International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

Automatic Diagnosis of Diabetic Retinopathy using Transfer Learning Approach

Nidhi Kamothi

Student, Dept. of Computer Engineering, Thakur College of Engineering and Technology, Maharashtra, India

---

Abstract - *Diabetic Retinopathy (DR) is a common eye disease and a major cause of blindness in diabetic patients. It is a disease in which the retina is damaged due to diabetes mellitus. It affects up to 80 percent of all patients who have had diabetes for 10 years or more. Detection and quantification of such mellitus from retinal images is tedious and requires expertise. Regular screening with fundus photography and timely intervention is the most effective way to manage the disease. Our proposed methodology aims to automatically diagnose the disease at various stages using Deep Learning (DL). For the evaluation and treatment of DR, stage classification is a crucial step. This paper presents the design and implementation of GPU accelerated Deep Convolutional Neural Networks (DCNN) to automatically diagnose and thereby classify high-resolution retinal images into 5 stages of the disease based on severity. We train this network using a high-end graphics processor unit (GPU) on the publicly available Kaggle dataset and demonstrate impressive results, particularly for a high-level classification task. State-of-the-art accuracy result has been achieved by VGG16 architecture, which demonstrates the effectiveness of utilizing DCNN for DR image recognition. On the data set of 2,500 images used, our proposed model achieves an accuracy of 71.6 % on 500 validation images.*

Key Words: **Deep Learning, Convolutional Neural Networks, Diabetic Retinopathy, Image Classification, Automatic Diagnosis, VGG16, VGG19**

1. INTRODUCTION

Diabetic Retinopathy also known as diabetic eye disease, is when damage occurs to the retina due to diabetes [1]. The WHO has declared that, in 2030, diabetes will be the most serious and 7th highest death-causing disease in the world [2]. If the disease is not treated properly at its early stages, the disease may get severe. The damage in the retinal blood vessel eventually blocks the light that passes through the optical nerves which makes the patient with DR blind.

In the area of ophthalmology, DL is performing a vital role to diagnose serious diseases including DR. Using the concept of DL, and then applying that knowledge to the medical field can help solve complex problems, save lives, and make the world a better place. If DR is diagnosed early, it can be managed using available treatments. Regular eye fundus examination is necessary because DR do not present any symptoms until late in disease. The retinal abnormalities in DR also include Haemorrhages (HM), "Cotton wool" spots, Microaneurysm (MA), Retinal neovascularization, Hard exudates, which are clearly shown in Figure 1[2].

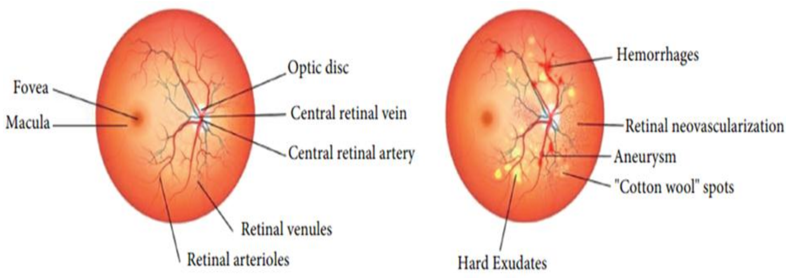

Fig -1: Normal eye and Infected eye

In recent years, DCNN have proved to be widely successful in the field of image classification, segmentation, and related tasks. The reason behind this is the sheer number of layers that enables the neural network as a whole to identify and extract complex features in the image. There is a need of automation in screening process due to the limited number of ophthalmologists to cover many diabetes patients and reducing burden on retina specialists. Automation can be carried out at two levels, first, to identify persons with DR, second, grading persons according to severity [3]. Table 1 shows DR stages and Severity.

In the ImageNet challenge for image classification, CNN based approaches have been recently dominating the leader board year after year. Well known CNN architectures like CNN [4], VGG16 [5], VGG19 [6] have repeatedly demonstrated high accuracies in image classification tasks.

Table -1: DR stages and Severity

DR Stages	Severity
No DR	No abnormalities
Mild NPDR	Microaneurysms occur. Microaneurysms are small areas of balloon like swelling in the retina blood vessels.
Moderate NPDR	Swelling and distortion of blood vessels
Severe NPDR	Many blood vessels are blocked, which causes abnormal growth factor secretion.
Proliferative DR	Growth factors induce proliferation of new blood vessels inside the surface of the retina, the new vessels are fragile and may leak or bleed, scar tissue from these can cause retinal detachment.

1.1 Problem Definition

DR can develop if you have type 1 or 2 diabetes and a long history of uncontrolled high blood sugar level. You'll be able to eventually lose your sight. Currently, detecting DR is a time-consuming and manual process that requires a trained clinician to examine and evaluate digital color fundus photographs of the retina. By the time human readers submit their reviews, often each day or two later, the delayed results cause lost follow up, miscommunication, and delayed treatment. However, Researches shows that progression to vision impairment are often slowed or averted if DR is detected in early stage of the malady. There is a need to detect DR at early levels, so that the damage to the retina can be minimized as it is a common persistent problem faced up to date.

2. DATASET DESCRIPTION

The dataset considered in this approach is publicly available on the Kaggle online platform which was provided by EyePACS [7]. It consists of fundus images obtained over a broad spectrum of imaging conditions. The dataset consists of 2,500 labelled color fundus retinal images belonging to five classes corresponding to the five stages of the disease as portrayed in below Table. The training was done on 2000 images. The test set consists of 500 images. A trained clinician has rated the presence of DR in each image on a scale of 0 to 4. Figure 2 shows the various stages of DR.

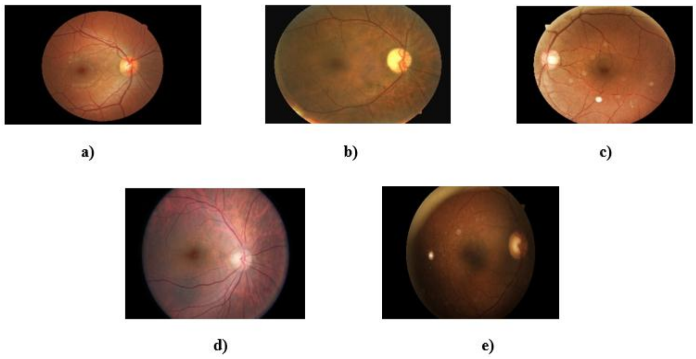

Fig -2: DR on a scale of 0 to 4

(a) Class 0 - No DR (b) Class 1 - Mild DR (c) Class 2 – Moderate DR (d) Class 3 - Severe DR (e) Class 4 - Proliferative DR

The class labels of the dataset are highly imbalanced i.e. more than 73% of the class are negative, which makes our model difficult to train. The images in the dataset come from different models and types of cameras, which can affect the visual appearance of left and right retinas. There is also noise in both the images and labels. Table 2 shows the distribution of image classes in the original dataset.

Table -2: Class Distribution in Original Dataset

Class	Name	No of Images	Percentage
0	No DR	33837	73.49%
1	Mild DR	3243	7.04%
2	Moderate DR	6888	14.96%
3	Severe DR	1146	2.49%
4	Proliferative DR	929	2.02%

3. PROPOSED METHODOLOGY

The Proposed approach is divided into two stages. First stage involves image pre-processing where the images are enhanced and converted to the desired format for the classification task. Further, different CNN architectures namely CNN, VGG16, and VGG19 are used for the classification task with the Deep learning approach.

A. Image Pre-processing

The dataset used is highly heterogeneous because the photographs are from different sources, cameras, resolutions, and have different degrees of noise and lighting [3]. To achieve high accuracy, we proceeded some Pre-processing steps as follows:

- **Denoising**

Image denoising is to remove noise from a noisy image, to restore the true image. Image denoising plays an important role in modern image processing systems.

- **Size Normalization**

Due to non-standard image resolutions, the training images could not be utilized directly for training. The image was scaled down to a fixed resolution size of 224 x 224 pixels to form a standardized dataset.

- **Color Normalization**

After the size of each image is normalized, its color must be tuned because different devices may produce images with different color temperatures, and the illumination conditions can vary.

B. CNN Architectures Experimented

Three CNN architectures namely CNN [4], VGG16 [3], VGG19 [5] are classification tasks. The details of these architectures are given below:

1) Convolution Neural Network (CNN)

In recent years, DL has been used extensively in a wide range of fields. In DL, CNN are found to give the most accurate results in solving real world problems. CNNs are the specialized ANNs that are designed to perform well with images [4]. In this paper, we have experimented with the CNN architecture by flattening the last layer and adding a two dense layers followed by a softmax layer. This model began to overfit on the train data after a few epochs which did not improve even when dropout was added after the dense layers. Further, the CNN model with maxpooling followed by a softmax layer had been used, which gave a kappa score of 0.25.

2) VGG16

VGG16-Net [5] is widely adopted for various medical image analysis applications. It includes 13 convolutional layers and three fully connected layers. Thirteen blocks of 3x3 convolutions followed by a Pooling layer are considered for training. Finally, Dense layer is applied on the output of the last block consisting of 256 neurons. The final layer is dense layer with five outputs - one corresponding to each category of the DR images. Our VGG16 network architecture yielded significant classification accuracy. The details of this architecture are shown in Figure 3.

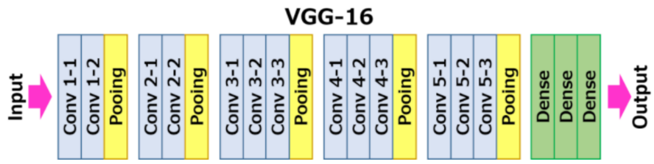

Fig -3: VGG16 architecture

3) VGG19

VGG19 is that type of neural network which is also specifically trained on more than a million images from the "ImageNet" database, but the difference between the VGG19 and the VGG16 is that, this type of network is 19 layers deep. The detailed information about VGG19 network and architecture can be extracted from [6].

4. IMPLEMENTATION

4.1 Data Collection

Data Collection is the basic step to collect the data for experiments because, without an image, no processing is possible. The images that are acquired are completely unprocessed. So, some pre-processing steps have proceeded. After these pre-processing, a total of 2,500 images were selected from dataset of Kaggle. From these 2,500 images, 80% of the images are used for training purpose and remaining 20% is used for testing the system. The decomposition of these 2,500 images are given below in Table 3.

Table -3: Composition of Dataset

Stage	Description	No of Images
0	No DR	400
1	Mild DR	400
2	Moderate DR	400
3	Severe DR	400
4	Proliferate DR	400

4.2 Data Augmentation

Having sufficient training data is the key for training a neural network successfully. To mitigate shortages in data and fully utilize the data that are available, certain data augmentation techniques are carried out in our proposed methodology.

We augmented our data through the following means:

- flip the image horizontally
- flip the image vertically
- randomly rotate the image in the range of [-15,15] degree
- randomly zoom in or out in the range of [0,0.15]

4.3 Model Selection

We deploy three state-of-the-art CNN architectures namely, CNN, VGG16 and VGG19, for DR stage classification. In model selection, we check the accuracy of each algorithm and get to know which algorithm will give us more accurate results for the prediction of DR.

4.4 Model Evaluation

Model evaluation metrics are required to quantify the model performance. The algorithm measures performance on different parameters like Accuracy, Confusion Matrix, Error Rate, Precision, Recall, F1 Score, Cohens Kappa Score.

4.5 Model Deployment

To evaluate our work in real clinical environments, we deploy our models on Web Application based platform and provide a diagnostic service to several hospitals in nearby cities via internet.

5. RESULTS AND DISCUSSION

The result obtained from the proposed method is compared with the state-of-the-artwork. The experiments are performed on "Google Colab" using GPUs. The deep learning package Keras was used with the TensorFlow machine learning back end library. The algorithm measures performance on different parameters like Accuracy, Error Rate, Confusion Matrix, Precision, Recall, F1 Score. We also use Quadratic Weighted Kappa as the performance metric. The final trained network achieved, 71.6% Accuracy, 0.284 Error rate, 71.8% Precision, 71.6% Recall, 71.2% F1 Score, and a quadratic weighted kappa value of 0.645.

Accuracy: The accuracy can be calculated in terms of positive and negative classes.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Error Rate: It is the number of all incorrect predictions divided by the total number of the dataset. The best error rate is 0.0, whereas the worst is 1.0.

Error Rate = 1 – Accuracy

Confusion Matrix: As the name suggests, it gives us a matrix as output and describes the complete performance of the model.

There are 4 important terms:
True Positives (TP): Predicted positive and are positive.
False positives (FP): Predicted positive and are negative.
True negatives (TN): Predicted negative and are negative.
False Negatives (FN): Predicted negative and are positive.

Precision: It is defined as the number of true positives divided by the number of true positives plus the number of false positives.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall/Sensitivity: It is the ratio of correctly predicted positive observations to the all observations in actual.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1 Score: It is a harmonic mean of precision and recall. It is also called the F Score or the F Measure.

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

Figure 4 shows the Comparison of model.

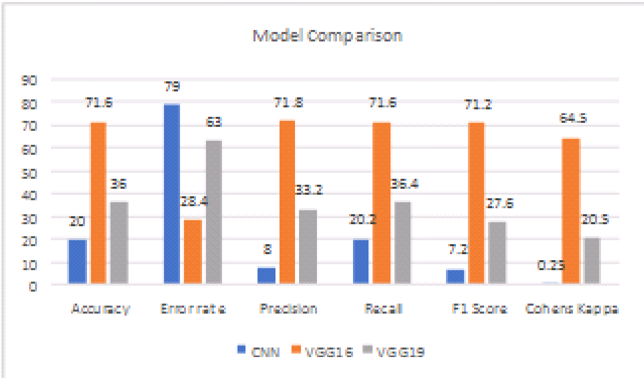

Fig – 4: Model Comparison

Performance analysis of the three CNN architectures are given in table 4.

Table -4: Performance Analysis of CNN architectures

CNN Architectures	Accuracy	Error Rate	Precision	Recall	F1 Score
CNN	20%	79	8%	20.2%	7.2%
VGG16	**71.6%**	**28.4**	**71.8%**	**71.6%**	**71.2%**
VGG19	36%	63	33.2%	36.4%	27.6%

Figure 5 shows the confusion matrix of VGG16 model.

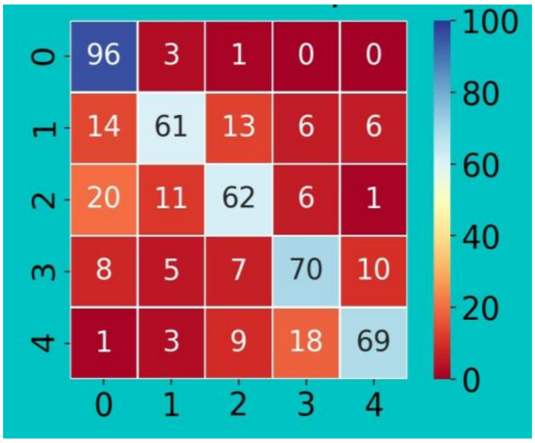

Fig -5: Confusion matrix of VGG16 model

6. CONCLUSION

The huge population of diabetic patients and the prevalence of DR among them have fostered a great demand in automatic DR diagnosing systems. This model presents the design, architecture and implementation of deep convolutional neural networks for automatic detection and classification of DR from color fundus retinal images. The images in the dataset were pre-processed to remove a major part of the background and the effects of lighting in different regions of the scan image. The pre-processed images are then trained with models based on three CNN architectures namely CNN, VGG16 and VGG19 using TensorFlow Keras. Among the three, the model based on the VGG16 architecture achieved the highest accuracy that is 71.6% with quadratic weighted kappa score of 0.645. Ensembling of various CNN models can be considered as a future work to achieve a better kappa score.

REFERENCES

[1] T. Chandrakumar, R. Kathirvel, "Classifying DR using DL Architecture", Int. J. Edu. Res. Technol., Vol. 5, no. 06, pp. 2278-0181, June 2016

[2] M. Mateen, J. Wen, Nasrullah, S. Song, S. Hayat, "Exudate Detection for DR Using Pretrained Convolutional Neural Networks", Hindawi, Vol. 2020, April 2020

[3] M. Nikhil, A. Rose, "DR Stage Classification using CNN", Int. Res. J. Eng. Technol., Vol. 06, no. 05, pp. 2395-0072, May 2019

[4] A. Krishnan, Derik Clive R, V. Bhat, P. Ramteke, and S. Koolagudi, "A Transfer Learning Approach for DR Classification Using Deep Convolutional Neural Networks", IEEE India Council Int. Conf., pp. 2325-9418, Dec 2018

[5] M. Hajabdollahi, R. Esfandiarpoor, K. Najarian, "Hierarchical Pruning for Simplification of Convolutional Neural Networks in Diabetic Retinopathy Classification", Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., pp. 1558-4615, Oct 2019

[6] S. Gupta, A. Panwar, S. Goel, A. Mittal, R. Nijhawan, A. Singh, "Classification of lesions in Retinal Fundus Images for Diabetic Retinopathy using Transfer Learning", Int. Conf. Inf. Technol., 2019

[7] Diabetic retinopathy detection: identify signs of diabetic retinopathy in eye images, https://www.kaggle.com/c/diabetic-retinopathy detection/data

[8] Z. Gao, J. Li, J. Guo, Y. Chen, Z. Yi, "Diagnosis of DR Using Deep Neural Networks", IEEE Trans. Ind. Electron., Vol. 7, pp. 2169-3536, Dec. 2019

[9] H. Pratt, F. Coenen, D. M. Broadbent, S. Harding, Y. Zheng, "Convolutional Neural Networks for DR", in Procedia Computer Science, Vol. 90, pp. 200-205, 2016

[10] R. Ghosh, K. Ghosh, and S. Maitra, "Automatic Detection and Classification of DR stages using CNN", Int. Conf. Signal Process. Integr. Netw., Vol. 2017, 2017

[11] M. Mateen, J. Wen, Nasrullah, S. Song, S. Hayat, "Automatic Detection of DR: A Review on Datasets, Methods and Evaluation Metrics", IEEE Trans. Ind. Electron., March 11, 2020

[12] X. Wang, Y. Lu, Y. Wang, W. Chen, "DR Stage Classification using Convolutional Neural Networks", Int. Conf. Inf. Reuse Integr. Data Sci., 2018

[13] S. Qummar, F. Khan, S. Shah, A. Khan, "A Deep Learning Ensemble Approach for DR Detection", IEEE Trans. Ind. Electron., Vol. 7, pp. 2169-3536, Oct. 2019

[14] H. Sharma, A. Singh, A. Chandel, P. Singh and A. Sapkal, "Detection of DR using Convolutional Neural Network", Int. Conf. Commun. Inf. Process., 2019

[15] M. Jebran, S. Gupta, "Detection of Lesions for DR By using Machine Learning Algorithms", Int. J. Recent Technol. Eng., Vol. 8, no. 5, pp. 2277-3878, Jan. 2020

BIOGRAPHIES

Nidhi Kamothi received BE degree in computer engineering from St John College of Engineering and Technology, India in 2017 and pursuing Master of Computer Engineering from Thakur College of Engineering and Technology, India.