IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Classification of Diabetic Retinopathy using Residual Neural Network

To cite this article: C Priyadharsini *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **925** 012033

Classification of Diabetic Retinopathy using Residual Neural Network

C Priyadharsini[1], J Jeyapriya[2], R Jagadeesh Kannan[3], G Bharadwaja Kumar[4] and Tulasi Prasad Sakriki[5]

[1]School of Computer Science & Engineering, Vellore Institute of Technology Chennai Campus, India,
[2,3,4,5]School of Computing Science and Engineering, VIT Chennai, India,
[1]priyadharsini.c27@gmail.com, [2]priyacseme@gmail.com, [3]jagadeeshkannan.r@vit.ac.in, [4]bharadwaja.kumar@vit.ac.in, [5]tulasiprasad.sariki@vit.ac.in

Abstract. Diabetic retinopathy (DR) in patients affects retina function owing to chronic elevated excessive blood glucose rates. The DR is a severe medical disorder. Patients with diabetes are easily prone to this medical complication which when not detected and treated at earlier stages, leads to vision loss. Globally, diabetes mellitus is the fifth leading cause of vision loss. So, active research is being conducted in this area to find novel ways in identifying the stages of DR. Specific image recognition methods and computer simulation algorithms were initially used to classify DR, but their usefulness was inadequate in real-time clinical practice. The evolution of deep learning models like convolution neural network performed better in identifying DR and non-referable DR compared to conventional machine learning models. Different variations of CNN architecture are being developed over the period, but more analysis and experimentation needs to be carried out to choose the appropriate architecture for detecting Diabetic Retinopathy. The aim of this research is to apply and understand how the performance of pre-trained Deep learning model – ResNet, a deep layered neural network performs to identify non-referable DR and different types of referable DR.

Keywords: CNN, Deep Learning, Diabetic retinopathy, Neural network, ResNet.

1. Introduction

T as a consequence of hyperglycaemia, a persistent diabetes occurs-elevated blood glucose rates. The presence of hyperglycaemia for a longer period is the starting point of DR development. The numbers of people prone to Diabetes globally are increasing at a very faster pace. In India alone, 69,9 million and 80 million people are projected to cross diabetic by 2025 and 2030. And across the globe, estimated population with diabetes retinal related disease is 382 million and in the year 2025 it might rise up-to 592 million [1]. A diabetes mellitus period in patients was the key predictor for assessing the incidence of diabetic retinopathy. After the duration of 10 years, 20% of Type I and 25% of Type II, 20% of Type I and 60% of Type II, 95% of patients with Type I or Type II acquire this disorder over the span of 30 years. Type I diabetic patients have absolute deficiency of insulin and Type II diabetic patients suffer from defect in progressive insulin secretion [2]. Females are more affected than males in the ratio of 4:3. Pregnancy, obesity, hypertension, hyperlipidaemia, smoking, anaemia are other risk factors associated with progression of DR [3].

Clinical studies show that the progression of DR can be slowed down by bringing blood lipids, blood glucose, blood pressure under optimal control when detected at the earlier stage. But the rate of diabetic population is rising at a rapid pace due to lack of awareness in one-third of the population. And another reason is, as per analysis by experts for 60 million diabetic populations prone to retinal disease, approximate number of Ophthalmologists and Retinal consultants are 12,000 and 3,500 respectively [1]. This is the reason why WHO categorises DR as a major retinal disease and expects attention from professionals and authorities [1, 20]. With advanced machine/ deep learning models and high computational power, computer professionals and engineers can make a significant contribution in identifying DR. This research aims to explore deep learning trends accessible and to explain how ResNet performs in the recognition of various stages in DR in a strong pre-trained model.

Organisation: The remaining research is structured as follows. The following. Part 2 offers a awareness of diabetes retinopathy. The related research articles are discussed in Section 3. Section 4 offers details and the need for ResNet architecture. The ResNet model's study findings are presented in section 5 and potential work is carried out in section 6. And last but not least, Section 7 ends this report.

2. Technologies used

DR is caused when lesions are formed as a reaction to change in structure of blood vessels or due to leakage of fluids like lipid in the retina. As the progression of DR takes place, different symptoms become prominent. This section discusses about features of DR and its classification of different stages [3][4].

2.1. Features of DR

To Major features associated with DR, identified depending upon the shape, colour, texture and size of the symptoms are as follows:

- Microaneurysms
- Retinal haemorrhages
- Hard exudates
- Cotton-wool spots
- Neovascularization

Microaneurysms are identical in nature and dot haemorrhages differ in duration. The microaneurysms, red spherical dots appear small. Blot and flame haemorrhages are big in size. The yellowish white waxy appearance is an attribute associated with hard exudate while cotton wool spots have white fluffy characteristics. If retinal blood vessels alter their shape to obstruct blood supply, there are additional arteries in various places in the eye. This condition of new blood vessel development is called neovascularization.

2.2. Classification of DR

DR is generally graded as a non-proliferative DR (NPDR) and proliferative DR (PDR) in two stages. The development of fresh blood vessels in the eye is correlated with growth here. Non-proliferative DR is a stage where different types of lesions develop in the retina because of micro vascular occlusion which differ in shape, structure, texture but there is no new development of blood vessels. DR phases of NPDR may be mild, moderate and extreme, depending on the level of sensitivity of the pupils. The current step of DR is PDR where blood vessels that contribute to blood leakage contribute to lack of vision as in figure 1. Chart 1. The retinal fundus photos displayed below at different DR levels.

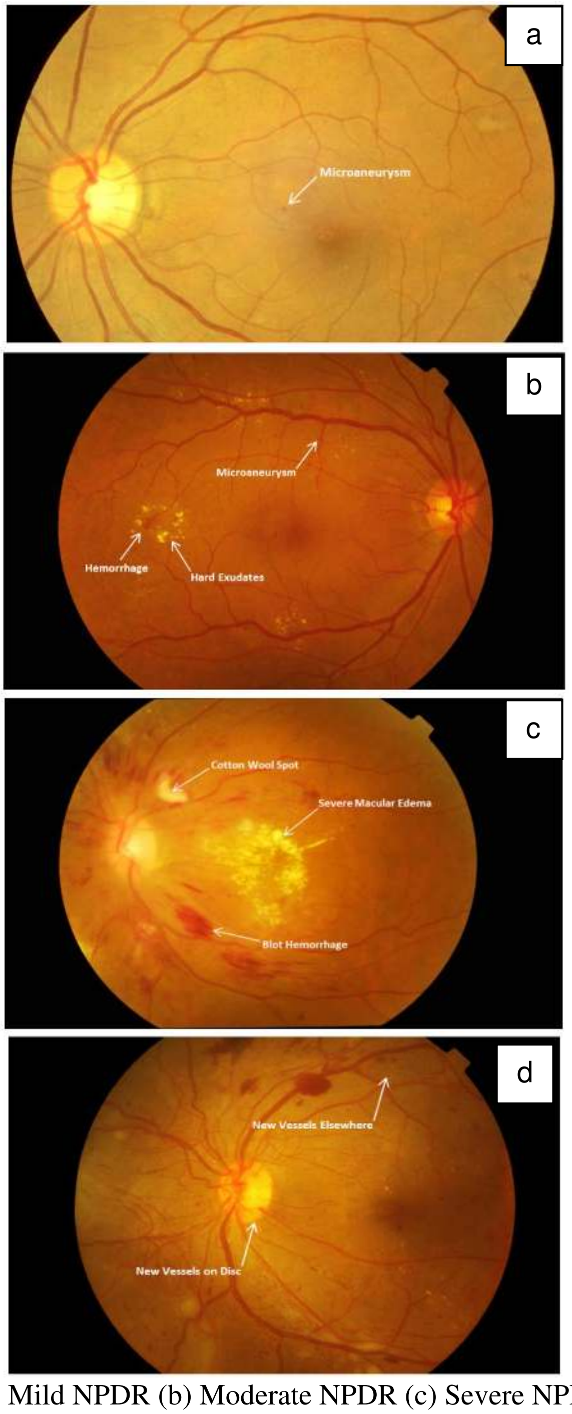

Figure 1. (a) Mild NPDR (b) Moderate NPDR (c) Severe NPDR (d) PDR

According to ETDRS study, non-proliferative DR is classified as follows as in table. 1 based on the severity level [3].

Table 1.non-proliferative DR classification

Grade	Features
R0: No abnormalities	No abnormalities
R1: Mild NPDR	At least one microaneurysm, Retinal hemorrhages, Hard or soft exudates
R2: Moderate NPDR	Strong exudates in at least 1 quadrant (Cotton wool spots) Microaneurysms and / or line and mark haemorrhages
R3: Severe NPDR	Any one of the following 3 features is present (Known as the 4-2-1 rule) • Intraretinal hemorrhages and microaneurysms in all four quadrants. • In 2 or more quadrants Venous Beading • Strong IRMA in 1 quadrant or more
R4: PDR	There are two of the features of the 4 2 1 law

3. Related works

Earlier the use of machine learning models was focussed towards detecting different types of lesions in DR related research works. The paradigm has been changed in the last years to introduce specific profound thinking models to assess the DR gravity point.

In a survey paper published during the year 2014 on ocular disease by Zhuo et al. [5], The author notes that much of his research has focussed on lesion identification relevant to DR. And there were few studies accessible for DR identification that included the DR classification in binary, i.e. reproducible or non-referable DR. Pal et al. [6] studied the implementation of algorithms in machine learning such as the Decision Tree, Classification in Naïve Bayes, Vector Support machines and neighboring K-retinopathy algorithms. The results obtained from applying ML algorithms on Messidor dataset showed SVM had an accuracy of 74.65 % which was highest among the other models.

In Messidor and Eyepacs datasets, Gulshan et al. [7] recorded for the first time the DL for DR recognition. Big picture databases of the fundus were used for monitoring of a deep CNN. The emphasis of the study was only on identification of DRs, but no comparable knowledge on innovative DRs or other DR phases was provided. Abr`amoff et al [8] developed IDx-DR X2.1, a device used for DR detection. The class of classification used in this device are rDR, vtDR and pDR which means referable, vision threatening and proliferative DR respectively. The first designed CNN architecture called AlexNet was used to identify the characteristics of lesions, optic disk and fovea to stage the DR affected retinal images. Lu et al [9]. in his paper on overview of artificial intelligence in ophthalmology, observed that CNN is the most used technique to classify ocular diseases. The author states the performance of DL goes down when more than one disease is included. Therefore, it is important to build systematic AI systems to diagnose general eye disorders in the modern world utilizing multimodal data.

To detect non-DR and DR affected retinal images, Gargeya et al. [10] used CNN for learning deep discriminative features and the result was visualized using heatmap. Features related to pixel height, width and field of view was added with 1024 features obtained after the average pooling layer in CNN and gradient boost classifier on top of it to perform binary classification.

GoogleNet architecture was modified by Takahashi et al. [11] to classify the severity level of DR. Deletion of top 5 layers, reduction in batch size, expansion of crop size are the major modifications applied to GoogleNet architecture. Two different approaches were used by the author in model design. One approach is usage of three coloured retinal images and other retinal image of only one colour for manual staging. The experiment was validated with 20-fold cross validation and compared with ResNet model.

Sequence of pre-processing steps were applied by Garcia et al. [12] before feeding the data on various CNN architectures for detecting DR. The average color of each pixel in the picture is measured and subtracted and the resolution is restated to 256 x 256. After pre-processing data augmentation is carried out as a next step. The results are analysed by applying the resultant dataset on different neural network architecture by varying the learning rates and number of layers.
Though CNN architectures perform better in identifying DR, the number of parameters calculated are huge. So, this paper aims at analysing the performance of ResNet, a neural network architecture where the number of parameter calculation is reduced but the network is deep.

4. Architecture of resnet model

4.1. Description

Krizhevsky et al. [13] shifted the paradigm from conventional machine learning model by proposing a deep conventional neural network architecture called AlexNet. Based on this architecture, the concept of simple design is the number of features that the model increases with rising layers in the neural network. Many models such as ZFNet, VGG16, GoogleNet were born of this idea.
To understand the impact of number of layers in the design of CNN on the performance of the model, He et al. [14] experimented and analysed two CNN models by varying the number of layers- one CNN with 20 layers and other with 56 layers with CIFAR-10 dataset. Figure 2 shows training error plot and test error plot plotted against respective errors vs number of iterations. It is observed that both training error and test error is high for the CNN model with 56 layers. The reason for high error rate is to depict it is difficult for all the models to optimise and not as a result of overfitting or adding more layers.

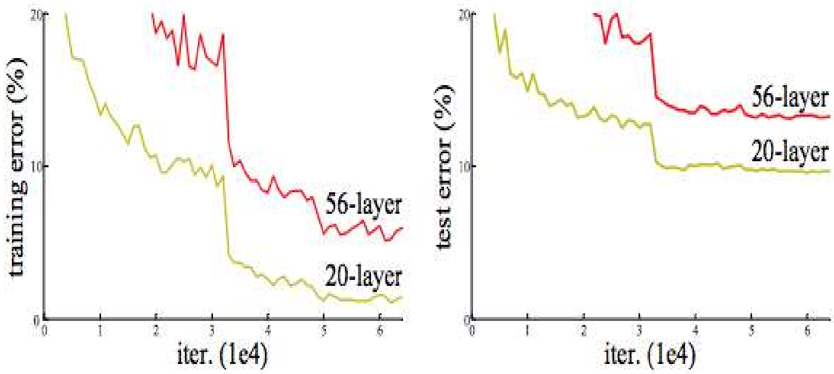

Figure 2. (a) shows the plot between training error percentage vs number of iterations for ResNet 20 and ResNet 56 layer, (b) shows the plot between testing error percentage vs number of iterations for ResNet 20 and ResNet 56 layer

To overcome this issue, residual block was introduced as a new neural network layer which works on the concept of skip or shortcut connections. This architecture was proposed by Microsoft called deep residual learning framework []. Figure.3 shows the structure of residual block.

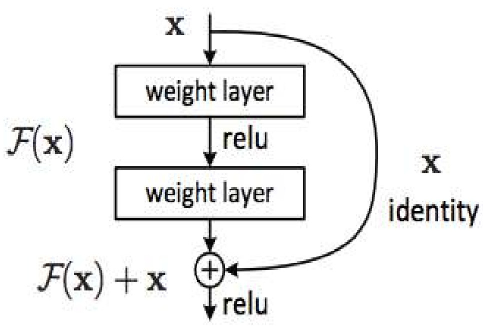

Figure 3. Structure of residual block

IOP Conf. Series: Materials Science and Engineering **925** (2020) 012033 doi:10.1088/1757-899X/925/1/012033

Skip connections improved the performance and mitigated the problem of adding deep layers since vanishing gradient problem is solved by adding the input to the convolution block output. Also, the skip connections ensured that the high layers do not perform less than the low layers by letting the model to learn an identity function.

4.2. Design principle

Different architectures [14] used for ImageNet are shown in figure 4. VGG-19, 34-layered single-scale networks and ResNet-34 are running at a pace of 19.6 billion, 3.6 billion and 3.6 billion per second. It can be noted that 19 layered VGG model has huge number of FLOPs compared to 34 layered ResNet model. The convolution layers in plain baseline model are assigned mostly with 3 x 3 filters. The basic principle of this model is when the output feature map is same or halved, then the number of filters is same or doubled respectively. In ResNet model, shortcut connection is introduced. When the dimensions of input and output is same, identity shortcut F(x{W}+x) is usedAs measurements change, then the projection workaround in ID feature is not used to suit measurements by padding neros for expanded dimensions after the description mapping or by using 1 x 1 convolution.

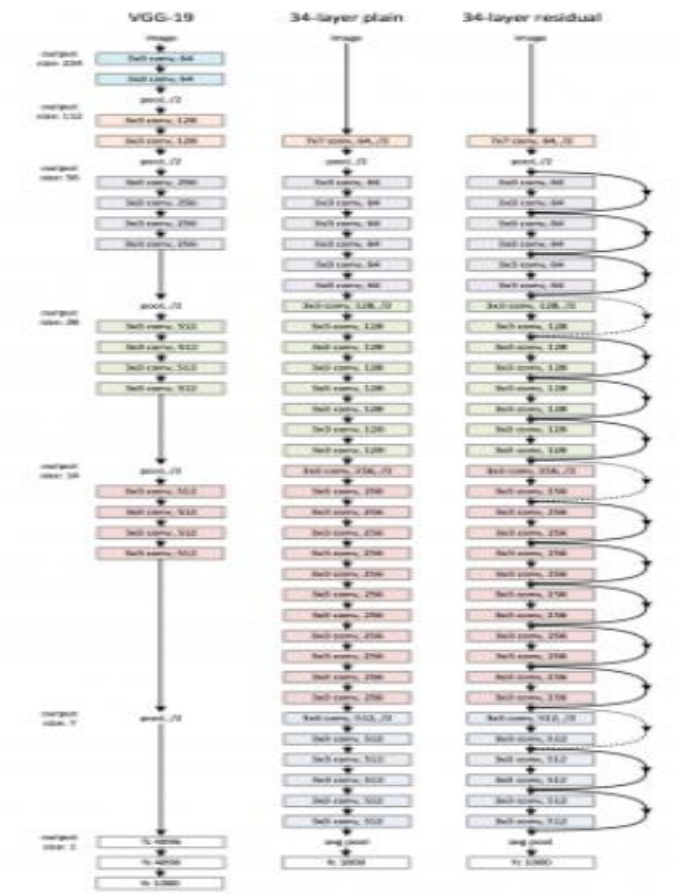

Figure 4. shows architecture of 34-layer residual model, 34-layer plain model and VGG-19 model.

In networks of small size like ResNet 18 and ResNet 34, each ResNet block is of depth 2 layers[15]. In ResNet model of 50 layers figure.6, 101 layers and 152 layers, the 2 layered blocks are replaced by three layered blocks which is shown in Figure .5. The number of FLOPs performed by this model is 3.8 x 10 9.

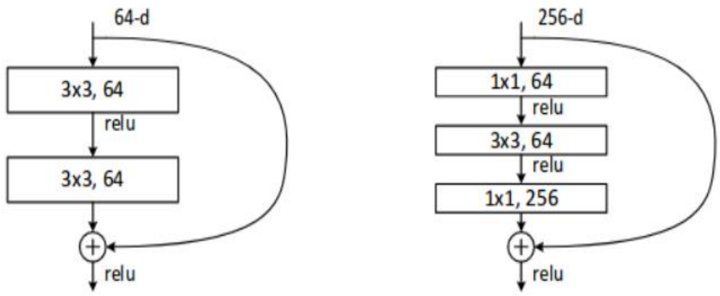

Figure 5. (a) 2 layered residual block and (b) 3 layered residual block

4.3. Specification

Below table tabulates the number of filters, each filter size and output size from convolution layers for ResNet 50 model:

Layer Name	Output size	50-layer
Conv 1	112 x 112	7 x 7, 64, stride 2
Conv 2_x	56 x 56	3 x 3, max pool, stride 2 3 x [1x1,64 3 x 3,64 1 x1,256]
Conv 3_x	28 x 28	3 x [1x1,64 3 x 3,64 1 x1,256]
Conv 4_x	14 x 14	3 x [1x1,64 3 x 3,64 1 x1,256]
Conv 5_x	7 x 7	3 x [1x1,64 3 x 3,64 1 x1,256]
	1 x 1	Average pool, 1000-d fc, softmax
FLOPs		3.8×10^{9}

Figure 6. ResNet 50-layer architecture

4.4. Implementation

In the present work, Fastai deep learning library has been used to retrain the ResNet models. Fastai has many pre-trains: resnet18, densenet121, densenet169, densenet201, densenet161, vgg16 bn, vgg19 bn, alexnet, resnet101, resnet152, squeezenet1 0, squeezezenet1 0. Learning transition is a methodology in which a model trained on a large dataset may be used and the model can then be applied to the dataset from another area. The theory is that the neural network can learn well in very large datasets and that the model would benefit from this experience, particularly if the data set is limited relative to a randomly initialized model. In the last part of the model, the number of classes inside the current dataset has to be modified.

The experiments in the present work have been carried out on NVIDIA K-80 Tesla GPU with 12 GB RAM. In the present work, we have experimented with Resnet34, Resnet50 and Resnet101 architectures. But, when we increase the batch size and use Resnet101, there was a memory limitation on GPU RAM. Resnet34 was not able to produce better results. Hence, we have confined to Resnet50 architecture. We varied the batch size and also optimized the learning rate in the present experiments. The detailed experimentation is given in the next section.

5. Experimental results and analysis

In this research paper, ResNet 50 model is applied on retinal images to analyse the efficiency of the model for two tasks:

- Binary classification: to identify referable and non-referable DR images.

- Multiclass classification: to identify non- referable DR and different stages of DR the retinal image is affected with.

The dataset used for this experiment is collected from Kaggle website as part of the APTOS 2019 Blindness Detection competition. The collection includes a broad variety of high-resolution retina photographs in different imaging conditions. Every topic has a left and right area. Images are labelled both left and right with a subject identifier [16]. For each photograph the clinician has measured with the following scale the occurrence of diabetic retinopathy for 0 to 4: 0- No DR, 1-Make-Up, 2-Mild, 3 – Severe, 4-DR Proliferative. The tests are conducted with 80% of the training photos and 20% of the evaluation photos. The data set's class distribution is seen below Figure .7.

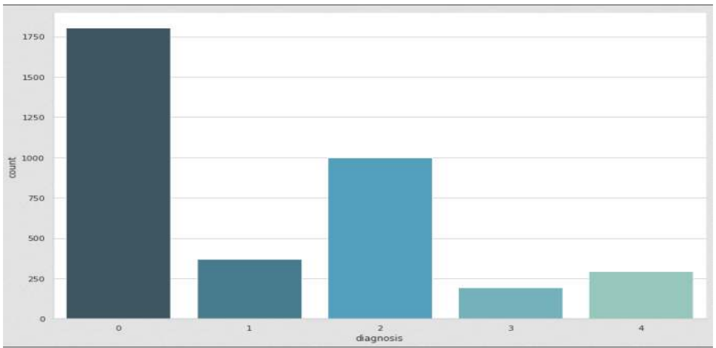

Figure 7. DR class distribution

With repeated experiments on the data, following value for the hyper-parameters are fixed. The number of epochs is set to five, batch size is fixed as 50. Apart from conventional method of tuning the hyper-parameters, for multi-class classification the input image size is changed, and the results are observed. The product efficiency is assessed by means of a uncertainty matrix analysis. Accuracy measure is defined as the ratio of positive and negative observations that are predicted same as the actual value to that of sum of all the possible combinations of negative and positive cases of actual and predicted value. The formula for accuracy is given by:

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

5.1. Binary Classification

Image Size = 224,224

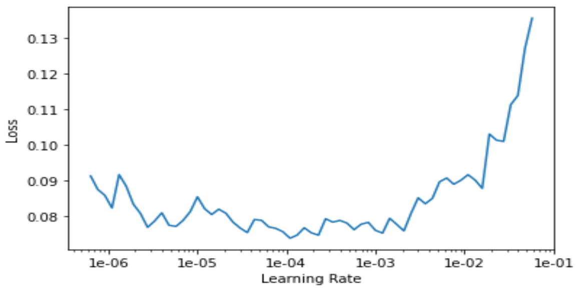

Figure 8. Loss function vs different learning rates

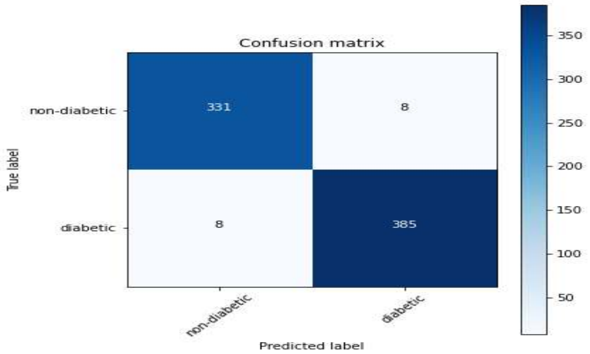

Figure 9. Confusion matrix

To identify referable and non-referable DR, the input image is resized to 224 x 224 and fed into the ResNet 50 model. The results are analysed by plotting a graph between loss function and learning rate for 50 experiments. From the plot Figure .8, we can find that loss is minimum when learning rate is at 5e-6,1e-5. The accuracy of the model at this learning rate is 0.9781420765027322 calculated from the confusion matrix in Figure .9.

5.2. Multiclass Classification
Two types of experiments are carried out by varying the input image size and analysed to check its effect over the performance of the model.

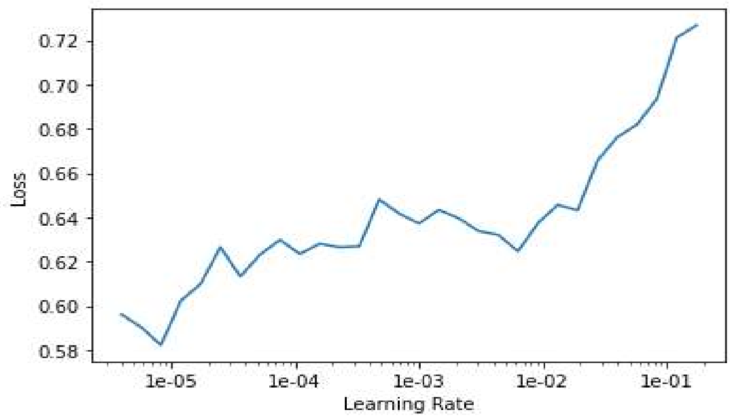

Figure 10. Loss function vs different learning rates

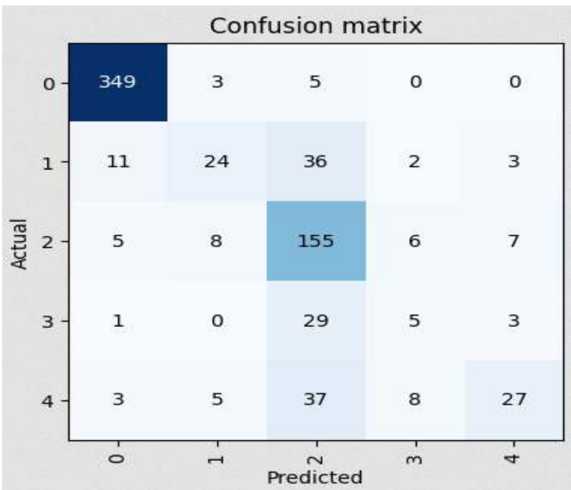

Figure 11. Confusion matrix

To identify DR images belonging to all the 5 categories, the image is resized to 100 x 100 and the results are analyzed for 50 times by varying the learning rates. From the plot figure .10 above, the best-chosen learning rates are 5e-6, 1e-5 and the accuracy estimated using confusion matrix figure. 11 is 0.7650. The reason for less accuracy is the count of images belonging to class one and four but predicted as the images with stage 2 of DR is more

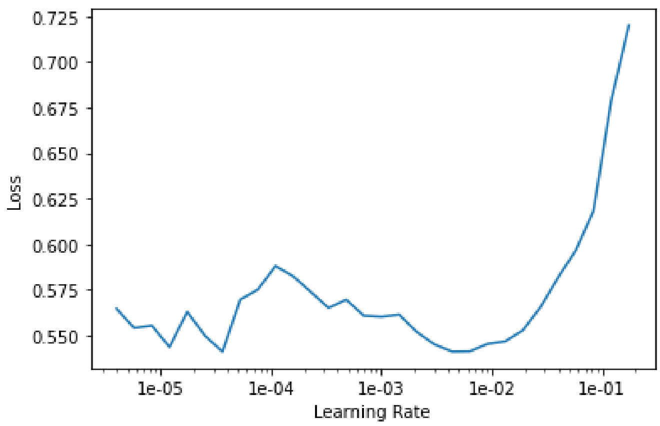

Figure 12. Loss function vs different learning rates

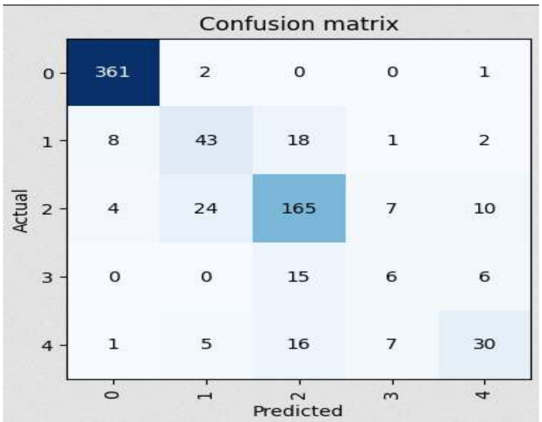

Figure 13. Confusion matrix

The accuracy is found to improve from 0.7650 to 0.8265 when the image size is increased from 100 x 100 to 224 x 244. The reason for this improvement in accuracy could be because of the skip or short connections in ResNet architecture which preserves the features related to DR found at the end of convolution layers as shown in figure 12 and figure 13.
To understand the efficiency of the model on the obtained Kaggle dataset for different classes of DR, additional steps of pre-processing or extraction of features or including hand-crafted features are not included before applying to ResNet model. It is found that ResNet models show promising accuracy of 97.8 % for binary classification. And for multiclass classification the accuracy is 82.65%. Table 2 shows the comparison table. Figure 14 shows the comparison graph.

Table 2. Comparison table

S.No	Methods	Accuracy (%)
[17]	Multiscale amplitude-modulation-frequency-modulation	92
[18]	Vector quantization	90
[19]	Support vector machine	86
Proposed system	**Residual Neural Network**	**97.8**

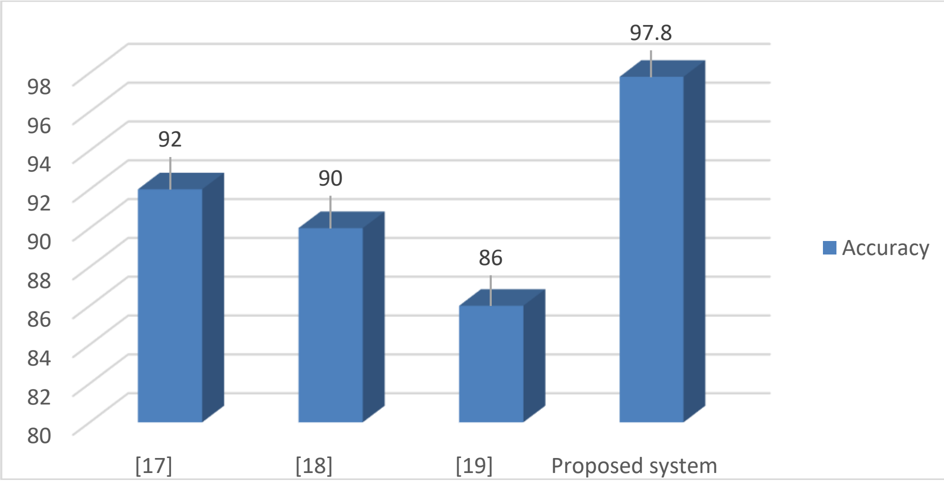

Figure 14. Comparison graph

6. Conclusion

The increase of diabetic retinopathy patients at an alarming rate pose a limitation for all the patients to undergo manual eye examination. Efficient algorithms to detect DR are the need of the hour to detect DR at different stages. So, that proper diagnosis can be planned at the earliest to avoid from vision loss. Deep learning models perform much better compared to conventional machine learning models in identifying DR. The ResNet model is implemented with the Kaggle data set to conduct binary and multiclass classification, in this research article, a deep layered neural network. From the experimental analysis, it can be inferred that a deep layered model – ResNet is able to perform better even on the dataset where additional steps to enhance the properties of the image is not carried out. Future work includes the current work can be extended to evaluate the performance by comparing different ResNet models, comparing ResNet model with other deep layered networks and compare the results of applying ResNet model after performing various image pre-processing operations.

References

[1] Suresh G, Pandey K, Sharma V 2018 World Diabetes Day 2018: Battling the Emerging Epidemic of Diabetic Retinopathy *Indian Journal of Ophthalmology* **66** 1652.
[2] Classification and Diagnosis of Diabetes, Diabetes Care 2015 Jan; **38**, S8-S16, 2015.
[3] Khurana AK Khurana B 2015 Comprehensive Ophthalmology: With Supplementary Book-Review of Ophthalmology *JP Medical Ltd*
[4] Wong TY, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh VC, Maia M, Mathenge W, Moreker S, Muqit MM, Resnikoff S. 2018 Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings *Ophthalmology* **125** 1608-22.
[5] Zhang Z, Srivastava R, Liu H, Chen X, Duan L, Wong DW, Kwoh CK, Wong TY, Liu J 2014 A survey on computer aided diagnosis for ocular diseases *BMC medical informatics and decision making* **14** 80
[6] Pal R, Poray J and Sen M 2017 Application of machine learning algorithms on diabetic retinopathy *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)* 2046-2051
[7] Gulshan V, Peng L and Coram M 2016 Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs *Journal of American Medical Association* **316** 2402
[8] Abramoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC and Niemeijer M 2016 Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning *Investigative Ophthalmology & Visual Science* **57** 5200–06

IOP Conf. Series: Materials Science and Engineering **925** (2020) 012033 doi:10.1088/1757-899X/925/1/012033

[9] Lu W, Yan T, Yue Y, Yiqiao X, Changzheng C 2018 Applications of Artificial Intelligence in Ophthalmology: General Overview *Journal of Ophthalmology* 1-15.
[10] Gargeya R and Leng T 2017 Automated identification of diabetic retinopathy using deep learning *Ophthalmology* **124** 962–969
[11] Takahashi H, Tampo H, Arai Y, Inoue Y and Kawashima H 2017 Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy, *Plos One* **12** 6
[12] Garc´ıa G, Gallardo J, Mauricio A, L´opez J and Del Carpio C 2017 Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images *In International Conference on Artificial Neural Networks ed*Springer **11071** 635–642
[13] Alex K, Sutskever I and Hinton EG 2012 ImageNet Classification with Deep Convolutional Neural Networks *In Advances in neural information processing systems* 1097-1105
[14] He K, Zhang X, Ren S and Sun J 2016 Deep Residual Learning for Image Recognition *In Proceedings of the IEEE conference on computer vision and pattern recognition* 770-778
[15] Rajan A, and Ramesh GP 2016 Glaucoma detection in optical coherence tomography images using undecimated wavelet transform, *Research Journal of Pharmaceutical Biological and Chemical Sciences* 7 878-85
[16] Shahada SAA, Hreiji SM, Atuduand SI, Shamsudheen S 2019 Multilayer Neural Network Based Fall Alert System Using IOT *International Journal of MC Square Scientific Research* **11** 1-15
[17] Agurto C, Murray V, Barriga E, Murillo S, Pattichis M, Davis H, Russell S, Abràmoff M and Soliz, P 2010 Multiscale AM-FM methods for diabetic retinopathy lesion detection *IEEE transactions on medical imaging* **29** 502-512
[18] Satyarthi D, Raju BAN and Dandapat S 2006 Detection of diabetic retinopathy in fundus images using vector quantization technique, *Annual IEEE India Conference* 1-4
[19] Barriga ES, Murray V, Agurto C, Pattichis M, Bauman W, Zamora G and Soliz P 2010 Automatic system for diabetic retinopathy screening based on am-fm, partial least squares, and support vector machines *IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 1349-1352
[20] Prabu S, Balamurugan V, and Vengatesan K 2019 Design of cognitive image filters for suppression of noise level in medical images *Measurement* **141** 296-301