International Research Journal of Pharmacy and Medical Sciences
ISSN (Online): 2581-3277

Diabetic Retinopathy Classification Using Deep Learning Architectures

Safeela P., Deepa T.
Dept. of Electronics and Communication Engineering MGMCEPS Malappuram, Kerala, India
Email address: safeelasamad@gmail.com, deepa.ec@mgmcet.in

Abstract— *Diabetes is the main cause for the blindness in the working stage of adults. It causes to blurry vision and loss of eye sight in diabetic patients. Early detection and treatment for this condition leads to prevention of Diabetic Retinopathy (DR). DR is an art and science of recording fundus images on the diabetic patient. Based on the severity of diabetes, DR can be classified into five stages- no DR, mild DR, moderate DR, severe DR, proliferate DR. Before classifying fundus retinal images of diabetic patient, images must be preprocessed. For improving and testing we can use the Kaggle dataset. Our experiments have been performed by using deep learning networks – multilayer perception and convolutional neural network. We report the results based on the accuracy such as validation accuracy and training accuracy. By using different optimizers and different activation functions we report the comparison charts of these two methods – multilayer perception and convolutional neural network.*

Keywords— *Diabetic retinopathy, Multilayer perception, CNN architrecture, Deep learning.*

I. Introduction

In the present modern days one of the main reasons for blindness is DR. Diabetic retinopathy is the result of untreated prolonged diabetes. Early detection of this condition helps to treatment and can prevent DR. DR has five stages – no DR (class 0), mild DR (class 1), moderate DR (class 2), severe DR (class 3), proliferate DR (class 4) as shown in table 1. These stages can be found on the severity of diabetes.

To perform the experiments, Kaggle datasets are used. The datasets from Kaggle have the property of varying sizes as well as different DR severity. So the preprocessing is needed for the classification of these datasets. By preprocessing techniques the images transform into fixed size. Geometric transformations are used as preprocessing technique. Multilayer perception network for the classification purpose on one side. And on another side we use convolutional neural network for classification of the fundus images. By comparing validation and training accuracies obtained from these two methods , we report the results. We use DL (deep learning) approach for multiclass classification.

II. Datasets

A. Kaggle Datasets

We propose Kaggle datasets has class labels corresponding to 5 stages of DR – class 0 label for no DR, class 1 for mild DR, class 2 for moderate DR, class 3 for severs DR and class 4 for proliferate DR. The images in the Kaggle datasets have very large size of 2MB of disk space. So the preprocessing of images is an important part of classification. Number of images from class in Kaggle dataset are shown in Table 1.

Total 35,216 images are available from Kaggle dataset for classification. Those images are large and varying sizes. From Table 1, 73.29 percentage of images are belonging to the class 0 that is no DR. Very small amount of images are belonging to the proliferate DR stage.

TABLE I. Number of images per class in Kaggle dataset

Class	Stage	Kaggle dataset
0	No DR	25810
1	Mild DR	2443
2	Moderate DR	2592
3	Severe DR	873
4	Proliferate DR	703
Total		35716

III. Preprocessing of Images

The fundus images are of large and varying sizes. To process dataset, we propose a preprocessing approach. By preprocessing techniques the large and varying sized images are transform into fixed size.

Preprocessing techniques include geometric transformation such as rotating, zooming, resizing etc. After preprocessing technique the size of the fundus images are of 300×300×3. Resizing of images is the main preprocessing technique here. For classification purpose, we need the fixed sized images for it proper function. By resizing the images we get the images of size 300×300×3, where 300×300 indicate length×breadth and 3 indicate depth of image.

IV. Deep Learning

Deep learning is an emerging field of machine learning. That is, it is a subset of machine learning where learning happens from past examples or experiences with the help of Artificial neural network. Deep learning uses deep neural networks, where the word ‘deep’ signifies the presence of more than 1or 2 hidden layers apart from the input and output layers.

Types of deep neural network

1) Artificial neural network
2) Multilayer perception
3) Recurrent neural network
4) Convolutional neural network

International Research Journal of Pharmacy and Medical Sciences
ISSN (Online): 2581-3277

V. Proposed Architectures for Classification

A. Convolutional neural network

CNN is an more accurate way for image classification problems. By using CNN architecture does not need an external feature extraction method. It extract features from images automatically. Figure 1 shows the structure of CNN.

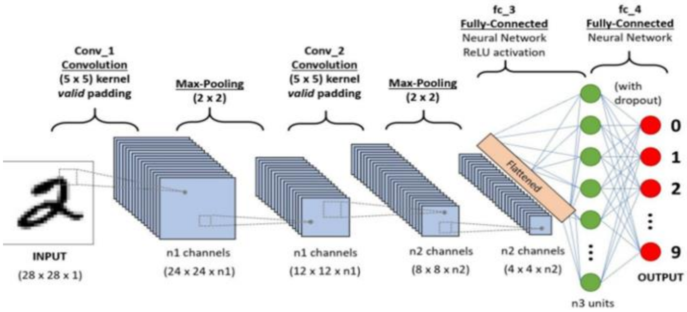

Fig. 1. Structure of CNN

1. *Convolutional layer*

Convolutional layer is the first layer of convolutional network. It can perform feature extraction by sliding the filter over the input image. The output or convolved feature is the element wise product of filters in the image and their sum for every sliding action. The output layer also known as the feature map , correspond to original images like curves, sharp edges, textures etc. In the case of networks with more convolutional layers are meant for extracting the generic feature while the complex parts are removed as the network deeper.

2. *Pooling layer*

Pooling layer is used to reduce dimensionality of image obtained from the convolutional layer. Primary purpose of this layer is to reduce the number of trainable parameters by decreasing spatial size of the image. So we can reduce computational cost. The most popular pooling technique, known as max pooling, uses the feature map's most important component as its input. Max pooling is then performed to give the output image with dimensions reduced to a great extent while retailing the essential information.

3. *Fully connected layer*

The completely linked layers are the final few layers that determine the output. The fully connected layer receives the output from the pooling layer after it has been flattened into a one-dimensional vector. The output layer has the same number of neurons as the number of categories we had in our problem for classification.

4. *Optimizers and activation functions*

a. Stochastic gradient descent optimizers

Stochastic gradient descent is an optimization method for unconstrained optimization problem.

b. Adam optimizers

In place of the conventional stochastic gradient descent method, Adam is an optimization technique that may be used to iteratively update network weights depending on training data.

c. RMSprop

The goal is to maintain a steady movement of the average gradient square.

We here using tanh and Relu activation functions.

B. Multilayer perceptron

Multilayer perceptron (MLP) is a type of forward network. It has three layers – input layer, output layer and hidden layer. The input layer takes input data to be processed. Classification is performed by the output layer. Number of hidden layers are situated between input and output layers. Similar to feed forward network in MLP the data flows in forward direction. That is from input to output layer. Pattern classification, recognition, prediction, and approximation are the main use cases for MLP.

VI. Experimental Results

By classifying the retinal images of diabetic patients by using these two architectures approaching deep learning, we get the results are as shown in Table 2,Table 3,Table 4 Table 5.

TABLE II. Number of hidden layers, activation function, optimizer, learning rate and epoch used for the MLP

No. of hidden layer	Activation function	Optimizer	learning rate	epoch
16	Tanh	Adam	Updtes	5
16	Tanh	Adam	Updtes	10
32	Tanh	Adam	Updtes	5
32	Relu	Adam	Updtes	5
32	Relu	Adam	Updtes	10
256	Tanh	Adam	Updtes	10
256	Tanh	SGD	0.001	10
256	Tanh	RMSprop	0.001	10

TABLE III. Accuracy and loss during training and validation based on table 2 while using MLP

Training loss	Training accuracy	Validation loss	Validation accuracy
1.2877	0.4902	1.2954	0.5034
1.2817	0.4902	1.2963	0.5034
1.2881	0.4902	1.2968	0.5034
1.4911	0.4902	1.4323	0.5034
1.3827	0.4902	1.3335	0.5034
1.2365	0.4883	1.2356	0.5034
1.2970	0.4907	1.3079	0.5034
1.3046	0.4843	1.3284	0.4954

TABLE IV. Number of hidden layers, activation function, optimizer, learning rate and epoch used for the CNN

1st CL	2nd CL	3rd CL	4th CL	Activation function	No. of hidden layer	Optimizer
16	32	64	X	Relu	16	RMSprop
16	32	64	X	Relu	16	RMSprop
16	32	64	256	Relu	16	RMSprop
16	32	64	256	Relu	256	RMSprop

TABLE V. Accuracy and loss during training and validation based on table 4 while using CNN

Epoch	Training loss	Training accuracy	Validation loss	Validation accuracy
5	0.8672	0.7066	0.8594	.6899
15	.7972	0.7207	.7676	0.7227
30	0.7870	.7259	.7992	0.7090
15	0.8008	0.7183	0.7795	.7063

While comparing the results with state of art methods in table 3, we observe that the multilayer perceptron outperforms 50 percent accuracy of classification.

However, training accuracy obtained from CNN architecture with 16 hidden layers is around 70. Maximum training accuracy can be obtained by using RMSprop optimizer and Relu activation function with 30 epochs.

VII. Conclusion

Large image sizes are the major issues while dealing with bio-medical image classification problems. On the basis of experiments conducted on images by CNN and MLP, we conclude that the results obtained by using CNN architecture is better. And large number of epochs improve performances. Thus we propose a novel approach to deal with large and varying size input data by CNN method.

References

[1] P. Mishra, S. Mishra, J. Nanda and K. V. Sajith, "Multilayer perceptron neural network (MLPNN) controller for automatic generation control of multiarea thermal system", *North American Power Symposium*,2011.

[2] D. Zhang and T. Zhou, Deep Convolutional Neural Network Using Trans- fer Learning for Fault Diagnosis, *IEEE Access*, vol. 9, 2021.

[3] N. Ahmed, T. Natarajan and K. R. Rao, Discrete Cosine Transform, *IEEE Transactions on Computers*, vol. C-23, no. 1, Jan. 1974.

[4] W. Zheng, "Intrusion Detection Based on Convolutional Neural Net- work" *International Conference on Computer Engineering and Applica- tion (ICCEA)*,2020.

[5] M. E. Mavroforakis and S. Theodoridis.' Support Vector Machine (SVM) classification through geometry", *13th European Signal Process- ing Conference*, 2005, pp. 1-4

[6] Z. Khan. "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture", *IEEE Access* 2021, vol. 9

[7] H. Qassim, A. Verma and D. Feinzimer, "Compressed residual-VGG16 CNN model for big data places image recognition", *IEEE 8th An- nual Computing and Communication Workshop and Conference (CCWC)* 2018, pp. 169-175

[8] L. J. Chipman, T. M. Orr and L. N. Graham, "Wavelets and image fusion", *International Conference on Image Processing,* 1995