International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Automated detection of Diabetic Retinopathy using VGG-16 architecture

Abhishek Deshpande[1], Jatin Pardhi[2]

[1]Student, Dept. of Chemical Engineering, Visvesvaraya National Institute of Technology, Maharashtra, India
[2]Student, Dept. of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, Maharashtra, India

---

Abstract - Diabetic Retinopathy is a disease that affects the retina and is caused by chronic diabetes. Early detection of this disease is important and will benefit a significant number of people. The Asia-Pacific Tele-Ophthalmology Society (A.P.T.O.S) 2019 Blindness Detection dataset that we have used as a training dataset for this model contains 3668 retinal images. Fundus photography technique was used to take these photographs. We have used the pre-trained Convolutional Neural Network (CNN) VGG-16 to detect the severity of Diabetic Retinopathy from the image. This model of ours was tested over 1728 completely new set of images which weren't included in the training dataset. The model was able to achieve a 74.58% accuracy rate. Over the course of the 30 epochs, the loss was 5.06 percent. Because of the use of the Categorical Cross entropy loss function, the loss was kept to a minimum. The ADAM optimizer assisted in training the model at a high speed and efficiency. The output is an integral value on the scale of 0-4 according to the severity of Diabetic Retinopathy. This model can assist doctors in detection of Diabetic Retinopathy at an early stage.

***Key Words*: Diabetic Retinopathy, APTOS, Convolutional Neural Network, VGG-16, ADAM Optimizer 1.**

1. INTRODUCTION

1.1 Diabetic Retinopathy (DR)

Diabetes mellitus is one of the most pressing public health concerns in the world. Diabetic retinopathy is the world's fifth most common cause of visual impairment and, as a result, the fourth most common cause of blindness. Cooperation between those responsible for diabetes management and those affected by diabetic retinopathy is the most important role of health systems in managing diabetes and avoiding permanent blindness from the disease. [1]

Diabetic retinopathy has four stages:

- *Mild Non-Proliferative DR (NPDR)* - Microaneurysms form at this point. Within the retina's small blood vessels, there are small areas of balloon-like inflammation. [2]
- *Moderate NPDR* - Blood vessels that nourish the retina are blocked at this stage. Within the retina, there are also haemorrhages. [2]
- *Severe NPDR* - More blood vessels are obstructed at this stage, depriving several areas of the retina of blood supply. The amount of haemorrhage in the retina also rises dramatically. [3]
- *Proliferative DR* - New and abnormal blood vessels develop on the surface of the retina in this advanced stage of DR. These new blood vessels are delicate and have a tendency to bleed, causing vision-threatening haemorrhage to fill the eye. They'll also turn into connective tissue, which will contract over time, causing the retina to detach and cause blindness. [3]

Because each stage has its own characteristics and properties, doctors may overlook some of them and thus make an erroneous diagnosis. As a result, the idea of developing an automatic solution for DR detection arises. With effective and timely treatment and eye supervision, more than half of the most recent cases of this disease could be avoided.

1.2 VGG-16 Model

VGG16 is a convolutional neural network(CNN) model. " VGG1-16 is one of the most successful vision model architecture. This model accomplishes 92.7% top-5 test precision on ImageNet dataset (Dataset having 15 million images of various different categories) which contains 14 million pictures having a place with 1000 classes. "[4]

VGG-16 Model Architecture shown in *Figure 1 [5]*

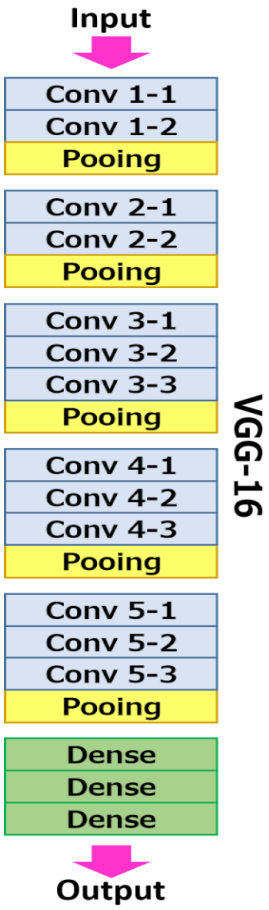

Figure 1 Architecture of VGG-16 Model

1.2.1 Architecture

VGG-16 architecture consists of different sets of convolutional layers and pooling layers. The input to convolution-1(convo) layers, has a fixed size of 224 x 224 RGB image. Convo 1-1 and convo 1-2 layer have 64 channels of 3 x 3 filter size and have the same padding. These filters have a very small receptive field. After convo1, convo2 layers there is a max pooling layer of stride 2 x 2. Similar to previous layers there are again two convolution layers (Con2-1 and Convo2-2) and followed by a pooling layer. Convolution-2 layer has 256 filter size and max pooling layer of 2 x 2 strides which is similar to previous layers. After this there are two sets of three convolution(convo) layers and max pool layer, Here, each filter has 512 (3 x 3) filter size and has similar padding. Spatial pooling is done by five max-pooling layers, which follow a portion of the conv. layers (not all the conv. layers are trailed by max-pooling). Max-pooling is performed over a 2 x 2 pixel window, with step 2. After convolution layers there are three fully connected(dense) layers. These layers have different depth in different architectures. The first two layers have '4096' channels each and third contains '1000' channels. For all networks configuration of fully connected layers remain the same.

1.2.2 Configuration

Different VGG Configurations are shown in figure 2

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

The above table shows different VGG configurations.

All these versions follow the same architecture; they just differ in depth. From the above table we can see that VGG-16 has two different versions i.e. version C and version D. In one version 3 x 3 filter size convolution is used while in other 1 x 1 else both the versions are almost similar.

2. RELATED WORK

Automated detection and classification models for Diabetic Retinopathy have been developed and studied in the past. Pratt et al. [6] trained a CNN using the SGD (Stochastic Gradient descent) method to classify Diabetic Retinopathy into five groups, and it achieved 95 percent precision, 75 percent accuracy, and 30 percent sensitivity using a training dataset including more than 70,000 retinal images.

Hagos et al. [7] used a pre-trained Inception-V3 framework to develop a model that classifies images into 2 classes of Diabetic Retinopathy and for this they used the Kaggle DR detection challenge dataset [8] which consisted of 2500 fundus photographs. They used SGD optimizer for their model and achieved an accuracy of 90.9% and a loss of 3.94%.

Garcia et al. [9] suggested a procedure that applied CNN to both the right and left eye images individually (Alexnet, VGGnet16, etc.). To improve the image contrast, the pre - processing and augmentation steps were applied to the dataset. They had a precision of 93.65 percent, a sensitivity of 54.47 percent, and an accuracy of 83.68 percent. DR stages, on the other hand, were not specifically categorized in their work.

The related work consists of two models that have done binary classification of images i.e they classified images into classes labelled DR, No DR. Although these models were successful in achieving a high accuracy they haven't been able to specify the severity of Diabetic Retinopathy. Our model was successful in specifying the severity of Diabetic Retinopathy from the fundus image that has been given to it as input. The one model from the related work that was able to classify the images into 5 classes has used a dataset consisting of 70,000 images and still achieved an accuracy of 75% this means that the training process for this model must have taken a lot of time and the fact that it still achieved an accuracy of 75% means that this model is highly inefficient. On the other hand our model required a dataset of 3668 retina images for training and achieved an accuracy of 74.58%.

3. METHODOLOGY

The Diabetic Retinopathy Detection Model's process flow diagram is shown in Figure 2.

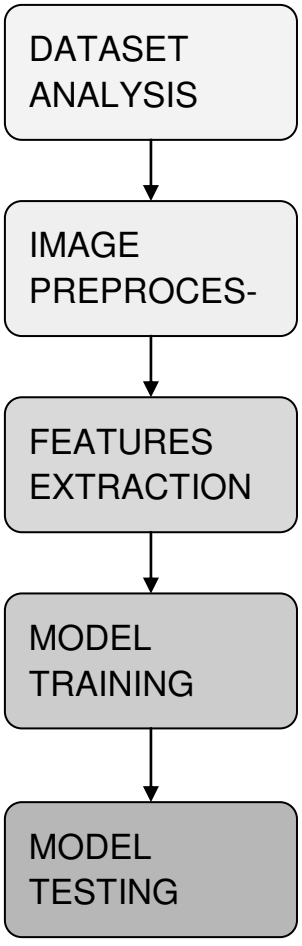

Figure 2 Process Flow Diagram

3.1 Dataset Analysis

We have used the data collected by the Asia-Pacific Tele-Ophthalmology Society (APTOS) in 2019 which is available on the kaggle platform as a dataset for the detection of Diabetic Retinopathy [10]. This dataset consists of 3668 retina images. These images have been obtained using the fundus photography technique. The fundus photography technique is exclusively used to find the abnormalities present in the eye. These images have been obtained under varied imaging circumstances. We have used the dataset of such images for the training and testing of our model. Each image in the dataset has been assigned an integral value on the scale of 0 to 4 according to the severity of the disease by a professionally trained clinician as shown in Table-1.

INTEGRAL VALUE	SEVERITY
0	No DR
1	Mild NPDR
2	Moderate NPDR
3	Severe NPDR
4	Proliferative DR

TABLE 1

The number of images under each label is represented by a histogram in Figure 3.

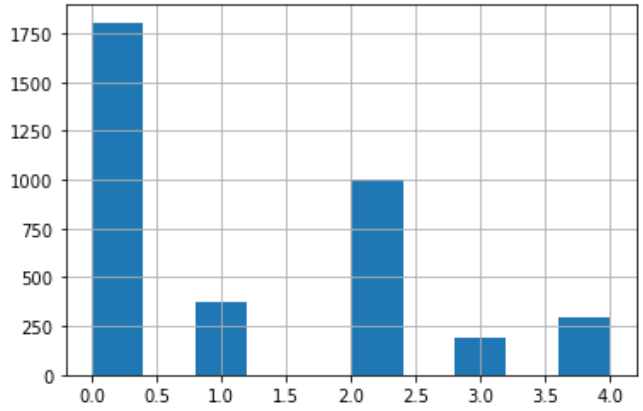

Figure 3 Dataset Analysis

According to the histogram in Figure 3 and Table-1 the training Dataset consists of just above 1750 images having No DR, 275 images having Mild NPDR, 1000 images having

moderate NPDR, just below 300 images having severe NPDR and just above 300 images having Proliferative DR.

3.2 Pre - processing of images and extraction of features

The framework of a Convolutional Neural Network is split into two parts: convolutional and classifier. The classifier part categorizes the input data based on the collected attributes, while the convolutional portion interacts with the data and extracts the distinctive characteristics of the different classes. Many Deep Learning models are made openly accessible in Keras that do well when trained on a small segment of the ImageNet dataset. These models' convolutional portions have already been trained on the ImageNet dataset to extract its distinguishing characteristics. [7]

The pre-trained convolutional portion of VGG-16 was used to extract the distinctive features of the images obtained from the fundus photography technique in this study.

Tensorflow keras was used to import the VGG16 framework. The image module was imported to pre-process the image object, while the preprocess_input package was imported to adjust the pixel values suitably for the VGG16 model. The numpy package was imported for array processing. The pretrained weights from the imagenet dataset were then loaded into the VGG16 model. An image of the retina from our dataset is shown in Figure 4.

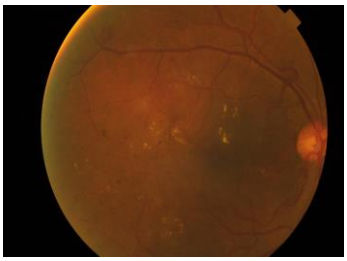

Figure 4

A sequence of convolutional layers are succeeded by multiple dense (or fully - connected) layers in the VGG16 model. Include_top lets the user determine whether or not the final dense layers should be included. False indicates that the final dense layers are excluded when the model is loaded. The feature extraction portion of the model is defined as the portion of the framework from the input layer to the last max pooling layer (labelled by 7 x 7 x 512), while the remainder of the framework is defined as the classification portion. We loaded the image that was given as input to the model with the dimensions 224*224 after defining the model. By converting the image PIL object to a NumPy array of pixel data, we were able to expand the 3-dimensional array to a 4-dimensional array with different sample sizes, rows, columns, and channels. For the VGG-16 model, the pixel values were adjusted suitably. The features of the photographs could then be extracted.

3.3 Model Training

We trained the model over 30 epochs. An epoch is one loop over the entire training dataset in convolutional neural networks. Training a neural network usually requires numerous epochs. The model was trained using ADAM optimizer. Optimizers are techniques or tools that adjust the characteristics of your neural network, such as the learning rate and also the weights, to minimize losses. ADAM is considered as the best optimizer as it is very efficient and takes very little time to train the model. The model is trained using a Cross Entropy loss function, a Softmax activated layer, and a batch size of 32.. We used a learning rate of 0.001, which is Adam's default parameter.

3.4 Model Testing

This model of ours was tested on a completely new set of 1728 retina images which were not included in the 3668 images on which the model is trained. While testing the model Image pre-processing and features extraction is done. The model labels the images in the output on a scale of 0-4 according to the severity of Diabetic Retinopathy (refer to Table-1). Our model was 74.58 % accurate in this regard. There occurred a loss of 5.06% over the course of 30 epochs.

Since Kaggle enables users to access and upload datasets and create projects/models on an internet based data science platform, we used it to download the VGG-16 pretrained model, train, and test our model.

RESULTS AND DISCUSSION

The model was able to detect the severity of Diabetic Retinopathy of the input images and labelled them according to it on the scale of 0-4 . The details of the model and the results achieved are given in Table-2.

SIZE OF TRAINING DATASET	3668 Images
SIZE OF TESTING DATASET	1728 Images
LOSS FUNCTION	Categorical Cross Entropy
OPTIMIZER USED	ADAM
CLASS MODE	Categorical

EPOCHS	30
LOSS	5.06%
ACCURACY	74.58%

TABLE 2

The results suggest that using this model can be an effective method for the detection of Diabetic Retinopathy at an early stage. They also suggest that such models can be developed for the detection of other medical conditions pertaining to the eye.

Doctors can overlook some of the characteristics and properties of each point, resulting in an incorrect diagnosis As a result, the concept of creating an automated DR detection solution emerges. More than half of the most recent cases of this disease could be prevented with appropriate and prompt care and eye monitoring.

Although this model might seem to be an easy way to detect the disease it cannot be an alternative to doctors as it is not 100% accurate thus this can only be used as a way to assist the doctors in the diagnosis of Diabetic Retinopathy.

CONCLUSIONS

In recent years, diabetes has been one of the fastest-growing diseases. According to multiple studies, a diabetic patient has a 30% risk of developing Diabetic Retinopathy (DR). If the disease is not diagnosed early on, it can lead to floaters, blurred vision, and eventually blindness. Manual diagnosis of these photos is time-consuming and complex and requires highly qualified specialists. We have successfully developed a Convolutional Neural Networks' model with a pre-trained VGG-16 framework that detects Diabetic Retinopathy and also gives information regarding the severity of the disease. We have been able to achieve an accuracy of 74.58% for the model. This model can be helpful in assisting doctors for a faster diagnosis of this disease. Similar models can be developed for diagnosis of other diseases especially the ones pertaining to the eye. This might be helpful in detecting these diseases at an early stage and avoiding permanent blindness.

REFERENCES

[1] Alexandr Pak, Atabay Ziyaden, Kuanysh Tukeshev, Assel Jaxylykova & Dana Abdullina | Duc Pham (Reviewing editor) (2020) Comparative analysis of deep learning methods of detection of diabetic retinopathy, Cogent Engineering, 7:1, DOI: 10.1080/23311916.2020.1805144

[2] Sridhar, Sashank & Sanagavarapu, Sowmya. (2020). Detection and Prognosis Evaluation of Diabetic Retinopathy using Ensemble Deep Convolutional Neural Networks. 10.1109/IES50839.2020.9231789.

[3] Nirmala, S. & Sharma, Purabi. (2017). Computer Assisted Methods for Retinal Image Classification. 10.4018/978-1-5225-0571-6.ch039.

[4] GeeksForGeeks https://www.geeksforgeeks.org/vgg-16-cnn-model/

[5] Neurohive

https://neurohive.io/en/popular-networks/vgg16/

[6] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding and Y. Zheng, "Convolutional neural networks for diabetic retinopathy", *Procedia Comput. Sci.*, vol. 90, pp. 200-205, Jan. 2016.

[7]https://arxiv.org/abs/1905.07203

[8] Kaggle

https://www.kaggle.com/c/diabetic-retinopathy-detection

[9] García G., Gallardo J., Mauricio A., López J., Del Carpio C. (2017) Detection of Diabetic Retinopathy Based on a Convolutional Neural Network Using Retinal Fundus Images. In: Lintas A., Rovetta S., Verschure P., Villa A. (eds) Artificial Neural Networks and Machine Learning – ICANN 2017. ICANN 2017. Lecture Notes in Computer Science, vol 10614. Springer, Cham. https://doi.org/10.1007/978-3-319-68612-7_72

[10] APTOS 2019 Blindness Detection dataset.

https://www.kaggle.com/c/aptos2019-blindness-detection