March - April 2020
ISSN: 0193-4120 Page No. 11671 - 11676

A Comparative Study on the Various Neural Network Approaches to Classify Diabetic Retinopathy

[1]Subhadeep Kundu, [2]Anuvab Biswas, [3]D Vanusha

[1,2]Bachelor of Technology, [3]Assistant Professor, [1,2,3]Department of Computer Science SRM IST
[1]sk3814@srmist.edu.in, [2]ab1387@srmist.edu.in, [3]vanushad@srmist.edu.in

Article Info
Volume 83
Page Number: 11671 - 11676
Publication Issue:
March - April 2020

Abstract

Diabetic Retinopathy affects more than 85% of patients with longstanding diabetes and is it is one of the primary causes of blindness for the age group 20-64. DR can be divided into two types: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). In recent years, more than 85% of people suffer from diabetic retinopathy due to lack of proper diagnosis or early prediction. Several methodologies have been proposed involving various core concepts to diagnose this issue. The present model classifies DR into five categories with integer values from zero to four. Deep CNN approach proves to be sufficiently efficient but defining the problem uniquely in order to ensure the occurrence of the disease still remains a problem. Even though CNNs are popular for their generally high accuracy rate, there are obstacles like computational complexities and processing time. Region Proposals are a way to solve this problem. Using Region based CNN approaches, the region of interest for the purpose can be detected. Another challenge is the unavailability of a universally or majorly accepted database of fundus images which makes it harder to get accurate results for the algorithms. One such database, the DRiDB proposes to overcome this obstacle. In this paper we try to present a comparative study of the various region based proposals and object detection methods that can be used to predict diabetic retinopathy and make it more accurate and efficient.

Article History
Article Received: *24 July 2019*
Revised: *12 September 2019*
Accepted: *15 February 2020*
Publication: *16 April 2020*

Keywords: *SVM, RCNN, Fast RCNN, Faster RCNN, RPN, SPPNet, YOLO, SSD, DRiDB*

1. Introduction

Diabetic retinopathy (DR) is an eye disease which damages the retina caused by diabetes mellitus. Diabetic Retinopathy is caused when the glucose in the blood increases causing the vision to get blurred ultimately leading to complete blindness. Excess glucose in the blood vessels may cause anomalies like microaneurysms, hemorrhages, hard exudates and cotton wool spots being developed during the various phases of diabetic retinopathy. The two types of DR are non proliferative diabetic retinopathy and proliferative diabetic retinopathy. Diabetic Retinopathy is classified into five categories with an integer value from 0 till 4. It is one of the main causes of blindness. Diabetic retinopathy affects up to 80 percent of today's population. At least 90% of new cases could be reduced with proper treatment and monitoring of the eyes. The longer a person has diabetes, the higher his or her chances of developing diabetic retinopathy. Diabetic retinopathy is mainly caused by high blood glucose level and blood pressure.

March - April 2020
ISSN: 0193-4120 Page No. 11671 - 11676

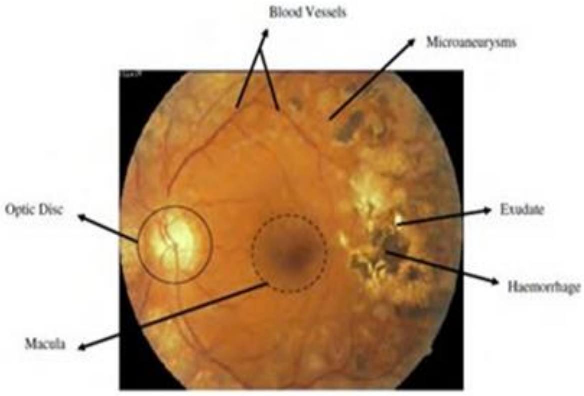

Figure 1: Diabetic Retinopathy

So to detect and predict DR various methods have been used that could achieve the output with certain accuracy. Most of the work done till date uses SVM and normal CNN. One method that was suggested to detect DR automatically gave an accuracy of 86.17%. This method used a CNN and also lacked the clinical data for training and validating the dataset. Although CNN proved to be sufficiently satisfactory, its computation power has always been a challenge. Future work could take into consideration various other proposals discussed below into consideration to deal with it. To solve this problem of computation region proposals could be used for CNN which will have less computational complexities. RCNN approach uses selective search method as it takes only certain required parts into consideration. To improve the efficiency of RCNN came into existence the concept of Fast RCNN which is based on SPPNet. Further research developed the concept of Faster RCNN which requires less computation due to lower resolution of feature image. But one problem with RCNNs was that it slowed down the algorithm during testing thus creating bottlenecks. Thus to solve the problem a new object detection method known as YOLO was developed that solved the problem of bottlenecks. Also all the databases lack enough datasets which also reduced the accuracy of the system. So in this paper we also discuss the usage of a universal database named DRiDB which could be used along with these methods to achieve higher accuracy.

2. Methods

In this section we will be discussing the various methodologies that could be adopted in order to increase the efficiency and accuracy in classifying diabetic retinopathy. This will provide a comparative study of various methods that have been developed till date and listing out their pros and cons and also highlighting how one method overcomes the shortcomings faced by its primitive one.

A. Primitive CNN

The first classification of diabetic retinopathy was done using SVM classifier. The images were passed into the network and sent through multiple convolutional and pooling layers and the output is the classification of the object to the class it belongs to. It was one of the earliest methods in prediction of early diabetic retinopathy. But this system lacked data resources and had a lot of computational complexities resulting from having to select numerous regions. The image is expected to have different aspect ratios and spatial ratios, like the portion of the image containing the object may vary and also the objects can be in different shapes. So to solve these complexities researchers came up with using the CNN model which proved to be better than the traditional SVM classifier. The performance of this system was directly related to the number of convolutional layers and pooling layers that has been made. This model classified an image under five categories with integer values ranging from 0 to 4. But to overcome the degradation of the system due to computation time and other complexities caused because of taking into consideration the entire images which was large to handle, came into existence the concept of region based proposals which can be implemented in order to predict diabetic retinopathy.

B. RCNN

The concept of regional proposals was introduced to overcome the shortcomings of the traditional CNN approach, with the aim of reducing the number of regions to be considered. This method uses a selective search algorithm for extraction of boxes from the images and check if these selected boxes contain the object or not. It predicts the offset values for increasing the accuracy. There are four patterns/parameters used for recognizing objects - varying scales, colours, textures and enclosure. This algorithm identifies these patterns from the image and proposes the regions depending on this. Similar regions are combined to form a larger region based on similarities in colour, texture, size, shape. RCNN too is effectively costly and slow because it extracts 2,000 regions for each and every image based on the algorithm and this is for a single image. For n images, it will amount to n*2000. Another problem with RCNN is that feeding the region proposals every time to the system slows down the process of training the model. It is slow because it performs a ConvNet forward pass for each proposed object, and does not share computation. Spatial pyramid pooling networks are a step forward which shares computation in order to increase the speed of RCNN. Also as the algorithm is fixed, there is no learning in the process.

C. SPPNet

Generally at the transition of the convolutional layer and the fully connected layer there is either a single or sometimes even no pooling layer. With more related research in this field, researchers came up with the idea of using multiple pooling layers with different scales and eventually named them Spatial Pooling Pyramid Network

March - April 2020
ISSN: 0193-4120 Page No. 11671 - 11676

(SPPNet). Also with SPP any size of input size can be fed and it is also preferred as it will increase the robustness of the system, thus reducing and improving the error rates. Compared with R-CNN, SPPNet processes the image at convolutional layers for only one time and that can be used to find out the CNN representation of every object we got by Selective Search while R-CNN processes the image at convolutional layers for 2k times since there are 2k region proposals. In this method the last pooling layer is replaced by a spatial pyramid pooling layer instead of traditional max-pooling to handle the deep network of images of varied sizes.

D. Fast RCNN

The author of RCNN, Rosh Girshick came up with an idea of running CNN just a single time for every image. The idea of Fast R-CNN is developed using the concept of SPPNet. SPPNet removes the crop/warp from the R-CNN, and replaces the last pooling layer before FC layer with SPP, and keeps the output image m*n parts independent of the resolution of input image. In this method the input image is fed to the CNN for generating the convolutional feature map therefore solving the problem of feeding the region proposals every time as in case of RCNN and hence making the process of training the model faster than normal RCNN. Also the accuracy and precision of Fast RCNN can be increased by using Inception v3 and ResNet 50 instead of the normal ZF/VGG. In addition to these, when large datasets are taken into account, it lacks computational power to give proper accuracy. But with the advent of technology and research, researchers developed a new technology which is an advanced version of Fast CNN known as Faster CNN. Fast CNN is relatively slow because of selective search and the time consuming process so they developed the Faster CNN in order to handle the difficulties created by selective search algorithm and make the system learn the region proposals instead.

E. Faster RCNN

The gradual advancement of CNN has minimised the computation time of the neural networks but caused a problem of exposing the complete region proposal computation creating a bottleneck. So to improve these researchers came up with the idea of replacing the selective search methodology with Region Proposal Network (RPN). It involves full image convolutional features with detection network, so that the network can learn itself and thus increase the efficiency of the network. Also due to this there is less computation involved due to lower resolution of feature images. In this model, above the convolutional features we put a RPN by constructing few convolutional layers that can at the same time regress bounding regions as well as objectness scores at every point on a rectangular grid. This method is used to make the training process faster than Fast RCNN. In this the detection algorithm remains the same as Fast RCNN but for training the network RPN is used that learns proposals over time and proposes the regions.

F. SSD

Acronym for Single Shot Multibox Detector, which are the key components in the architecture of the system. ‘Single Shot’ refers to the tasks of localising objects and classifying it in a singular forward pass in the network. ‘Multibox’ is a method developed for bounding box regression. The ‘Detector’ is the component to classify the objects detected. SSD has been de- signed for real-time object detection. It is an improvement on the Faster-RCNN method by eliminating the regional proposal network. The loss of accuracy from this is accounted for by the usage of multiscale features and default boxes (equivalent to anchors in Faster-RCNN). SSD predictions can be divided into two categories - positive matches or negative matches. Localization loss is calculated which is the difference between the ground truth box and the predicted boundary box. Only those predictions are considered in this step which belong to positive matches to get as close as possible to the exact value. One downside of SSD is that it performs worse than Faster R-CNN for objects of small scale. A tradeoff exists between accuracy and speed when more default boundary boxes are considered. Accuracy of SSD can be improved by using multi- scale feature maps.

G. YOLO

For YOLO, detection is a basic regression concept that takes an image as an input and process it in order to learn the class probabilities and the bounding box coordinates. In YOLO images have to be run only once, so this algorithm works very fast and in real time. Another main difference is that YOLO sees the entire image at once instead of looking at only the generated region proposals as mentioned in the previous methods. YOLO uses the complete image at the time of its training and testing so this system encodes entire information related to the class and its appearance. Fast RCNN makes mistakes with background patches but YOLO reduces this error rate to at least half. Also YOLO works on generalized representations of objects so it does not face any problem when it is given any new domain of images or new sets of unexpected images. During making predictions, it involves spatial diversity. But this model faces a problem of spatial constraints during predictions of bounding box. It faces problems when it is subjected to images with unusual aspect ratios or configurations because it works on bounding boxes, so it finds it difficult to predict small objects which are represented to the system as a group. Also a major problem with this model is that it treats the error in the loss function equally in case of small as well as large bounding box, because a small error in a large box can have minimum effect but that same error in a small bounding box can have a large impact.

3. Dataset Improvements - DRIDB

All the work that has been done till date was done using any one of the databases from DRIVE, STARE, Messidor, ImageRet, ARIA. But the problem with these databases is that none of these contain both annotated pathologies like hemorrhages, microaneurysms, soft and hard exudates, and normal fundus structures such as blood vessels and optic disc. Another significant problem with the above stated databases is that most of them are subject to manual annotation bias because of being annotated by only one person. So researchers came up with the idea of developing a comprehensive database which will overcome all the shortcomings faced by these databases. In view of that they developed a database that contains all the images properly annotated by a minimum of five experts per patient image with categorization of grade of the disease of each patient image. They named this database Diabetic Retinopathy Image Database (DRiDB). One of the previous mentioned databases mostly consists of images of patients who does not have any hemorrhages, so when it is used with an algorithm which has detection of hemorrhages, it is expected to give poor performance.

To overcome this DRiDB was developed using manually segmented blood vessels because this can compare the different algorithms that can be used for classification of diabetic retinopathy with more reliability. This database can also be worked upon to increase the size and improve the quality of the database. Also the number of experts who annotate the images can also be increased in order to get better accurate results.

4. Discussions

The above stated are few of the concepts that could be used to increase the accuracy of diabetic retinopathy. Most of the works till date are mostly based on models using SVM and primitive CNN. The accuracy of CNN has increased over SVM but also there are certain loopholes and shortcomings that need to be addressed. The biggest challenge with the SVM model was the non-detection of microaneurysms in a few retinal data but this was later solved with the CNN model. With the advancement of technology a lot of improvement has been made on the primitive CNN model to improve the accuracy rate make the training process faster.

Table 1: A Comparative Study of the Various Methods Discussed

CHARACTERISTICS	CNN	RCNN	Fast RCNN	Faster RCNN
Algorithm	Divides images into regions which further divided into classes	Selective Search for regions	Selective Search with SPPNet for regions	RPN
Architecture	Region based technique not applicable	Bounding Box based Regression	Bounding Box based Regression	Anchor box based network and Bounding Box based Regression
Input to Network	Entire image is fed	Region proposals hence convolutional operation many times per image	Entire image to generate convolutional feature map and then extracting the regions	Entire image to generate convolutional feature map and then extracting the regions
Learning Regions	Entire image learning so slower process	No learning of region proposals	No learning of region proposals	Separate network for learning and hence predicting Region Proposals
Pooling	Max Pooling and Fractional Max Pooling	ROI Pooling	ROI Pooling and Spatial Pyramid Pooling	ROI Pooling
Computation Time	Lot of regions (Includes unnecessary regions) to predict hence more time	High Computational time due to each region fed separately and Selective search	Better than RCNN because of SPP but still high due to Selective Search	Faster because of learning and predicting network
Image Resolution	Independent of Image Resolution	Independent of Image Resolution	Independent of Image Resolution	Lower resolution feature image hence faster computation
Training Process	Very Slow	Very Slow	Slow (improved by SPP)	Faster (no selective search, hence no bottleneck)

The next step could be implementing these techniques in the field of diabetic retinopathy detection. Modern RCNN techniques have prove to improve the error rates and make the training process more effective and faster. As we discussed earlier region proposals could be a step forward to increase the accuracy and efficiency of the model. The most advanced region proposals are expected to increase the efficiency of the model by a larger amount. These models are expected to reduce the computational problems associated with its previous versions. In this paper we have also covered many object detection techniques like YOLO, SSD that can be used in detection of diabetic retinopathy. The object detection algorithms work very fast even with large datasets so it appears to happen in real time. Another major problem that most of the researchers who have worked on diabetic retinopathy faced was the availability of large amounts of data for training the network. Also the problems with the databases are that most of them contain images annotated by only one expert. So to overcome this problem we have suggested using DRiDB, a database which was developed in order to address the problems faced by the earlier databases. This database is expected to increase the accuracy of any other above models because of the availability of large number of datasets and quality of the images because each of the images in this database is mostly annotated by a minimum of five experts, unlike the previous databases that have been used in training the earlier models.

5. Conclusion

In this paper we have addressed various methods that have been used in the detection of diabetic retinopathy and also we propose new technologies that could be used in order to achieve better accuracy and efficiency. Region proposals and object detection methodologies could be a step forward to these. We also highlight the various problems and shortcomings of the previous methods. The methods proposed show a gradual advancement of technology in terms of efficiency. It gives a comparative study of the methods showing how one method is better than the others and lists the individual benefits and drawbacks of each method. This will help to distinguish which method will be more suitable for the analysis. It gives an overall idea about the various convolutional neural network techniques and highlights the problems these could solve which were faced by researchers while using the earlier methods of detection. It also addresses the usage of universal database (DRiDB) which claims to solve the various problems faced while using previous datasets like limited dataset size and annotated images by single expert. Overall this paper will help to develop better and more efficient models in the future in order to detect diabetic retinopathy with better efficiency.

6. Acknowledgement

The paper and the research behind it would not have been possible without the immense support from our University, SRM Institute of Science and Technology. We are grateful to the university for helping us in all possible ways starting from making all the paid journals needed for research available to guide us in the right direction. We also would like to thank the reviewers from our college who provided constant insights into the research making it more resourceful. Their expertise in the field has helped us improve the quality of the paper in many ways and rectify the errors. We would like to mention Towards Data Science and Medium for providing various informational and insightful blogs on various RCNN and Object Detection Techniques. We thank the various authors of the papers mentioned in the references section which we referred for coming to various conclusions in our research. Last but not the least; we would like to thank our colleagues who had always been a constant support throughout the research.

References

[1] Yung-Hui Li, Nai-Ning Yeh, Shih-Jen Chen and Yu-Chien Chung, ”Computer-Assisted Diagnosis for Diabetic Retinopathy Based on Fundus Images Using Deep Convolutional Neural Network”, Hindawi, Mobile Information Systems, Volume 2019

[2] Pavle Prentasic, Sven Loncaric, Zoran Vatavuk, Goran Bencic, Marko Subasic, Tomislav Petkovic, Lana Dujmovic, Maja Malenica-Ravlic, Nikolina Budimlija, Raseljka Tadic, ”Diabetic Retinopathy Image Database (DRiDB): A new database for diabetic retinopathy screening programs research”, 2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA)

[3] Yogesh Kumaran, Chandrashekar M. Patil, ”A Brief Review of the Detection of Diabetic Retinopathy in Human Eyes Using Pre-Processing & Segmentation Techniques”, International Journal of Recent Technology and Engineering (IJRTE), December 2018

[4] Murugan. V, Vijaykumar V.R and Nidhila. A, ”A Deep Learning RCNN Approach for Vehicle Recognition in Traffic Surveillance System”, International Conference on Communication and Signal Processing, April, 2019

[5] Enrique V. Carrera, Andres Gonzalez, Ricardo Carrera, ”Automated detection of diabetic retinopathy using SVM”, 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)

[6] Kong Fansheng , Wang Zhulin, ”Based on image pixel gray value feature extraction of PHM research”, Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference

[7] P. Ramesh Kumar , S.S. Dhenakaran, ”Pixel

March - April 2020
ISSN: 0193-4120 Page No. 11671 - 11676

based feature extraction for ear biometrics", 2012 International Conference on Machine Vision and Image Processing (MVIP)

[8] Russell A. Poldrack, "Region of interest analysis for fMRI", Social Cognitive and Affective Neuroscience, Volume 2, Issue 1, March 2007

[9] Sunil Prasad Jaiswal, Vinit Jakhetiya , Anil Kumar Tiwari , Ashutosh Singla, "A lossless image prediction algorithm using slope estimation and least square optimization", 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings

[10] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", arXiv:1506.01497 [cs.CV]

[11] Yun Ren ,Changren Zhu, Shunping Xiao, "Object Detection Based on Fast/Faster RCNN Employing Fully Convolutional Architectures", Hindawi, Mathematical Problems in Engineering, Volume 2018

[12] Juan Du, "Understanding of Object Detection Based on CNN Family and YOLO", IOP Publishing Ltd Journal of Physics: Conference Series, Volume 1004, conference 1

[13] Ross Girshick, "Fast R-CNN", arXiv:1504.08083 [cs.CV]

[14] Xiaolong Wang, Abhinav Shrivastava, Abhinav Gupta, "A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection", arXiv:1704.03414 [cs.CV]

[15] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", arXiv:1506.02640 [cs.CV]