Paper No: SC-11 **Smart Computing**

A tree structure-based classification of diabetic retinopathy stages using convolutional neural network

M. S. H. Peiris*
Department of Mathematics
Eastern University, Sri Lanka, Sri Lanka
shashi.hiru15@gmail.com

S. Sotheeswaran
Department of Mathematics
Eastern University, Sri Lanka, Sri Lanka
sotheeswarans@esn.ac.lk

***Abstract* - Detection, and classification of medical images have become a trending field of study during the last few decades. There is a considerable amount of vital challenges to be overcome. Ample work has been carried out to provide proper solutions for those key challenges. This study was carried out to extend one such medical image classification process to classify the stages of Diabetic Retinopathy (DR) images from colour fundus images. The study proposes a novel Convolutional Neural Network (CNN) architecture which is considered to be one of the most trending and efficient forms of classification of DR stages. Initially, the pre-processing techniques were employed to the DR fundus images with Green channel extraction and Contrast Limited Adaptive Histogram Equalization (CLAHE). The data augmentation strategy was utilised to increase training images from the DR images. Finally, Feature extraction and classification were carried out by using the proposed CNN architecture. It consists of a 14 layered CNN model, which continues three main classifications. In this proposed classification, the images were classified into a tree structure based binary classification as No_DR and DR at the beginning, and then the DR images were again classified into two classes, namely Pre_Intermediate and Post_Intermediate. Moreover, those two classes were again separately classified into Mild, Moderate, and Proliferate_DR, Severe, respectively. The Kaggle is one of the benchmark dataset repositories which was used in this study. The proposed model was able to achieve accuracies of 81%, 96%, 84%, and 97% for the above-mentioned classifications, respectively.**

Keywords - CLAHE, classification, CNN, diabetic retinopathy, green channel

I. INTRODUCTION

Detection and classification of medical images or medically-related objects in an image play an essential role as medical images are full of different characteristics which are absent in standard images. Preprocessing, segmentation, feature extraction, detection, classification, and prediction are some of the key challenges associated with medical image processing. Diabetic Retinopathy (DR) is the leading cause of vision loss and preventable blindness in grown-ups aged 20-74 years globally. The normal retina and diseased retina are shown in Figure 1.

Fig. 1. Normal retina and DR

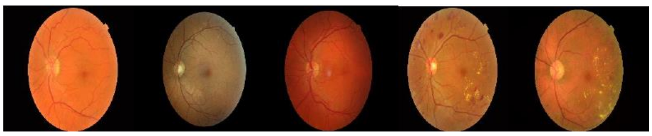
The main risk factor for the development of diabetic retinopathy is long-term diabetes which causes damage to blood vessels in the retina from high blood glucose levels [1]. DR can be classified into five stages as *No apparent retinopathy, Mild Non-Proliferative Diabetic Retinopathy (NPDR), Moderate NPDR, Severe NPDR,* and *Proliferative Diabetic Retinopathy*. Visual loss can be prevented up to 90% with the proper management of DR [2].

Non-proliferative retinopathy (also named background retinopathy) emerges first and creates increased capillary permeability, microaneurysms, haemorrhages, exudates, macular ischemia, and macular edema (thickening of the retina resulted from fluid leakage from capillaries). Proliferative retinopathy progresses after non-proliferative retinopathy and is more critical; it may point to vitreous haemorrhage and traction retinal detachment [3].

The medical features of Diabetic retinopathy are as follows:

- Microaneurysms are the tiny swellings on the walls of blood vessels inside the retina that are caused due to absence of the Pericyte. These are the earliest clinically visible changes. Microaneurysms eventually rupture to form haemorrhages deep within the retina [4].
- Haemorrhages appear as large spots on the retina.
- Hard exudates form when protein drips from blood vessels, and they are wavy and yellow or white deposits of protein.
- Cotton wool spots form when leakage of blood vessels blocks the vessels. An eye with more than six cotton wool spots is generalised as a pre proliferative state [5].

Figure 2 depicts the sample pictures from each class of the DR stages mentioned above.



(a) (b) (c) (d) (e)

Fig. 2. Classification of DR stages

(a). No_DR (b). Mild (c). Moderate (d). Severe (e). Proliferate DR

When focusing on the detection of DR, there are several methods used. The existing architectures of CNN such as VGG16 [6], InceptionNetV3 [7], and AlexNet [8] can be cited as examples. Many Convolutional Neural Network [9 – 13] models are developed to achieve a successful classification. To begin the treatments for DR, it is crucial to diagnose and classify its stages. Therefore, it is a complex task for the Ophthalmologists to diagnose and classify DR as per the stages since the manual feature extraction is a time-consuming and less accurate process. Moreover, it requires expert skills. Thus detailing the fundus retinal images with computer-aided systems paves the way to an effective and accurate improved methodology rather than manual performance.

The objective of this study is to address the classification of the stages of Diabetic Retinopathy images with the use of a Hierarchical Convolutional Neural Network technique which initially classify the DR and No-DR images then the classified DR images will be classified as Pre_Intermediate and Post_Intermediate. Moreover, those two Pre_Intermediate and Post_Intermediate classes were again separately classified into Mild, Moderate, and Proliferate_DR, Severe, respectively.

The rest of the paper is ordered as follows. In Section II, different techniques that are related to DR classification are summarised. The background of this work is explained in section III. In Section IV, the proposed methodology is described in detail. Section V contains the experimental setup and the testing results obtained. Finally, Section VI is allocated for the conclusion and future extensions.

II. PREVIOUS WORK

In [9], an automated diagnosis system was developed to recognise retinal blood vessels, and a multi-class classification of DR was carried out. Green channel extraction and contrast limited adaptive histogram equalisation (CLAHE) were carried out as the preprocessing techniques. After preprocessing the images, feature selection was done followed by feature extraction. Finally, the images were classified using the Support Vector Machine (SVM) classifier. Two publically available datasets were used for this work. DIARETDB1 with 130 images where 42 mages for training and 88 images for testing and DIARETDB0 with 89 images where 28 images for training and 61 images for testing were used. The method proposed here obtained an accuracy of 93.6% and a sensitivity of 90.6% for all 219 images. It would be clearer if they could include the size of the used images in this paper.

In [10], detection of blood vessels, identification of the haemorrhages, and classification of DR into three classes were the main objectives taken into consideration. The images were classified as normal, moderate, and non-proliferate DR. 65 images of normal (30), moderate (23), and non-proliferate DR (NPDR) (12) were used from the STARE dataset with the dimension of 576×768. Green channel extraction and Adaptive histogram equalisation were used as the preprocessing techniques. A 3×3 median filter was operated to remove the noise. The matched filtered image was converted to binary equivalent with a global threshold value. Then binarization was carried out using a matrix. The images were then augmented. The classification was finally carried out using the Random Forest technique based on the area and perimeter of the blood vessels and haemorrhages. The normal class with 20 training images and ten testing images achieved an accuracy of 90%. The moderate class with 15 training images and eight testing images achieved an accuracy of 87.5% and the severe NPDR class with four training images and eight testing images achieved an accuracy of 87.5%.

In [11], the authors have proposed a customised CNN architecture to classify diabetic retinopathy (DR) images. One thousand two hundred coloured fundus images were used from the Messidor dataset, where 840 are used for training images, and 360 are used for testing. Images were preprocessed by cropping to remove the black background and then resizing to 224×224, and the quality was adjusted using the histogram equalisation technique. Four CNN models were used where three were from pre-trained models such as AlexNet, VGG16, and SqueezeNet, and the remaining one was newly proposed. The performance of the classification of DR images of the newly proposed five-layered model was compared with the pre-trained models. In the proposed model, four separate kernels with size 3×3 were convolved in the first layer to extract features. Also, the image was zero-padded along by two. A pooling layer was also included in the first layer, and this layer reduces the calculations of the convolution layer and optimizes the time. The five-layered model produces a sensitivity of 98.94%, specificity of 97.87 %, and accuracy of 98.15%. It would be more effective if they could clarify the number of classes to which the images belonged and could use a higher number of images for testing and training.

In [12], the authors have considered the InceptionV3 architecture to classify Diabetic Retinopathy (DR). The dataset was taken from the famous Kaggle dataset which contains 35126 images. A five-class DR classification was done by splitting the dataset as 80% for training and 20% for testing with the input size as 299×299. Random scaling, resizing and centre cropping was done as preprocessing. The proposed model consisted of Inception V3 architecture and pre-trained on ImageNet as it can accelerate the process of training and also Inception V3 has a better performance on ImageNet. The architecture of the proposed model consists of five layers: Convolutional 2D layer, batch normalization layer, pooling layer, concatenate layer, and fully connected layer. Stochastic gradient descent (SGD) was used as the optimizer. Data augmentation was used with an early stop for 15 iterations to overcome the overfitting. Finally, the system was evaluated using 7023 test images. The system had achieved remarkable performance with an accuracy of 80% and a kappa score of 0.64.

In [13], the authors had employed a group of Convolutional Neural Networks (CNN) as a stage classification of Diabetic Retinopathy (DR). A fine-tuned three architectures; AlexNet, VGG16, and InceptionNet V3 were used to train the images. A total of 166 images from the Kaggle dataset were chosen to train the models. A five-class classification was done in this work. The images in the dataset were resized to pixels of 227×227, 224×224, and 299×299 for AlexNet, VGG16, and InceptionNet V3 respectively. The models AlexNet, VGG16, and InceptionNet V3 gained significant accuracies of 37.43%, 50.03%, and 63.23% for the dataset respectively. Higher rates of the accuracy of results have been achieved by the

InceptionNet V3 architecture. It would be effective if the authors could use a higher number of images to train and test these models.

III. BACKGROUND

A. *Diabetic retinopathy*

Diabetic Retinopathy is a related disease that is derived from Diabetes. The damage of the small blood vessels of the retina is the leading cause of it. Moreover, retinal blood vessels break down, leak or block. It affects the transportation of oxygen and nutrients inside the retina, causing vision loss over time. The presence of blockages, growth of abnormal blood vessels on the retinal surface increases the probability of bleeding leakages. These will result in vision blurring to vision loss over time.

B. *Machine learning*

Machine learning is a subfield of artificial intelligence where computers were made to learn from the data fed to them. It gives computers the ability to digest more data and reprogram themselves to execute a particular task with increasing precision. Then machines learn to perform a task more accurately through trials and errors. Machine learning usually uses several algorithms along with different tools to improve the prediction of desired outcomes [14]. Machine learning can be classified as supervised, unsupervised, and reinforced based on the algorithm it implements [15].

C. *Convolutional Neural Network (CNN)*

The neural network plays a major role in this report's work for the classification of Diabetic Retinopathy. Neural networks function similarly to the neurons in the human brain. It is important to note that all the neurons do not activate at once. Neurons are activated as per the signals received to carry out a particular task inside the body. This phenomenon is exactly used as neural networking in deep learning. CNN is formed of a set of layers that are stacked together. Each layer in the architecture owns a convolutional operator. Usually, a neural network inputs data process them with multiple neurons, and then outputs the results through an output layer [16]. Feature extraction and a fully connected layer are the two main parts of a basic CNN architecture. The convolution tool used to separate and identify the various features is known as the feature extraction, and the fully connected layer predicts the classes of the images using the features extracted in the previous layers.

IV. METHODOLOGY

The proposed tree structure-based binary classifications of DR are illustrated in Figure 3.

A. *Preprocessing*

Foremost in the experiment, the green channel was extracted from the procured images after centre cropping them to the size of 140×205. Then those images were subjected to **C**ontrast **L**imited **A**daptive **H**istogram **E**qualization (CLAHE). Figure 4 shows the preprocessed image samples.

B. *Data augmentation*

After obtaining the green channel and CLAHE processed images, those were subjected to data augmentation. The basic parameters used in this augmentation are flipping left, flipping right and rotation of 180^0 as shown in Figure 5. The other data augmenting parameters like shearing and zooming were not used since they did not have much impact on feature identification. The augmented images were saved separately and then were fed to the model. Data augmentation played a major role in extending the dataset to 49000 images. Datasets of images around 35000 were mostly found in the existing research works and that paved the way to derive an image set of 49000 images for this proposed work.

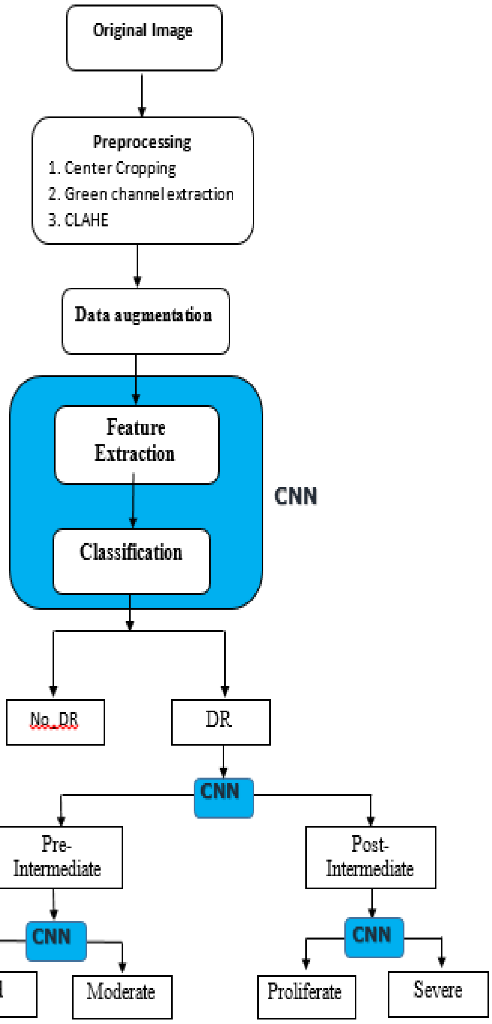


Fig. 3. Proposed methodology

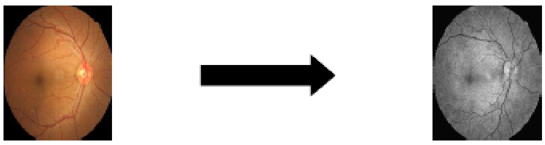


Fig. 4. Sample of RGB and pre-processed images

Fig. 5. Samples of data augmentation

C. Proposed CNN Architecture

The proposed CNN model consists of a 14 layered architecture as shown in Figure 6. It contains four Convolution 2D layers of the same format, each followed by a max-pooling layer. Then a flatten layer is present. Next, there are two dense layers, followed by a dropout layer for each. The Softmax classification layer is present at last. The learning rate of 0.01 was used for each convolution layer due to the use of more epochs while training.

When moving deep inside the layers, the first two convolutional 2D layers are of kernel size (3,3) with a sum of 16 filters per layer. The padding 'same' is used here to receive the output with equal dimensions as the input. The ReLU activation function is used to overcome the gradient vanishing problem. The default stride (1,1) is used in addition to the above-mentioned. Each layer is followed by a max pooling layer with default values. The third Convolutional 2D layer is of kernel size (3,3) with 32 filters. The padding ‘same’ is used with the default stride (1,1) and the ReLU activation function is used to activate the neurons [17]. A max pooling layer is followed by this layer.

The fourth convolutional 2D layer is of kernel size (3,3) and 64 filters are available. The padding ‘same’ is used in this layer as well with the ReLU activation function. A Max pooling layer is followed as stated above. Then a flatten layer is used to convert the data which comes from the above layers into a one-dimensional array for inputting it to the next layer. Next, there are two Dense hidden layers each followed by a Dropout layer (0.5). The two Dense hidden layers consist of 128 units per layer with a ReLU activation function. The final layer is the classification layer with the number of classes considered for the classification and Softmax as the activation function. The model was compiled with 50 epochs, a batch size of 32, a learning rate of 0.01, and “Adam” as the optimizer.

V. Experimental Setup

This section provides a brief description of the training and testing images, and the experimental setup of Diabetic retinopathy classification with the obtained testing results.

A. Dataset

The dataset was used in this work from the Kaggle dataset repository [18] which was illustrated in Figure 7. There are a number of datasets available for diabetic retinopathy in Kaggle. The dataset which was used for this piece of work consists of 35126 fundus images. These images were of size 224×224 and were centre cropped to 140×205 to remove the black background. The objectives of cropping the images were to remove the black background as much as possible while preserving the majority of the retinal vessels. The number of images in the original dataset is given in Table I. Data augmentation is used to increase the number of images in each level of classification.

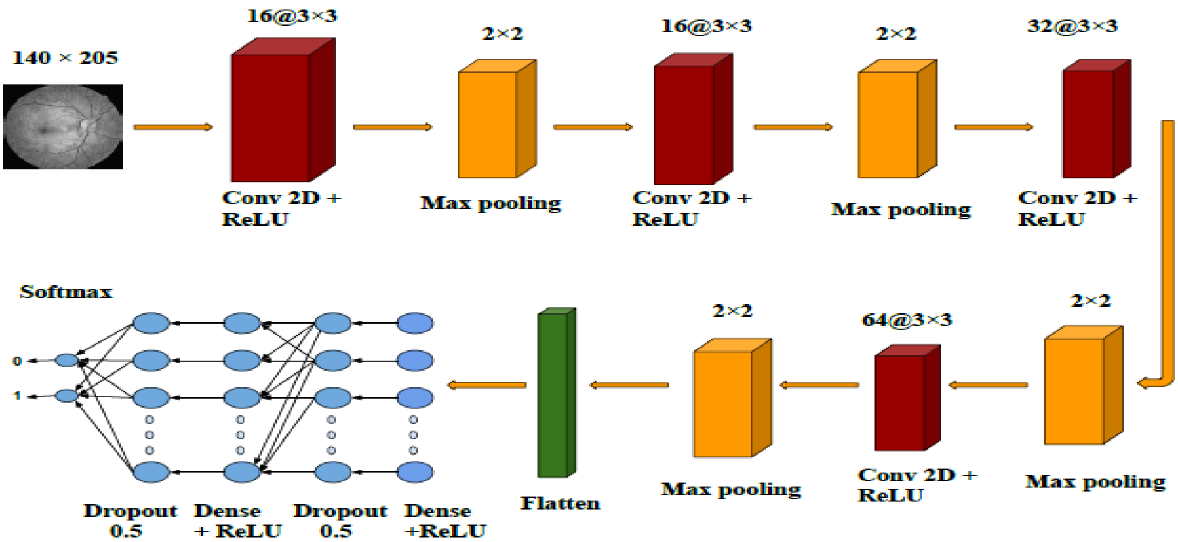


Fig. 6. Visualisation of the proposed CNN architecture

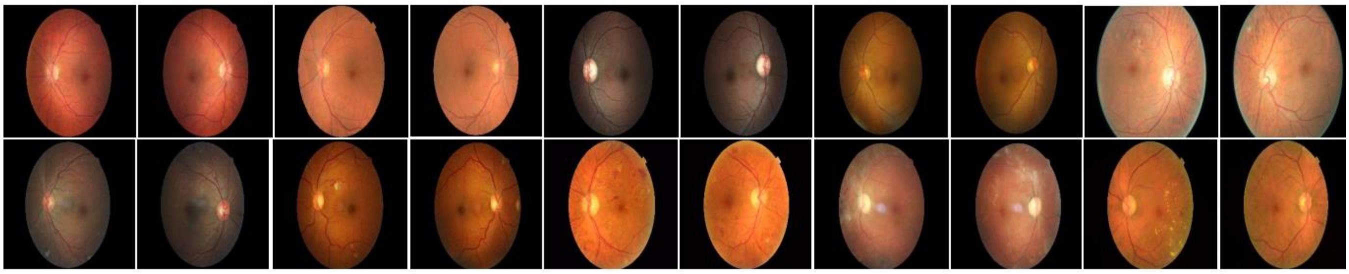


Fig. 7. Some sample images of the dataset

TABLE I. THE ORIGINAL DATASET IS IN DETAIL

Class ID	Class Name	Number of images
0	No_DR	25810
1	Mild	2443
2	Moderate	5292
3	Severe	873
4	Proliferate_DR	708

B. *Tree based classification*

Here, the results for the continued binary classifications were obtained. The Level 1 classification started with an image set of 49000 images and the Level 2 started with 24500 images per class. Finally, both Level 3(A) and Level 3(B) started with 12250 images per class. The order of the classification and results are displayed in Figure 8 and Figure 9, respectively.

C. *Testing results*

The model was trained and tested with images on the basis of 80% for training and 20% for testing. Accuracy, Precision, Recall and F1-score were also employed by obtaining the results in this work. We report the particular equations for the above parameters as follows:

$$\text{Accuracy} = (TP + FP) / \text{Total} \quad (1)$$

$$\text{Precision} = TP / (TP + FP) \quad (2)$$

$$\text{Recall} = TP / (TP+FN) \quad (3)$$

$$\text{F1 score} = 2 \times (\text{Recall} \times \text{precision}) / (\text{Recall} + \text{Precision}) \quad (4)$$

Where, TP - true positive, FP - false positive, TN - true negative, FN - a false negative. The average results of the classification for 50 epochs are reported in Table II.

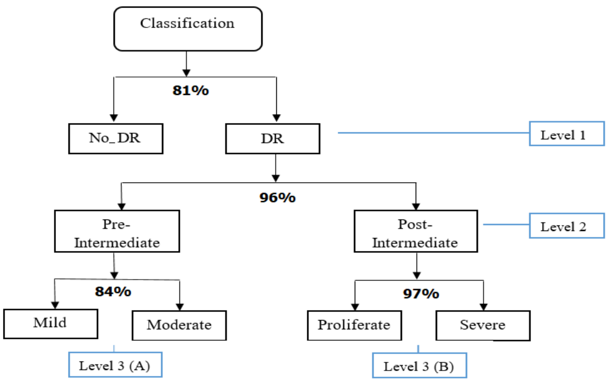
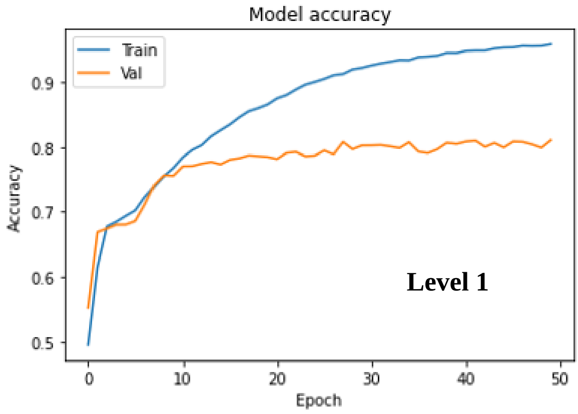
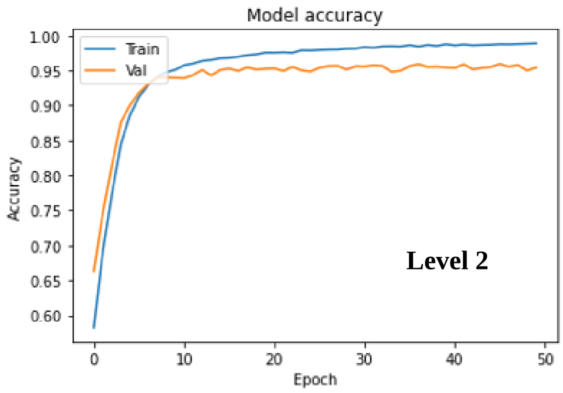


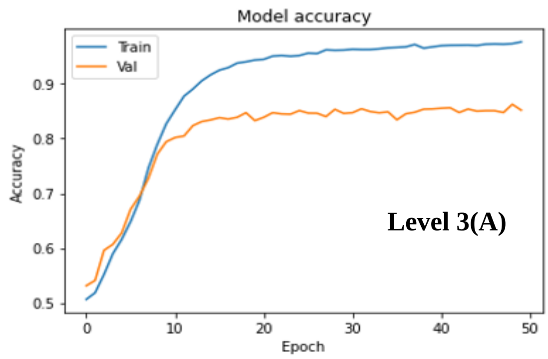
Fig. 8. Levels of the classification

TABLE II. AVERAGE RESULTS OF THE CLASSIFICATION WITH 50 EPOCHS

Classification Level	Avg. Precision	Avg. Recall	Avg. F1-score	Train Accuracy	Test Accuracy
Level 1	0.65	0.64	0.63	0.9574	0.8111
Level 2	0.70	0.58	0.50	0.9896	0.9571
Level 3(A)	0.66	0.64	0.62	0.9764	0.8396
Level 3(B)	0.96	0.96	0.96	0.9957	0.9737







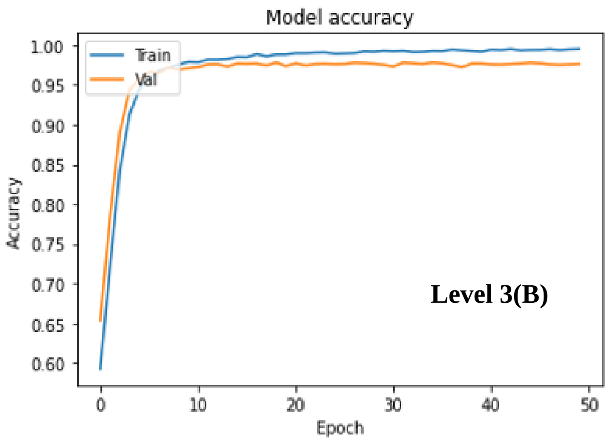


Fig. 9. Model accuracy for each level against epochs

All the experiments were carried out using a Virtual Machine (VM) from Microsoft Azure [19].

VI. Conclusion

In this piece of work, we have illustrated a proposed CNN architecture to classify Diabetic Retinopathy stages with a novel classification tree-based structure that continues with binary classifications. Moreover, the use of preprocessing techniques, Green channel extraction, CLAHE, and Data augmentation played a major role in achieving better accuracies. Centre cropping of all the images to the specified dimensions made it easy to remove the black background of the fundus images as much as possible. It was found out that the removal of the eye borders does not affect the feature extraction since a majority of the features are extracted from the retinal vessels present. The selection of the VM on training the models made a huge impact on gaining more accuracy. Hence, it can be concluded that this study which we proposed has been able to propose a model for the classification of Diabetic Retinopathy and has achieved worthy results for the novel classification approaches. While concluding the achieved results from this piece of work, it was able to achieve the particular accuracies of 81% for level 1, 96% for level 2, 84% for level 3(A), and 97% for level 3(B) on the proposed model. Deep learning approaches provide better results than geometrical approaches [20] of medical images. The expected future work of this particular study is to be stretched to enhance this model with a novel idea of classification and compare it with the bag-of-features approach.

References

[1] S. Vujosevic et al., "Screening for diabetic retinopathy: new perspectives and challenges", The Lancet Diabetes & Endocrinology, vol. 8, no. 4, pp. 337-347, 2020.

[2] L. Wu, "Classification of diabetic retinopathy and diabetic macular edema", World Journal of Diabetes, vol. 4, no. 6, p. 290, 2013.

[3] W. Wang and A. Lo, "Diabetic Retinopathy: Pathophysiology and Treatments", International Journal of Molecular Sciences, vol. 19, no. 6, p. 1816, 2018.

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

[5] A. Mahdjoubi, Y. Bousnina, G. Barrande, F. Bensmaine, S. Chahed and A. Ghezzaz, "Features of cotton wool spots in diabetic retinopathy: a spectral-domain optical coherence tomography angiography study", International Ophthalmology, vol. 40, no. 7, pp. 1625-1640, 2020.

[6] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images", International Journal of Scientific and Research Publications (IJSRP), vol. 9, no. 10, p. 9420, 2019.

[7] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818-2826, 2016.

[8] Z. Yuan and J. Zhang, "Feature extraction and image retrieval based on AlexNet", Eighth International Conference on Digital Image Processing (ICDIP 2016), 2016.

[9] P. Adarsh and D. Jeyakumari, "Multiclass SVM-based automated diagnosis of diabetic retinopathy", International Conference on Communication and Signal Processing, 2013.

[10] K. Verma, P. Deep and A. Ramakrishnan, "Detection and classification of diabetic retinopathy using retinal images", Annual IEEE India Conference, 2011.

[11] Mobeen-ur-Rehman, S. Khan, Z. Abbas and S. Danish Rizvi, "Classification of Diabetic Retinopathy Images Based on Customised CNN Architecture", 2019 Amity International Conference on Artificial Intelligence (AICAI), pp. 244-248, 2019.

[12] H. Chen, X. Zeng, Y. Luo and W. Ye, "Detection of Diabetic Retinopathy using Deep Neural Network", 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), pp. 1-5, 2018.

[13] A. Samanta, A. Saha, S. Satapathy, S. Fernandes and Y. Zhang, "Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset", Pattern Recognition Letters, vol. 135, pp. 293-298, 2020.

[14] O. Simeone, "A Very Brief Introduction to Machine Learning with Applications to Communication Systems", IEEE Transactions on Cognitive Communications and Networking, vol. 4, no. 4, pp. 648-664, 2018.

[15] M. Kang and N. Jameson, "Machine Learning: Fundamentals", Prognostics and Health Management of Electronics, pp. 85-109, 2018.

[16] S. Albawi, T. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network", International Conference on Engineering and Technology (ICET), 2017.

[17] N. Gupta, P. Bedi and V. Jindal, "Effect of Activation Functions on the Performance of Deep Learning Algorithms for Network Intrusion Detection Systems", Proceedings of ICETIT, pp. 949-960, 2019.

[18] Dataset:https://www.kaggle.com/sovitrath/diabetic-retinopathy-2015-data-colored-resized.

[19] "Data Science Virtual Machines | Microsoft Azure", Azure.microsoft.com, 2021. [Online]. Available: https://azure.microsoft.com/en-us/services/virtual-machines/data-science-virtual-machines/. [Accessed: 25- Jun-2021].

[20] D. V. D. S. Abeysinghe and S. Sotheeswaran, "Novel computational approaches for border irregularity prediction to detect melanoma in skin lesions," International Research Conference on Smart Computing and Systems Engineering (SCSE), pp. 216-222, 2020, doi: 10.1109/SCSE49731.2020.9313042