International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37733

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Literature Survey on Diabetic Retinopathy Classification Using Deep Learning

Alfiya MD. Shaikh[1], Sherica L. Menezes[2]
[1, 2]Computer Engineering Department, Goa Engineering College, Farmagudi, Ponda, Goa, India

Abstract: Diabetic retinopathy (DR) is a medical condition that damages eye retinal tissues. Diabetic retinopathy leads to mild to complete blindness. It has been a leading cause of global blindness. The identification and categorization of DR take place through the segmentation of parts of the fundus image or the examination of the fundus image for the incidence of exudates, lesions, microaneurysms, and so on. This research aims to study and summarize various recent proposed techniques applied to automate the process of classification of diabetic retinopathy. In the current study, the researchers focused on the concept of classifying the DR fundus images based on their severity level. Emphasis is on studying papers that proposed models developed using transfer learning. Thus, it becomes vital to develop an automatic diagnosis system to support physicians in their work.

I. INTRODUCTION

Diabetic retinopathy (DR) is a medical condition that damages eye retinal tissues. Diabetic retinopathy leads to mild to complete blindness. It has been a leading cause of global blindness. Diabetic Retinopathy is not noticeable in the early stages. When DR is discovered, it is often the time when the disease is more serious. So, whether or not diabetics a retinal examination should be performed regularly. Whereas the identification of the disease is not visible to everyone, it requires a doctor with professional skills to diagnose. This brings up another issue the human resources, that as we know vary across different regions. In remote areas, there is likely to be no relevant doctors, which will lead to cases that cannot be detected. While in areas with human resources, there are also examples of different interpretations of the same picture.

Therefore, in summary, we urgently need to use medical image recognition approaches to assist in the judgment of this disease. Once successful, it can greatly reduce the consumption of human resources and provide a powerful reference for doctors to make an accurate diagnosis. The identification and categorization of DR take place through the segmentation of parts of the fundus image or the examination of the fundus image for the incidence of exudates, lesions, microaneurysms, and so on. This research aims to study and summarize various recent proposed techniques applied to automate the process of classification of diabetic retinopathy.

In the current study, the researchers focused on the concept of classifying the DR fundus images based on their severity level. Emphasis is on studying papers that proposed models developed using transfer learning. Thus, it becomes vital to develop an automatic diagnosis system to support physicians in their work.

II. LITERATURE SURVEY

This section constitutes a summarization of previous work done by several researchers to classify diabetic retinopathy based on the transfer learning approach. As we comprehend that the field of deep learning has been continuously evolving moreover architectures based on deep learning are being implemented and participated in the ImageNet competition. All of the following summarized work has used some of these models pre-trained on ImageNet Dataset.

K. Shankar *et al.* in their work titled automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model [1] focus on the detection and classification of DR fundus images based on severity level using a deep learning model. Their proposed method involves various processes namely pre-processing, segmentation and classification. Their proposed method begins with a pre-processing stage in which unnecessary noise in the edges was removed. They believed that the green channel offers better information related to optical nerves and relevant characteristics of the retina. So, the images were segregated into three RGB channels. In the next stage, histogram-based segmentation occurs to extract the useful regions in green colour for further processing. This stage involved two levels. In the first level, the main colours present in the images were recognized and the segmented image regions were constructed from them. The Histogram H was generated in the first phase or level and the complete peaks were identified. In the second level, the merging process occurs based on the sizes of the segmentation region to minimize the number of regions in the segmented image. In this stage, the whole Image *I* pixels were not designated towards the segmented regions. Those *I* pixels that could not be assigned to any segmentation region were dealt with in the next stage of the process.

To combine any voxels peak of H, a second variable τ was utilized. The τ value depends upon the maximal disparity. The standard value of τ=50 is recommended. The whole peaks that were identified were viewed as predominant. To infer a predominant result, the determination of the value of τ is significant. For each image, the unrivalled τ values were found. To decrease too much segmentation and to think about unassigned pixels, blending was performed.

Then, in the final stage, Synergic Deep Learning (SDL) model was applied to classify the DR fundus images to various severity levels. The proposed SDL model includes a set of three major components namely input layer, k DCNN components, and C^2_k Synergic Network (SN). The SN includes a model of a fully connected structure. The SDL model receives n input data that undergoes arbitrary selection from the training set. To maintain the uniform size of the images, they were resized to 224×224×3 using bicubic interpolation.

To initialize each DCNN component (residual network) ResNet−50 was used. These components were then trained and were aimed at computing the group of variables θ that reduces the cross-entropy loss. The justification for the presented SDL model was carried out on the Messidor DR Dataset. This dataset contains 1200 colour fundus images and was categorized into four categories.

The image without any symptoms was considered a healthy retina. The image which had some microaneurysms represents stage 1 whereas the image with some microaneurysms as well as haemorrhages denotes stage 2. The images that indicate more microaneurysms, as well as haemorrhages, were placed under stage 3. The evaluation parameters used to investigate the performance of the presented SDL model were sensitivity, specificity, and accuracy. The model achieved a sensitivity of 0.9854 and a specificity of 0.9938 for the Messidor dataset.

R. Patel and A. Chaware [2] developed a model to classify Diabetic Retinopathy fundus images based on a transfer learning approach using the MobileNetV2 model. MobileNetv2 was used for extracting features from the given set of retina images. In this work, the Image dataset was taken from the Kaggle APTOS 2019 Blindness Detection competition. In the pre-processing step, they used image padding to preserve the information present at the border pixels of the images.

MobileNetv2 has 155 layers including the classification layer. Initially, they loaded the MobileNetv2 network excluding the top classification layer. The model was then customized by adding the GlobalAveragePooling layer and softmax classifier layer on the top of the base model. They froze the convolutional base, thereby preventing the weights from being updated during training.

To further improve the performance of the model the weights of the top few layers of the base network were fine-tuned. From the 87th layer of the MobileNetV2, they unfroze the layers. The model was compiled with categorical_crossentropy loss and Adam optimizer with a learning rate of 0.00002. Their proposed model attained training accuracy improvement from 70% to 91% and validation accuracy increased from 50% to 81% after fine-tuning.

The authors of [3] presented a model based on deep learning using VGG-16 as a pre-trained network for fine-tuning to classify the fundus images of the retina into the five levels of severity. The dataset used was from the Kaggle competition which contains 35,126 images of both the left and right eyes [12]. The dataset was obtained from EyePACS which is a free platform for screening DR. Each image was first pre-processed by normalizing, centering, and cropping it to a size of 512 x 512 pixels.

To mitigate the huge imbalance in their dataset, they augmented the dataset by randomly rotating images between 0 to 360 degrees, flipping them horizontally or vertically, and changing their brightness, contrast, hue, and saturation levels. Before the augmentation step, however, the dataset was split into training and test sets; each test set had 200 images for each class. The remaining 34,126 images were again split into training and validation sets with 100 images for each class. In the end, augmentation was done on the remaining images to create ~100,000 images, wherein the classes had approximately equal numbers of training examples.

During the training process, three convolutional layers in the last block (Block-5) were made trainable. The three dense layers in the VGG-16 model were removed and replaced by a dense layer with 128 neurons and a rectified linear unit (RELU) activation function. The model was executed either for 8 hours (30 epochs) or 8 consecutive epochs wherein no improvement in the validation loss invoked the Early Stopping mechanism. They used Learning Rate Scheduling wherein if there is no reduction in validation loss for 3 epochs, the learning rate was reduced by a factor of 10.

To evaluate the performance of the model, the values of precision, recall (i.e., sensitivity), F1-score, specificity, and false-positive rate were computed. Their proposed model attained area under the curve (AUC) as 0.80, sensitivity as 80%, and cross-entropy specificity as about 65% and accuracy as 74%.

Y. Wu et al. [4] introduced a model in which an autonomous network was used for diabetic retinopathy recognition based on the migration learning approach. The dataset used in the paper was obtained from Kaggle's official website and was identical to the dataset used in [3]. The dilemma with the diabetic retinal data used in this paper was that the images in each of the classes were extremely unbalanced, and if the class imbalance is severe, it will cause a hindrance in the learning process.

To overcome the issue of imbalance in the data, they employed data enhancement techniques. The method used in this paper for data enhancement is "oversampling" the data category or class with a small number in the original data set. The chosen data enhancement tool is Augmentor, which amplifies the original image. Data was augmented using operations random mirror flipping, random cropping, random rotation, random contrast, random chromatic aberration, and random brightness.

The platform used in this paper is Keras which is a high-level neural network API written in pure Python based on Tensorflow, Theano, and CNTK backends. The approach taken in this research is Feature-based Transfer Learning. They have used pre-trained models such as VGG19, InceptionV3, Resnet50 trained on the ImageNet dataset. The pre-trained model was used for the feature extraction task wherein corresponding structures and weights were directly applied to the dataset.

The training method used in this paper was to import the pre-trained models, replace the output layer, that is, the softmax classification layer, and then apply the remaining network to a new data set as a fixed feature extractor. After this training step, the weights of some layers starting from the model were kept unchanged, and the remaining layers were retrained to obtain new weights.

The experimental results show that the accuracy rate becomes higher and higher with the complexity of the network structure, so the performance of the InceptionV3 network was the best. It was also observed that for the same InceptionV3 network, the accuracy rate increased with the increase in the number of training epochs, indicating that it is necessary to increase the number of training epochs appropriately. The classification accuracy achieved by this method was 0.60.

Hathwar and Srinivasa [5] explained a deep learning solution for automated grading of DR from retinal fundus images. They employed two Convolutional Neural Network (CNN) architectures Inception-ResNet-V2 [8, 9] and Xception [10] that have achieved state-of-the-art performance in image recognition tasks. These two networks were trained to compare their performance for the task of grading DR severity. They conducted their experimentation on the EyePACS dataset which includes 35,124 images with sufficient variability like the varying resolution of images, the artifact(s) in the images, out-of-focus shots, overexposed and underexposed images.

They also used the IDRiD (Indian Diabetic Retinopathy Image Dataset) that contains 413 annotated images. It is the first dataset representative of the Indian population. The images were captured and graded by retinal specialists at an Eye clinic in Nanded, Maharashtra, India. Training on one dataset and testing on another show the invariance of proposed models to the specific datasets.

The EyePACS dataset contains fundus images with varying resolutions which might introduce overhead during the training process, thus all images were resized to 512 x 512 pixels with the FOV (visible region of the fundus image) centered in each image. EyePACS and IDRiD datasets also exhibit a large variation in lighting. Therefore they employed local average magnitude subtraction to overcome the variation in lighting between images of these two datasets. As a result, the local average color gets mapped to 50% grey. Also, to remove the boundary effect, the FOV was clipped to 85% diameter of the original FOV.

In their presented model they used the 35,124 images from EyePACS for training each of the models and the 413 images from IDRiD for evaluation. A separate subset of EyePACS with 1000 images was utilized as the validation set. This validation set was used to fine-tune the learning rate if the validation loss did not drop at the end of few epochs. The training was stopped when the cross-entropy loss on this smaller dataset plateaued.

The dataset they chose was heavily skewed since 73.5% (25808 images) were annotated as grade 0 with the remaining 9316 images distributed among the 4 severity grades. To resolve the issue they utilized real-time data augmentation, in which the original data was synthetically modified to create new data for training. Random transformations employed to generate synthetic data in real-time were:

1) *Rotation:* This makes the model more robust as Retinopathy lesions can appear in any orientation.
2) *Zooming:* Aids in learning scale-invariant features.
3) *Shearing:* Synthetically generates lesions of different shapes.
4) *Flipping:* Ensures positional invariance of Optic Disc and other retinal features in determining DR severity.

The different model parameters used in this paper are as follows: Adaptive moment estimation (Adam) optimizer to recompute the network weights after every iteration, learning rate α initialized to 0.0003, exponential decay rates for moment estimates, and $\beta1$ and $\beta2$ initialized to 0.9 and 0.999 respectively. Also, if the cross-entropy loss did not improve on the validation set after few epochs of training, the learning rate was decreased by a factor of 0.8 with a lower bound of 0.00005 sets on the learning rate.

The training stopped when validation loss did not improve for several epochs, indicating convergence of the model. They employed early stopping mechanisms in order to prevent overfitting and to improve the generalization of the CNN models. That is the training process stopped when the training accuracy exceeded 90%. Their best-performing model was Xception that achieved a sensitivity of 94.3% and specificity of 95.5% and quadratic weighted kappa score (k^2) of 0.88 for grading DR severity.

In the medical imaging field, it is difficult to have an adequate amount of image datasets to train a Deep Convolution Neural Network from scratch. Training a network from scratch requires not only extensive computational and memory resources but also large datasets. Training a network with a small dataset quite often leads to overfitting problems. Transfer learning is a solution for this. There are several models pre-defined on the ImageNet dataset having good top-5 accuracy. Some of these models have been studied in this paper for the same problem statement that is severity classification of diabetic retinopathy using deep learning. The future work of this paper is to study more models and implement a model with the best top-5 accuracy for the ImageNet dataset.

REFERENCES

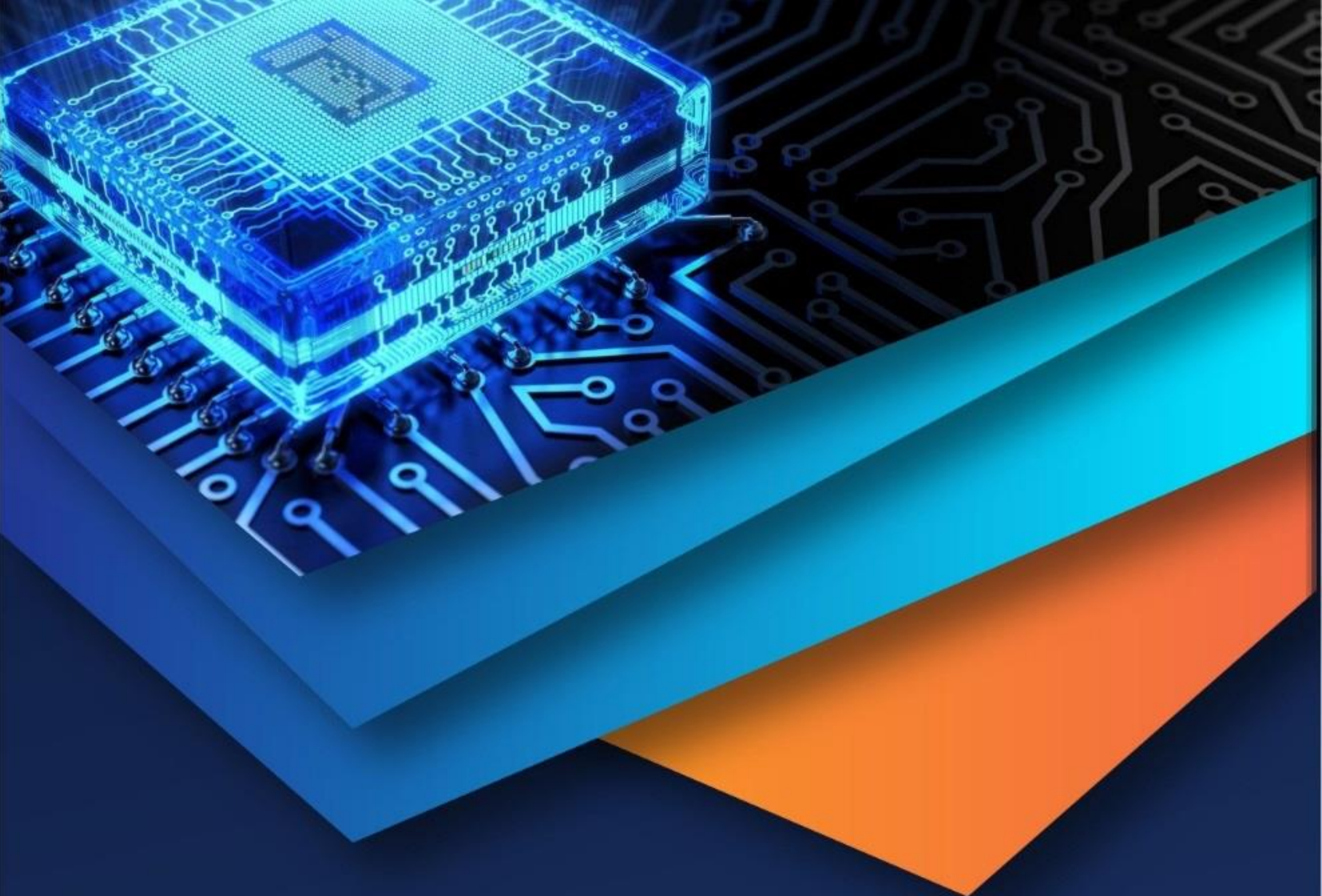
[1] K. Shankar, Abdul Rahaman Wahab Sait, Deepak Gupta, S.K. Lakshmanaprabu, Ashish Khanna, Hari Mohan Pandey, “Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model”, Pattern Recognition Letters, Volume 133, ISSN 0167-8655, 2020, Pages 210-216, DOI: 10.1016/j.patrec.2020.02.026.

[2] R. Patel and A. Chaware, "Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-4, DOI: 10.1109/INCET49848.2020.9154014.

[3] N. B. Thota and D. Umma Reddy, "Improving the Accuracy of Diabetic Retinopathy Severity Classification with Transfer Learning," 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS), 2020, pp. 1003-1006, DOI: 10.1109/MWSCAS48704.2020.9184473.

[4] Y. Wu and Z. Hu, "Recognition of Diabetic Retinopathy Based on Transfer Learning," 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 2019, pp. 398-401, DOI: 10.1109/ICCCBDA.2019.8725801

[5] S. B. Hathwar and G. Srinivasa, "Automated Grading of Diabetic Retinopathy in Retinal Fundus Images using Deep Learning," 2019 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 2019, pp. 73-77, doi: 10.1109/ICSIPA45851.2019.8977760.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY