International Journal of Research in Engineering, Science and Management
Volume-3, Issue-8, August-2020
journals.resaim.com/ijresm | ISSN (Online): 2581-5792 | RESAIM Publishing

Detection of Diabetic Retinopathy Using Deep Learning Techniques

Jeevika Rai[1*], K. A. Anagha[2], K. Bhoomika[3], M. K. Prakruthi[4]
[1,2,3]*Student, Department of Computer Science and Engineering, SJB Institute of Technology, Bengaluru, India*
[4]*Assistant Professor, Department of Computer Science and Engineering, SJB Institute of Technology, Bengaluru, India*
*Corresponding author: jeevikarai13@gmail.com

Abstract*: Diabetic retinopathy is one of the most compromising complexities of diabetes that can lead to vision impairment and even irreversible blindness if left untreated. For the treatment to be a success early detection is one of most important challenges. Unfortunately, physical diagnose by trained eye physician would require large amount of time to diagnose this disorder individually furthermore, the specific recognizable proof of the diabetic retinopathy stage can be famously precarious from the fundus pictures. Convolutional neural systems (CNN) have been applied effectively in numerous neighboring activities and for discovery of diabetic retinopathy itself. In this paper, we propose an automatic deep-learning-based method for identifying the phases of diabetic retinopathy from the human fundus images. Here, the use of three models i.e. VGG, DenseNet and EfficientNet for the comparison of the efficiencies. Our system utilizes CNN alongside denoising to distinguish highlights like micro-aneurysms and hemorrhages on the retina. We prepared this system utilizing a top of the line GPU on the openly accessible Kaggle dataset utilized for preparing and approval. The standard exactness measurements and quadratic weighted Kappa score is utilized to assess grouping capacity of utilized models we utilized that is determined between the anticipated scores and scores provided in the dataset. The finest tested model achieving an accuracy approximately 97% in detecting the retinopathy and assessing its stage is the DenseNet model with size 300x300. Moreover, system obtained decent Kappa scores for EfficientNet, DenseNet and VGG models respectively over a dataset of 3500 images. The accomplished outcomes demonstrated that deep learning algorithms can be effectively utilized to take care of this difficult issue.

***Keywords*: Convolutional neural network, Deep learning, DenseNet, Diabetic retinopathy, Kappa score.**

1. Introduction

On the previous researches there is a scope for the design of classifier to detect the type of diabetic retinopathy, this provide a better and more reliable results for the patients, so that more patients can be diagnosed and cured. In line with this, diabetic retinopathy identification is very useful in encouraging good quality in diabetic retinopathy diagnosis. There is a requirement for automated in acknowledgment of diabetic retinopathy frameworks so that the noise during diagnosis and treatment can be minimized. Therefore, this project will initiate a model for diabetic retinopathy detection which is consistent, efficient and cost effective by exploring the technology of image processing techniques. Diabetic retinopathy is one of the most compromising intricacies of diabetes that prompts perpetual visual deficiency whenever left untreated. One of the fundamental difficulties is early detection, which is significant for treatment achievement. Tragically, the specific recognizable proof of the diabetic retinopathy stage is famously precarious and requires master human understanding of fundus images. Rearrangements of the detection step is urgent and can help a huge number of individuals. Convolutional neural networks (CNN) have been effectively applied in numerous adjoining subjects, and for detection of diabetic retinopathy itself. In any case, the significant expense of huge marked datasets, just as irregularity between various specialists, block the presentation of these methods.

Over the most recent couple of years, deep learning (deep CNN) have made surprising accomplishments in a lot of computer vision and image classification outperforming all the image analysis methodologies. It can model high level of data abstraction relative to specific prediction task. A multi-layer network is developed for automating feature designs in CNN. Each layer in the deep architecture performs a non-linear transformation on the outputs of the previous layers, so that the data is represented in the form of hierarchy of features from low-level to high-level. Enormous CNNs are utilized to effectively handle profoundly complexed image recognition, image classification in medical imaging too. Most are used in the field of medical applications, Eg.: MRI analysis, dermoscopic and standard image analysis. One of the chief points of interest of deep learning algorithm is a capacity to learn portrayal of exceptionally dimensional information. As it were, deep neural algorithms can learn separate data features during the training period and to join this information in a type of weights and biases the neural network parameters. Therefore, with just a little pre-processing deep neural network can operate on fundus images, like cropping or resizing or little alterations done on the images of the datasets.

In this paper, we propose an automatic deep- learning-based technique for diabetic retinopathy stage detection from the available fundus images. The main objective of this project is to "classify the images provided as input into the class or type of diabetic retinopathy using the techniques of Deep Learning" and also produce the result stating the patient has which class of diabetic retinopathy. Also, to carry out optimization study for arriving at an efficient algorithm and to evaluate and validate the performance of the proposed model.

2. Related Work

A ton of work has been finished by specialists in the recognition of Diabetic Retinopathy, contingent upon their examination territory and their field of intrigue. The work identifying with clinical sciences and AI show that scientists have executed a lot of machine leaning models, but the comparative study for deep learning models are still lacking where diabetic retinopathy is concerned. Here the work done is a way to deal with explore different learning profound learning strategies for diabetic retinopathy location while additionally considering the outcomes and discoveries for different other AI techniques for DR

The paper proposed by Kulwinder Mann and Sukhpreet Kaur [1] utilized the ideas of ANN to recognize the ailment a t a beginning phase by portioning of veins in the retina in their paper. This was used as a base paper for further research. The author showed up at an outcome that utilizing ANN could get great outcomes and used for detection in an early stage but they used only certain images retrieved from the database.

The paper [2] "Automated detection of diabetic retinopathy in digital retinal images by Usher, D & Dumskyj, M & Himaga, M & Williamson, Tom & Nussey, S & Boyce, J." proposed the use of ANN. After pre-pre-processing of images characterization was a two-stage process containing grouping of every applicant injury as obvious lesion or noise utilizing an artificial neural system and order of pictures and patients as ordinary or unusual utilizing scientific standards. Excluding ungradable images, sensitivity and specificity for detection of any retinopathy on an image by image and a patient by patient basis using the 773 evaluation patients are summarized as 70.8% and 46.3%. The features detection has been exceptionally powerful yet gives next to no proficiency of these calculations in clinical practice

In the paper [3] "Convolutional Neural Networks for Diabetic Retinopathy by Harry Pratta, Frans Coenen, Deborah M Broadbent, Simon P Hardinga, Yalin Zhenga" proposes a CNN way to deal with analysing diabetic retinopathy from the computerized fundus pictures and to precisely arrange its seriousness and furthermore use data augmentation method to recognize complex highlights associated with the order. Through the colour fundus pictures to distinguish the nearness and essentialness of numerous little highlights alongside a mindboggling grading framework it requires experiences clinicians which can make it difficult and time-consuming due to which this approach was proposed. This proposed CNN model achieves an accuracy up to 75% and 95% sensitivity based on the dataset of 80,000 images out of which 5,000 were validation images.

[4] A Fully Convolutional Neural Network Base Structured Prediction Approach Towards the Retinal Vessel Segmentation by Avijit Dasgupta, Sonam Singh present the utilization of the advantages of the combination of structured predictions and convolutional neural networks and formulate the segmentation task as multi-label inference. Their proposed model dependent on convolutional neural system accomplishes a more prominent exhibition and altogether beats the condition of workmanship for automatic retinal blood vessel division with an exactness of 95.33% and 0.974 AUC score. They assessed the exhibition of the proposed model on an openly accessible DRIVE [5] dataset.

The paper [5] Automatic Screening of Diabetic Retinopathy Images with Convolution Neural Network Based on Caffe Framework by Yuping Jang, Huiqun Wu and Jiancheng Dong aimed on using CNN to automatically screen Diabetic retinopathy fundus images. Histogram equalization and Data Augmentation were the pre-processing techniques mentioned in this paper. Caffe Framework were used for construction and training of the CNN model used here. The model was trained used 8,626 images among which 1,925 fundus images were used for validating the model into patients having DR and patients that are non-DR ones with an accuracy of 75.70% on 1,925 test fundus images.

In the paper "Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy Using Fundus Image paper" by Kele Xu, Dawei Feng and Haibo Mi [6] center around utilizing Deep convolutional neural systems for pictures grouping for diabetic retinopathy fundus pictures. Since the deficiencies of little picture datasets have been perceived, data augmentation technique which include translation, stretching, rotation, flipping using label-preserving transformation is practiced to artificially enlarge the dataset lessening the overfitting on the picture information and furthermore increment the presentation of the model created. The model beat the outcomes got by utilizing classical old style draws near by obtaining an accuracy of 94.5%.

In the paper "Diabetic Retinopathy Detection via Deep Convolutional Networks for Discriminative Localization and Visual Explanation" by Zhiguang Wang & Jianbo Yang [7] a deep learning model center around the visual-interpretable feature which is achieved by adding regression activation map (RAM) on the convolutional network after the global averaging pooling layer in the above mentioned model. This model was evaluated on the Kaggle dataset consisting of 35126 images split into training and validation datasets based on 9 to 1 ratio where the results are submitted to Kaggle obtaining a Kappa score performance measure 0.85034 better than the benchmark.

3. Design and Implementation

A. *System design*

On the previous researches there is a scope for the design of classifier to detect the type of diabetic retinopathy, this provide a better and more reliable results for the patients, so that more patients can be diagnosed and cured. In line with this, diabetic retinopathy identification is very useful in encouraging good quality in diabetic retinopathy diagnosis. There is a requirement for mechanized in acknowledgment of diabetic retinopathy systems so that the noise during diagnosis and treatment can be minimized. Therefore, this project will initiate a model for diabetic retinopathy detection which is consistent, efficient and cost effective by exploring the technology of image processing techniques.

Diabetic retinopathy is one of the most undermining confusions of diabetes that prompts perpetual visual impairment whenever left untreated. The challenge in diabetic retinopathy is the detection of it in its early stages which is the key for treatment achievement. The detection of definite phase of diabetic retinopathy (DR) is precarious and needs human master translation of the pictures of the eyes. Improvement of the discovery step is urgent and can help a large number of individuals. Convolution neural systems (CNN) have effectively been applied in numerous contiguous subjects, and for finding of diabetic retinopathy. In any case, the significant expense of large named datasets, just as irregularity between various specialists, hinder the presentation of these techniques. This paper focuses on developing a robotized model for the identification of diabetic retinopathy and order the images given as input to a type of diabetic retinopathy it belongs to.

As shown in the below figure 3.1, the work flow starts with the data set collection. In first step input images are collected, that is, the data set from Kaggle repository. The images in the data set is then pre-processed. Pre-processing is done by data augmentation, that is, various methods like cropping, flipping are applied on the images to increase the number of images. Once the pre-processing is completed the pre-processed images can be either an inculcate(train) image or a examine(test) image. The train set of images is utilized for inculcating(training) the model and the test image set is the input given by the user to test the accurateness of the model.

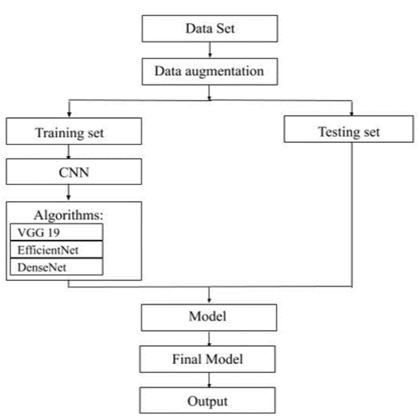


Fig. 1. Workflow of detection of diabetic retinopathy

Train image is used for training or while modelling the dataset and the test image is the input image given by the user to detect the type of note. The input images are passed through a number of steps and layers. The output states the level of diabetic retinopathy. If diabetic retinopathy is not fund, then the output displayed will be 0 and if present the level of the disease from 1-4.

B. Implementation

1) Dataset

The data set for diabetic retinopathy is taken from Kaggle website with 5 class of labels, that is, normal, mild, moderate, severe and proliferative which are labelled from class 0 to 4(0 being normal and 4 being proliferative). The color images taken is resized and cropped to have a maximum size of 1024px. The images present in the Kaggle website is taken from different clinics with a variety of cameras to have different variations in the dataset. The dataset is separated into two sets- the training and testing set. The images with their labels are used for training the model which can then predict the different stages of DR from the test set to a highest possible accuracy.

2) Data augmentation

Even though the dataset taken contains a variety of images it sometimes might not be just enough, thus the data we collect can be used effectively using data augmentation. Data augmentation is a way in which we increment the sum and assorted variety of information. Collecting new data can sometimes be difficult so instead of collecting new data we can rather change the effectively present information. For training models using deep learning requires a very large amount of data and it’s not achievable to gather thousands or millions of information or images, so then for rescue comes data augmentation. The performance of a deep convoluted neural network improves with the amount of training data. Data augmentation is a method of artificially creating new training data from a previously existing training data.

3) Convolution neural network (CNN)

Deep learning is a subset or just in simple words a part of machine learning which mainly works on the idea of counterfeit neural system and the machine learns through unsupervised, semi-supervised or supervised. Deep learning designs are used in various fields like pattern recognition, medical image classification, processing of natural language and so on. The various deep learning design includes convoluted neural network, deep neural network, deep belief networks and recurrent neural network. A neural network comprises of different layers between the information and its yield layer. Depending on the need and the kind of output required the number of layers can fluctuate in the system or network. The various such layers of the network are:

- Convolution layer: This layer is the principle building square of a convolution neural network. The input consists of a set of learnable channels. The input to this layer is a tensor (describes the relationship between various objects in the vector space) image with parameters consisting various properties of the images like number of images, width, height and the depth. After passing the image the output is a feature map that consists of number of images, map width, map height and map channels.

- Pooling: Usually a pooling layer is embedded in the middle of progressive convolutional layer in the CNN. Its function is to diminish the size of the representation, reduce the number of parameters and the calculation which all in turn controls over fitting in the network. The CNN can include global or the local pooling. Global pooling combines all the neurons in the system of a convolutional layer. The local pooling acts on small clusters like 2x2. Pooling can be of two types, that is, the max pooling and an average pooling. The maximum pooling returns greatest incentive among each group of neurons from the past layer. The average pooling restores an average incentive among each bunch of neurons from the past layer.
- Fully connected layer: The neurons from the past layer is associated with all the neurons of the completely connected layer. Usually this layer is the last elements of deep neural network. The output layer, that is, the last layer is the softmax layer or a single neuron called the sigmoid neuron. Depending on the problem being solved, the output can be either binary or a multiclass classification.

C. *Models of CNN*

1. *VGG19:* VGG 19 is a CNN that has 19 deep layers. A fixed image size image consisting RGB value is given to this network. This network uses kernel size of 3x3. Padding is also performed to maintain the resolution of the image. Pooling performed in this network is the max pooling which is trailed by an activation function called the rectified linear unit also known as ReLU. ReLU returns the output directly for all input positive qualities and zero for negative qualities. Due to the activation function this model takes less time to train the model and hence easier and better performance. After this the model contains three fully connected layers. Among the three, two of them are of size 4096 and the third layer consists of 1000 channels. The final layer is the softmax layer that basically gives the output.
2. *DenseNet:* As the name suggests the denseNet is a next step to increase the dep of the convolution neural network. DenseNets take lesser parameter than the CNN. Each layer outputs a feature map from the input that it receives from its previous layers. There is a transition layer between various dense blocks that concatenates all the feature map. Dense block returns various feature maps of same size which then the transition layer concatenates. Even though concatenating can increase the number of input channels a bottleneck property helps decrease the number of parameters and an extra compression can be performed in the transition layer to further reduces the number of feature maps output.
3. *EfficientNet:* “Rethinking model scaling for CNN”. To efficiently scale CNN to increase its performance, certain ways to do this is to increase either of the following-image width, image resolution or the depth of the network. In efficientNet compound scaling is the technique used which aims at scaling all the three parameters, that is, the image width, resolution and the depth of the network. To uniformly increase the network dimensions a compound scaling co-efficient is used. So the compound scaling method used in the efficientNet increases the performance which in turn leads to better model efficiency and accuracy.

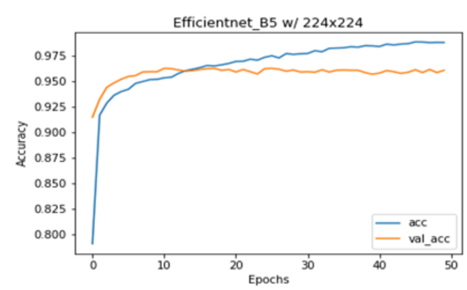
4. Results

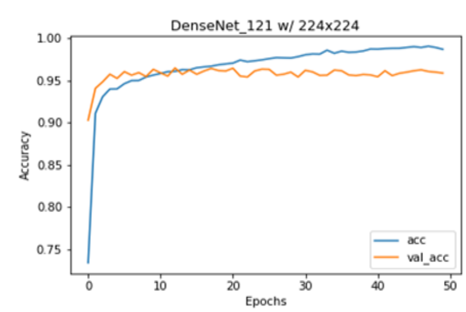
A. *Performance evaluation*

Confusion matrix which is a synopsis of expectation results on a classification problem. The quantity of inaccurate and right expectations is summed up with check esteems and separated by each class. This is the way in to the confusion matrix. The confusion matrix shows the manners by which the model utilized for characterization may be befuddled when it makes predictions. It gives the insight to the mistakes being made by a classifier and also more critically the kinds of errors that are being made.

Accuracy or classification accuracy is the value we get when we divide number of predicitions that are correct by the complete number of predictions. If the classes in the dataset are imbalanced, that is, if any one of the classes in the dataset is more than the others then the accuracy calculation might be inappropriate. Due to the imbalance even the inaccurate model might generate higher accuracies.

The accuracies where first seen for the input images of size 224x244. The accuracies of all the three models for input 224x224 in shown in the fig. 2. DenseNet shows the best accuracy followed by VGG19 and EfficientNet.





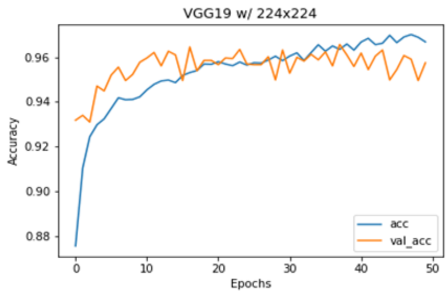


Fig. 2. Accuracies of DenseNet, VGG19, EfficientNet for 224x224

To further increase the accuracy and better performance of the model, eliminate the least accurate one, that is, efficientNet and increased the input image size to 300x300 are tested for denseNet and VGG19. The accuracy is shown in the figure below (fig. 2). This did increase the exactness of the model and the best accuracy was seen for denseNet of approximately 97%.

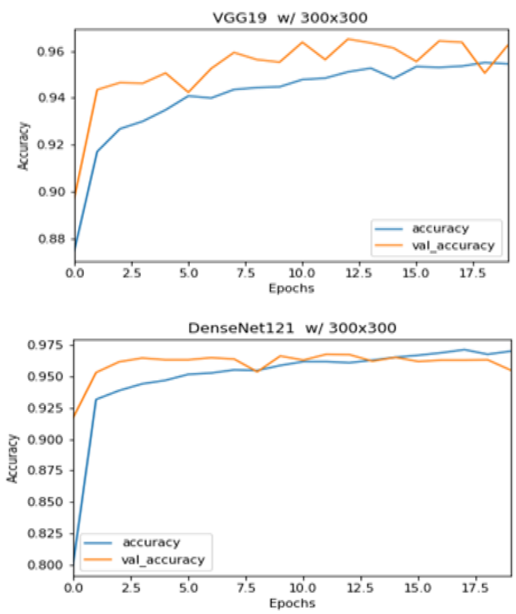


Fig. 3. Accuracies of DenseNet and VGG19 for 300x300

5. Conclusion

This project focuses on determining of the diabetic retinopathy, based on various features such as the micro aneurysms, hard exudates, soft exudates or cotton wool spots, hemorrhages, neovascularization, macular edema. Examination can be made with the assistance of machine learning algorithm, in which we train the framework dependent on the historical backdrop of the pictures put away in the database, and test the current picture to decide, regardless of whether the test picture comes in the class of diabetic retinopathy or not, on the off chance that it does, at that point we decide its stage. An examination can be made with the current frameworks; machine learning decreases the computational time. Hence, the treatment can begin faster.

The outcomes got from this venture shows that the proposed model is quite accurate is recognizing the phases of diabetic retinopathy. The results also showed that denseNet model yields faster and better results when compared to efficientNet and VGG19 model of the CNN. The accuracy obtained by denseNet is approximately 97%.

References

[1] Mann, Kulwinder & Kaur, Sukhpreet. (2017). Segmentation of retinal blood vessels using artificial neural networks for early detection of diabetic retinopathy. AIP Conference Proceedings.
[2] Usher, D & Dumskyj, M & Himaga, M & Williamson, Tom & Nussey, S & Boyce, J. (2004). Automated detection of diabetic retinopathy in digital retinal images: A tool for diabetic retinopathy screening. Diabetic medicine: A journal of the British Diabetic Association. 21. 84-90.
[3] Harry Pratt, Frans Coenen, Deborah M. Broadbent, Simon P. Harding, Yalin Zheng, Convolutional Neural Networks for Diabetic Retinopathy, Procedia Computer Science, Volume 90, 2016, Pages 200-205.
[4] Dasgupta, Avijit & Singh, Sonam. (2016). A Fully Convolutional Neural Network based Structured Prediction Approach Towards the Retinal Vessel Segmentation.
[5] Jiang, Yuping & Wu, Hui-Qun & Dong, Jay. (2017). Automatic Screening of Diabetic Retinopathy Images with Convolution Neural Network Based on Caffe Framework. 90-94.
[6] Xu, Kele & Feng, Dawei & Mi, Haibo. (2017). Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy Using Fundus Image. Molecules.
[7] Wang, Zhiguang & Yang, Jianbo. (2017). Diabetic Retinopathy Detection via Deep Convolutional Networks for Discriminative Localization and Visual Explanation.