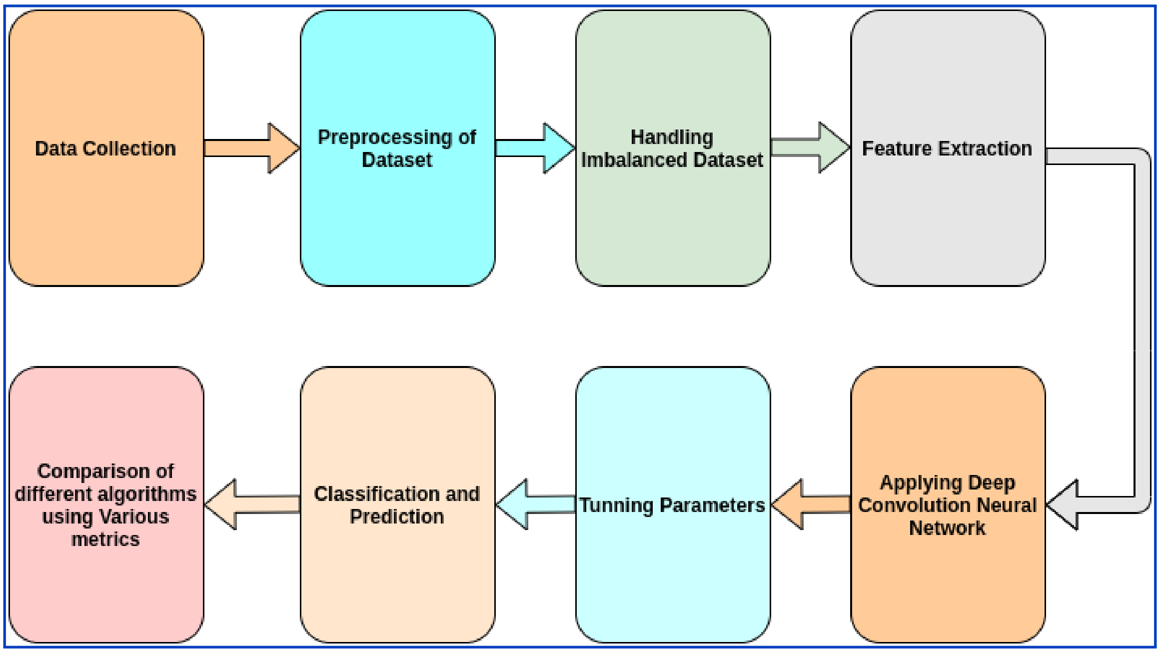

Figure 3: Diabetic Retinopathy Disease Detection Methodology

3.2 Data Collection

Diabetic retinopathy detection dataset is collected from public repository of kaggle. The training data consists of 35,126 images. These high resolution retinal images has different visual appearance either left or right. The size of training data is about 36GB, which is very large.

3.3 Data Preprocessing

Data preprocessing is very important step, to train the model more precisely and accurately. In the dataset the image files are separated with labels, so the first step we have performed in this dataset is to map the image with their respective labels. As the dataset contains high resolution retinal images and every image has different resolutions this may cause the learning model to train inaccurately. To overcome this issue we have transformed every image to 32 X 32 pixel of fixed resolution. As this is the real world data Images may contain artifacts, be out of focus, underexposed, or overexposed. we have taken care of such issues in our proposed work.

3.4 Data Exploratory Analysis

The dataset basically contains and images and their respective labels. As we know that, training data contains 35,126 retinal images. These images can sub-categorized into their respective levels. We found 5 levels, whose values lies between 0 to 4 that represents normal (without DR), mild and moderate (less sever-ity), severe Non-Proliferative Retinopathy (NPDR) and Proliferative Retinopathy (PDR). Data visualization can be carried out in this way shown in Figure 4. The levels are the target features for predictive modelling.

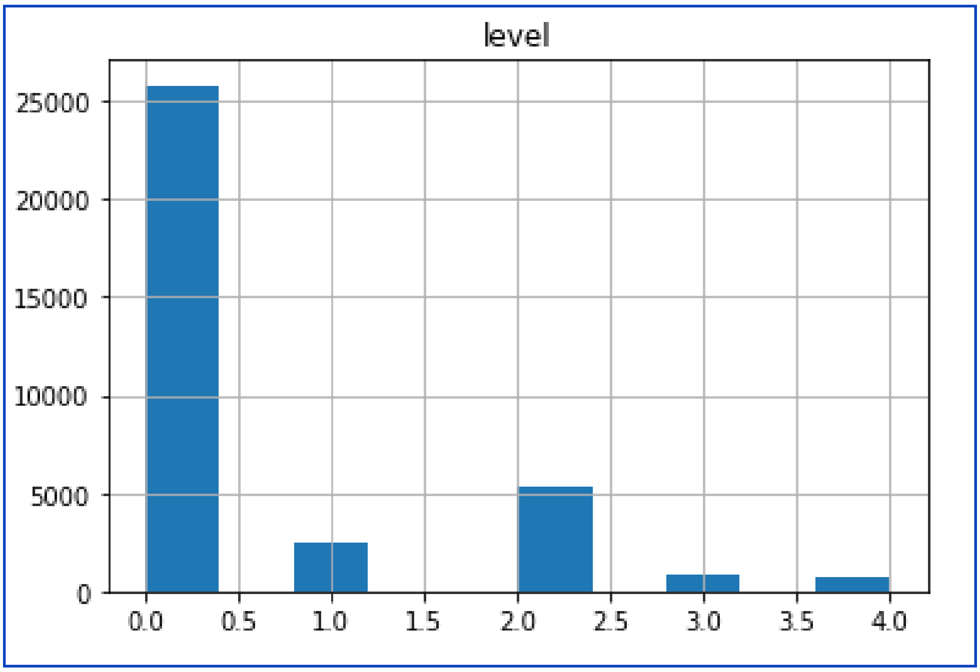

Figure 4: Count of Levels for Retinal Images (Imbalanced Dataset)

This above dataset described in Figure 4 contains the imbalance data. We will perform the same experiment over the balanced dataset as well. To do this we have reduced the number of samples of images to train the model. The graph can be shown in Figure 5. As the number of retinal images for label 3 and 4 were very less, we have reduce the maximum number of samples to 2500. The number of samples Non-Proliferative Retinopathy (NPDR) and Proliferative Retinopathy (PDR) are very less but still it will be helpful for models to train these images in order to predict the diabetic retinopathy.

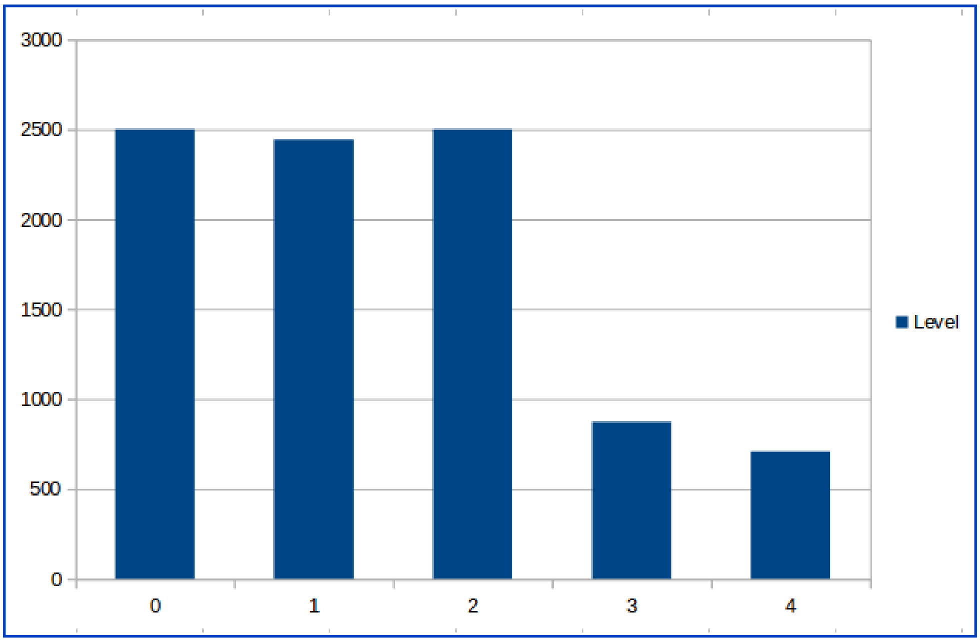

Figure 5: Count of Levels for Retinal Images (Balanced Dataset)

3.5 Project Design Process Flow

The project design process for detection of diabetic retinopathy consists of a business logic for the classification of different stages of disease. These steps involves, Data collection, feature extraction, pre-processing and transformation of data. These images will be trained using various deep learning models and results will be visualized with the help of python language. Project design process flow as shown in Figure6

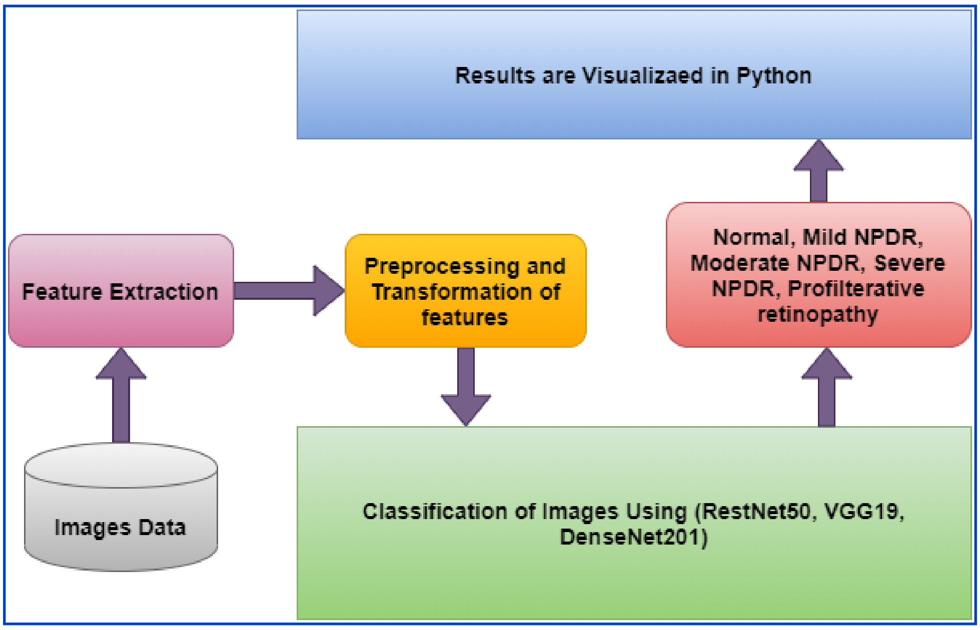

Figure 6: Project Design Process Flow of Diabetic Retinopathy Detection

4 Implementation, Evaluation and Results of Diabetic Retinopathy Disease Detection

4.1 Introduction

This section discusses the implementation, evaluation and results of different models used to detect the diabetic retinopathy disease. The different transferred learning based methods has been executed with pre-trained architectures such as ResNet50, VGG19 and DenseNet201. Transfer learning holds the pre-trained model weights, optimized number of layers and can perform feature extraction with the help of network layers. Diabetic retinopathy detection problem consist of multiple levels/stages, it can considered as a multiclass classification problem.

4.2 Implementation, Evaluation and Results of ResNet50

ResNet stands for Residual Neural network, its a kind of deep neural network which can train the data with 150+ layers. The ResNet Neural network can be implemented with the help of tensorflow and keras libraries. Our model uses 70% data for training and 30% data for testing purpose. Python language is used for implementation.

4.2.1 Evaluation and Results (Over Imbalance Dataset)

Total 35,126 retinal images has been trained over the TPU, where the number of Epochs is 10 with batch size 32. As due to the limited resources capacity of our personal system. We have used computing capacity provided by Google colab to process such large data on TPU. To evaluate the performance, proposed model uses metrics such as accuracy, loss, precision, recall and F1-score. The confusion matrix for ResNet50 is shown in Figure 7.

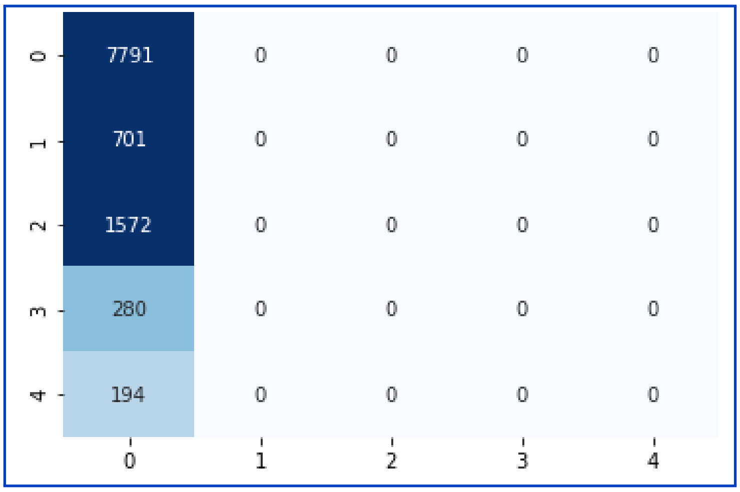

Figure 7: Confusion Matrix for ResNet50 (Imbalanced Dataset)

The accuracy achieved by ResNet50 is 73.98% with the loss of 0.84. Whereas, the precision, recall and f1-score for label 0 is 0.74, 1 and 0.85 respectively. All other label values are approximately 0. Same scenario it to be noted in Figure 7confusion matrix where highest number of true positives are found for label 0, whereas all other label values are found to be null or 0 . This problem persists because of imbalanced labeled data. Figure 4 shows that the number of images for label 0 are more as compared to other labels.

4.2.2 Evaluation and Results (Over Balanced Dataset)

In this experiment, we have tried to balance the dataset labels. Because as from the result of previous confusion matrix we could check that model was unable to identify the other levels except 0 (Normal). All the values for mild, moderate, Severe NPDR and Proliferative retinopathy is almost 0. To solve this issue we have reduced the number of images to form a balanced dataset. The number of samples taken for balanced dataset can be found in Figure 5. The maximum accuracy by ResNet50 model for balanced diabetic retinopathy dataset achieved is 32.38%. This accuracy has been achieved by training 2500 maximum samples with 40 number of epochs. Graph shown in Figure 8.

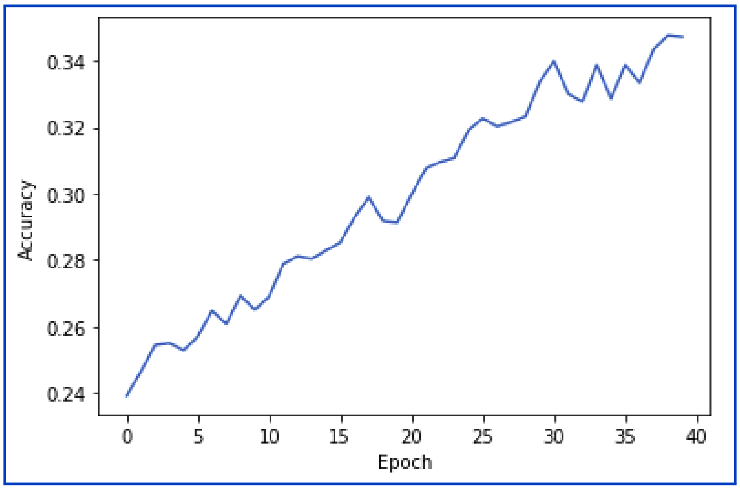

Figure 8: Accuracy Graph for ResNet50 (Balanced Dataset)

As we can observe that as increasing the number of epochs increases the accuracy and reduces the loss associated with it. The minimum loss calculated for 40 epochs by ResNet50 is 1.46 for balanced dataset. The graph for loss is shown in Figure 9.

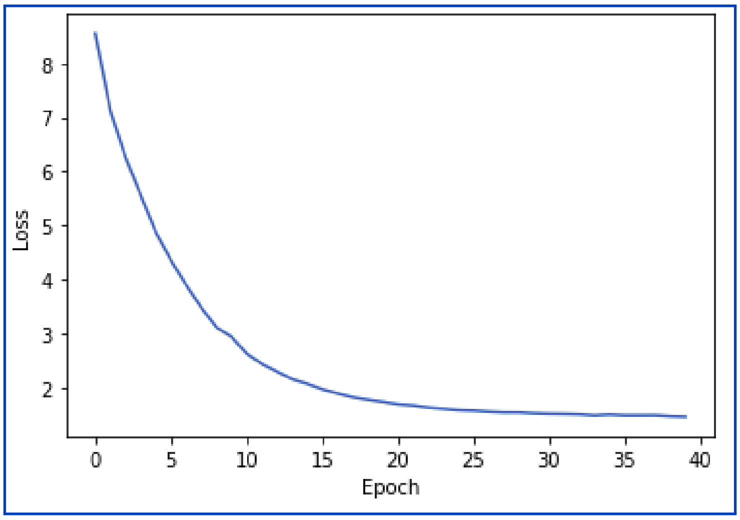

Figure 9: Loss Graph for ResNet50 (Balanced Dataset)

To analyze the results in more details we have plotted a confusion metrics. Confusion matrix will help us to find out the relationship between predicted and actual values for each target attributes. When we compare the confusion matrix of balance dataset with imbalanced data using ResNet50, we can observe that model is able to identify all the labels. Although accuracy is not that good but that can be improvised by adding more number of training samples. The Confusion matrix using ResNet50 for balanced data is shown in Figure 10.

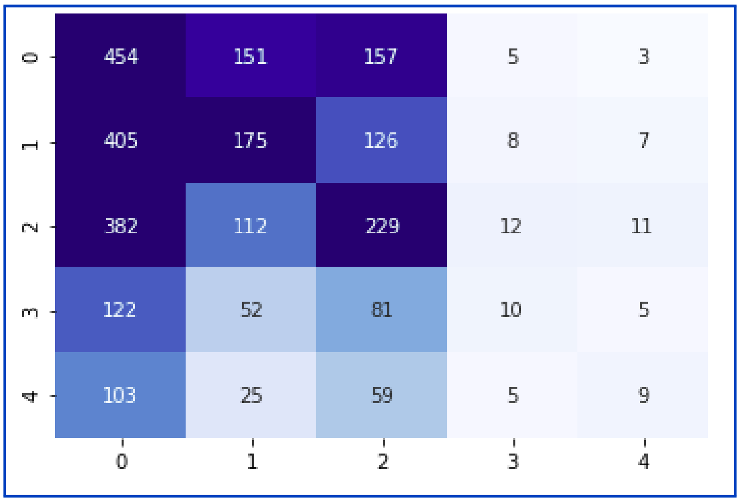

Figure 10: Confusion Matrix for ResNet50 (Balanced Dataset)

4.3 Implementation, Evaluation and Results of VGG19

VGG19 uses 3 X 3 convolution layers and is characterized by its simplicity. These convolution layers are stacked on each other with increasing depth. Volume size can be reduced with the help of max pooling. The number 19 shoes the number of weight layers in the network. Similar to all other models in VGG19 dataset is divided in the training and testing set with the ratio of 70:30.

4.3.1 Evaluation and Results (Over Imbalance Dataset)

Total 35,126 retinal images has been trained over the TPU, where the number of Epochs is 10 with batch size 32. To evaluate the performance, proposed model uses metrics such as accuracy, loss, precision, recall and F1-score. The confusion matrix for VGG19 is shown in Figure 11.

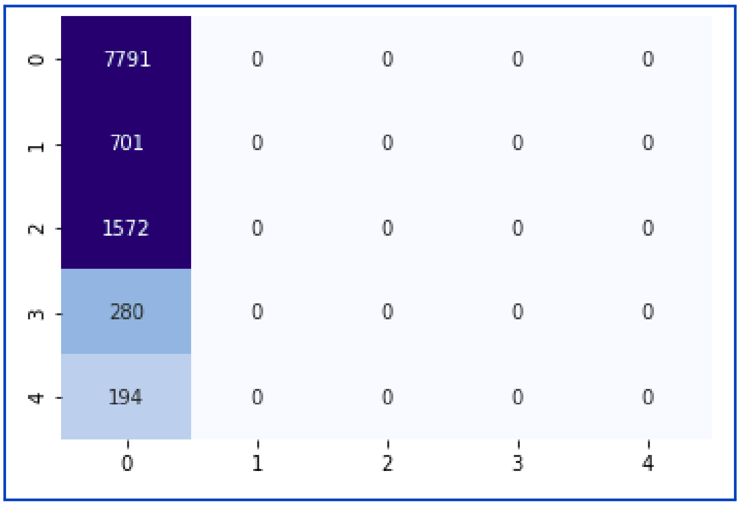

Figure 11: Confusion Matrix for VGG19 (Imbalanced Dataset)

After training the model for 10 epochs we achieved accuracy of 73.93%. While having a glance at confusion matrix we can observe that Most of the predicted values of other lables is 0. This is due to the imbalance dataset which can be observed from exploratory analysis of dataset in Figure 4.

4.3.2 Evaluation and Results (Over Balanced Dataset)

This experiment has been performed for balanced dataset using VGG19. As from the result of previous confusion matrix we could check that model was unable to identify the other levels except 0 (Normal). All the values for mild, moderate, Severe NPDR and Proliferative retinopathy is almost 0. To solve this issue we have reduced the number of images to form a balanced dataset. The number of samples taken for balanced dataset can be found in Figure 5. The maximum accuracy by VGG19 model for balanced diabetic retinopathy dataset achieved is 27%. This accuracy has been achieved by training 2500 maximum samples with 40 number of epochs. Graph shown in Figure 12.

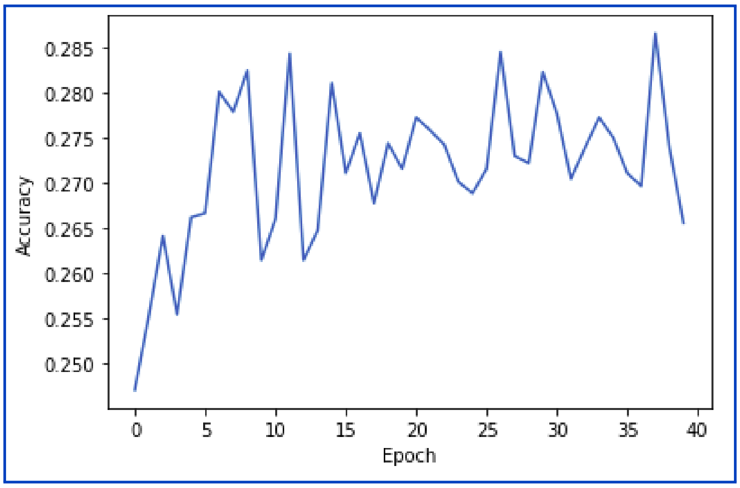

Figure 12: Accuracy Graph for VGG19 (Balanced Dataset)

In the VGG19 we can observe that after 4 epochs there is no reduction in loss. Even there is not much increase in accuracy between epoch 10 and epoch 40. The VGG19 uses only 19 number of weights to train the model. The model does not learn much in detail. When we train this over imbalanced dataset accuracy was good so training on more image can definitely help in achieving better accuracy. The loss in respect to epoch with the help of graph is shown in Figure 13.

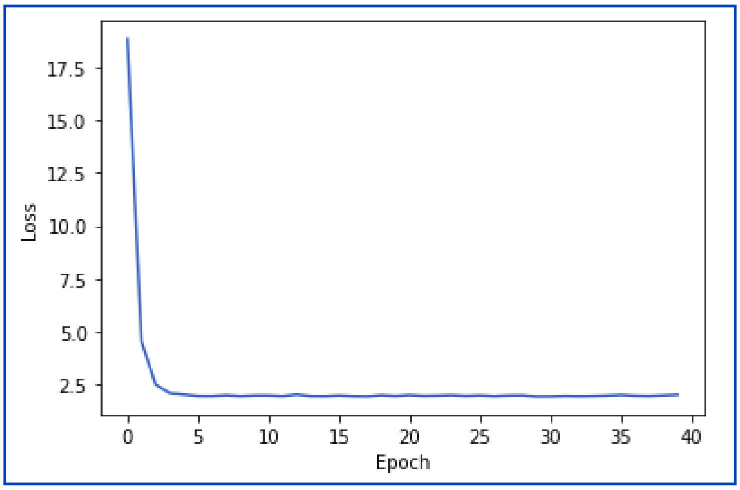

Figure 13: Loss Graph for VGG19 (Balanced Dataset)

The results of confusion matrix is better with respect to imbalance dataset. Model is able to identify the almost all the labels. But still it needs more training data to identify the images correctly. Confusion matrix for VGG19 over balanced dataset is shown in Figure 14.

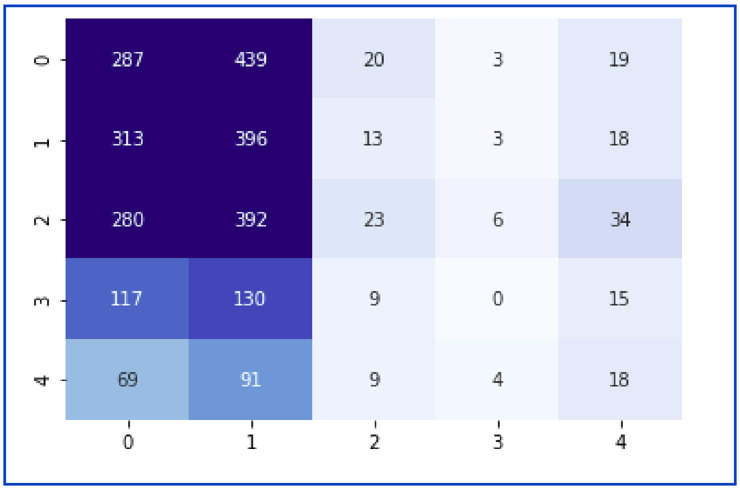

Figure 14: Confusion Matrix for VGG19 (Balanced Dataset)

4.4 Implementation, Evaluation and Results of DenseNet201

To increase the depth of deep convolution neural network DenseNet201 was proposed. DenseNet is based on the concept of ResNet but the difference between both is DenseNet do not sum incoming feature maps with output feature map. Instead it concatenate them. DenseNet-201 is CNN that is 201 layers deep. Similar to all other models in DenseNet dataset is divided in the training and testing set with 70:30 ratio.

4.5 Evaluation and Results (For Imbalanced Dataset)

Total 35,126 retinal images has been trained over the TPU, where the number of Epochs is 10 with batch size 32. To evaluate the performance, proposed model uses metrics such as accuracy, loss, precision, recall and F1-score. The confusion matrix for DenseNet201 is shown in Figure 15.

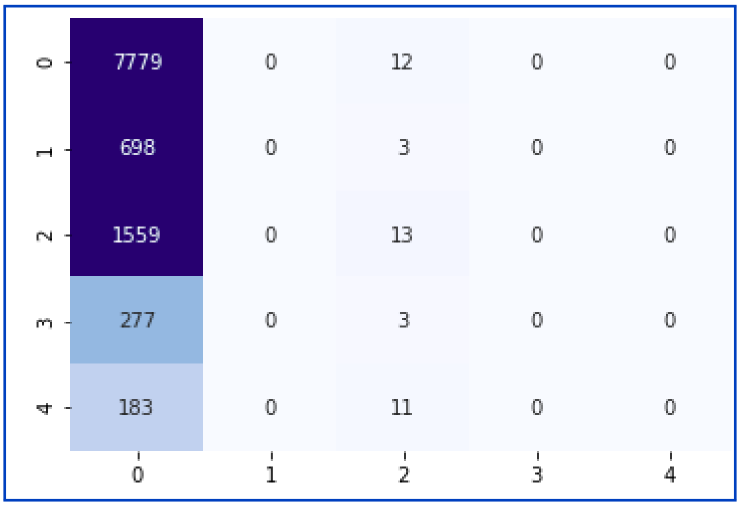

Figure 15: Confusion Matrix for DenseNet201

The maximum accuracy has been achieved is 73.94% over the 10 epochs. with the precision, recall and f1-score of 0.74, 1 and 0.85. As the confusion matrix in the Figure 15 indicates the some integer values for label 0 and label 2. It means DenseNet201 is able to identify the label2 images as well. From the Figure 4 it is also clear the label 2 has the second highest value in the dataset. Therefore DenseNet201 model is able to predict that type of images.

4.6 Evaluation and Results (For Balanced Dataset)

This experiment has been performed for balanced dataset using DenseNet201. As from the result of previous confusion matrix we could check that model was unable to identify the other levels except 0 (Normal). All the values for mild, moderate, Severe NPDR and Proliferative retinopathy is almost 0. To solve this issue we have reduced the number of images to form a balanced dataset. The number of samples taken for balanced dataset can be found in 5. The maximum accuracy by DenseNet201 model for balanced diabetic retinopathy dataset achieved is 34.4% which is best among all the models. This accuracy has been achieved by training 2500 maximum samples with 40 number of epochs. Graph shown in Figure 16.

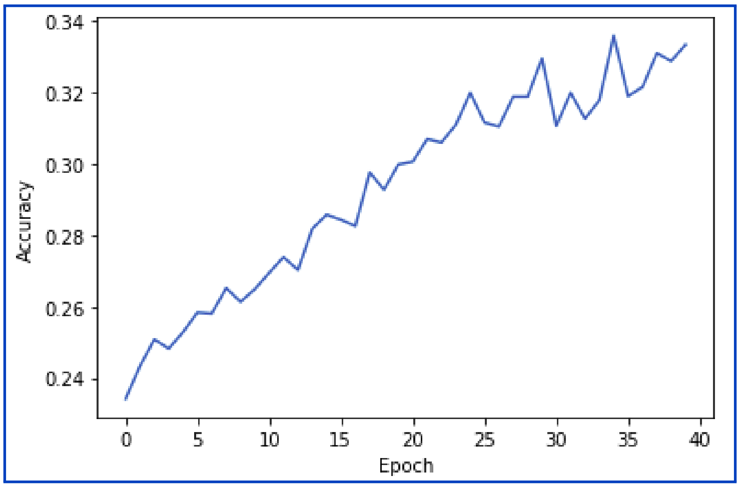

Figure 16: Accuracy Graph for Densenet201 (Balanced Dataset)

DenseNet201 model has provided the best accuracy as compared to other models. We can observe that there is significant reduction in the loss after every epochs. The accuracy of the model can be improved by increasing the number of epochs also by training the model over more sample images. The DenseNet201 model learns much in detail. When we train this over balanced as well as imbalaced dataset. The loss in respect to epoch with the help of graph is shown in Figure 17.

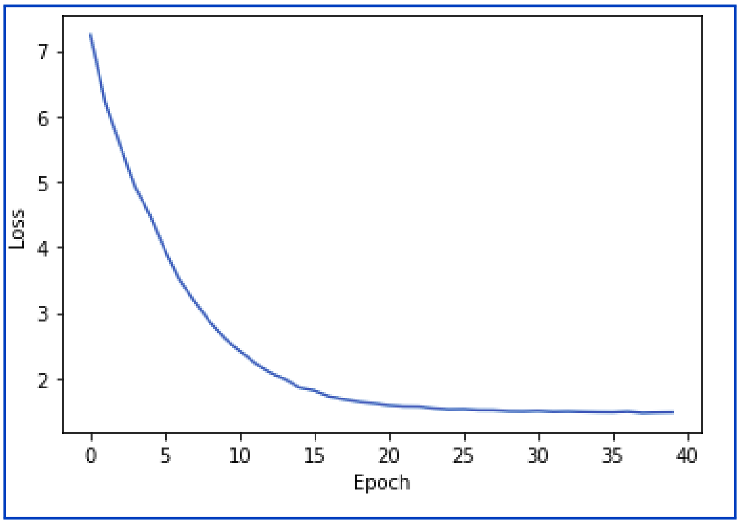

Figure 17: Loss Graph for DenseNet201 (Balanced Dataset)

The results of confusion matrix is better with respect to imbalance dataset. But still it can identify the label 3 and 4. As the training data of label 3 and label 4 is not much. But accuracy achieved by DenseNet201 is more as compared to other algorithms Still it needs more training data to identify the images correctly. Confusion matrix for DenseNet201 over balanced dataset is shown in Figure 18.

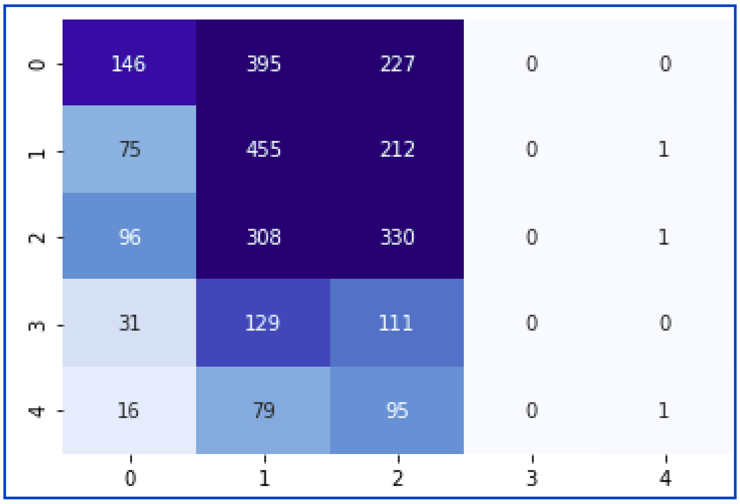

Figure 18: Confusion Matrix for DenseNet201 (Balanced Dataset)

4.7 Comparison of Developed Models & Conclusion

In this work, we have performed a comparative analysis between the 3 Deep convolution neural network models which are RestNet50, VGG19, and DenseNet201. These models have been applied for both Balanced and imbalanced dataset. For imbalanced dataset the accuracy, precision, recall and f1-score of every model is nearly same. The main issue here is every model is only able to predict value for label 0 as the number of training data for label 0 was very high. While observing the confusion matrix for all the models for imbalanaced dataset it has been noted that VGG19 and ResNet50 are unable to predict the values of label 2. Whereas, the DenseNet201 model was able to correctly predict the values of label 2 for imbalanced dataset. It is also to be noted that our dataset contains 5292 samples for label 2. Still results are not that better in numbers. But when we compare the results with other models Dense201 results are better but still accuracy is almost same. When we performed same experiment over balanced dataset, we observed reduction in the accuracy to a great extent. The low accuracy is due to the small size of training data, the number of samples very only 2500 when we balanced the dataset. Also we observed that DenseNet201 has achieved maximum accuracy of 34.4%, which is best among all the model but densenet was unable to predict the label 3 and label 4 as they were less in the number. Model VGG19 and RestNet are able to identify more number of images for label 3 and label 4 but their accuracy is less as compared to DenseNet201 Model. If we can generated the number of samples for training data by maintaining a balanced dataset, we can increase accuracy and decrease the loss to a great extent. The precision recall and f1-score for each model in described in Figure 19 and Figure 20.

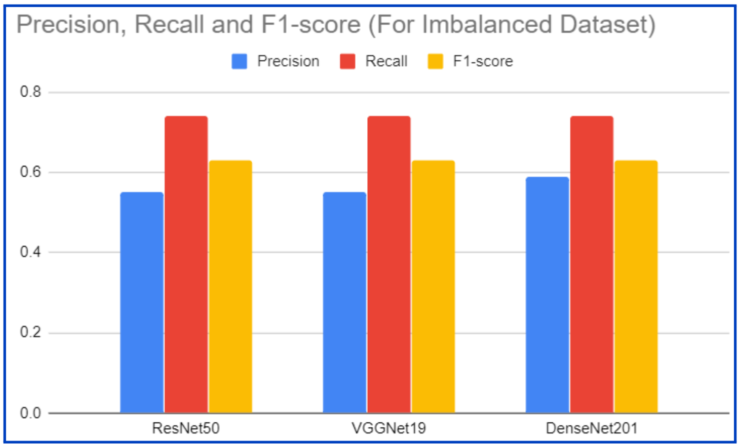

Figure 19: Precision, Recall and F1-Score for Imbalanced Dataset

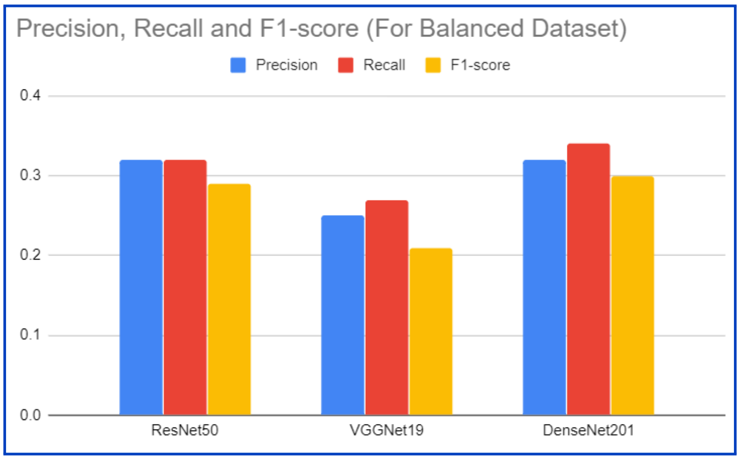

Figure 20: Precision, Recall and F1-Score for Balanced Dataset

5 Conclusion and Future Work

To detect the diabetic retinopathy disease and its multiple stages, we have used deep learning convolution neural network with different predefined architectures. The dataset has been tested with different deep learning models which are ResNet50, VGG19 and DenseNet201. Although every model has the similar accuracy around 74% for imbalanced dataset. by deep diving into the results we found that the accuracy of 74% is only due to label 0 images. Mostly models are unable to predict the images with other labels because the size of training data is very less as compared to label 0 data. Due to the imbalance dataset, behaviour of every model is approximately similar. Another experiment has been performed by reducing the size of training dataset to 2500 samples for making balanced dataset. Here we achieve the maximum accuracy of 34.4% with the help of DenseNet201. Due to the limited computing capacity of system we could not perform more operations on images also training the image dataset requires high processing GPUs and TPUs. Performing both the experiments we have observed that we can increase the accuracy of model to a great extent by increasing the training data and number of epoch. In future work, we will train the more dataset to achieve the better accuracy also in order to deal with imbalance dataset Various techniques can be used to deal such as data augmentation which includes Image flipping, smoothing, contrast adjustment, rotations, cropping etc. Adjusting the network parameters can also helps us to reduce the loss and increase in accuracy.

Acknowledgement

I would like to thank Dr.Catherine Mulwa for all the guidance and feed-backs throughout the supervision sessions for doing this project. I would also like to thank my Mother, Father and Sister for all the support and trust in me.

References

Abràmoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C. and Niemeijer, M. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning, *Investigative ophthalmology & visual science* **57**(13): 5200–5206.

Amin, J., Sharif, M. and Yasmin, M. (2016). A review on recent developments for detection of diabetic retinopathy, *Scientifica* **2016**.

Bellemo, V., Lim, Z. W., Lim, G., Nguyen, Q. D., Xie, Y., Yip, M. Y., Hamzah, H., Ho, J., Lee, X. Q., Hsu, W. et al. (2019). Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in africa: a clinical validation study, *The Lancet Digital Health* **1**(1): e35–e44.

Chetoui, M., Akhloufi, M. A. and Kardouchi, M. (2018). Diabetic retinopathy detection using machine learning and texture features, *2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, IEEE, pp. 1–4.

Doshi, D., Shenoy, A., Sidhpura, D. and Gharpure, P. (2016). Diabetic retinopathy detection using deep convolutional neural networks, *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, IEEE, pp. 261–266.

Gargeya, R. and Leng, T. (2017). Automated identification of diabetic retinopathy using deep learning, *Ophthalmology* **124**(7): 962–969.

Ishtiaq, U., Kareem, S. A., Abdullah, E. R. M. F., Mujtaba, G., Jahangir, R. and Ghafoor, H. Y. (2019). Diabetic retinopathy detection through artificial intelligent techniques: a review and open issues, *Multimedia Tools and Applications* pp. 1–44.

Kumaran, Y. and Patil, C. M. (n.d.). A brief review of the detection of diabetic retinopathy in human eyes using pre-processing & segmentation techniques.

Lam, C., Yi, D., Guo, M. and Lindsey, T. (2018). Automated detection of diabetic retinopathy using deep learning, *AMIA Summits on Translational Science Proceedings* **2018**: 147.

Lim, G., Lee, M. L., Hsu, W. and Wong, T. Y. (2014). Transformed representations for convolutional neural networks in diabetic retinopathy screening, *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Neural Network Technique for Diabetic Retinopathy Detection (2019). *International Journal of Engineering and Advanced Technology Regular Issue* **8**(6): 440–445.
URL: *https://www.ijeat.org/wp-content/uploads/papers/v8i6/E783506851.pdf*

Nguyen, Q. H., Muthuraman, R., Singh, L., Sen, G., Tran, A. C., Nguyen, B. P. and Chua, M. (2020). Diabetic retinopathy detection using deep learning, *Proceedings of the 4th International Conference on Machine Learning and Soft Computing*, pp. 103–107.

Ogunyemi, O. and Kermah, D. (2015). Machine learning approaches for detecting diabetic retinopathy from clinical and public health records, *AMIA Annual Symposium Proceedings*, Vol. 2015, American Medical Informatics Association, p. 983.

Paranjpe, M. J. and Kakatkar, M. (2014). Review of methods for diabetic retinopathy detection and severity classification, *International Journal of Research in Engineering and Technology* **3**(3): 619–24.

Patel, P., Sharma, K. and Gaudani, H. (2016). Diabetic retinopathy detection systems : A review.

Prabhjot Kaur, Somsirsa Chatterjee, D. S. (2019). Neural network technique for diabetic retinopathy detection.

Rakhlin, A. (2018). Diabetic retinopathy detection through integration of deep learning classification framework, *bioRxiv* p. 225508.

Subhashini, R., Nithin, T. and Koushik, U. (n.d.). Diabetic retinopathy detection using image processing (gui).

Xu, K., Feng, D. and Mi, H. (2017). Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image, *Molecules* **22**(12): 2054.

Zago, G. T., Andreão, R. V., Dorizzi, B. and Salles, E. O. T. (2020). Diabetic retinopathy detection using red lesion localization and convolutional neural networks, *Computers in biology and medicine* **116**: 103537.