Detection of Diabetic Retinopathy using Binocular Convolutional Neural Network

D. Nithya[1], M. Ramya[2], S. Shuruthi[3]

[1]Assistant Professor, Computer Science and Engineering, VCET, Tamilnadu, India. Email: d.nithyacse1990@gmail.com

[2]Student, Department of Computer Science and Engineering, VCET, Tamilnadu, India. Email: ramyamuthuraman25@gmail.com

[3]Student, Department of Computer Science and Engineering, VCET, Tamilnadu, India. Email: shuruthy03@gmail.com

Abstract - Diabetic retinopathy (DR) is considered as one of the vital causes of visual loss worldwide. However, DR is difficult to be detected in early stages and therefore the procedure is also time-consuming even for skilled experts. Therefore, a computer-aided diagnosis method supports deep learning algorithms is proposed to automated diagnose of the referable diabetic retinopathy (RDR) by classifying color retinal fundus photographs into two grades. In the proposed model, a convolutional neural network model with Siamese-like architecture is trained with transfer learning technique. It differs from previous workbecause the proposed model accepts inputs as binocular fundus images and learns their correlation to help making prediction. A locality under the receiver operating curve (AUC) of 0.951 is obtained by the proposed binocularmodel, which is 0.011 over that obtained by existing monocular model on comparison. Theeffectiveness of the binocular design can further be verified by a binocular model for five-class DR detection is additionally trained and evaluated ona 10% validation set. The result shows that it achieves a kappa score of 0.829 which is over that of existing non-ensemble model.

Keywords - Neural Network, Retinopathy, Binocular, Siamese-like, Fundus.

I. INTRODUCTION

Diabetic retinopathy (DR) can be a standard complication of diabetes. It is associated with retinal vascular damage caused by long-standing Diabetes Mellitus. DR has become one of the vitalcauses of blindness and vision impairment worldwide, since 0.4 million cases of blindness and 2.6 million cases of severe vision impairment globally are also attributed to it in 2015. In fact, the impairment of vision as a result of DR is also controlled or averted if it's detected earlier and treated in time. However, many patients miss the best time for treatment since there are few signs or symptoms at the primary stage of DR. Also, observation and evaluation to fundus photographs are essential for the diagnosis of DR, for which procedure is also time-consuming even for skilled experts. Therefore, computer-aided automated diagnosis approaches have been a great help in clinical detection of DR during a very short time. A deep learning based method which is inspired by the diagnostic process of human ophthalmologists is proposed. It automatically classifies the fundus photographs into 2 types – with or without RDR. This work does not adopt fundus images of single eye as input like most previous works. But, a unique Siamese-like CNN model is built with weight-sharing layers supported Inception V3, which is prepared to accept fundus images of both eyes as inputs and outputs the classification results of every eye at the identical time. For the better adaption of the model, those binocular fundus images are paired and pre-processed correspondingly and then fed into the network. The proposed binocular is compared with the existing monocular model that directly transferred from original Inception V3 with relevance. Theclassification results of both models are evaluated with the AUC score. It shows that proposed binocular model outperforms the monocular model by a margin. Besides, to further prove the effectiveness of the binocular design, a binocular model for the five-class DR detection task is also trained and evaluated. The result shows that, on a tenth validation set, the binocular model achieves a kappa score of 0.829 which is above that of existing non-emsemble model. Finally, the comparison between confusion matrices obtained through models with paired and unpaired inputs is performed and it shows that the binocular architecture does improve the classification performance.

II. SCOPE AND OBJECTIVES

- The binocular model makes use of Siamese-like architecture.
- It aims to spot the identified area within the earlier stages in order that the timely treatment may be provided.
- The binocular model outperforms the monocular model by a margin.
- It detects the disease by the differences within the pixels within the input.

III. EXISTING SYSTEM

Shahin developed a system that can automated classify retinal fundus images into those with or without proliferate diabetes retinopathy. In that system, the pathological features such as blood vessels and exudates are extracted using morphological extraction. Jaafar, on the other hand, proposed an automated algorithm, which mainly consists of two parts: 1. The top-down Segmentation 2. A co-ordinate system. Casanova introduced an algorithm of random forest to discriminate eye images with or without DR. It gave results with the accuracy of more than 90% and assesses the DR risk based on graded fundus photographs and systemic data. Most of works either rely on the variables or much effort manually measured by experts into extracting handcrafted features with image processing approaches.

A. Drawbacks of Existing System

- High Complexity
- More Human Efforts.
- Unstability

IV. Proposed System

- The binocular model accepts the preprocessed images to search out the affected areas.
- The preprocessed images help in standardizing the pictures and modify the pictures to be acceptable for the network.
- The Siamese-like Structure helps to search out the similarity between two input images of the identical person. The problem of the image dataset is that it's too small for a deep learning model to solve a medical image recognition problem with high accuracy. Therefore, besides of the pre-processing steps, multiple image augmentation steps are further imposed on the data set in order to improve the generalization performance of the proposed model. But before the augmentation process, it should be noted that the original fundus images contains many physiological information of patients.

A. Advantages of Proposed System

- High Accuracy
- Less Complexity.

V. Modules

A. Image Acquisition

Image Acquisition could be a process of getting an Input Image for the method of automatic detection of Diabetic retina using Digital Image Processing.

B. Pre-Processing

Pre processing could be a common name for operations with the pictures at rock bottom level of abstraction both input and output is the input images. The aim of pre processing is an improvement of image data that suppress unwanted image data distortions or enhance the some image features important for the further processing. Four categories of image pre-processing methods in keeping with the dimensions of pixel neighborhood that's used for the calculation of recent pixel brightness:

- Pixel brightness transformations
- Geometric transformations
- Pre-processing methods that use a neighborhood of the processed pixel,
- Image restoration that requires knowledge about the entire image.
- If pre processing aims to correct some degradation in the image, the nature of a prior information is important:
    - Knowledge about the nature of the degradation; only very general properties of the degradation are assumed,
    - Knowledge about the properties of the image acquisition device, the nature of noise (usually its spectral characteristics) is sometimes known.
    - Knowledge about objects that are searched for in the image, which may simplify the pre-processing very considerably .If knowledge about objects is not available in advance it can be estimated during the processing.

C. SIFT: (Scale Invariant Feature Transform)

SIFT keypoints of objects are first extracted from a bunch of reference images and stored using a database. An object is recognized as a replacement image by individually comparing each feature from the new image to the current database and finding candidate matching features supported Euclidean distance of their feature vectors. From the complete set of matches, subsets of keypoints that agree on the item and its location, scale, and orientation in the new image are identified to be as good matches. The determination of consistent clusters is performed rapidly by using an efficient hash table implementation of the generalised Hough transform.

Each cluster of three or more features that agree on an object and its pose is then subject to further detailed model verification and subsequently outliers are discarded. Finally the probability that a particular set of features indicates the presence of an object is computed, given the accuracy of fit and number of probable false matches. Object matches that pass of these tests could even be identified as astute with high confidence. For any object in the very picture, interesting points on the item could even be extracted to provide a "feature description" of the item. This description, extracted from a training image, can then be accustomed to identify the item when attempting to locate the item during a test image containing many other objects. To perform reliable recognition, it's important that the features extracted from the training image should be detectable even under changes in image scale, noise and illumination. Such points usually lie onhigh-contrast regions of the image, like object edges. Another important characteristic of those features is that the relative positions between them within the initial scene shouldn't change from one image to another image. Features located in articulated or flexible objects would usually not work if any change within their internal geometry happens between two images within the set being processed. However, in practice SIFT detects and uses the way larger number of features from the images, which reduces errors caused by these local variations within the common error of all feature matching errors. SIFT can identify objects even among clutter and under partial occlusion, because the SIFT feature descriptor is invariant to uniform scaling, orientation, illumination changes, and partially invariant to affine distortion. This section summarizes the first SIFT algorithm and mentions some competing techniques available for beholding under clutter and partial occlusion. The SIFT descriptor relies on image measurements in terms of receptive fields over which local scale invariant reference frames are established by local scale selection.

D. Segmentation using Expectation and Maximization Segmentation

The process of partitioning of images into multiple frames or images constitutes Image Segmentation. The segmentation is applied is to represent image into something that is more meaningful and easier to analyse. Objects and boundaries (lines, curves, etc.) in images are typically located using Image Segmentation. In other words, a label is assigned to every pixel in an image such that pixels with the same label have certain characteristics in common As a result, a set of segments that are obtained collectively cover the entire image, or a set of features extracted from the image. Every pixel in an image or region is identical in some characteristic or computed property, such as color, intensity, or texture. Adjacent regions differ with respect to the same characteristic(s) significantly.

E. Expectation maximization

Maximum Likelihood (ML) estimate in the presence of missing or hidden data is computed using the EM algorithm. It is an efficient iterative procedure. In ML estimation, the model parameter(s) are estimated for which the observed data are the most likely. EM algorithm consists of two processes for all iterations: The E-step, and the M-step. In the Expectation, or E-step, the observed data and current estimate of the model parameters to used to compute or estimate missing data. This is achieved using the conditional expectation, explaining the choice of terminology. In the M-step, the likelihood function is maximized. It is done under the assumption that the missing data are known. The estimation of the missing data from the E-step is used in place of the actual missing data. Convergence is guaranteed since the algorithm is assured to improve the likelihood at each iteration.

F. Feature Extraction

In machine learning, pattern recognition and image processing, feature extraction is essential part of the entire process. It starts from an initial set of measured data and derive values

(features) that are informative and non-redundant, facilitating the subsequent learning and generalization steps, eventually leading to better human interpretations. Feature extraction is related to dimensionality reduction.

When the data that is given to an algorithm is too large to be processed or suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images as pixels), then a reduced set of features can be taken (also named a feature vector). Determining a subset of the initial features is called feature selection.

The relevant shape information contained in a pattern is called Feature Extraction. This makes the task of classifying the pattern easy by a formal procedure. In character recognition system, Feature extraction is done after the preprocessing phase. The primary task of pattern recognition involves assigning one of the possible output classes for the input pattern that is taken. This process can be divided into two general stages: Feature selection and Classification. Feature selection is critical for the whole process. Criteria to choose features given by Lippman are: “Features should contain information required to distinguish between classes, be insensitive to irrelevant variability in the input, and also be limited in number, to permit, efficient computation of discriminant functions and to limit the amount of training data required”.

To achieve High performance, single most important factor is feature selection of a feature extraction method. Feature extraction has been given as “extracting from the raw data information that is most suitable for classification purposes, while minimizing the within class pattern variability and enhancing the between class pattern variability”. Thus, selection of a suitable feature extraction technique needs to be done with utmost care and input should be taken into consideration. Taking into consideration all these factors, it becomes essential to look at the various available techniques for feature extraction in a given domain, covering vast possibilities of cases.

Diabetic retinopathy is a damage of the retina. It is considered to be one of the serious consequences of the diabetes. Early detection of diabetic retinopathy is extremely vital to prevent premature visual disability and blindness. Diabetic Retinopathy (DR) is a complication of diabetes and is found to be the leading cause of blindness and visual loss. The retinopathy caused by the diabetes can be detected early by the regular eye examination. The diabetic retinopathy screening helps to detect whether the individuals require follow-up referral for further treatment or not. Thus, an accurate and robust retinal screening system is required in order to assist the retinal screeners so that the retinal images are effectively classified with high confidence.

By using MATLAB, various morphological processing techniques are applied to extract and detect the features such as blood vessels, micro aneurysms and exudates. Based on the results of area computation of features, they are classified into normal, mild, severe. At the end, this classification helps to identify the stages of non-proliferative diabetic retinopathy.

The proposed developed system could be a benchmark for the development of other retinopathy signs' detection systems, such as for Microaneurysms, Retinal Detachment and Macular Degeneration. As a conclusion, employing the Maximization Segmentation in the image processing can help to produce a more reliable screening system. In addition, it helps to achieve the overall aim of the screening, which is the early detection of sight threatening diseases and ensuring a timely treatment so that the vision loss can be prevented.

VI. Implementation

A. Pre – Processing

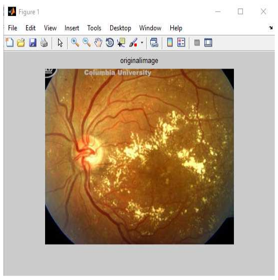

Fig. 1 Original Image

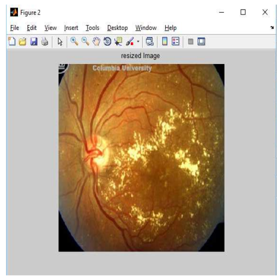

Fig. 2 Resized Image

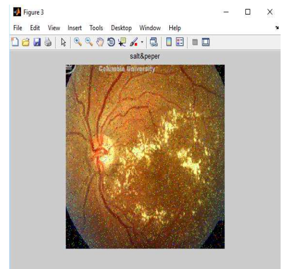

Fig. 3 Salt & Pepper

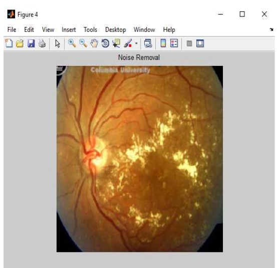

Fig. 4 Noise Removal

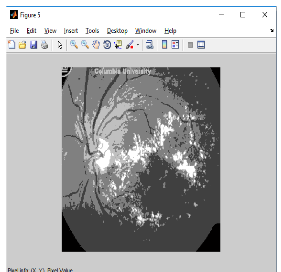

Fig. 5 Scaling

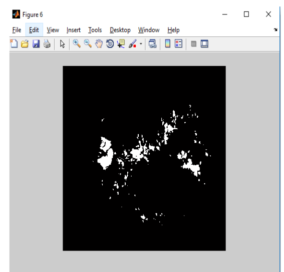

Fig. 6 Feature Extraction

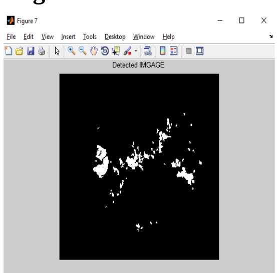

Fig. 7 Detected Image

VII. Conclusion

Diabetic retinopathy is a damage of the retina. It is considered to be one of the serious consequences of the diabetes. Early detection of diabetic retinopathy is extremely vital to prevent premature visual disability and blindness. Diabetic Retinopathy (DR) is a complication of diabetes and is found to be the leading cause of blindness and visual loss. The retinopathy caused by the diabetes can be detected early by the regular eye examination. The diabetic retinopathy screening helps to detect whether the individuals require follow-up referral for further treatment or not. Thus, an accurate and robust retinal screening system is required in order to assist the retinal screeners so that the retinal images are effectively classified with high confidence.

By using MATLAB, various morphological processing techniques are applied to extract and detect the features such as blood vessels, micro aneurysms and exudates. Based on the results of area computation of features, they are classified into normal, mild, severe. At the end, this classification helps to identify the stages of non-proliferative diabetic retinopathy.

The proposed developed system could be a benchmark for the development of other retinopathy signs' detection systems, such as for Microaneurysms, Retinal Detachment and Macular Degeneration. As a conclusion, employing the Maximization Segmentation in the image processing can help to produce a more reliable screening system. In addition, it helps to achievethe overall aim of the screening, which is the early detection of sight threatening diseases and ensuring a timely treatment so that the vision loss can be prevented.

REFERENCES

[1] G. Quellec, K. Charria Re, Y. Boudi, B. Cochener, and M. Lamard, “Deep image mining for diabetic retinopathy screening,” Medical Image Analysis, vol. 39, pp. 178–193, 2017.

[2] R. Gargeya and T. Leng, “Automated identification of diabetic retinopathy using deep learning,” Ophthalmology, vol. 124, no. 7, pp. 962–969, 2017.

[3] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818– 2826.

[4] R. Casanova, S. Saldana, E. Y. Chew, R. P. Danis, C. M. Greven, and W. T. Ambrosius, “Application of random forests methods to diabetic retinopathy classification analyses,” Plos One, vol. 9, no. 6, p. e98587, 2014.

[5] S. J. Pan, Q. Yangetal., “A survey on transfer learning,” IEEE Transactions on knowledge and data engineering, vol. 22, no. 10, pp. 1345–1359, 2010.

[6] A. Ahmad, A. B. Mansoor, R. Mumtaz, M. Khan, and S. H. Mirza, “Image processing and classification in diabetic retinopathy: A review,” in European Workshop on Visual Information Processing, pp. 1–6, 2015.

[7] E. M. Shahin, T. E. Taha, W. Al-Nuaimy, S. E. Rabaie, O. F. Zahran, and F. E .A. El-Samie, “Automated detection of diabetic retinopathy in blurred digital fundus images,” in Computer Engineering Conference, 2013, pp. 20–25.

[8] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” pp. 448-456, 2015.

[9] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.

[10] B. Graham, “Kaggle diabetic retinopathy detection competition report,” https://kaggle2.blob.core.windows.net/forummessageattachments /88655/2795/competitionreport.pdf/,August6,2015,accessed May 20, 2018.