A Transfer Learning Approach for Diabetic Retinopathy and Diabetic Macular Edema Severity Grading

Nida Nasir¹, Neda Afreen²*, Ranjeeta Patel³, Simran Kaur¹, Mustafa Sameer⁴

¹ Research Institute of Science and Engineering (RISE), University of Sharjah, Sharjah 27272, UAE
² Department of Computer Engineering, Jamia Millia Islamia, New Delhi 110025, India
³ School of Electrical Engineering, KIIT Deemed to be University, Bhubaneshwar 751024, India
⁴ Indian Institute of Information Technology and Management, Kerala 695581, India
⁵ Department of Electronics and Communication Engineering, National Institute of Technology Patna, Bihar 800005, India

Corresponding Author Email: afreenneda1@gmail.com

https://doi.org/10.18280/ria.350608

Received: 20 October 2021
Accepted: 10 December 2021

Keywords: diabetic macular edema, diabetic retinopathy, transfer learning, ResNet50

ABSTRACT

Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) are complication that occurs in diabetic patient especially among working age group that leads to vision impairment problem and sometimes even permanent blindness. Early detection is very much needed for diagnosis and to reduce blindness or deterioration. The diagnosis phase of DR consumes more time, effort and cost when manually performed by ophthalmologists and more chances of misdiagnosis still there. Research community is working on to design computer aided diagnosis system for prior detection and for DR grading based on its severity. Ongoing researches in Artificial Intelligence (AI) have set out the advancement of deep learning technique which comes as a best technique to perform analysis and classification of medical images. In this paper, research is applied on Resnet50 model for classification of DR and DME based on its severity grading on public benchmark dataset. Transfer learning approach accomplishes the best outcome on Indian Diabetic Retinopathy Image Dataset (IDRiD).

1. INTRODUCTION

In the medical care field, therapy of disease is more viable when detected at a beginning phase. Diabetes disease occurs when level of glucose rises due to insufficient insulin. Diabetes influences retina, kidneys, heart and nerves. According to the report of WHO diabetes affected 423 million people in 2014 across worldwide which may go to 700 million in 2050 [1].

Diabetic Macular Edema (DME) and Diabetic Retinopathy (DR) are complication that occurs in diabetic patient especially among working age group. DR is an intricacy of diabetes that leads to expand the retina veins and to spill liquids and blood. DR leads to vision impairment problem [2]. Diagnosis of DR at early stage might be founded on practical conversion in electroretinography (ERG), retinal vein type and retinal blood stream [3]. The probability of DR increases in diabetes patients who experience the ill effects of the disease for an extensive stretch. The screening of retina at regular time is important for patients to analyze and for DR treatment at an initial phase to mitigate the possibility of vision loss [4]. DR is distinguished by the presence of various kinds of sign on a retina picture. These signs are listed as [5] Haemorrhages (HM), Soft exudates, hard exudates (EX) and Microaneurysms (MA) in Table 1.

Detection is very crucial for the diagnosis at early stage and to reduce blindness or deterioration. Manual diagnosis takes more time and is costly. On the other hand, automated detection makes the process easy. To address such issue and to maximize the result, this study explored the transfer leaning based Resnet50 network for DR and DME detection. The main contribution of this study is to implement transfer learning technique to get improved DR and DME classification result based on severity grading on the publicly available dataset as transfer learning use less computation and time as compared to traditional machine learning.

Table 1. Lesion and their description

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemorrhages</td>
<td>Enormous mark upon the retina having dimension more than 125 µm</td>
</tr>
<tr>
<td>Microaneurysms</td>
<td>Initial condition of DR shows up like little red round specks upon the retina and size is under 125µm and sharp edge</td>
</tr>
<tr>
<td>Soft exudates</td>
<td>Inflammation of the nerve fibre, having round or oval shape</td>
</tr>
<tr>
<td>Hard exudates</td>
<td>Dazzling yellow mark upon the retina</td>
</tr>
<tr>
<td></td>
<td>Keen edge present in the retina’s external layer</td>
</tr>
</tbody>
</table>

DR is graded into five different stages relying upon the appearance of these signs which are no DR, mild DR, moderate DR, severe DR and proliferative DR [6], that are concisely mentioned in Table 2 [7]. In the stage of mild DR patients have at least one MA in which close inspection and monitoring is very important, in moderate stage patients could have hemorrhages or MAs in one to three retinal quadrants and it is essential to visit the doctor every 6 to 8 months, in severe DR patients have intraretinal hemorrhages, venous beading in
two or more quadrants and it should be monitored using both macular OCT and fluorescein angiography to detect any DME and in the proliferative stage patients require immediate referral to a retina specialist. Retina image having some abnormalities and its grading into different stages are illustrated in Figure 1 and Figure 2.

DR detection based on automated methods saves time and cost and shows better proficiency compared to manual analysis. The chance of misdiagnosis is more when performed manually and more effort is also required in this method compared to automated technique. Early detection is crucial for diagnosis and to reduce blindness or deterioration. Ongoing researches in Artificial Intelligence (AI) have set out the advancement of deep learning technique which comes as a best technique to perform analysis and classification of medical images. And here Resnet50 model performs classification of DR and DME based on severity grading on public benchmark dataset IDRiD and achieves better accuracy for training and testing. The proposed challenge design provides an avenue for determining precise DR severity status as well as the potential to deliver precise measures for lesions, which may aid in follow up studies to notice changes in the retinal atlas.

The paper is arranged in the following manner: Section 2 introduces literatures of related work. Section 3 depicts the details of working model methodology, Section 4 shows our dataset along with exploratory outcomes and Section 5 follows conclusions.

![Figure 1. Retina images with grading levels (i) normal, (ii) mild, (iii) moderate, (iv) severe and (v) proliferative DR](image1)

<table>
<thead>
<tr>
<th>DR severity grading</th>
<th>Signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>No DR</td>
<td>Absence of any abnormalities</td>
</tr>
<tr>
<td>Mild DR</td>
<td>Only MA present</td>
</tr>
<tr>
<td>Moderate DR</td>
<td>Signs lying between MA and severe DR</td>
</tr>
<tr>
<td></td>
<td>Can be any of them:</td>
</tr>
<tr>
<td>Severe DR</td>
<td>Number of intraretinal HM in each four quadrants is more than 20.</td>
</tr>
<tr>
<td></td>
<td>Venous beading in more than two quadrants.</td>
</tr>
<tr>
<td>Proliferative DR</td>
<td>Noticeable intraretinal microvascular issue in more than one quadrant neovascularization</td>
</tr>
<tr>
<td></td>
<td>Pre-retinal/Vitreous HM</td>
</tr>
</tbody>
</table>

2. RELATED WORK

Early work on DR depends on manual features to quantify the veins and optic circle and measuring the presence of irregularities, for example, Haemorrhages, Microaneurysms, Hard exudates and soft exudates. After that grading was performed by hand crafted features using different machine learning methods such as support vector machines (SVM), k-nearest neighbor (KNN). Lachure et al. [8] suggested a method for detecting microaneurysms and exudates from retina which is considered as first sign of DR with the help of K-nearest neighbor and support vector machine classifier. After preprocessing, feature extraction using GLCM is performed for further classification. SVM classifier gets better result over KNN. Patton et al. [9] performed exploration and laid out standards for retinal image examination and discussed about the technique for detecting retinal landmarks and signs related with DR. Jordan et al. [10] gives a short prologue to quantitative strategies for the examination of fundus pictures with an attention on recognizing retinal signs and examining of retinal disease based on automatic techniques. Acharya et al. [11] suggested a method for grading DME into stages by applying feature extraction techniques without any segmentation process. Features are mainly shape, size, texture, brightness, color, pixels clusters information in domain of spatial and or transform. Feature extraction is performed using Top-hat filtering, Radon transform, discrete cosine transform and discrete wavelet transform and classification is done through support vector machine and decision tree classifier.

2.1 Deep learning techniques

In deep learning method, network directly learns low level depiction and high level parameters from data directly and lowers the need of human intervention for feature engineering. Hacisofaoglu et al. [12] presented a method based on GoogleNet, AlexNet, ResNet50 and CNN based framework to increase the result of DR recognition in cell-phone based and conventional fundus camera retina pictures. Retraining of these frameworks are performed on various datasets such as Messidor, EyePACS, IDRiD to examine the outcome of deploying images from different group of single, multiple, cross datasets. These techniques achieve high accuracy on independent test datasets. Alyoubi et al. [13] presented the new cutting edge strategies for DR color fundus pictures location and deep learning procedures for classification. Moreover, the DR datasets for the color fundus retina surveyed. Contrast testing issues that need more examination are additionally...
taken by them. Li et al. [14] presented a new method for grading DR and DME jointly using cross disease attention network. It investigates the inward connections between the diseases with just picture level oversight. They designed two attention modules separately to learn disease dependent and disease specific features and then integrate these for grading DR and DME to increase the outcome of grading. They use Messidor and IDRiD dataset for evaluation. Reguant et al. [15] proposed a method where they initially perform visualization using CNN to find the innate image features engaged with the CNN's accountability process. Then, they fundamentally break down those provisions regarding generally known pathologies in particular hemorrhages, microaneurysms and exudates, and other visual segments. And additionally examine different CNNs by thinking about what picture highlights they get at learning time for predicting their clinical importance. The analyses are done on openly accessible fundus EyePACS and DIARETDB1 datasets. While different CNNs bring out predictable characterization results, the pace of getting image features conflict among models can be pretty much higher as 70%.

Here, we opted to classify the most recent papers on the basis of whether or not they used deep learning in the ambience of DR. Researchers can improve the performance of the DR referral system by using artificial intelligence to create more robust diagnosis system that can offer quantitative information for many diseases that meets international clinical relevance requirements. Some of the shortcoming of traditional technique is overcome by machine learning specifically deep learning techniques because of its automatic feature extraction technique. Discussed recent literature concludes that deep learning can be efficiently used for retinal lesion detection and grading.

3. METHODOLOGY

This segment introduces the overall structure of employed deep learning architecture with the help of transfer learning technique. Deep learning (DP) is a subset of Artificial Intelligence (AI) which takes motivation from the structure of human brain. DL has different layers of hierarchy consisting of non-linear processing phases for pattern classification as well as for unsupervised features learning. DL has many applications in medical image analysis including segmentation, image registration, retrieval and classification. DL learns the structure by extracting relevant features from the network using training set images [16]. This learning ability definitely reduces the requirement for engineering special features, but the approach based on full end-to-end training of DL commonly require a critical number of samples. Training is important part of any DL framework where the network wants more than thousand numbers of images to be feed into to learn pixel esteems and edges [17]. However, datasets which are available publicly have not sufficient number of retinal images to train such a large network from beginning, so transfer learning approach comes into picture to utilize the pretrained networks.

Training DL framework is exceptionally required in terms of managing computational resources and data requirement [18]. Training strong and precise model can be little bit difficult regarding clinical data. Medical images are exceptionally difficult to acquire because of the involvement of vital curation, lawful issues and annotation. It is conceivable anyway to use models that are already trained on huge datasets, for example, ImageNet which transfers the acquired information to another model regardless of whether the application fields contrast. Transfer learning does precisely improve the learning part of model by transferring the information from one to another which are already pretrained [19]. Knowledge transfer for medical imaging area from ImageNet [20] helps the network to easily find out the low level features from image. In retinal image dataset, detection of DR and DME can be performed by fine tuning the model which means retraining the model for new task which will increase the overall accuracy and speed in comparison to training the model from beginning [21].

3.1 Network architecture of ResNet50

We use ResNet50 architecture for classification and feature extraction problem. It is one of the famous DL architecture which is used to perform classification task. It is trained on large ImageNet dataset comprises millions of images and classify them into thousand number of classes [22]. ResNet50 have 177 layers and there is introduction of skip connection which takes the input from preceding layer and feed into the succeeding layer without any changes as mentioned in Figure 3. ResNet50 implements 1×1 layer of convolution to lessen the overall computation, which in turns decrease the size of the model and also the numbers of parameters. In transfer learning, the last three layers of pretrained network substitutes by fully connected, softmax and classification layer. First few layers are frozen by fixing the learning rate zero and remaining layers are updated during the training process [23]. The network is retrained on retina image using learning rate 1e-4, batch size of 20 for 100 to 200 epochs depends on image size followed by drop out layer with dropout rate 0.5 and Adam optimizer [24] for optimization. Loss function sparse categorical loss is employed for DR and DME grading and to find particular features of disease [25].

![Figure 3. Transfer learning approach for grading DR/DME](image)

4. EXPERIMENT DATASET, RESULTS AND COMPARISONS

4.1 Dataset (IDRiD)

The IDRiD dataset was made from genuine clinical test acquired at an eye hospital situated in Nanded, a place inside Maharashtra, India. It collects the retinal photograph of a person having diabetes by focus on macula with the help of Kowa VX-10α fundus camera. The images captured by camera have field view of 50° and resolution of 4288×2848 pixels and are stored in jpg format. The final dataset contains 516 images having distinct pathological conditions of 5 DR classes and 3 DME classes [26]. The grading of DR into five classes is based on the disease severity and DME is annotated according to the least distance among the macula and the hard exudates position [27]. There are 413 training images and 103
testing images respectively [28]. The statistics labeling inside IDRiD dataset is given in Table 3. Three kind of ground truth are associated with the IDRiD dataset which are given below:

1. **Annotation on pixel level** - This kind of annotation is used in locating particular lesions inside a picture and to divide and find out regions of interest from data. Dataset comprises of 81 color fundus picture with DR sign and One hundred sixyt four without any sign of DR. The color fundus image are present in .jpg file format and separate binary masks are also there in .tif file format for each lesion type and a special optic disc mask (OD) for all 81 color fundus image. These annotation is important in research for analyzing the computation of lesion segmentation inside the image [29].

2. **DR and DME disease Grading** - It comprises of information that provides the risk factor related to whole image. The clinical specialist graded the whole 516 images having various pathological stage of DR and DME [29]. Grading of all images on DR (0-4) and DME (0-2) severity level is accessible in the CSV record. Training and testing data comprises of 413and 103 images respectively by keeping desired ratio of disease stratification.

3. **Optic disk and Fovea center co-ordinates** - The co-ordinates of fovea center and OD are listed for whole 516 image data and the markup is present as a CSV record.

**Table 3. Statistics of labels in the dataset with label counts and relative value**

<table>
<thead>
<tr>
<th></th>
<th>DR</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>134,26%</td>
<td>18,3,5%</td>
<td>36,7,0%</td>
<td>5,1,5,0%</td>
<td>4,0,8%</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0,0%</td>
<td>0,0%</td>
<td>24,4,6%</td>
<td>4,0,8%</td>
<td>2,0,4%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0,0%</td>
<td>0,0%</td>
<td>140,27,1%</td>
<td>116,22,4%</td>
<td>33,6,4%</td>
<td></td>
</tr>
</tbody>
</table>

### 4.2 Performance metrics

There are numerous performance standard measures for quantifying classification performance of Deep Learning techniques. The widely used standard measures are area under curve (AUC), accuracy, specificity and sensitivity. Accuracy is measured as the number of images in percentage which are correctly classified. Specificity measures the normal images in percentage that classified as normal and sensitivity measure the abnormal image percentage which are classified as abnormal by classifier. AUC draws in form of graph which is plotted between specificity and sensitivity.

\[
\text{Accuracy} = \frac{TN + TP}{TP + FP + FN + TN} \tag{1}
\]

\[
\text{Sensitivity} = \frac{TP}{FN + TP} \tag{2}
\]

\[
\text{Specificity} = \frac{TN}{FP + TN} \tag{3}
\]

where, TP is True Positive which shows the instances disease images which are correctly classify into disease. True Negative (TN) is the instances of non-diseased image which classified under non-diseased. False Positive (FP) is the instances of non-diseased image but classify into diseased while False Negative (FN) is the instance of diseased image but are classify into normal. Confusion matrix is used to demonstrate these instances which is shown in Figure 4.

**Table 4. Model parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Rate</td>
<td>1e-4</td>
</tr>
<tr>
<td>Epochs</td>
<td>100/200</td>
</tr>
<tr>
<td>Optimization</td>
<td>Adam Optimizer</td>
</tr>
<tr>
<td>Loss Function</td>
<td>Sparse categorical loss</td>
</tr>
</tbody>
</table>

**Table 5. Training and testing accuracy of DR and DME**

<table>
<thead>
<tr>
<th>Grading</th>
<th>Training Accuracy</th>
<th>Testing Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR</td>
<td>98.5%</td>
<td>53.4%</td>
</tr>
<tr>
<td>DME</td>
<td>95.8%</td>
<td>74.76%</td>
</tr>
</tbody>
</table>

**Figure 4. Confusion matrix**

**Figure 5. (a) Accuracy curve (b) loss curve**
Training and validation accuracy and loss of DR grading is given in Figure 5.

Training and validation accuracy and loss of DME grading on 100 epoch is given in Figure 6.

![Figure 6](image-url)

**Figure 6.** (a) Accuracy curve (b) loss curve

5. CONCLUSIONS

To advance research in clinical area especially for diabetic retinopathy diagnosis, employment of transfer learning for DR and DME classification based on its grading is performed. Diabetic Retinopathy is a complication that occurs in diabetic patient especially among working age group that leads to vision impairment problem and sometimes even permanent blindness. Early detection is crucial for diagnosis and to reduce blindness or deterioration. Ongoing researches in Artificial Intelligence (AI) have set out the advancement of deep learning technique which comes as a best technique to reduce blindness or deterioration. Ongoing researches in Artificial Intelligence (AI) have set out the advancement of deep learning technique which comes as a best technique to

**REFERENCES**


