Research Project

MSc Research Project

Data Analytics

Suvrojitmoy Ghosh

Student ID: x19179090

School of Computing

National College of Ireland

Supervisor: DR. Rashmi Gupta

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Suvrojitmoy Ghosh		
Student ID:	X19179090		
Programme:	MSC Data Analytics	**Year:**	2019-2020
Module:	MSC Research Project		
Lecturer:	DR. Rashmi Gupta		
Submission Due Date:	17/8/2020		
Project Title:	Research Project		
Word Count:	8418	**Page Count:** 21	

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: 17/08/2020

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):	

DETECTION OF DIABETIC RETINOPATHY USING DEEP LEARNING

SUVROJITMOY GHOSH

X19179090

Abstract

Diabetic retinopathy is a disease caused due to untreated diabetes for a long period of time. This disease is increasing world-wide and is a major reason for blindness. The effective way of cure for this disease is to detect it at an early state. Even though there are different types of screening techniques but there is still scope for improvements. One of the way to improve the accuracy is by concentrating more on the factors present in the eye which helps in determining the stages of diabetic retinopathy, and secondly by using pre-trained deep convolution deep neural network which can reduce the cost of training. In this paper we have experimented with different types of pre-processing techniques and their effects on enhancing the model performance. Finally the enhanced images were classified using pre-trained models. Total six pre-trained models were used out of which three were fine-tuned namely ResNet-50, Inception-v3 and InceptionResNet-v2 and the rest were used as conventional transfer learning models namely Inception-v4, DenseNet-169 and DenseNet-201. A brief comparison was drawn between the fine-tuned models and the conventional models, using the evaluation metrics like accuracy, validation accuracy and model loss. From the above mentioned experimentation we can conclude that pre-processed images along with fine-tuned pre-trained models are the best combination for diabetic retinopathy classification.

1. INTRODUCTION:

1.1 Background

In recent years diabetic retinopathy has the most number of attention all over the world. Human eye is the most delicate part of the human body, without which there will be no visualisation. Diabetic retinopathy is one such disease which can lead to blindness. It is a serious threat to the human beings, which destroys the blood vessels present in the sensitive tissues of the eyes. An estimation of the World Health Organization shows that around 170 million of the world population is affected with diabetic retinopathy, and is expected to reach 366 million by the end of 2030 (Islam, et al., 2019). Every year people suffering from diabetic retinopathy is increasing, for which this has been an area of serious concern. The eyes get abnormalities during diabetic retinopathy such as microaneurysms, hard exudates, soft exudates, hemorrhages, development of cotton wool spot and so on (Krishnan, et al., 2018). There are two different stages of diabetic retinopathy proliferative and non-proliferative. Research has shown prolonged untreated diabetes leads to proliferative diabetic retinopathy that causes blindness. The only possibility that this disease can be cured if it is detected at an early stage.

1.2 Motivation

In general diabetic retinopathy detection is very time consuming process due to which the treatment is delayed and thus leading to blindness of the patient. Early diagnosis of this disease has estimated over 90% of the patients suffering from diabetic retinopathy can be cured. Now there are two main methods for detecting diabetic retinopathy, either manually or through computer aided systems. To detect the disease manually the doctors has to be very expert and this process is very time consuming as well. On many occasions this process has proven to be inefficient especially if a large number of images are presented. Manual diagnosis also requires more resources, thus proving its disadvantages over computer aided diagnosis. On the other hand computer aided diagnosis is very efficient in detecting a small indication of diabetic retinopathy, without any difficulty and without any resources (Doshi, et al., 2020). Therefore it eliminates the need of manual labour for detecting diabetic retinopathy. In recent time artificial intelligence and deep learning and cross application of artificial intelligence with other disciplines has flourished in medical field. Artificial intelligence has two main advantages of early detection of diabetic retinopathy; less amount of human error, with minimum workload compared to manual diagnosis and it is more efficient in detecting lesions in the retina. Artificial intelligence has a huge impact on detecting diabetic retinopathy either with machine learning techniques or with deep learning.

Although a good amount of work has already been conducted, and a huge progression has been made towards detection of diabetic retinopathy, but still a large number of people are suffering from the disease. Even today there are so many places where diagnosis is still not possible due to lack of technology and resources thus leading to delay of treatment. It has also been observed that classification of the disease has often been misleading for which patients are still suffering thus pointing towards the scope of improvement (Li, et al., 2020). Therefore we urgently need to use medical image recognition machines, which will help in classification of the disease, consume less human resources and provide a powerful reference for the doctors.

1.3 Research Objectives

Table 1- Research objectives

RESEARCH OBJECTIVES	DESCRIPTION	METRICS
Objective1	Review of diabetic retinopathy detection and identification of gaps in research.	
Objective2	To visualize the effect of different pre-processing techniques for image enhancement.	
Objective3	Fine-tuning of pre-trained models like ResNet50, Inception-v3 and InceptionResNet-v2	Accuracy, validation accuracy and model loss
Objective4	Pre-trained models which were not fine-tuned like Inception-v4, DenseNet-169 and DenseNet-201	Accuracy, validation accuracy and model loss
Objective5	Comparison between the models which were fine-tuned with the models which were not fine-tuned	

1.4 Research question

Question1: How well can image pre-processing techniques along with transfer learning reduce the model learning time and improve the classification performance?

Question2: How effectively fine-tuned transfer learning model performs as compared to a conventional transfer learning model for image based classification?

2. RELATED WORK:

Numerous number of work has been conducted in the field of research for image based classifications. Here in this project we have divided the literature review in two different sections. The first part is based on exploring the architectures which helped us to determine the different classifiers used for this research. The second part is based on the previous researches conducted in the field of Diabetic Retinopathy

2.1 LITERATURE REVIEW OF IMAGE BASED CLASSIFICATION MODELS:

In the year 2020 the author (Li, et al., 2020) conducted a research on bearing fault diagnosis on Time-frequency graphs using InceptionResNetV2 and deformable convolutional networks. Now it has been observed that due to the fixed geometric structures under difficult working condition, it is very difficult for a normal Convolutional Neural network to capture the fault features in order to diagnose fault. To overcome these challenges a novel approach of combining InceptionResnetV2 with a deformable convolutional network also known as DEIN was applied. Here in this method the deformable convolution networks were used instead of the basic convolutional layers for some specific layers. An auxiliary classifier along with a main classifier were used for the classification result, as well as to adopt the non-rigid characters and large receptive field in time-frequency graph. In during the experiment one-dimensional signals were converted to Time-frequency graphs, which helps to determine the useful features during the training process. To further verify the generalization ability of the model, cross-over tests based on two separate datasets were applied. The model achieved an accuracy of 99.87% and 94.52% precision score, which were better compared to the previous state-of-the-art of CNN models.

The author (Kumar, et al., 2019) here conducts a research on identifying the dog breeds using deep learning. This was a competition where various different dog images of ImageNet were provided without specifying their breeds in order to identify them into different categories. Deep learning architectures were used in order to learn statistical patterns which will enable it to learn different breeds of dogs. The model train itself with the different features extracted from the images and represent them numerically. Initially the images were divided into numerous patterns, and a training batch was set accordingly. An algorithm was used to split the dataset, and combine the descriptors, to extract the image channel information, which would act as input for the neural network. Finally a convolutional network was designed to identify the breeds of the dog. Different architectures were used like InceptionV3, InceptionResNetV2, Resnet50 and Resnet101 among which Resnet101 was the best model with an accuracy of 90.26%.

The author (Liu, et al., 2018) conducts an experiment for early gastric cancer classification on magnifying narrow-band imaging images using transfer learning with convolutional neural network. In this experiment fine tuning of Convolutional neural network with transfer learning is applied to automatically classify the M-NBI images into two different groups' namely gastric images and ECG images. The paper explores four different aspects of transfer learning on image classification performance namely basic architectures of deep Convolutional neural network, basic architecture, total number of fined tuned layers and the network size. Three different classification models were selected to conduct the experiment, which were VGG16, InceptionV3 and InceptionResNetV2. The results clearly shows that transfer learning of deep convolutional neural network is better compared to any other methods. Among the other models used in the experiment InceptionV3 outperformed with an accuracy of 98%, sensitivity of 98.1% and specificity of 98.9%.

In the year 2017, the author (Xia , et al., 2017) worked on a research which aimed at classifying various different images of flowers with the help of transfer learning. InceptionV3 was the only transfer learning model which was used for this experiment in order to improve the accuracy. Two different datasets namely Oxford-17 and Oxford-102 were included to conduct the experiment. Oxford-102 contained 102 different types of species, with each species containing 40-258 range of images, whereas the Oxford-17 contains 17 different types of flower species with each species containing 80 images. Significant improvements in both the training and validation accuracy were observed. The training and the validation accuracy for the Oxford-17 dataset were 100% and 99%, and for the Oxford-102 dataset the training and the validation accuracy were 100% and 95%.

The author (Emara, et al., 2019) conducted a research where an existing InceptionV4 was modified for an imbalanced skin cancer classification. The main idea behind the modification of the network architecture, was to prove that not always an ensemble complex model is necessary for image based classification. The dataset which was used to conduct the experiment was HAM10000 dataset. The model architecture was enhanced by feature reusing obtained from long residual layers, where the feature extracted from previous layers were concatenated with high-intensity layers in order to increase the classification performance. Sampling approach was involved to deal with the imbalance nature of the dataset. The model achieves an accuracy of 94.7% on the testing data which was obtained from the International Skin Imaging Collaboration (ISIC) dataset.

The research shows the importance of data augmentation for image based classification problem. The author (Ayan & Unver, 2018) conducted a research on classification of skin lesion via deep learning using augmentation. Melanin cancer is a very unique type of cancer which can only be cured if it is detected during its early stages. Deep learning architectures were used to distinguish between benign lesions and malignant lesions. Now for a deep learning architecture to be successful a good amount of training data is required. The dataset selected for conducting this experiment was insufficient, therefore Data Augmentation technique was used to create synthetic images from the existing images. This technique is very useful to create a powerful classifier from a very small amount of data. The experiment was conducted by training the network with both augmented images and with non-augmented images to detect malignant skin lesions. It was observed that the network trained with augmented images yield better result than the network trained without augmenting the data.

In this research the author (Rogers, et al., 2019) made use of adversarial artificial intelligence for overhead image classification models. The aim of the experiment was to compare the already existing deep neural networks with the advanced adversarial artificial intelligence, and also to focus on overhead imagery especially of the ships. Publicly available dataset in kaggle was used for this experiment, and models like ResNet50, DenseNet201 and InceptionV3 were used to detect ships in overhead imagery. The adversarial artificial intelligence was used to misclassify the image contents, so that the perturbations can be recreated in the physical world. The prime focus was to create physical condition which can reduce the accuracy within the network. Even though there is existence of military applications for this kind of research, but the general findings can be applied on any artificial intelligence overhead classifications. This research will explore the vulnerabilities as well as visualize the vulnerabilities of the existing neural network.

The author (da Nóbrega, et al., 2018) had conducted a research using CT scan images of the lungs to classify the lung nodule using deep transfer learning. This research was motivated by the success of deep learning in the field of image classification. The aim of the research was to explore the performance of deep transfer learning for classifying malignancy nodules using CT lung images. The CNN architectures which were involved in the research were VGG16, VGG19, MobileNet, Xception, InceptionV3, ResNet50, Inception-ResNet-V2, DenseNet169, DenseNet201, NASNetMobile, and NASNetLarge. All this networks were used as a feature extractors to process the lung images. Further these extracted features were classified using Multilayer Perceptron (MLP), Support vector machine (SVM), K-Nearest Neighbors and random forest. To evaluate the classification performance the metrices which were used were Accuracy, Area under the curve, True positive rate, f1-score and Precision. ResNet-50 along with SVM-RBF was considered the best model among the other models involved in the experiment. The model gave an accuracy of 88.41% and AUC score of 93.19%. Thus deep transfer learning proved to be a relevant strategy in terms of medical image classification.

2.2 LITERATURE REVIEW OF DIABETIC RETINOPATHY IMAGE CLASSIFICATION MODELS:

The author (Qiao, et al., 2020) conducted a research using deep learning algorithms to detect diabetic retinopathy using early diagnosis system and microaneurysms for non-proliferative diabetic retinopathy. The main aim of the experiment was to detect microaneurysms in the retinal fundus images using convolutional neural network along with deep learning. Graphical processing unit (GPU) was involved in this experiment to accelerate the network in order to get a high performance with low latency inference. Images were classified into two different classes namely normal and infected using semantic segmentation algorithm. This algorithm helps to detect the microaneurysms based on the image pixels. Therefore this will help the ophthalmologists to grade the disease as early non-proliferative diabetic retinopathy, moderate non-proliferative diabetic retinopathy and severe proliferative diabetic retinopathy.

The author (Li, et al., 2020) conducted a research for grading both Diabetic macular edema and Diabetic retinopathy using Cross-Disease attention network. The aim of the experiment was to grade both Diabetic retinopathy as well as Diabetic macular edema by exploring the internal relation between the diseases using Cross-disease network. Two of the specific contributions of the research were disease-dependent attention module which helps in finding the internal relationship between the two diseases and disease specific selective module which can specifically learn useful features for every individual disease. The two attention modules were integrated within a deep network to extract features and also to maximize the overall performance. The model was further evaluated on two publicly available datasets namely, 2018 IDRiD challenge dataset and Messidor dataset. The best result was achieved for the IDRiD dataset and the model also outperforms other methods on Messidor dataset.

In the year 2020 the author (Zhu, et al., 2020) conducted a research on the patients with Type2 Diabetes and diabetic retinopathy in china to identify the tortuous branching arterioles, venules and retinal main within the patients. For this research 495 retinal images of the patients suffering from Type2 diabetes were taken by using ophthalmoscope. Based on the theory of best-fit curves in the roto-transational group the images were extracted from the main retina, branching arterioles and venules in the whole retina depending on the tortuosity. It was observed that both the venular tortuosity and the retinal arteriolar were remarkable features in terms of assessing the level of severity for the diabetic retinopathy patients. It is also useful for assessing individuals with high risk of renal disease in diabetes.

The author (Patel & Chaware, 2020) conducted a research on detecting the different stages of diabetic retinopathy using fine-tuned MobileNetV2 with transfer learning. In this research a pre-trained network MobileNetv2 was used to extract features from the given set of images. The layers of the model were further modified by adding globalaveragepooling and a softmax classifier, to classify the images into multiple classes. In during the initial stage only the stacked layers were trained thus preventing the pre-trained weights from updating, and to further enhance the performance of the model, few of the weights in the model were fine-tuned. The diabetic retinopathy dataset from kaggle was used to evaluate the model, total 2929 retinal fundus images were used for training, 733 images were used for validation and a total of 1928 images were used for testing. The end result shows that fine tuning a network was very effective since it helped in increasing the accuracy from 70% to 91%. Similarly the validation accuracy increased from 50% to 81%. The model loss was approximately same which indicates that the model was a perfect fit.

In the year 2020 the author (Suresh, et al., 2020) conducted a research using the content based image retriever technique to develop an effective method to retrieve diabetic retinopathy. The proposed methodology broadcast a relationship between the intensity of the pixels and the reduced features. Intensities between the inter-plane relationships are computed using the key pixels from an edgy image. The centre pixel is identified using some selected points to determine the local binary pattern. The obtained result was enhanced by using the proposed approach, by compressing the resultant metrics, and random bin selection was used to reduce the features. The STARE dataset was used to conduct the experiment, where the average precision rate was incremented by 42.04% compared to the other methods.

The author (Trisha & Israj, 2020) conducted a research on automatic detection of diabetic retinopathy with intensity based optical disk detection. The aim of the research was to determine the three very important regions of the eye namely optic disc, macula and the exterior boundary of the eye. The analysis is based on this three regions, now the desired area of the retinal image is segregated by determining the boundary of the retinal image. Since macula is the darker region of the retinal image, therefore it is necessary to remove the unwanted dark region in order to determine macula. The images are then converted to grey from RGB, histogram for the segregated area is prepared to discover the threshold from that histogram. Then a binary image similar to that of the threshold value is determined, along with the optic disc since both the optic disc as well as the exudates have similar intensities. The result obtained shows us that the important regions were successfully identified.

The author (Doshi, et al., 2020) conducted a research on implementing downscaling algorithm and deep learning to classify diabetic retinopathy. The aim of this research was to explore several downscaling algorithms before feeding the data into a deep learning network for the final classification. In order to improve both the training as well as testing accuracy two datasets were combined together, which were diabetic retinopathy image data from kaggle and Indian diabetic retinopathy image dataset also known as IDRiD. Self-crafted image pre-processing was done before feeding the data into a neural network. A multi-channel transfer learning model InceptionV3 was used for classifying the images. The model was evaluated using accuracy, sensitivity and specificity, and it was observed that the model outperforms the previous state of the art methods.

The author (Sugasri.M , et al., 2020) conducted a research for early detection of diabetic retinopathy using a screening system. There are three different types of lesions which are extremely beneficial to determine diabetic retinopathy stages, these are haemorrhages, micro aneurysms and exudates. The aim of this experiment is to identify the different stages of diabetic retinopathy lesions using a super pixel algorithm. Skin locus segmentation was performed before feeding the images into a machine learning algorithm. Finally Support vector machine (SVM) along with linear kernel was used to classify the images into different stages of diabetic retinopathy. The model was evaluated by using accuracy metrics. The proposed technique was able to achieve a 97% classification accuracy, thus proving it to be a successful model in determining the stages of diabetic retinopathy.

The author (Kajan, et al., 2020) conducted a research by using pre-trained deep neural network to detect diabetic retinopathy. The aim of this research is to classify diabetic retinopathy symptoms automatically using pre-trained neural networks. The data for this research was gathered from EyePacs diabetic retinopathy database. There were about 5 million images out of only 35,126 were considered to conduct the research. The database was divided into 5 different diabetic retinopathy categories. The training and testing images were selected in a 4:1 ratio. The images were then pre-processed using different techniques and then finally were classified using deep neural networks. Four different network architectures were considered for this research namely VGG-16, ResNet-50 and Inception-v3. Among all these network Inception-v3 was considered the best with an average classification accuracy of 70.29% and sensitivity and specificity score above 90%.

The author (Reddy , et al., 2020) conducted a research to classify diabetic retinopathy into different stages by using an ensemble machine learning model. The aim of the research was to classify diabetic retinopathy by combining different machine learning models. The model comprises of Random forest classifier, Decision tree classifier, Logistic regression classifier, and K-Nearest neighbour classifier and Adaboost classifier. The dataset used was taken from UCI repository, the images used were firstly normalized using the min-max normalized method. After normalization the dataset was feed to the ensemble model and was then evaluated against individual machine learning models. The end result shows a clear dominance of the ensemble classifier over the other machine learning models. The ensemble model outperforms the machine learning models in terms of all the evaluation metrics. Thus the experiment was successful, stating that ensemble classifier are far better than the individual machine learning models.

In the year 2020 the author (Behera & Chakravarty, 2020) conducted a research on classification of diabetic retinopathy using support vector machine. As a matter of fact diabetic retinopathy is a disease which can cause blindness and it is very important that the disease is diagnosed at an early stage. Therefore a regular screening is necessary, even though there are several automatic screenings are available but there is still scope for improvement in accuracy of prediction. One of the most important part to consider in terms of increasing the accuracy is Exudates. In this research the author uses two pre-defined feature extraction techniques which are as speeded up robust feature (SURF) and scale invariant feature transform (SIFT) to extract the exudates regions from the retinal fundus images. Feature matrix was used to store the exudates of each retinal fundus images. These features were then feed to a Support vector machine classifier (SVM) for further classification. The model was then evaluated using the sensitivity score, for a 100 set of test images the sensitivity score was 94%.

The author (Lands, et al., 2020) conducted a research on classifying diabetic retinopathy using deep learning algorithms. The aim of this research is to speed up the disease selection using deep learning algorithms. The dataset used was taken from kaggle which had thousands of retinal images taken under different imaging conditions. As a part of image pre-processing Gaussian blur subtraction and image augmentation were performed. These pre-processed images were then feed to a deep neural network for further classification. The network architectures which were used are ResNet-50, DenseNet-121 and DenseNet-169. The model was built in such a manner that it can be used to develop system with user friendly interface, it will also be able to detect object and segment capillaries in the retinal images. The model was evaluated using accuracy, validation accuracy, loss and validation loss. DenseNet-169 was the best model compared to the other model which gave an accuracy of 95%, validation accuracy of 90% and the loss was 0.1143, whereas the validation loss was 0.2176.

The author (Zhang, et al., 2020) conducted a research using hyperparameter tuning of deep learning algorithms to classify diabetic retinopathy from regular fundus images. The research includes an automated hyperparameter tuning of inception-v4 model to detect diabetic retinopathy. The images used were pre-processed using Contrast limited adaptive histogram equalization (CLAHE), along with a histogram based segmentation process. These pre-processed images were then feed inside the inception-v4 model and hyperparameter tuning was applied to extract the features from the segmented images. The segmented images were further classified using Multi-layer perceptron (MLP). The dataset which was involved for this research was (Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology) MESSIDOR. The result obtained was evaluated based on accuracy, sensitivity and specificity, the model achieved an accuracy of 99.49%, sensitivity of 99.83% and specificity of 99.68%. Thus the proposed approach outperforms all other methods in a significant way.

2.3 Conclusion

From the above mentioned reviews it is evident that comparative analysis between different pre-processing techniques and also between fine-tuned transfer learning model and conventional transfer learning model has never been performed. All the previous work were conducted using the Kaggle dataset, MESSIDOR dataset and the IDRiD data set, whereas we have conducted the experiment on a new dataset named DeepDRiD, which was released on March 2020. The model was evaluated using Accuracy, Validation accuracy and model loss.

3. RESEARCH METHODOLOGY:

Here the research is based upon cross-industry standard process for data mining or (CRISP-DM). CRISP-DM is very useful when modelling clinical system and applying data mining, due to its stepwise procedure and general applicability it is preferred by the industry. CRISP-DM is considered as a well-planned methodology since it explains all the phases of the project along with the relationship between each task. Our research follows the following six measures that a CRISP-DM methodology consists of, these are business understanding, data understanding, data preparation, modelling, evaluation and deployment. The Figure 1 gives a description about the methodology followed by the research work

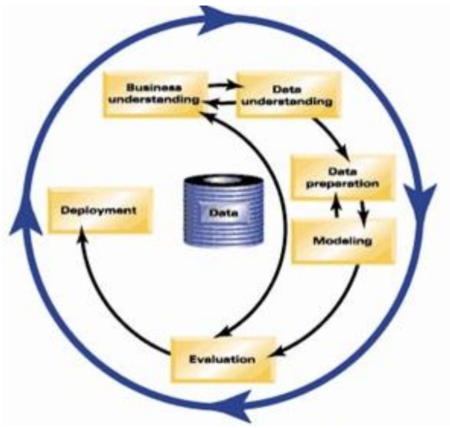

Figure 1- Steps involved in research methodology[1]

[1] https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html

3.1 DATA UNDERSTANDING:

The data used to conduct this research was taken from[2], it consists of regular fundus retinal images which were originally collected from nationwide screening for compilation for diabetes and also from the Chinese university hospitals. A total of 2000 images were obtained from 500 patients. The images had a resolution of 1956*1934 pixels and these images were also verified by experts who had identified that the images present in the dataset are of good quality, clinically relevant and not duplicated. These images were further sub-divided into five different levels of diabetic retinopathy namely no retinopathy (represented as '0'), Mild non-proliferative diabetic retinopathy (represented as '1'), Moderate non-proliferative diabetic retinopathy (represented as '2'), severe non-proliferative diabetic retinopathy (represented as '3') and Proliferative diabetic retinopathy (represented as '4'). Below are the figures of all the classes of diabetic retinopathy represented as Figure 2

Figure 2-No Retinopathy

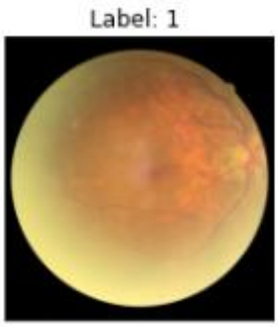

Figure 2a- Mild non-proliferative Retinopathy

Figure 2b-Moderate non-proliferative Retinopathy

Figure2c-Severe non-proliferative Retinopathy

[2] https://isbi.deepdr.org/

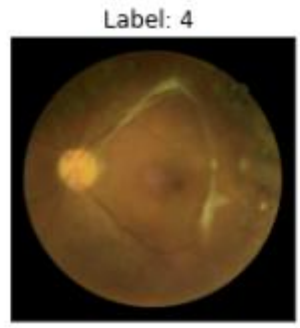

Figure 2d-Proliferative Retinopathy

3.2 DATA PREPARATION:

As mentioned the dataset consists of 2000 images which were further classified into five different categories, according to the diabetic retinopathy levels. The data was then split into training, testing and validation. The training set consists of 60% of the total data, whereas the validation and the testing data had 20% each. The dataset consisted of images which were in JPG format, both the training and validation data comes along with a CSV file which had all the relevant information regarding the image labels. Therefore we had to map all the images with their respective labels. Another major issue was that every image in the dataset had different resolutions. To avoid this situation the images were rescaled and then transformed into a fixed scale. Since these were real world images, so the images can be underexposed, overexposed or may even be out of focus, we have taken care of every situation in our proposed work. The distribution for both the training and validation set is given in Figure 3

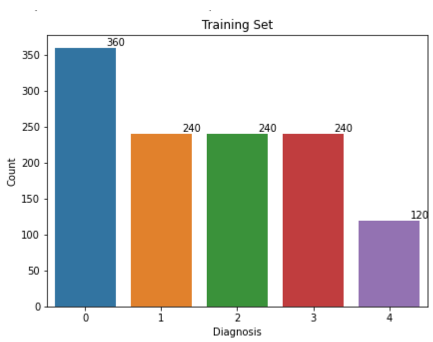

Figure3a- Training data distribution

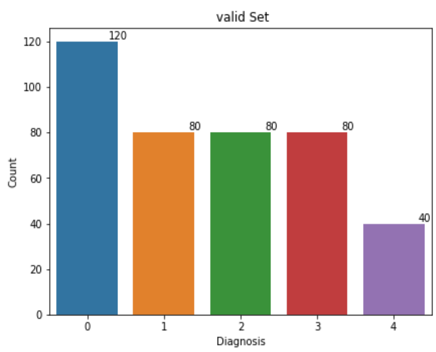

Figure 3b- Validation data distribution

3.3 DATA AUGMENTATION:

In real world problems even a small amount of data collection is near to impossible and sometimes it is very expensive, especially in medical imaging. To make the most out of very little data is a key skill while conducting a data science research project. The data that was considered for our project is comparatively less than real world datasets. Now to overcome this problem the data was augmented with different types of augmentation techniques. Augmentation is a very useful technique for increasing the size of the data by creating synthetic data over the existing dataset. This technique is very useful since it prevents the model from getting the exact same picture twice, overfitting and also helps the model to generalize better.

Data Augmentation was performed using *keras.preprocessing.image.ImageDataGenerator* class. This class allows the following steps:

- The image data can be configured with random transformations and normalization operations.
- It helps to create generators of the augmented batches using *.flow_from_dataframe (data, labels)* or *.flow_from_directory (directory*). These generators were further used with the keras models that accepts generators as inputs, *fit_generator*, *predict_generator* and *evaluate_generator*.

Below are the following augmentation steps which were involved in the research:

Rescale - It is a process by which we will multiply the data with a value before conducting other pre-processing techniques. The original image has RGB coefficients which has a range of [0-255], since this value is too large to process therefore the values were rescaled with a factor of 1/255 so that the values are between 0 and 1.

Rotation_range - It is a value by which the images are randomly rotated within a given range. The values are in degrees and it ranges from [0-180], for this research we have considered a range of [15].

Width_shift and height_shift range - within this process the images are randomly translated either vertically or horizontally. Both the width_shift_range and the height_shift_range were considered [0.1] for this research.

Horizontal_flip and Vertical_flip - this process generally refers to flipping the images horizontally and vertically.

Zoom_range - It refers to randomly zooming inside pictures. For this research a zoom range of [0.9-1.25] was considered.

Shear_range - This process is randomly used for shearing transformation. For this research a range of [0.01] was considered.

Fill_mode - While rotating the images we have to consider the blank pixel spaces generated in the augmented images. This places were filled using the nearest pixel values.

ZCA_Whitening - This process generally refers to the whitening of the data which is as close as the original data.

Brightness_range - The image data can be augmented by either brightening the images or by darkening the images. It is always ideal to train the model with both bright and dark images, so that the model can generalize better. Values less than 1.0 results in a dark image, whereas values greater than 1.0 refers to a bright image. For this research a value of range [0.5-1.5] was considered.

3.4 DATA PRE-PROCESSING:

The main aim in this phase is to perform some image enhancement techniques in order to extract some useful information from the image. Originally we have applied different types of pre-processing techniques so that we can enhance the performance of the model, and after doing research with numerous different techniques, finally some of the techniques were selected for the project which actually helped to enhance the model performance. Therefore the data pre-processing part is divided into two sections, the first part is based on the techniques which were used during the trial phase of the research and the second part consists of the techniques which were implemented in the final project. All the techniques were applied using the CV2 library.

3.4.1 TRIAL PHASE TECHNIQUES:

Below mentioned Figures 4, 5, 6, 7, 8 and 9 are the techniques which were initially used to conduct the research but these techniques were not able to enhance the performance of the model.

APPLYING MEDIAN-BLUR AND THEN REDUCING THE IMAGE RADIUS

Figure 4a- Original image

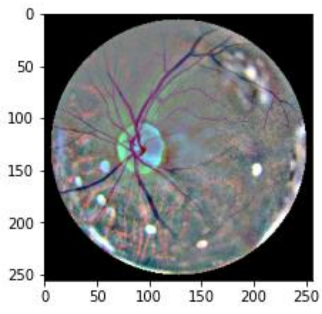

Figure 4b-Image with median-blur

SELECTIVE GAMMA CORRECTION

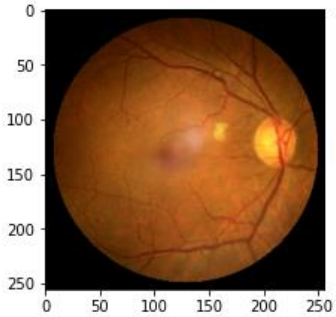

Figure 5a-Original image

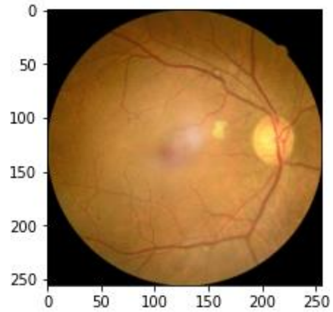

Figure 5b-Image with gamma correction

CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALISATION (CLAHE)

Figure 6a- Original image

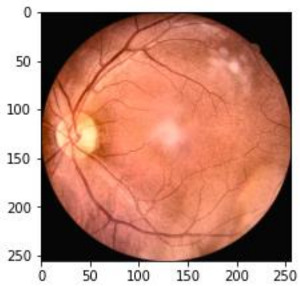

Figure 6b-Image with CLAHE

CONTRAST STRETCHING

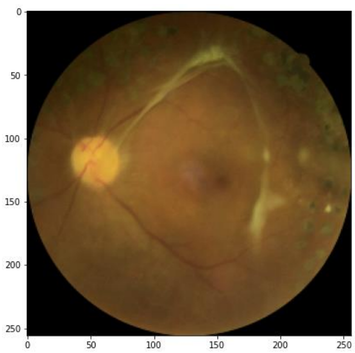

Figure 7a-Original image

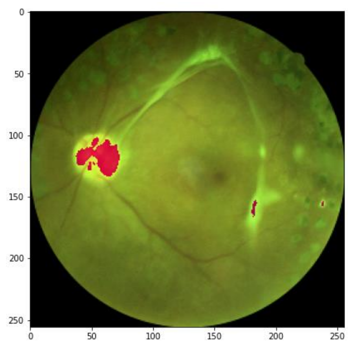

Figure 7b-Contrast stretched image

HISTOGRAM NORMALIZATION

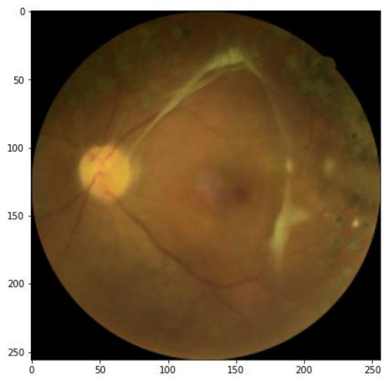

Figure 8a-Original image

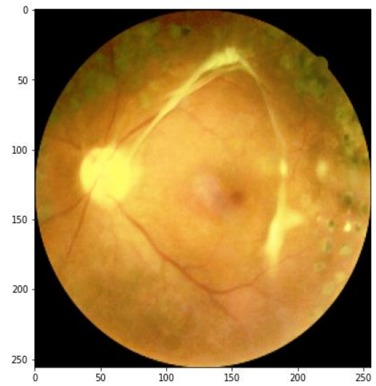

Figure 8b-Image with normalized histogram

KRISCH FILTER

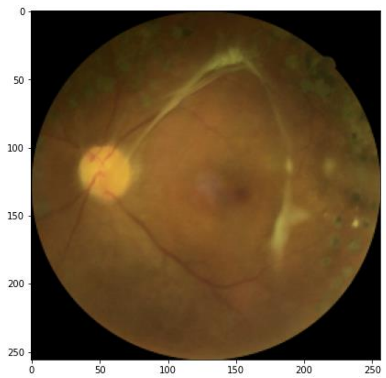

Figure 9a-Original image

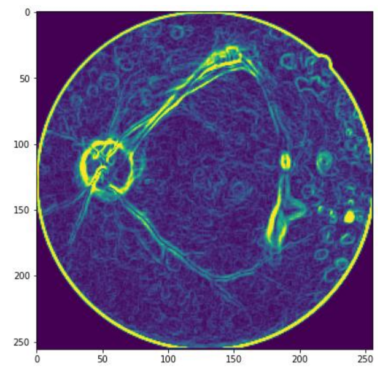

Figure 9b-Image with Krisch Filter

3.4.2 FINAL PHASE TECHNIQUES:

In this phase those techniques are mentioned which actually had a huge impact on the model accuracy and were retained for the final implementation. The technique that we have used for this research is also known as Ben Graham’s pre-processing. This technique involves the following steps

GRAYSCALE

Firstly the images were converted to grayscale in order to enhance the features present within the images, so that the model can learn better.

CIRCULAR-CROP

After converting the images to grayscale, the images were cropped from the centre in order to remove the excess black borders present within the image.

GAUSSIAN-BLUR

Gaussian blur is a technique typically used in image pre-processing for blurring an image by using a Gaussian function. The aim of this function is to reduce image noise and detail. After using circular-crop, Gaussian blur was used to enhance the image features. The final image after all this pre-processing steps is given below in Figure 10

Figure 10a-Original image

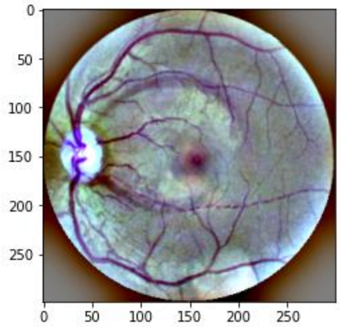

Figure 10b-Image after pre-processing

4. PROJECT DESIGN PROCESS FLOW

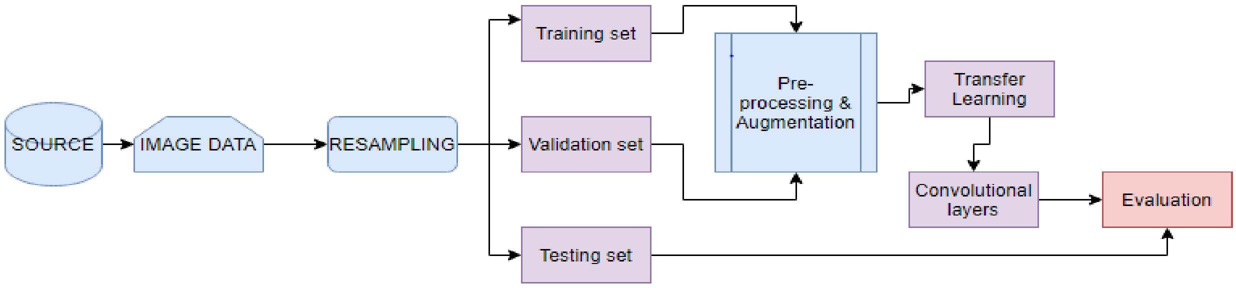

Figure 11-process flow diagram

5. DESIGN SPECIFICATION:

5.1 RESNET50

ResNet-50 is also referred to as a CNN (Convolutional Neural Network) architecture, which has got 50 layers. The network is trained on more than millions of images from the ImageNet database. This network can classify up to 1000 different classes. Pre-trained networks have two types of weights associated with them, with_top (meaning the dense fully connected layer) and without_top (removing the dense fully connected layer). For this research we have considered the without_top weight and replaced the fully connected layer with the layers of our own choice (which is mentioned in the implementation section). We have fine tune the ResNet-50 network, now fine-tuning generally refers to a process where a pre-trained network is taken and used to perform a task using its existing knowledge. Here in this situation all the existing layers of the network was freeze and only the added layers were trained. Then finally all the existing layers were unfreeze and were trained along with the added layers.

5.2 INCEPTION-V3

Inception-v3 is a convolutional neural network which has got 48 deep layers and is trained with more than one million images from the ImageNet dataset. This layer has got 11 inception modules, and each module consists of convolutional functions and pooling layers and RELU (rectified linear unit) as an activation function. The model has weights with_top and without_top, we have considered the without_top weights to conduct the experiment. The fully connected layers were replaced with some selected layers. This model was also fine-tuned in a similar way like the previous model.

5.3 INCEPTIONRESNET-V2

Inception-ResNet-v2 is again a convolutional neural network architecture which is even deeper than Inception-v3 and has got 164 deep layers. This network is also trained on millions of images and can classify up to 1000 classes. It consists of both inception modules as well as residual modules. The architecture of this model is considered to be more accurate than any previous state of art models. Similar to the other network, weights without_top was also considered and further layers were added to the model. The network was also fine-tuned the similar way like RseNet-50.

5.4 INCEPTION-V4

Inception-V4 architecture has evolved from googleNet, it is also known as a convolutional neural network architecture, which has both the inception as well as the residual layers. This network has got more inception layers than Inception-v3, and can be trained without partitioning the replicas, with memory optimization and backpropagation. For the residual version of the inception networks, cheap Inception blocks were used. After each inception block filter expansion layers were added, which are generally used for dimensionality reduction. We have implemented several versions of the residual block but only two of them are mentioned over here. Firstly Inception-resnet-v1 which matches the computational cost of Inception-v3 and secondly Inception-Resnet-v2 which has a similar computational cost as inception-v4. Pre-trained weights without_top were loaded in the model, and the architecture was further enhanced by adding fully connected layers. Finally the model was used as a classifier.

5.5 DENSENET-169

Research has shown that convolutional network can be more accurate and efficient if the connections are shorter between layers close to the input and those close to the output. Thus introduction of Densenet made it possible to achieve deeper convolutional network which connects each layer to every other layer in a feed forward fashion. DenseNet-169 is one of the models in the DenseNet group designed to perform image based classification. It is better than DenseNet-121 in terms of size and accuracy. Originally the model was trained on pytorch, and the authors converted them to caffe format. This model was also pre-trained on Imagenet dataset, and also consists of weights with and without top. The research was conducted using the no_top weights and the fully connected layers were replaced with layers suitable for classification and finally the model was used to classify the images into different classes.

5.6 DENSENET-201

DenseNet-201 is another convolutional neural network in the DenseNet group which is 201 layers deep. This model is better than DenseNet-169 in terms of size and accuracy. The network has been trained on more than millions of images from the ImageNet database. Pre-trained weights without top was considered in order to assign weight, and the top layers were replaced with best selected layers. The model was further used for image classification.

6 IMPLEMENTATION:

The research work was conducted using python language and keras library which uses tensorflow as backend. Firstly the data was transformed into a shape of [299, 299, 3]. Augmentation and pre-processing techniques were used to increase the size of the data and to reduce noise from the data. The data was then processed through pre-trained image classifiers, ResNet50, InceptionResNet-v2 and Inception-v3 (all these networks were fine-tuned), Inception-V4, DenseNet-169 and DenseNet-201 (used as transfer learning approach). After doing some research with different layers, optimizers and regularizers we were able to get the best possible combination, which were implemented for all the above mentioned networks.

Environment Setup

This research was implemented using Google Colab, the data was stored in Google drive from where it is being fetched in colab (which is a jupyter notebook environment), that requires no additional support and runs entirely on cloud. Python version of 3.6 was used, and GPU (graphics processing unit) was used to process the data.

Layers

GlobalaveragePooling2D - This layer computes an average value of all the values taken as tensor for each input channel across the entire matrix.

Dense - It is a regular layers of neurons, the neurons present in this layer receives input from all other neurons in the previous layer and are densely connected.

Experimentation with optimizers

Initially we have conducted experiments with different optimizers like

RAdam- Also known as Rectified Adam is a classic optimizer which provides a dynamic adjustment to the learning rate based on the effects of variance and momentum during training.

SGD- Also known as stochastic gradient descent, it is considered as an iterative method to optimize the object function. In dimensional optimization problem it helps to reduce the computational burden.

Adam- Adam is an adaptive learning optimizer specifically designed for deep neural network, it is a combination of RMSprop and stochastic gradient.

After experimentation **Adam** was selected for the final implementation of all the models, with a learning rate of [0.0001].

Regularizing techniques used in the model architecture

Dropout- This technique is generally used for reducing over-fitting in neural networks, it refers to the dropping of units. For this research a dropout of [0.5] was used.

Early Stopping- This process is used to monitor the performance of the validation set and stop training when the performance is no longer increasing. For this research we have monitored the validation loss and the mode was set to minimum.

Loss Function

This method is used to evaluate how well the algorithm fits the dataset, if the prediction is totally inappropriate then the output will be higher number and if it is perfect then the output will be lower number. Initially we have experimented with **categorical_crossentropy**, **binary_crossentropy** and **sparse_categorical_crossentropy**, since not much of changes were observed therefore **categorical_crossentropy** was implemented for the final experiment.

Activation Function

An activation function is a function which is added to the neural network in order to help the network to learn the complex pattern within the data. Both **ReLU** and **Softmax** activation functions were used to conduct the research.

Other factors implemented

Epochs- It is basically refers to the number of passes the training dataset completed by the machine learning algorithm. A total of 50 epochs were used to conduct the experiment.

Batch-size- This refers to the number of training samples used for one iteration, it is often recommended that lower the batch-size better the learning phase. Batch size of 16 was used for training the model.

ReduceLROnPlateau- It has often been observed that reducing the learning rate by a factor of 2-10 had benefit the model in terms of increasing the accuracy, this callback monitors a particular property (validation-loss in here), and if no improvement is seen for a patience number of epochs then the learning rate is reduced.

7 EVALUATION, RESULTS & DISCUSSIONS:

This section gives a detail analysis of the evaluations and the results achieved while conducting the research. For this research pre-trained models were used in different forms to compare the results, and to verify which is more suitable for image based classification. For this research six transfer learning models were considered namely Resnet-50, Inception-v3, InceptionResNet-v2, Inception-v4, DenseNet-169 and DenseNet-201, out of which Inception-v3, InceptionResNet-v2 and ResNet-50 were fine-tuned.

7.1EVALUATION

The model was evaluated based on the following metrics, which are briefly described:

Accuracy – It is calculated as the total number of correct predictions by the total number of predictions.

Validation Accuracy- Validation Accuracy or val_acc is the accuracy of the validation data. It generally refers to the accuracy of a set of samples which were hidden from the model, thus showing that how much the model can generalize outside the training set.

Model loss – Model loss is also referred to as log loss or logarithmic loss is used as a metric for evaluation during each learning iteration.

7.2 RESULTS

Table 2- Model Accuracy Table

MODEL NAME	ACCURACY	VAL_ACC
ResNet-50	94.97%	75.10%
Inception-v3	93.03%	84.40%
InceptionResNet-v2	89.45%	87.00%
Inception-v4	87.08%	46.50%
DenseNet-169	93.08%	65.25%
DenseNet-201	95%	66%

Table 3- Model Loss Table

MODEL NAME	LOSS	VAL_LOSS
ResNet-50	0.1289	0.6808
Inception-v3	0.1766	0.5273
InceptionResNet-v2	0.2341	0.3615
Inception-v4	0.7717	0.9214
DenseNet-169	0.0803	0.4788
DenseNet-201	0.0481	0.4641

In this experiment the validation accuracy and the validation loss was monitored, since we were more focused on the generalizing ability of the model. It was observed that InceptionResNet-v2 has the best validation accuracy of 87% and a validation loss of 0.3614, which is least among other models. From the accuracy table it is evident that Inception-v4 has the lowest validation accuracy of 46.50% and the highest validation loss of 0.9214 compared to other models. ResNet-50 and Inception-v3 performed better than DenseNet-169 and DenseNet-201. Therefore it can be concluded that out of the 6 models InceptionResNet-v2 is the most accurate model in terms of generalization.

7.3 DISCUSSIONS

The main aim of the research was to implement different types of pre-trained models in different ways, and compare them. The architecture implemented involves ResNet-50, Inception-v3, InceptionResNet-v2, Inception-v4, DenseNet-169 and DenseNet-201. A thorough study was conducted, based on which the models were selected. The models were trained on 1200 regular fundus images, where inceptionResNet-v2 achieved the best results. The balance between the training and validation accuracy of the InceptionResnet-v2 was such that it can be used in the future for further improvements. The model was tested using 400 regular fundus images. Due to the limitations in resources we were not able to fetch more images for the training set which would have helped to create an even better model.

8 CONCLUSION AND FUTURE WORK:

To conclude the research, all the stages of diabetic retinopathy were classified with a high accuracy and with a very limited computation power. The pre-processing steps were conducted very efficiently which helped the model to train better with less time consumption. In the end a comparison was drawn between the six transfer learning models out of which three were fine-tuned. The model which were fine-tuned are ResNet-50, Inception-v3 and InceptionResNet-v2. It was observed that the fine-tuned models had a better generalization accuracy compared to the models which were not fine-tuned namely Inception-v4, DenseNet-169 and DenseNet-201. It was also observed that InceptionResNet-v2 was the best model compared to the others with a validation accuracy of 87%, and the model with the least validation accuracy was Inception-v4 which was 46.50%. Thus Inception-v4 was found as a drawback to detect diabetic retinopathy.

Since diabetic retinopathy detection is getting very popular, in future there should be more different types of image structures which will make the detection even more challenging. Secondly other different types of dataset should be combined in order to build an even better model. A combination of different machine learning techniques should be used in order to create a reliable and robust model. The study was based on supervised learning and for future unsupervised learning approach can benefit for identifying hidden patterns of lesions and exudates.

ACKNOWLEDGEMENT:

I would like to thank Dr.Rashmi Gupta for all her support, guidance and feed-backs throughout the supervision sessions, without which the project would have been incomplete. I would also like to thank my Mother and Father for their immense love support.

References

Xia , X., Xu, C. & Nan, B., 2017. *Inception-v3 for Flower Classification.* s.l., IEEE.

Ayan, E. & Unver, H. M., 2018. *Data Augmentation Importance for Classification of Skin Lesions via Deep Learning.* s.l., IEEE.

Behera, . M. K. & Chakravarty, S., 2020. *Diabetic Retinopathy Image Classification Using Support Vector Machine.* s.l., IEEE.

da Nóbrega, R. V. M., Peixoto, S. A., da Silva, . S. P. P. & Filho, P. P. R., 2018. *Lung Nodule Classification via Deep Transfer Learning in CT Lung Images.* s.l., IEEE.

Doshi, N., Oza, U. & Kumar, P., 2020. *Diabetic Retinopathy Classification using Downscaling Algorithms and Deep Learning.* s.l., IEEE.

Emara, T., Afify, H. M., Ismai, F. H. & Hassanien, A. E., 2019. *A Modified Inception-v4 For Imbalanced Skin Cancer Classification Dataset.* s.l., IEEE.

Islam, N. et al., 2019. *An image guideddiabetic retinopathy detection technique using attention-based deep learning scheme.* s.l., IEEE.

Kajan, S., Goga, J., Lacko, K. & Pavlovi ´ cov ˇ a´, J., 2020. *Detection of Diabetic Retinopathy Using Pretrained Deep Neural Networks.* s.l., IEEE.

Krishnan, A. S. et al., 2018. *A Transfer Learning Approach for Diabetic Retinopathy Classification Using Deep Convolutional Neural Networks.* s.l., IEEE.

K. S. et al., 2020. *Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification.* s.l., IEEE.

Kumar, R., Sharma, M., Dhawale, K. & Singal, G., 2019. *Identification of Dog Breeds Using Deep Learning.* s.l., IEEE.

Lands, A. et al., 2020. *Implementation of deep learning based algorithms for diabetic retinopathy classification from fundus images.* s.l., IEEE.

Li, S. et al., 2020. *A Novel Method of Bearing Fault Diagnosis in Time-Frequency Graphs Using InceptionResnet and Deformable Convolution Networks.* s.l., IEEE.

Liu, X. et al., 2018. *TRANSFER LEARNING WITH CONVOLUTIONAL NEURAL NETWORK FOR EARLY GASTRIC CANCER CLASSIFICATION ON MAGNIFIYING NARROW-BAND IMAGING IMAGES.* s.l., IEEE.

Li, X. et al., 2020. *Cross-Disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading.* s.l., IEEE.

Patel, R. & Chaware, A., 2020. *Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy.* s.l., IEEE.

Qiao, L., Zhu, Y. & Zhou, H., 2020. *Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms.* s.l., IEEE.

Reddy , G. T. et al., 2020. *An Ensemble based Machine Learning model for Diabetic Retinopathy Classification.* s.l., IEEE.

Rogers, C. et al., 2019. *Adversarial Artificial Intelligence for Overhead Imagery Classification Models.* s.l., IEEE.

Suresh, L., Chandran , S., Vijayan , D. & E R, V., 2020. *An Efficient method to Retrieve Diabetic Retinopathy Images using CBIR Technique.* s.l., IEEE.

S., V., P. & S. s. B., 2020. *Screening System for Early Detection of Diabetic Retinopathy.* s.l., IEEE.

T. & I., 2020. *Intensity Based Optic Disk Detection for Automatic Diabetic Retinopathy.* s.l., IEEE.

Zhu, S. et al., 2020. *Tortuosity of Retinal Main and Branching Arterioles, Venules in Patients With Type2 Diabetes and Diabetic Retinopathy in China.* s.l., IEEE.