Grenze International Journal of Engineering and Technology, Special Issue

Diabetic Retinopathy Identification

Bhaven Naik[1], Sakshi Nihatkar[2], Yash Makwana[3] and Chitra Bhole[4]
[1-4]K J Somaiya Institute of Engineering and Information Technology/Department of Computer Engineering, Mumbai, India
Email: bhavenkumar.n@somaiya.edu, sakshi.nihatkar@somaiya.edu, yash.makwana@somaiya.edu, cbhole@somaiya.edu

***Abstract*—In recent years, there is an increasing enthusiasm in the healthcare research community for the automated diagnosis of various health related diseases. One of the major reasons behind this is that the medical assistance and infrastructure available at hand is unable to match the exponential rate at which the diseases are proliferating. This is where technologies like Deep Learning come into play. Medical Science has various branches, one such being the field of Ophthalmology. It is a branch of medicine that primarily deals with the diagnosis and treatment of eye disorders, which if not treated appropriately and in a timely fashion may lead to loss of vision. One such leading cause of blindness is Diabetic retinopathy, also known as diabetic eye disease. It is a medical condition in which long-term poor control of diabetes mellitus causes damage to the retina. The traditional approach of diagnosis requires the doctors to analyse the colour fundus images of retina in order to determine the level of severity of disease. However, this approach is extremely time-consuming and requires highly-experienced eye specialists for accurate analysis. But the required resources like infrastructure and skilled professionals is not only insufficient but also expensive in developing countries like India. Due to these factors, automated diagnosis is the need of the hour. Although there are existing models based on deep learning; we aim to develop it using Transfer Learning. Thus, our model intends to reduce the diagnosis time, facilitate deployable options and achieve higher accuracy.**

***Index Terms*— Diabetes, Diabetic Retinopathy (DR), Transfer Learning (TL), Deep Learning (DL), Convolutional Neural Network (CNN), Keras.**

I. INTRODUCTION

Diabetes is a growing challenge in India with estimated 8.7% diabetic population in the age group of 20 and 70 years [1]. According to the Indian Heart Association, India is projected to be home to 109 million individuals with diabetes by 2035 [2]. Diabetic Retinopathy is the most prevalent side effect of diabetes which affects nearly 18% of the diabetic population of India. Diabetic Retinopathy is getting frequent among the Indian diabetic population because of limited centres as well as limited ophthalmologists, according to experts India has about 12,000 ophthalmologists (approximately 3500 trained retina specialists) against 60 million diabetic patients facing derivates eye diseases of it. In the traditional approach for the analysis of the fundus image, the image is broken down into four quadrants and in each quadrant a count for each of the following defects is maintained, the defects are microaneurysms, haemorrhage's, neovascularization, no neovascularization.

Grenze ID: 01.GIJET.6.2.508

The count from each quadrant is summed up and tested against a pre-defined medical range too define an appropriate severity to that fundus image, this makes procedure a tedious and time consuming and also a certain level of expertise is a must.We aim to optimize pretrained models through transfer learning, ones that are implemented on the concept of Deep learning (CNN). The model would accept the microscopic images as an input and predict the presence of the severity of the image, which would be then presented to the end user through an intuitive graphical interface.

A. Literature Review

ZhentaoGao; JieLi; JixiangGuo; YuanyuanChen; ZhangYi; JieZhong[1]

Dataset construction was the key aspect of the project, a set of 4476 images labelled with the help of three different clinical departments in Sichuan Provincial People’s Hospital while a slightly different annotation criteria was used from the other available datasets, firstly the progression from severe NPDR to mild PDR is sometimes obscure from Fundus images so, if necessary, a fluorescein angiography is proper for a more precise diagnosis. The second is that the suggested treatments are same for mild and moderate NPDR patients or severe NPDR and mild PDR patients. Thus, it is convenient to group them together for clinical practice. This created a significant difference between the other available datasets as they were just labelled with severity of DR. Colour normalization was then performed on the images after the shape and size was tuned effectively, in this process mean value of each of the values from RGB channels were calculated and truncated above 255 and the values of each pixel were divided by the mean and again multiplied with means of each channel calculated for a set of 1000 fundus images, this lead to an enhancement of the images in the dataset and certain augmentation techniques too were carried out, The base model used is the Inception V3 network which includes the CNN and its pooling, the Inception A/B/C/D/E modules are blocks that combine convolution layers with different kernel sizes. When using this base model directly, the output units in the last SoftMax output layer are reduced to four as we only have four degrees of DR severity. In the end, we concatenate the features of the global average pooling layers from the four Inception-V3 models into one vector and feed this vector to a SoftMax output layer. The three advantages of this strategy. Firstly, more free parameters can give the model better capacity. Secondly, the four models can focus on different locations of a fundus image and learn location related lesions such as diabetic macular edema. Thirdly, since the original Inception-V3 model is designed for images of 299 by 299 pixels, when the input image is enlarged too much, the feature maps inside the model will be enlarged too, thus the depth of the model may no longer be optimal.

Jiawei Xu, Xiaoqin Zhang, Huiling Chen, Jing Li, Jin Zhang, Ling Shao, (Senior Member, Ieee), And Gang Wang [2]

Diabetic retinopathy (DR) is one of the most common microvascular complications and its early detection is critical for the prevention of vision loss. Recent studies have indicated that microaneurysms (MAs) are the hallmark of DR. After extracting the MAs from two successive images, image registration methods can be applied to both images for the detection of new, unchanged, and resolved MAs. To our best knowledge, early attempts on the image pair registration aimed to find the corresponding position in successive images to compare the changed and unchanged MA’s. In our image analysis method, we adopt the Generalized Dual Bootstrap-ICP image registration (GDBICP) to register the MAs candidates from baseline and follow-up images. GDBICP is an automated image registration algorithm; it extracts and matches the key points from image pair. However, the detection of MAs relies on trained clinicians and relatively expensive software. Moreover, manual errors often lower the accuracy of this detection. We also analysed the potential weight of pathological risk factors leading to the MAs turnover, which could provide an alternative guidance for the progression of DR than traditional detection methods. In conclusion, this study provides a novel and non-invasive detection technique for early diagnosis of diabetic retinopathy with a competitive accuracy.

Lei Zhou, Yu Zhao, Jie Yang, Qi Yu, Xun Xu [3]

As a weakly supervised learning technique, multiple instance learning (MIL) has shown an advantage over supervised learning methods for automatic detection of diabetic retinopathy (DR): only the image-level annotation is needed to achieve both detection of DR images and DR lesions, making more graded and de-identified retinal images available for learning. Retinal images coming from mass screening may have various image resolution, illumination and contrast. Normalizing these factors helps promote the deep learning process. However, the performance of existing studies on this technique is limited by the use of handcrafted features. Before making a classification for the whole retinal image, we attempt to detect its DR lesions first. Image patches are regularly extracted from the pre-processed image. Then they are fed into the CNN-based patch-level classifier to estimate their DR probabilities. Specifically, a pre-trained convolutional neural network is adapted to achieve the patch-level DR estimation, and then global aggregation is used to make the classification of DR images. Further, the authors propose an end-to-end multi-scale scheme to better deal with the irregular DR lesions. DR lesions may have an irregular shape and size, and usually a multi-scale scheme is good for handling such variance. Here we propose an end-to-end multi-scale MIL framework in order to achieve a robust estimation of the DR regions and further improve the detection performance. For detection of DR images, they achieve an area under the ROC curve of 0.925 on a subset of a Kaggle dataset, and 0.960 on Messidor. For detection of DR lesions, they achieve an F1-score of 0.924 with sensitivity 0.995 and precision 0.863 on DIARETDB1 using the connected component-level validation.

B. Convolutional Neural Network

A Convolutional Neural Network (CNN/ConvNet) is a class of deep neural networks, an algorithm which is explicitly described to deal with images as an input, assign importance (learnable weights and biases) to various aspects/objects in the image and able to contrast one from the other. The pre-processing required in a ConvNet is significantly subjacent when contrasted with other classification algorithms, making CNN to be distinct, as in the traditional algorithms the filters were hand-engineered. Algorithms used in training CNN are analogous to studying for exams using flash cards. First, you draw several flashcards and check if you have mastered the concepts on each card. For cards with concepts that you already know, discard them. For those cards with concepts that you are unsure of, put them back into the pile. Repeat this process until you are fairly certain that you know enough concepts to do well in the exam, similar is the case with CNN, it takes image as an input passes it through a series of convolution layers with certain filters, then pools the features most desired and continue the process to generate fully connected layers. Activation functions are then applied to classify the image with a certain probabilistic value, which decides the label it is given in the end.

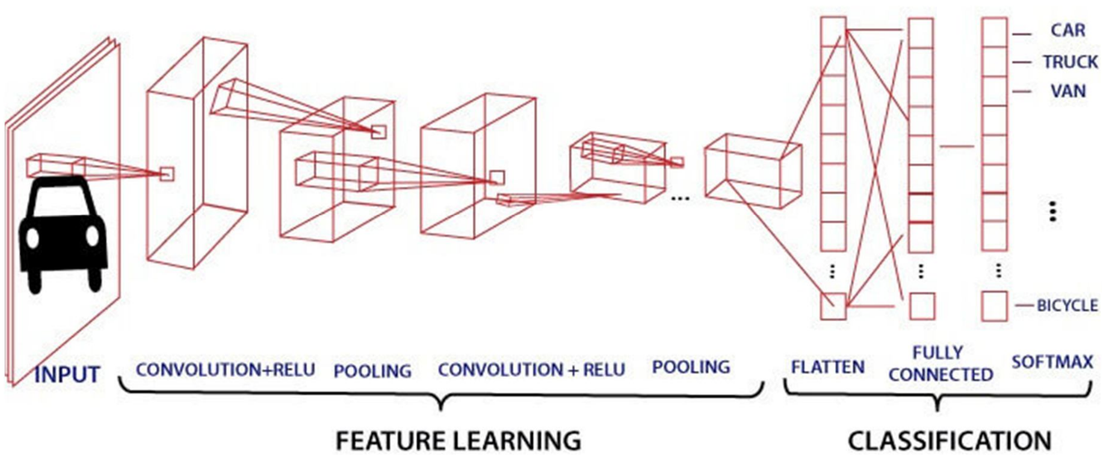

Figure 1. Convolutional Neural Network

The images can be from various spaces that exist such as RBG, Grayscale, CMYK, etc. Once the images reach dimensions of a higher order, the computation at this level gets exhaustive. ConvNet reduces the images into a form which is easier to process, without losing its essence (features), which are important for getting a better prediction, this makes it efficient to make scalable systems with huge datasets.

Convolution is the first layer that receives a signal from the input image, it is a process responsible for capturing the low-level features such as an edge, colour, gradient orientation, etc, the network tries to label the input signal by referring to what it has learned in the past. It is a mathematical operation that takes two inputs like an image matrix and a filter or kernel. The convolution of the image matrix when multiplied with the filter matrix results in the 'Feature Map' matrix. The output signal is not dependent on where the features are located, but simply whether the features are present or not, hence making an object recognizable irrespective of its position. The stacked features map is then passed through a ReLu layer, which removes the negative entries obtained in the matrices, which is to introduce linearity for easier computation purposes.

The Pooling Layer too is responsible for reducing the spatial size of the convolved feature, to decrease the computational power required to process the data through dimensionality reduction, extracting dominant features which are rotational and positional invariant, thus cultivate the process of effectively training of the model. Max Pooling returns the largest maximum value possible from the portion of the image covered by the kernel (rectified feature map), it also performs noise suppressant, it discards noisy activation altogether. In this layer, we flatten the matrix into vector and feed into a fully connected layer like a neural network, with the help of this FC layers, we merge extracted features by the model. We use a SoftMax activation to exhibit the outputs as per the desired categories.

C. Transfer Learning

Human learners appear to have inherent ways to transfer knowledge between tasks. That is, recognize and apply relevant knowledge from previous learning experience when we encounter new tasks, more the relation between the new task and our previous task, the more easily we can master it, similar is the case of Transfer Learning.

Transfer Learning involves the approach in which knowledge gained in one or more source tasks is transferred and used to improve the learning of related target task, in layman terms you have learned to differentiate between rotten potato and fresh potato. But now with your learning experience, you have to differentiate between rotten tomato and fresh tomato. In practice very few people train the entire Convolutional Network from scratch because it is relatively rare to have an entire dataset of sufficient size, instead in Deep Learning it is common to make use of a pre-trained model and leverage the features extracted by that network and then apply these learned features for another target task at hand. Leveraging the experience of these pre-trained models saves time, the computation power and the expense of costlier GPUs required for faster computation. Transfer Learning has two major scenarios, using the ConvNet as a fixed feature extractor and fine-tuning the ConvNet, in the feature extraction scenario, we start with a pe-trained model and only update the final layer weights from which we derive predictions, while in the latter scenario, we start with a pre-trained model and update all of the models parameters for our new task.

II. METHODOLOGY

A. Existing System – Significance

The choice of going ahead with the process was binary regarding the selection of model for the prediction purposes. The options were, either build a convolutional neural network from scrap or use one of the transfer learning scenarios discussed. Building from scrap would require an extensive amount of resources as well as the time to train the network, but this approach would deliver a fantastic accuracy given the model is designed for that specific task, the latter option posed a lesser number of requisites and gave finer results. The Transfer Learning model had already been trained on a large benchmark dataset with millions of classes, but was generalized in nature, also given the time constraint for project, transfer learning was the efficient way to go ahead. ImageNet dataset, which consists of 1.2 million images with 1000 categories, many different ConvNets are trained on this dataset and these are the models which are used as pretrained models, by using the weights that were obtained., the last layer that is the fully connected layer, which is responsible for the drawing predictions from the outputs of the above layers.

B. Proposed System

In the proposed system, we will be providing a diagnosis system which will identify the level of severity from 0 to 4 in a matter of seconds, that will result in early diagnosis. We'll have identification of all severities comprised into a single model which will learn from the CNN. Also, we will provide a diagnostics report which will help to identify number of patients with a particular level of severity, which in turn will help the doctors to analyse which level of severity is popular among the patients and be prepared with the necessary treatment on hand. Our system will learn from thousands of images belonging to the different severities. We will try to even out our existing dataset, since there are more images of a particular level compared to others, either by creating duplicate images or collecting more images for the same.

III. CONCLUSIONS

The proposed system overcomes the drawbacks of the existing system and provides results of more levels of severity. The use of transfer learning provides more flexibility to work with and supports in improving the accuracy of the system. Future work may include identifying patterns on the retina for different severities.

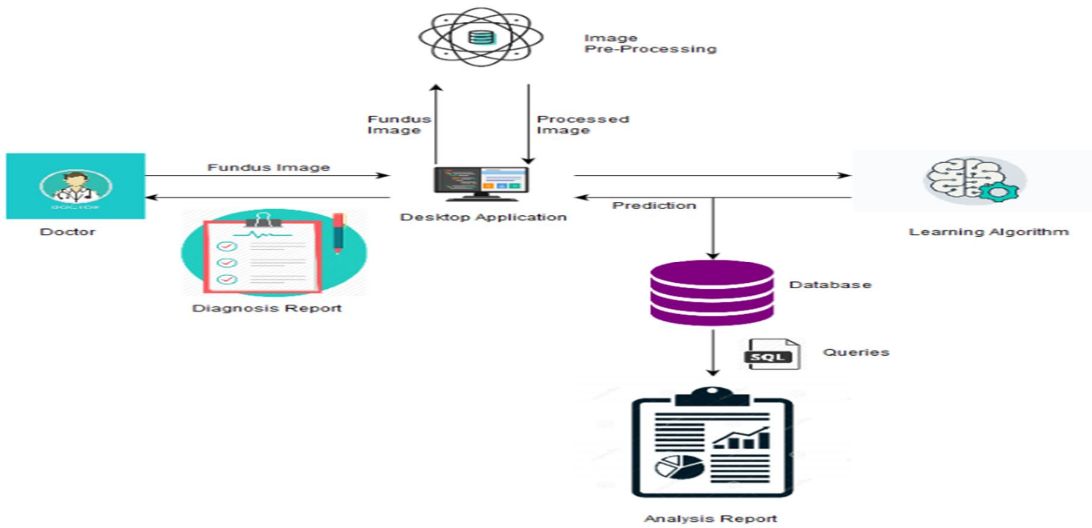

Figure 2. Proposed System

ACKNOWLEDGMENT

We would like to thank Prof. Chitra Bhole (Department of Computer Engineering) for her patience, encouragement, guidance and cooperation provided in the form of unparalleled facilities.

REFERENCES

[1] Zhentao Gao; Jie Li; Jixiang Guo; Yuanyuan Chen; Zhang Yi; Jie Zhong, "Diagnosis Of Diabetic Retinopathy Using Deep Neural Networks," 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)., 2019, pp. 3360-3370.

[2] Jiawei Xu, Xiaoqin Zhang, Huiling Chen , Jing Li, Jin Zhang, Ling Shao, (Senior Member, Ieee), And Gang Wang, "Automatic Analysis Of Microaneurysms Turnover To Diagnose The Progression Of Diabetic Retinopathy," IEEE Access, Volume 6, 2018, pp. 9632-9642.

[3] Lei Zhou, Yu Zhao, Jie Yang, Qi Yu, Xun Xu, "Deep Multiple Instance Learning For Automatic Detection Of Diabetic Retinopathy In Retinal Images," IET Image Processing, Volume 12 4th Issue, 2018, pp. 563-571.

[4] Ramasubramanian Bhoopalan, Selvaperumal Sundaramoorthy, "Efficient Approach For The Automatic Detection Of Haemorrhages In Colour Retinal Images," IET Image Processing, Volume 12 9th Issue, 2018, pp. 1540-1544.

[5] Shuang Yu, Di Xiao and Yogesan Kanagasingam, "Automatic Detection Of Neovascularization On Optic Disk Region With Feature Extraction And Support Vector Machine" 2016 38th Annual International Conference of the IEEE Engineering in Medical and Biology Society (EMBC)., 2016, pp. 1324-1327.

[6] Jack Kanski and Brad Bowling, "Kanski's Clinical Ophthalmology 8th Edition A Systematic Approach", 2015.

[7] https://emedicine.medscape.com/article/1225122-overview

[8] http://www.adcis.net/en/third-party/messidor/