Vol. XIV & Issue No.02 February - 2021
INDUSTRIAL ENGINEERING JOURNAL

DETECTION OF DIABETIC RETINOPATHY USING TRANSFER LEARNING METHOD AND MICRO ANEURYSMS DETECTION USING FASTER R-CNN

Manojkumar S B

Dr. H S Sheshadri

Abstract

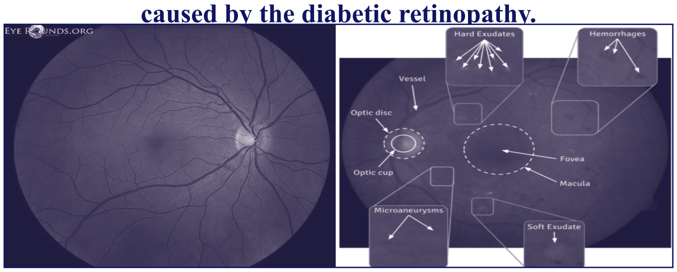
Prolong Diabetic Retinopathy (DR) disease leads for visual loss all over world. Regular screening and an early detection and treatment for this disease is necessary. Among many types and classifications of diabetic retinopathy (DR), Microaneurysms (MA's) perceives during the early stage of diabetic retinopathy and hence it is very difficult process to detection the Microaneurysms (MA's). During the period of one decade, many approaches applied for the detection and identification of the DR using some mathematical and image processing algorithms, features extraction techniques, and artificial neural network classification. In which few are failed during the pre-processing stage, feature extraction stage, vessels extraction and during the classification stage. In this paper it is clearly identified Microaneurysms during the early stage by developing the regions of a retina image to show the specific region of interest in terms of its severity level by collected the large data from kaggle, DIARETDB1dataset and then pre-trained the models and applied to deep learning classifiers using faster R-CNN model and can achieve very high performance on Diabetic detection and getting the performance merit of 82% accuracy.

***Keywords*: *Diabetic Retinopathy (DR), Micro aneurysms (MAs), R-CNN, Transfer Learning*.**

1. INTRODUCTION

Diabetes is a common disease in both developed and developing countries, and nearly 98 million people in India may have this disease around 2030[13]. Prolong years of diabetes affecting to retina leads to Diabetic retinopathy (DR). Basically, DR affects blood vessels of retina. DR is one of the leading cause of blindness for middle age adults in the world today [8], and around half of Indian's with diabetes have this disease to some extent. A well-known challenge for DR is that during the early stages it has no warning sign, even for diabetic macular edema. Thus, it is very much important that DR can be detected before it cross early stage. Unfortunately, in present situation DR detection in early stage is nearly infeasible. It requires an ophthalmologist to analyze digital color fundus photographs of retina, and DR is identified by locating the lesions associated with vascular abnormalities affected by diabetes, figure 1 shows the Optic disc and abnormal findings I the eye fundus caused by the diabetic retinopathy. Even though the present solution of ophthalmologist identifying DR disease is effective, but it is time consuming and chances of false positives and highly relies on the expertise of well training practitioners. In order to overcome this problem, in the past few years considerable efforts have been put on developing an automated solution for DR detection.

Figure 1: Normal and abnormal findings in the eye fundus caused by the diabetic retinopathy.



Among them automated detection of DR is the one which consists of two parts: feature extraction and detection/ prediction algorithm [9]. Feature extraction is mainly focused as standard machine learning algorithms can be directly used as the detection/prediction algorithm. It is effective to some extent but also suffer from several imperfections. Primarily, as reviewed in this Section, the extracted features are all hand crafted features. Therefore, these features are mainly depend on the parameters of the tools used in feature extraction and they are sensitive to the quality of fundus photography, like object view, artifacts, exposedness, noise, out-of-focus, etc. Secondarily, feature extraction may be a separate task instead of embedded into the complete DR detection framework. The above mentioned features extraction methods can be considered as the universal image feature extraction methods that are applicable to most computer vision tasks, and they are not obligate to the specific task, e.g., In this paper DR detection task is to be considered. It is worth noting that color fundus photography is more challenging than the standard scene or object images that most image feature extraction methods were developed based on, since the key signals are often very tiny in fundus photography and they often look separate from noise and artifacts. Thus, these two challenges make it highly desirable to develop a systematically feature representation approach to effectively characterize the novel approach of features that to be with respect to the DR detection.

2. LITERATURE SURVEY

Deng, L and Yang, J et.al., [2][12] the convolutional neural network (CNN) has achieved tremendous success in computer vision area. It can model high-level abstractions in data relative to specific prediction task. In CNN, a multiple layers network is built up for automating feature design. Specifically, each and every layer in deep architecture performs a non-linear transformation on the response of the previous layer, so that the data are represented by a hierarchy of features from low-level to high-level. The main attribute of the CNN is conducting different processing units o f the signals. Then, the deep learning architecture allows multiple layers of these processing units to be stacked, so that this deep learning model can characterize the salience of signals in different scales.

Silberman, N et.al,[8] in CNN, feature extraction and prediction algorithm are merged as a single model. Thus, the extracted features own more discriminative power, since the entire CNN model is trained under the supervision of output labels. Briefly telling, the features extracted by the CNN are task dependent and non-handcrafted.

Navoneel Chakrabarty[6] In this paper, we also adopt CNN as the key predictive algorithm, but aim to develop a more efficient CNN architecture that is particularly useful for large-scale dataset. Specifically, the CNN we built has no fully connected layer and only have convolutional and pooling layers. This setting significantly reduces the number of parameters (fully- connected layers often bring more parameters than convolutional layers in the conventional CNN) and provides better conditions for interpretability of neural network as presented below.

Shaoqing Ren et.al.[7]We show in experiments that with less parameter and no fully-connected layers the proposed CNN architecture can achieve the comparative prediction performance. The key advantage of the proposed network structure is that it can provide a regression activation maps (RAM) of input image to show the contribution score of each pixel of input image for DR detection task. This RAM output, to some extent, somehow mitigates the well-know nun-interpretable shortcoming of CNN as a black box method. We believe that this RAM output make the proposed solution more self-explained and can motivate the practitioners to trace the cause of the disease for every patient.

3. ALGORITHM

Description:

In the first phase we classify the fundus retina image into proliferative, non-proliferative mild, non-proliferative moderate and non-proliferative severe. For classification task we use inception V3 model and perform training using transfer training. In the second phase, we detect Exudates and localize exudates. Detection of exudates and micro aneurysm is achieved by Faster RCNN network.

4. METHODOLOGY

Figure 2: shows the block diagram of diabetic retinopathy classification using transfer learning process.

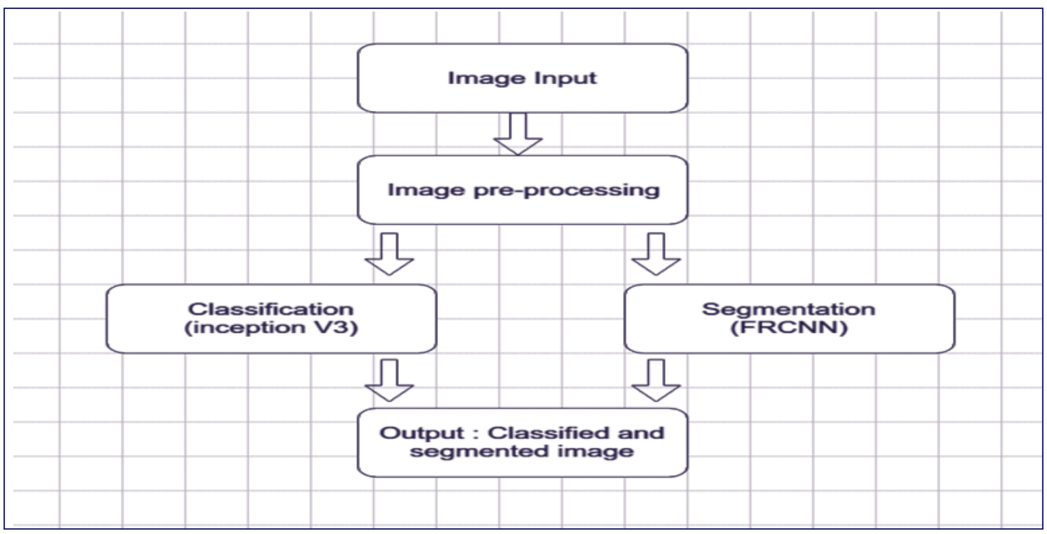
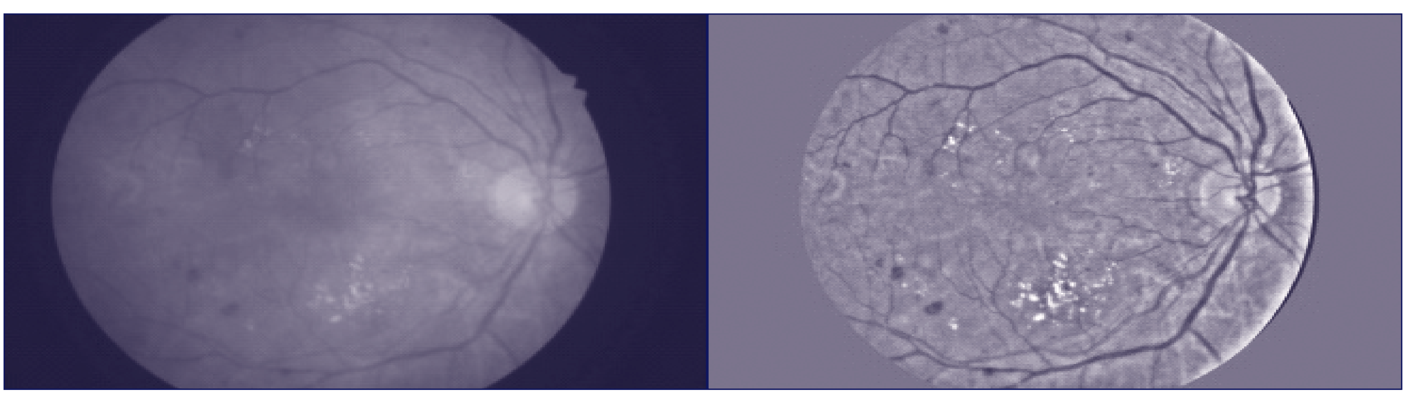


Figure 3 : (a) Input image **(b) Input image after pre-processing**



4.1 Image Pre-processing

The Hierarchical data set format of every images were converted into the pre-processing, augmentation, and training. The Pre-processing stage consists of several steps [3]: The original fundus images were compressed by an Otsu's method to isolate the circular colored image of the retina. Then the fundus images were normalized by subtracting the minimum pixel intensity from each channel and dividing by the mean pixel intensity to represent pixels in the range from 0 to 1. Figure 3 shows the image pre-processing stage.

4.2 Data Pre-processing

- Rescale input images to 300x300 or 500x500 pixels.
- Subtract the local average color, and the local average gets mapped to 50% gray scale.
- Compress the fundus images to 90% to eliminate the "boundary effects".

4.3 Data Augmentation

This step is to increase the variability of training data for enhancing performance.

- Randomly scale the size of images by ±10%.
- Randomly rotated by between 0 and 360 degrees.
- Randomly skew the images by ±0.2.
- For testing, the images are just rotated randomly.

4.4. Classification

4.4.1. Transfer learning

The first phase analyses all the images on disk and calculates and caches the bottleneck values for each of them [7]. The outer layer has been trained to output a set of numbers which is good enough for the classifier which is used to distinguish between the classes it's been taken to recognize. That is it has to be a clear and compact summary of the images, since that image contain enough information for the classifier to make a good choice in a very small set of values. The reason our final layer retraining can work on new classes is that it turns out the kind of information needed to distinguish between all the 1,000 classes in ImageNet[5] is often also useful to distinguish between new kinds of objects.

4.4.2. Training

After completion of transfer learning process, the training process of the top layer of the network will start. We observed a series of outputs steps, each step showing training accuracy, validation accuracy, and the cross entropy. The training accuracy shows how much percent of the overall images used in the current training batch were labelled with the correct class. The validation accuracy shows the precision on a randomly-selected group of selected images from the different set. The key difference is that the training accuracy is based on images that the network has been able to learn from so the network can over fit to the noise in the training data. An accurate measure will be measure by the performance of the network by which it measure its performance on a data set not contained in the training data-- this is measured by the validation accuracy. If the train accuracy is high but the validation accuracy remains low that means the network is over fitting and memorizing particular features in the training images that aren't helpful more generally. The Cross entropy is a loss function which gives a glimpse into how better the transfer learning process is progressing. The training's objective is to make the loss as small as possible, so you can tell if the learning is working by keeping an eye on whether the loss keeps trending downwards, ignoring the short-term noise.

4.4.3. Inception v3 Model

Figure 4: Inception v3 model

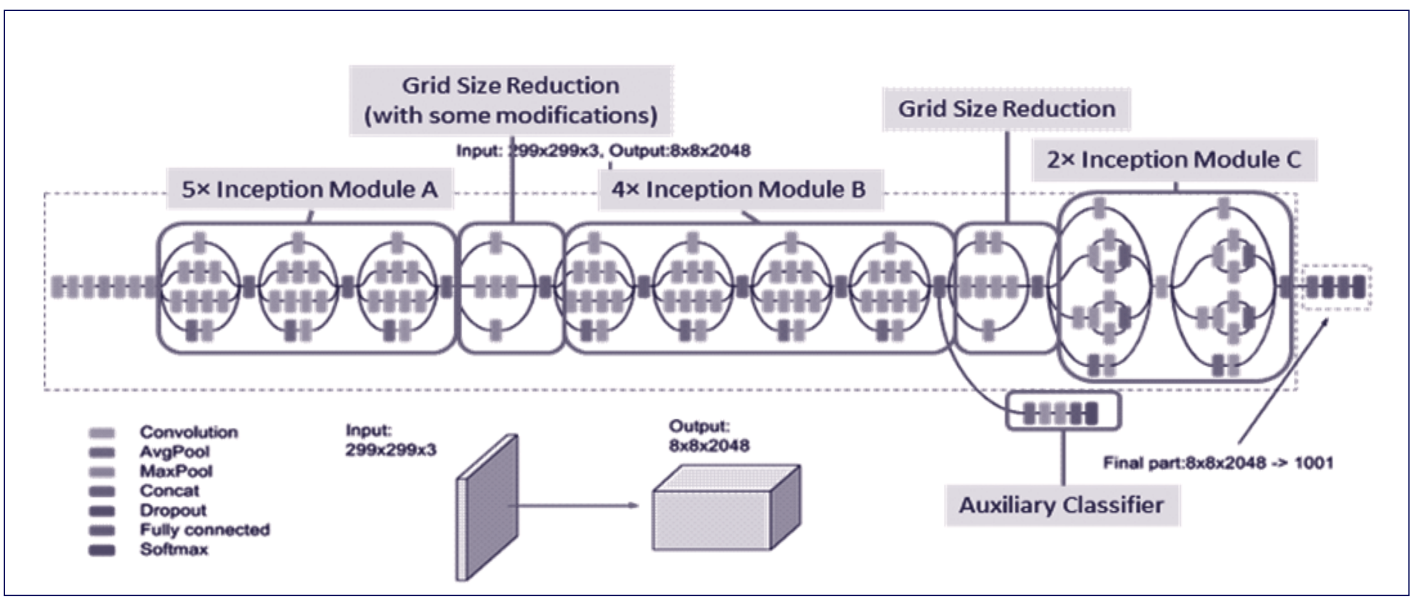


Figure 4 shows the Inception v3 model which is one of the widely-used image recognition model and that has been shown to attain greater than 78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years now a day.

4.4.4. Classification Architecture

Figure 5: Classification architecture

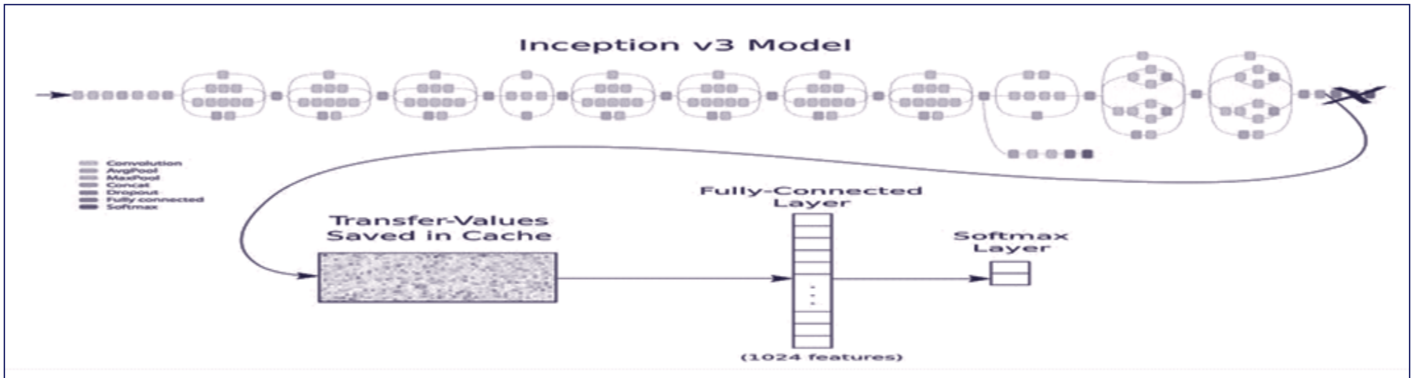


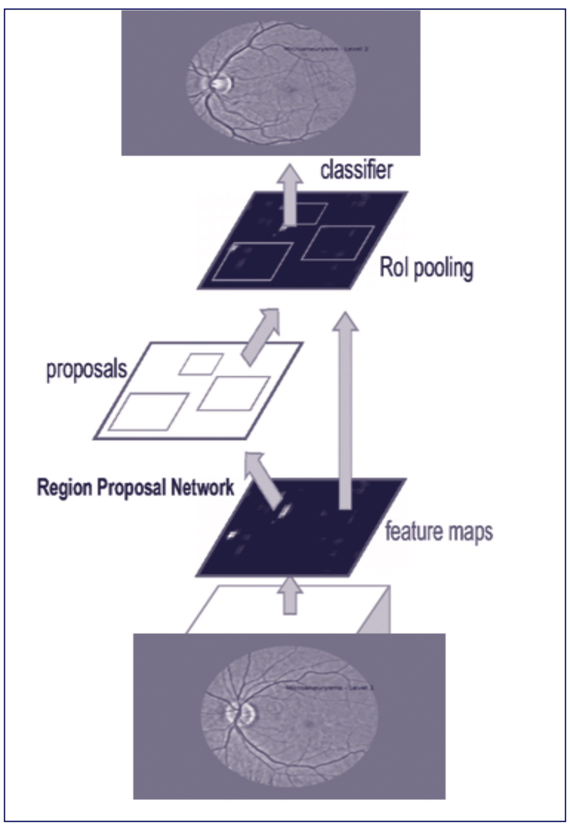
Figure 5 shows the classification architecture of the inception v3 model. Here transfer learning allows you to retrain the final layer of an existing model, resulting in a significant decrease in not only training time, but also the size of the dataset required. One of the most famous models that can be used for transfer learning is Inception V3. As mentioned above, this model was originally trained on over a million images from 1,000 classes on some very powerful machines. Being able to retrain the final layer means that you can maintain the knowledge that the model had learned during its original training and apply it to your smaller dataset, A highly accurate classifications without the need for extensive training and computational power will be obtained as a result of this work.

5. MICRO ANEURYSMS DETECTION AND LOCALIZATION

5.1. Faster RCNN

The Faster RCNN consists of two modules. Primary module is a deep fully convolution network that proposes regions, and the secondary module is the Fast RCNN detector [1] that uses the proposed regions. The complete system is a single, unified network for object detection. With the help of new popular terminologies of neural networks with 'attention' mechanisms, the RPN module clearly informs the Fast RCNN module to where to consider.

Figure 6: Faster RCNN for micro aneurysms localization



5.2. Region Proposal Networks

A Region Proposal Network (RPN) considered an image of any size as an input and outputs as a set of the rectangular object proposals network, each having an objectness score. To generate region proposals, we use a smallest network over the convolutional feature map output by the last shared convolutional layer. This smallest network takes as input an n × n spatial window of the input convolutional feature map.

5.3. Data Preparation

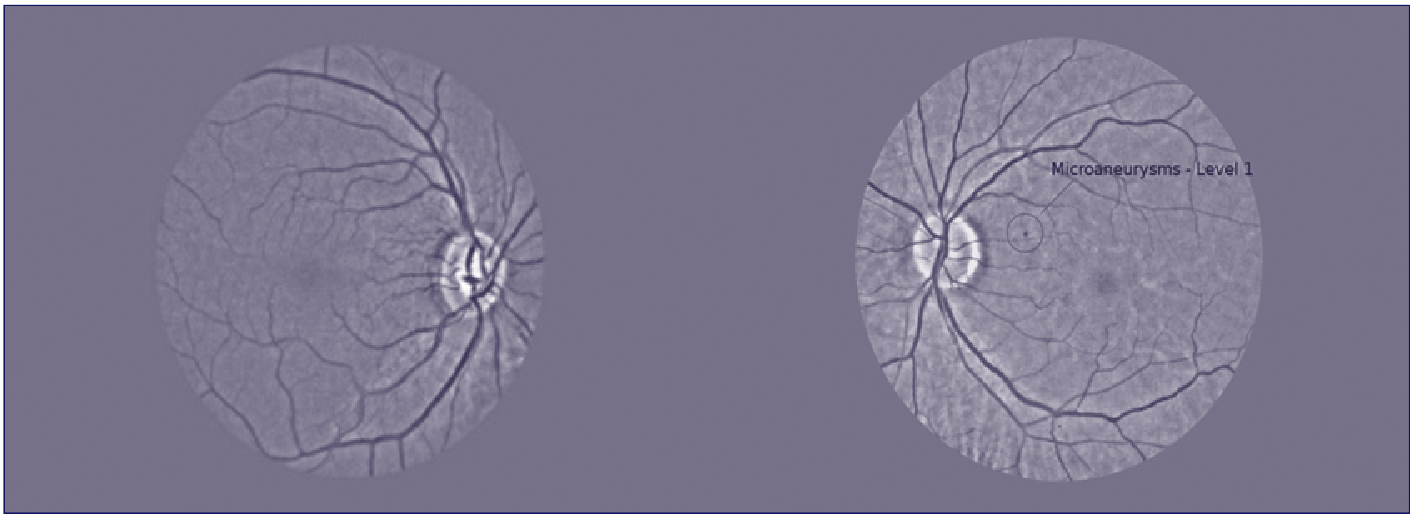
We used 80 images from Indian Diabetic Retinopathy Image Dataset for manual annotation. We created a text file, each line describing path of image, Bounding box [x1, y1, x2, y2] and label name. E.g. image file path, x1, x2, x3, x4, micro aneurysms.

5.4. Faster RCNN for Exudates localization

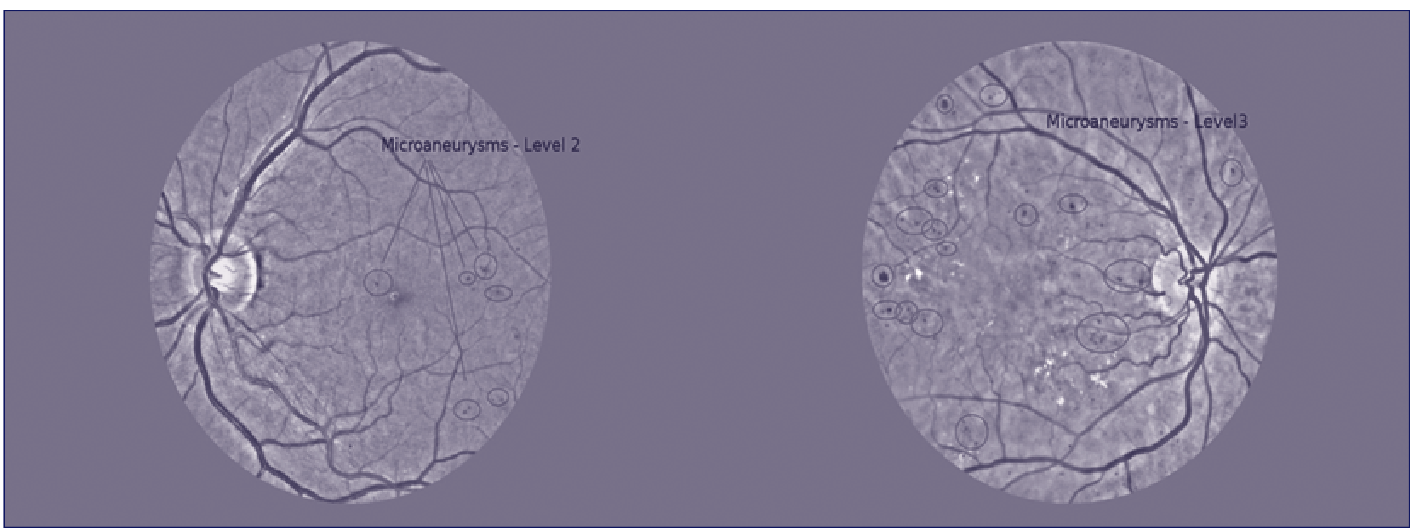
Figure 6 shows the faster RCNN for exudates localization. Here faster RCNN network receives pre-processed image as input. The network is configured to 32 ROI's, anchor box scales of {128,256,512} and anchor box ratios of {[1:1],[1:2],[2:1]}

6. RESULTS AND ANALYSIS

Figure 7: (a) Normal (level 0) result, (b) mild (level 1), (c) moderate (level 2) and (d) severe (level 3) results.



(a) **(b)**



(c) **(d)**

Figure 7 shows the detection of diabetic retinopathy classifications like normal (level 0) results and non-proliferated results mild (level 1), moderate (level 2) and severe (level 3) stages of micro aneurysms. And figure 8 shows the sensitivity and specificity chart of Normal, NPDR (Non-proliferated Diabetic Retinopathy) stages Mild (NPDR-1), Moderate (NPDR-2) and Severe (NPDR-3) along with Proliferated Diabetic retinopathy (PDR) and accuracy chart calculated from the below table-1.

Figure 8: Plot of sensitivity and severity.

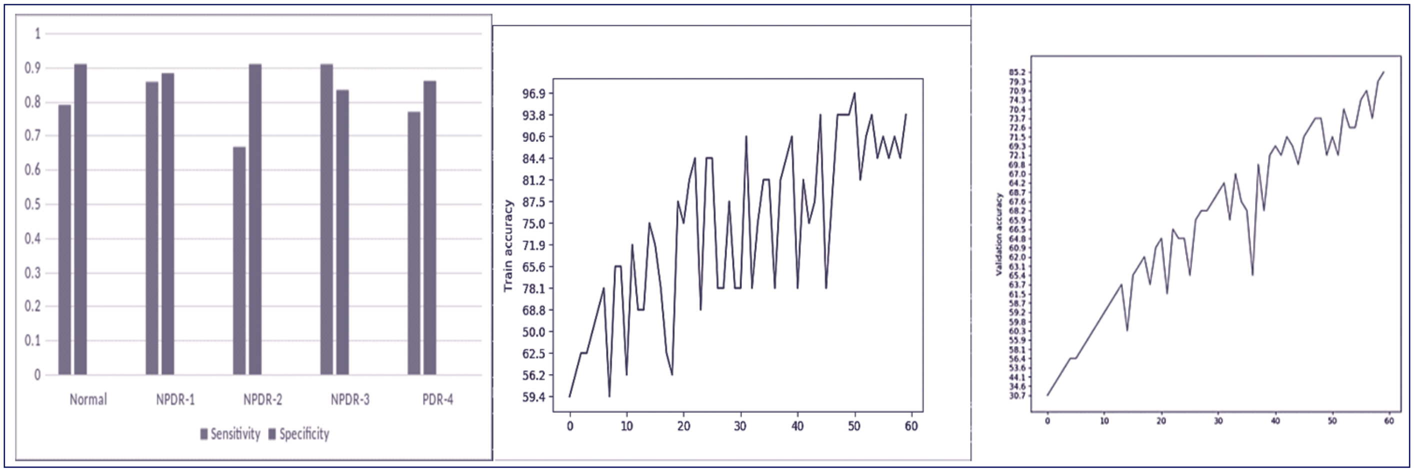


Table 1: Diabetic retinopathy raw image classification accuracy.

Diabetic retinopathy symptoms	Level	count	test data	positive results	classification accuracy
Micro aneurysms	1	< 5	1524	1260	82.67%
	2	> 5 &< 15	1482	1104	74.49%
	3	> 15	1520	1200	78.74%
Hemorrhages	1	< 5	1524	1260	82.67%
	2	> 5 &< 15	1482	1104	74.49%
	3	> 15	1520	1200	78.74%
Neovascularization	3	>1	1520	1200	78.74%
Exudates	3	> 1	54	45	83.33%

7. CONCLUSIONS

Practically, ophthalmologist can identify DR by the presence of lesions associated with the vascular abnormalities caused by the disease. While this method is more efficient, its resource demands are high. In this work, we provided a deep learning model that includes RPN and Faster RCNN to identify and analyze the exudates of diabetic retinopathy. The Faster-RCNN provide the robust interpretability of the proposed detection model in a single pass by monitoring the pathogenesis so that this work can be consider as an assistant for the ophthalmologist. With this features, this methodology can still yield the more competitive performance of DR detection and compared with the state of the art methods. In future, we would consider extending this method to other medical application problems.

8. FUTURE SCOPE

In this paper, deep learning method (R-CNN) are used instead of MATLAB algorithms like SVM so accuracy is reduced to 82% compare to SVM because of huge dataset used in this paper, which is the only limitation of this paper and even though we achieved more accuracy compare to others. In future if less dataset having few stages we used calculated then we can increase the accuracy more than 90%.

ACKNOWLEDGEMENTS: The authors would like to sincerely thank Dr. B. K. Narendra, Principal, BGSIT, BG Nagar, for providing the necessary facilities and support and Dr. M. B. Anandaraju, Professor and Head of the Department, Electronics and Communication Engineering BGSIT, BG Nagar, and all staffs of BGSIT and PESCE Mandya, for their support and valuable suggestion to complete this paper.

REFERENCES

[1]Darshit Doshi ; Aniket Shenoy ; Deep Sidhpura ; Prachi Gharpure, Diabetic retinopathy detection using deep convolutional neural networks 2016 International Conference on Computing, Analytics and Security Trends (CAST) IEEE 2017. DOI 10.1109/CAST.2016.7914977.

[2]Deng, L. 2014. A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Transactions on Signal and Information Processing.

[3] Diabetic retinopathy image dataset, [online] Available: https://www.kaggle.com/c/diabetic-retinopathy-detection/data.

[4]Dosovitskiy, A., and T.Brox. 2016. Inverting visual representations with convolutional networks. In CVPR.

2015. Diabetic Retinopathy Detection—Kaggle: https://www.kaggle.com/c/diabetic-retinopathy-detection.

[5]Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In NIPS.

[6] Navoneel Chakrabarty, "A Deep Learning Method for the detection of Diabetic Retinopathy", Electrical Electronics and Computer Engineering (UPCON) 2018 5th IEEE Uttar Pradesh Section International Conference on, pp. 1-5, 2018.

[7]Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks Jan 2016 version, v3, arXiv:1506.01497.

[8] Silberman, N.; Ahrlich, K.; Fergus, R.; and Subramanian,L. 2010. Case for automated detection of diabetic retinopathy. In AAAI Spring Symposium: Artificial Intelligence for Development. AAAI.

[9]Sopharak,A.;Uyyanonvara,B.;andBarman,S.2009. Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering. Sensors 9(3):2148–2161.

[10]Wang, Z.; Yan, W.; and Oates, T. 2016. Time series classification from scratch with deep neural networks: A strong baseline. arXiv preprint arXiv:1611.06455.

[11]Wu,D.;Zhang,M.;Liu,J.-C.;andBauman,W.2006.Onthe adaptive detection of blood vessels in retinal images. IEEE Transactions on Biomedical Engineering53 (2):341–343.

[12] Yang, J. B.; Nguyen, M. N.; San, P. P.; Li, X. L.; and Krishnaswamy, S. 2015. Deep convolutional neural networks on multi-channel time series for human activity recognition. In IJCAI.

[13]https://www.indiatoday.in/education-today/gk-current-affairs/story/98-million-indians-diabetes-2030-prevention-1394158-2018-11-22.

AUTHORS

Manojkumar SB, Research Scholar, PET Research Foundation, PES College of Engineering, Mandya - 571 401, Karnataka, India
Email: manojkumarsb.ece@gmail.com

Dr. H S Sheshadri, Professor, Dept. of ECE, PES College of Engineering, Mandya - 571 401, Karnataka, India
Email: hssheshadri@gmail.com