

Generic Model For Diabetic Retinopathy Detection Using Deep Learning

¹Narava Sai Chaitanya, ²V.Veerendra Subhash

^{1,2}Dept. of Computer Science & Engineering, KIET, Kakinada, AP, India

Abstract

Diabetic Retinopathy is a condition which occurs most commonly in patient with type 1 or type 2 diabetes. It affects the blood vessels of eye and delay in treatment can cause loss of vision. With the current state of the art deep learning technology, image classification can be performed with an accuracy as high as that of a human being. The idea behind this paper was to develop a highly accurate and reliable multi-class deep learning model which can detect the class of severity of diabetic retinopathy in a patient given an image of retinal fundus. The Aptos 2019 dataset was used for training the deep learning model. The proposed model also considered the high class-imbalance in the used dataset. Results – Our model achieved 99.18% categorical accuracy in training set and 75.68% categorical accuracy in validation set.

Keywords

Diabetic Retinopathy (DR), Deep Learning, Convolutional Neural Networks, EfficientNetB3, Artificial Intelligence

I. Introduction

Over time, diabetes creates an eye deficiency called as Diabetic Retinopathy (DR) and causes major loss of vision. The symptoms that can originate in the retinal area are augmented blood vessels, fluid drip, exudates, hemorrhages, and micro aneurysms.

In this paper, a deep learning model to detect the severity of diabetic retinopathy is proposed. The developed model takes retinal fundus image as input and outputs the class/severity of diabetic retinopathy in a patient.

Stages of Diabetic Retinopathy:

Table 1: Different stages of DR

Stages of DR	Description	Label
Normal	Without any abnormalities	0
Mild NPDR	Presence of micro aneurysms only	1
Moderate NPDR	Micro aneurysms are present but in a smaller amount.	2
Severe NPDR	Prominent intra retinal micro vascular abnormality in one or more regions	3
PDR	Vitreous/pre-retinal haemorrhage Neovascularization	4

No Diabetic Retinopathy, mild Diabetic Retinopathy, moderate Diabetic Retinopathy, severe Diabetic Retinopathy, and proliferate Diabetic Retinopathy are the five stages of diabetic retinopathy. The chances of going blind due to diabetic retinopathy significantly reduces if it is diagnosed and treated at early stages.

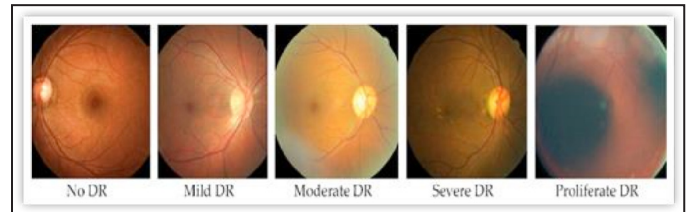


Fig. 1: Stages of diabetic retinopathy

Convolutional Neural Network (CNN):

A convolutional neural network (also known as CNN or ConvNet) is a feed-forward neural network that processes data in a grid-like structure to evaluate images.

They are analogous to traditional ANNs in that they are comprised of neurons that self-optimize through learning. Each neuron will receive an input and perform an operation just like in ANN.

The last layer will contain loss functions associated with the classes. CNNs are heavily used for image classification, pattern recognition and object detection.

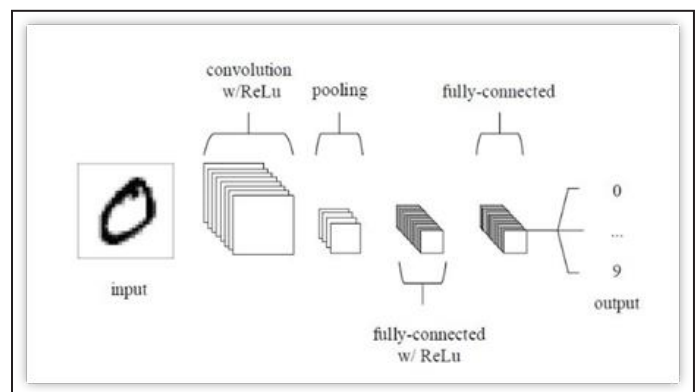


Fig. 2: A simple CNN architecture, comprising of five layers

II. Relatedworks

In the existing system, many researchers have worked on creating highly accurate deep learning model for detecting the stage of diabetic retinopathy, however most of them used small datasets to train their deep learning model. Lot of the work also did not account for the class imbalance in the datasets they used.

Quellec et al. [2] suggested an automated detection technique in which three CNNs (AlexNet and two additional networks) were utilized to detect micro aneurysms, hemorrhages, soft and hard exudates from Kaggle, DiaretDB1, and E-ophtha datasets (private) datasets.

In Kaggle and E-ophtha datasets, the disease was classified into two categories: referable and non-referable DR, with ROC values of 0.954 and 0.949, respectively.

Jiang et al. [3] provided a model in which the dataset was classified as referable diabetic retinopathy or non-referable diabetic retinopathy using three pre-trained CNNs (Inception-v3, ResNet152, and Inception-ResNet-v2). The photos were augmented and scaled before CNN training, and the models were

then combined using the Adaboost technique. Adam optimizer was used, and the system obtained 88.21 percent accuracy with an AUC value of 0.946.

Zago et al. [4] proposed a technique for detecting Diabetic Retinopathy and non-diabetic retinopathy images based on the probability of red lesion patches using two CNNs (pretrained VGG16 and a CNN). This model was trained on the DIARETDB1 dataset and tested on IDRiD, Messidor, Messidor-2, DDR, DIARETDB0, and Kaggle. On the Messidor dataset, the model had the best results, with a sensitivity of 0.94 and an AUC of 0.912.

Pratt et al. [5] developed a method for classifying the Kaggle dataset photos into five classes based on the DR severity levels using a CNN with ten convolutional layers, eight max-pooling layers, three fully connected layers, and a softmax classifier. The images were colour adjusted, scaled, and L2 regularization and dropout techniques were utilized to reduce over fitting during the preprocessing step. The approach yielded a 95 percent specificity, a 75 percent accuracy, and a 30 percent sensitivity.

Gulshan et al. [6] suggested a method for detecting diabetic macular edema (DME) and diabetic retinopathy using ten CNNs (pre-trained Inception-v3). The CNN model was tested using the Eyepacs-1 and Messidor-2 datasets. The images were first normalized, scaled, and fed into the CNN model. In two of the datasets used, the model has a specificity of 93% and a sensitivity of 97.5 percent and 96.1 percent in the Eyepacs-1 and Messidor-2 datasets, respectively.

Wang et al. [7] used the Kaggle dataset to test the performance of three different CNNs (pre-trained VGG16, AlexNet, and Inception-v3) in detecting the 5 classes of DR. For all three pre-trained CNN models, the dataset images were downsized to different sizes, and they achieved accuracy of 63.23 percent, 50.03 percent, and 37.43 percent in Inception-v3, VGG16, and AlexNet, respectively.

M. T. Esfahan et al. [8] employed ResNet34, to categorize DR photos from the Kaggle dataset into normal or DR images. The Gaussian filter, weighted addition, and image normalization were all used in the image preprocessing. They claimed a sensitivity of 86 percent and an accuracy of 85 percent.

Mobeen-ur-Rehman et al. [9] used custom CNN architecture and pre-trained models (including AlexNet, VGG-16, and SqueezeNet) to detect the DR levels in the MESSIDOR dataset.

Two convolutional layers, two max-pooling layers, and three FC layers made up their custom CNN. Their model achieved an accuracy of 98.15 percent, specificity of 97.87 percent, and sensitivity of 98.94 percent.

H. Jiang et al. [10] used three pre-trained CNN models (Inception V3, Inception-Resnet-V2, and Resnet152) to categorize their own dataset as referable DR or nonreferable DR. The Adam optimizer was used to update CNN's weights during their training. The Adaboost optimizer was used to combine these models. The accuracy was 88.21 percent, and the area under the curve (AUC) was 0.946.

M. Abramoff et al. [11] combined a CNN with an IDX-DR device. The Messidor-2 dataset, containing 1748 images, was subjected

to data augmentation. To detect DR lesions, their multiple CNNs were combined using a Random Forest classifier. The images were divided into three categories: no DR, referable DR, and severe DR. Their model reported an AUC of 0.980, a sensitivity of 96.8%, and a specificity of 87.0 percent. However, the images belonging to the mild DR stage were considered as no DR, and the five DR phases were ignored.

Attia et al. [12] took a survey to look at DR classification methods, with a general emphasis on deep learning techniques and a strong emphasis on traditional methods.

Gupta and Chhikara [13] reviewed DR detection techniques such as Adaboost, Random forest, SVM, and others, gradually demonstrating the gap that these traditional techniques present in terms of learning more disease-related features because some publicly available datasets have poor contrast and image quality, these comparisons are based on the quality of the fundus image.

Alyoubi et al. [14] reviewed a total of 33 papers that use deep learning for DR classification and emphasise the importance of continuous improvements to deep learning models given the global increase in diabetes cases. The authors also emphasised the use of data augmentation in model training to reduce overfitting.

Valarmathi and Vijayabhanu [15] discussed recent state-of-the-art (SoTA) CNN variants for DR classification while highlighting inconsistency in model evaluation metrics in the literature

Shamshad et al. [16] provides a thorough understanding of how transformers work for various medical imaging goals such as segmentation, classification, detection, and reconstruction. According to the survey, transformer-based research for medical imaging peaked around December 2021, with more than 40 recent publications. According to the survey, 73 percent of papers published in 2021 use vision transformers for segmentation tasks, while 27 percent of papers published between 2012 and 2015 use CNNs.

Gargeya Rishab and Leng Theodore [17] used customised deep CNN model based on the deep residual learning principle. The model's output was 0 for no DR and 1 for DR of any severity level. The training set was subjected to pre-processing techniques such as rotational invariance, contrast invariance, and brightness adjustment.

Garima Gupta et al. [18] presented a random forest-based classifier for segmenting true haemorrhages and distinguishing vessels from haemorrhages in [18]. On 191 images obtained from 58 diabetic subjects with varying degrees of pathological severity, the classifier achieved 82 percent sensitivity with 10 fold cross-validations. The main finding was that sensitivity increases for candidates with large haemorrhages, and variability in morphological and other image features such as appearance, colour, and texture increases confidence in identifying true haemorrhage candidates.

Daniel ShuWei Ting et al. [19] developed a deep learning system (DLS) for screening diabetic retinopathy and related eye diseases in [20]. Over 76,370 retinal images from Singapore's community-based national diabetic retinopathy screening programme (SIDRP 2010–2013) were used to train the system.

The primary validation set used was 71,896 images from the ongoing DR screening programme SIDRP 2014–15, including some low-quality ungradable images. The sensitivity was 90.5 percent, the specificity was 91.6 percent, and the AUC was 0.936 for detecting referable diabetic retinopathy.

VarunGulshan et al.[20] achieved excellent results by developing a deep CNN-based system trained on 128,175 retinal images for the development dataset. The algorithm’s performance was assessed using two different validation sets, EyePACS-1 and Messidor-2, which included 9963 and 1748 images, respectively. The images from the training and validation sets were graded multiple times by an ophthalmology panel. With very high sensitivity and specificity, AUCs of 0.991 and 0.990 were obtained on the Eyepacs-1 and Messidor-2 validation sets, respectively.

III. Proposed System

In the proposed we used EfficientNetB3 [1] CNN architecture. EfficientNetB3 [1] is a convolutional neural network architecture and scaling method that uses a compound coefficient to scale all depth/width/resolution dimensions evenly. The EfficientNetB3 [1] scaling method consistently scales network width, depth, and resolution with a set of predefined scaling coefficients, unlike approach, which scales these factors randomly.

We chose (300 x 300 x 3) as the input shape and for output, we created a dense layer of 5 neurons.

Firstly, we found out that the APTOS 2019 dataset contained a huge class imbalance (as show in Fig - 3) and would cause the model to underperform.

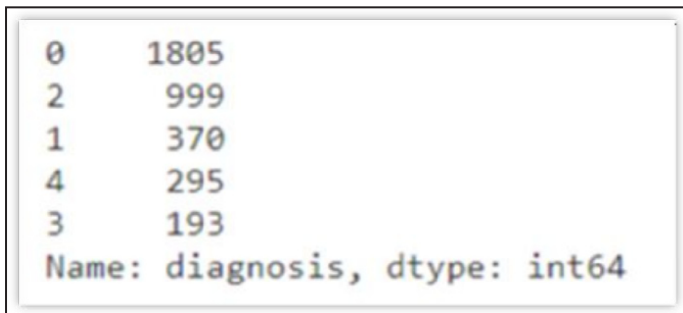


Fig - 3: APTOS 2019 class distribution

To avoid this problem, we sampled images from each class to make the dataset more balanced. The distribution of classes after sampling can be found in Fig - 4.

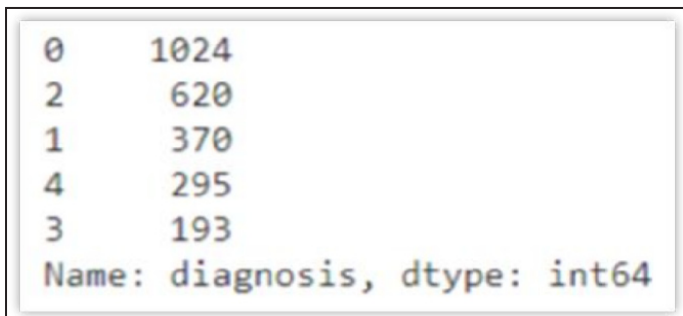


Fig - 4: APTOS 2019 dataset class distribution after normalizing

The total number of images after sampling was 2502. These 2502 images were then segregated into training and validation dataset. The train to validation dataset ratio was 80:20. Gaussian blur and

circle cropping techniques were used to preprocess the retinal images. We also used Image augmentation technique to increase the size of dataset and help improve the model’s performance.

For CNN model, we used EfficientNetB3 [1] pre-trained model with imagenet weights. The EfficientNetB3 [1] model was followed by one global average pooling layer and a dense layer with sigmoid activation.

The metrics used for the model were categorical accuracy, precision, recall, AUC and F1 score. The callbacks used for the model ReduceLRonPlateau and ModelCheckpoint. Binary crossentropy loss function was used to train the model. Also, Adam optimizer was used with 0.001 learning rate. The batch size used was 32.

IV. Proposed Architecture

In the proposed system, we used EfficientNetB3 CNN architecture.

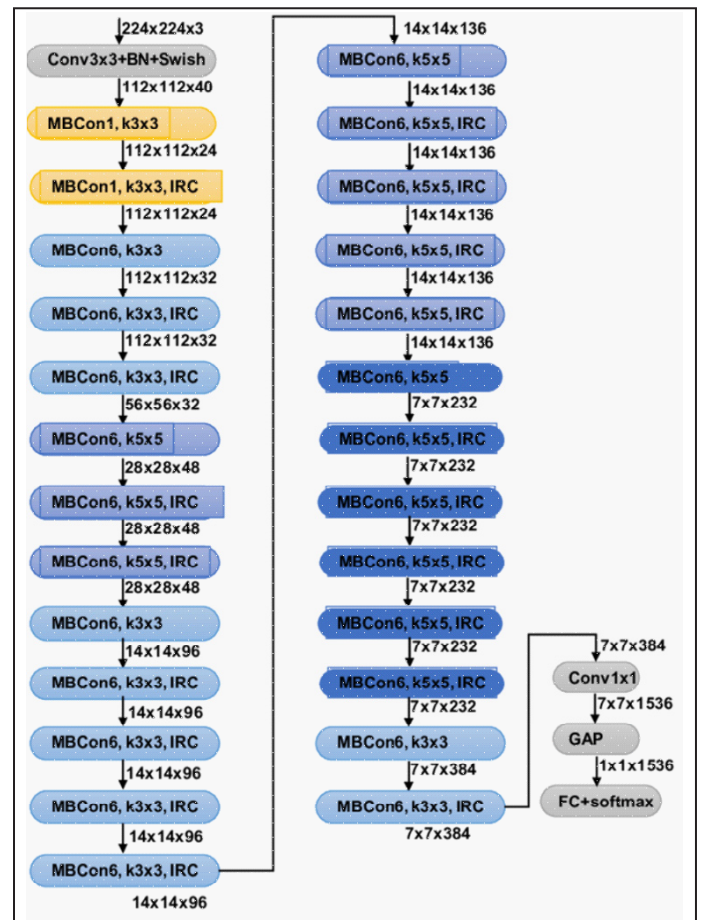


Fig - 5: EfficientNetB3 Architecture

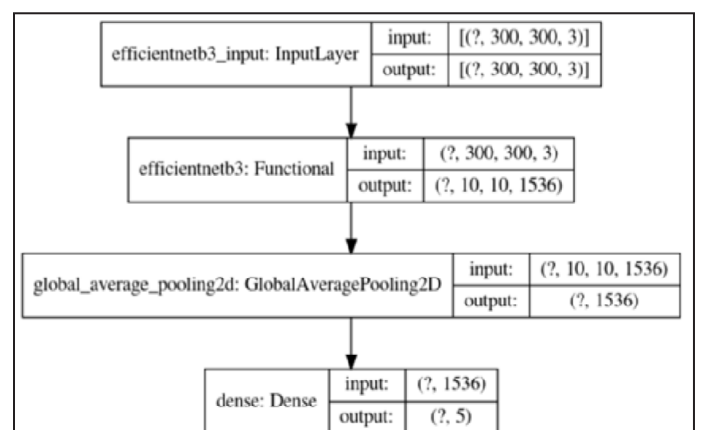


Fig - 6: Proposed Model Architecture

V. Modules And Description

A. Data cleaning and preparation:

This module is used to collect data and normalize it for training purposes. The APTOS 2019 dataset was used for training this model. The dataset was sourced from Kaggle. A huge class imbalance was observed in the above mentioned dataset. To avoid underperformance of model due to class imbalance, dataset was rebalanced and normalized for improving the deep learning model’s performance.

B. Model training:

This module is used to train the deep learning model with the given data set. We used transfer learning approach to reduce the model training time. We used EfficientNetB3 [1] and studied its performance on the given dataset.

C. Evaluating model:

This module evaluates the trained model for its accuracy and other metrics. Metrics like categorical accuracy, precision, recall, AUC and F1-score were used to monitor the performance of the EfficientNetB3 [1] model.

VI. Results

Our proposed model achieved 99.18% categorical accuracy, 99.19 % precision, 99.31% recall, 99.25 F1 - score in training set and 75.68 categorical accuracy, 76.31% precision, 75.00 % recall, 75.95 F1-score in validation set.

```
Epoch 49/50
50/50 [====] - ETA: 0s - loss: 0.0055 - categorical_accuracy: 0.9925 - precision: 0.9937 - recall: 0.9918 - auc: 1.0000 - f1_score: 0.9928
Epoch 49/50: val_categorical_accuracy did not improve from 0.77626
50/50 [====] - ETA: 0s - loss: 0.0055 - categorical_accuracy: 0.9925 - precision: 0.9937 - recall: 0.9918 - auc: 1.0000 - f1_score: 0.9928 - val_loss: 0.5434 - val_categorical_accuracy: 0.7614 - val_precision: 0.7659 - val_recall: 0.7588 - val_auc: 0.9064 - val_f1_score: 0.7653
Epoch 50/50
50/50 [====] - ETA: 0s - loss: 0.0061 - categorical_accuracy: 0.9918 - precision: 0.9919 - recall: 0.9931 - auc: 1.0000 - f1_score: 0.9925
Epoch 50/50: val_categorical_accuracy did not improve from 0.77626
50/50 [====] - ETA: 0s - loss: 0.0061 - categorical_accuracy: 0.9918 - precision: 0.9919 - recall: 0.9931 - auc: 1.0000 - f1_score: 0.9925 - val_loss: 0.5888 - val_categorical_accuracy: 0.7568 - val_precision: 0.7631 - val_recall: 0.7588 - val_auc: 0.9199 - val_f1_score: 0.7595
```

Fig. 7: Model training

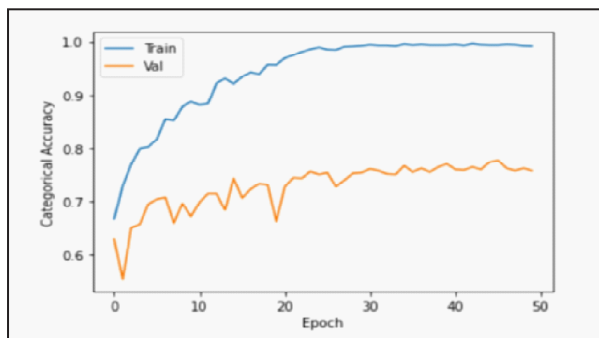


Fig. 8: Categorical accuracy graph

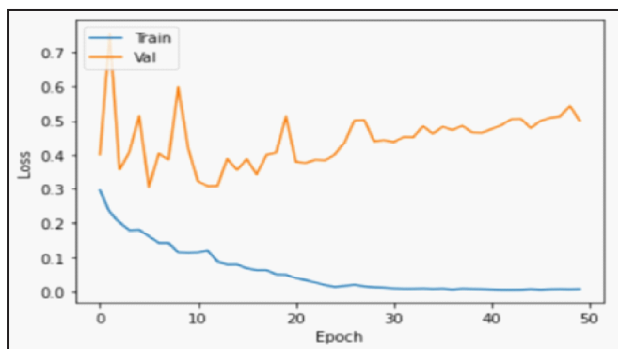


Fig. 9: Loss graph

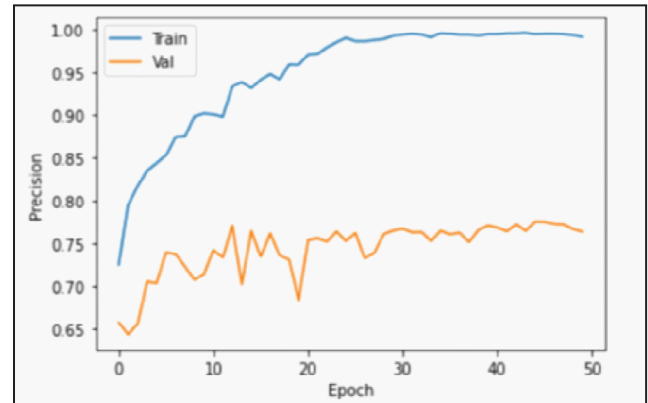


Fig. 10: Precision graph

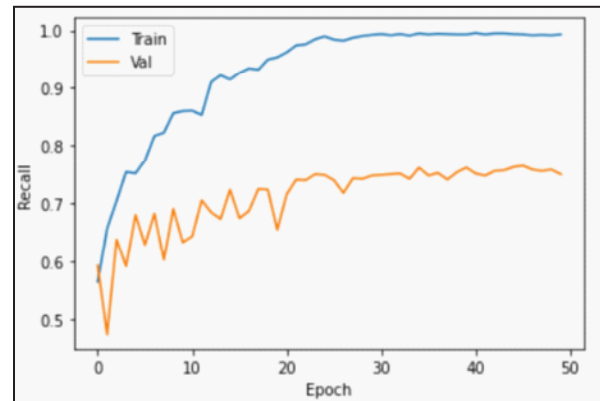


Fig. 11: Recall graph

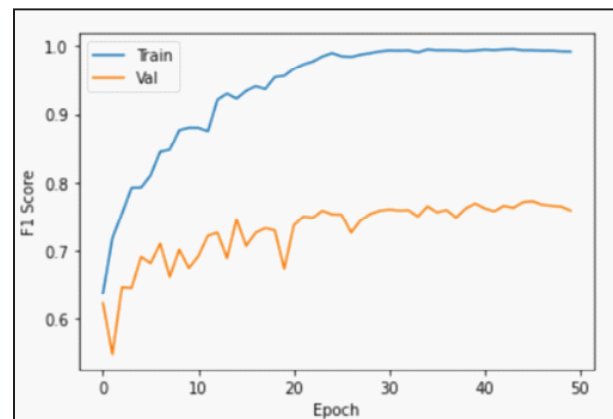


Fig. 12: F1 - score graph

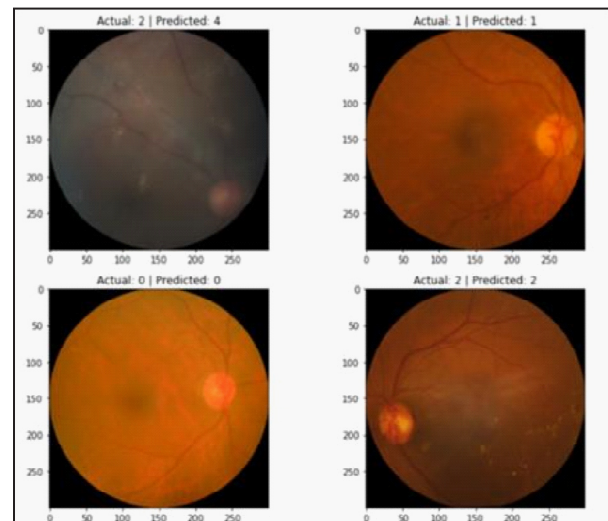


Fig. 13: Sample predictions

VII. Conclusions

Vision loss can be avoided by detecting and treating diabetic retinopathy early. The use of a CNN model to recognize and classify fundus images will aid ophthalmologists in their attempts to eradicate diabetic retinopathy-related vision loss. In this paper, we presented an EfficientNetB3 [1] model for the severity detection of diabetic retinopathy. We used an EfficientNetB3 [1] model which was pre-trained on image net dataset. The APTOS 2019 dataset was used to train our network. The proposed model is 99.18% categorical accuracy on training set and 76.31% categorical accuracy on test set.

VIII. Future Enhancement

Our proposed system was trained on 2500 high resolution images. Multiple datasets can be used to further improve the model's performance. Also other CNN architectures like EfficientNetB7 [1] can be used.

References

- [1] Mingxing Tan, Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", arXiv:1905.11946v5 [cs.LG]
- [2] Quéllec, G., Charrière, K., Boudi, Y., Cochener, B. and Lamard M, "Deep image mining for diabetic retinopathy screening", *Med Image Anal.*, 2017, 39:178–193
- [3] Jiang, H., Yang, K., Gao, M., Zhang, D., Ma, H. And Qian W, "An interpretable ensemble deep learning model for diabetic retinopathy disease classification" in 41st Annual International conference of the IEEE engineering in medicine and biology society (EMBC), 2019, pp. 2045–2048
- [4] Zago, GT., Andreao. RV., Dorizzi B. and TeatiniSalles EO, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks", *Computers in Biology and Medicine*, 2020
- [5] Pratt, H., Coenen, F., Broadbent, DM., Harding, SP. and Zheng, Y, "Convolutional neural networks for diabetic retinopathy", *ProcediaComputSci*, 2016, 90:200–5
- [6] Gulshan, V., Peng, L., Coram, M. et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs", *JAMA*, 2016, 316(22):2402–2410
- [7] Wang, X., Lu, Y., Wang, Y. and Chen WB, "Diabetic retinopathy stage classification using convolutional neural networks" in International Conference on information Reuse and Integration for data science, 2018, p. 465–71
- [8] Esfahani MT, Ghaderi M, Kafiyeh R. Classification of diabetic and normal fundus images using new deep learning method. *Leonardo Electron J PractTechnol* 2018; 17(32):233–48.
- [9] Mobeen-Ur-Rehman, Khan SH, Abbas Z, Danish Rizvi SM. Classification of diabetic retinopathy images based on customised CNN architecture. In: *Proceedings - 2019 Amity International Conference on artificial intelligence, AICAI 2019*; 2019. p. 244–8.
- [10] Jiang H, Yang K, Gao M, Zhang D, Ma H, Qian W. An interpretable ensemble deep learning model for diabetic retinopathy disease classification. In: 2019 41st Annual International conference of the IEEE engineering in medicine and biology society (EMBC); 2019. p. 2045–8.
- [11] Abramoff MD, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *InvestigOphthalmol Vis Sci* 2016;57(13):5200–6.
- [12] A. Attia, Z. Akhtar, S. Akrouf, and S. Maza, "A survey on machine and deep learning for detection of diabetic RetinopathyY," *ICTACT J. Image*
- [13] A. Gupta and R. Chhikara, "Diabetic retinopathy: Present and past," *Proc. Comput. Sci.*, vol. 132, pp. 1432–1440, Jan. 2018.
- [14] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Informat. Med. Unlocked*, vol. 20, Jan. 2020, Art. no. 100377. [15]S. Valarmathi and R. Vijayabhanu, "A survey on diabetic retinopathy disease detection and classification using deep learning techniques," in *Proc. 7th Int. Conf. Bio Signals, Images, Instrum. (ICBSII)*, Mar. 2021, pp. 1–4.
- [16] F. Shamshad, S. Khan, S. W. Zamir, M. H. Khan, M. Hayat, F. S. Khan, and H. Fu, "Transformers in medical imaging: A survey," 2022, arXiv:2201.09873.
- [17] GargeyaRishab, Leng Theodore. Automatic identification of diabetic retinopathy using deep learning. *Ophthalmology* 2017;124(7):926-9
- [18] Gupta Garima, Ram Keerthi, Kulasekaran S, Joshi Niranjana, SivaprakasamMohanasankar, Gandhi Rashmin. Detection of retinal hemorrhages in the presence of blood vessels. In: Chen X, Garvin MK, Liu JJ, editors. *Proceedings of the ophthalmic medical image*