Comparative Analysis Between Algorithms forDetection of Diabetic Retinopathy

Tejas Amrutkar[1], Sahaj Choudhary[2], Vidisha Bardekar[3], Shivani Ahire[4],Prof. Smita Chaudhari[5],

Department of Information Technology

K. K. Wagh Institute of Engineering, Education, and ResearchNashik, India

Email: [1] amrutkartejas123@gmail.com, [2]sahajc97@gmail.com , [3]vidibardekar@gmail.com, [4]shivaniahire0606@gmail.com, [5] s.n.chaudhari@kkwagh.edu.in

***Abstract*—Diabetes-The most common disease abounding around 422 million people worldwide. As the sugar levels in the body rise it affects the body in several negative ways, and one of the most menacing effect of it is diabetic retinopathy.Diabetic retinopathy is mainly caused by damage to the blood vessels in the tissue at the back of eye. The most intimidating effect of having diabetic retinopathy includes formation of small wool-like spots on the retina which can lead to irreversible blindness. From the past 10 years the number of people getting this disease has increased conspicuously. Therefore it is very important to detect diabetic retinopathy as early as possible.Among the most common and easy way to detect this diseaseis to have a proper examination of pictures of the back of theeye and rate them for disease presence and severity. There are existing methods to detect diabetic retinopathy by using image processing and different machine learning algorithms. Here we describe algorithms Convolutional Neural Network (CNN) along with InceptionNet and Colour and Edge Directivity Descriptor Algorithm (CEDD), Edge Histogram Descriptor(EHD) along with Absolute Distance Algorithm to predict Diabetic Retinopathy progression by means of deep learning (DL) .Our aim is to uncover the best method to detect diabetic retinopathy among these pairs of algorithms. The experiments are performed using the dataset of Kaggle and Messidor.**

Keywords—Diabetic Retinopathy, Convolutional Neural Net- work, Colour and Edge Directivity Descriptor Algorithm, deep learning, Edge Histogram Descriptor

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Diabetic Retinopathy is the main disease that can causeblindness in people. Analyzed by WHO that around 135 mil- lion people were affected by diabetes mellitus and the number of people affected by diabetes will be increased in 2025 [1]. The prevalence of DR in diabetic population is as high as 24.7- 37.5%. People suffering from diabetes have a higher risk of developing DR as the elevated glucose levels can damage the blood vessels [2].Diabetic Retinopathy is categorized into two stages, non-proliferative diabetic retinopathy (NPDR) and pro-liferative diabetic retinopathy (PDR), where the initial NPDR stage characterized by the appearance micro aneurysms cause appearance of haemorrhage[1]. NPDR is further divided into Mild, Moderate and Severe stages. Where the mild stage has one micro aneurysm (MA), a small circular red dot at the end of blood vessels[3]. In moderate stage MAs rapture into deeper layers and form a flame-shaped haemorrhage in the retina[3]. The severe stage contains more than 20 intraretinal haemor- rhages in each four quadrants, having definite venous bleeding with prominent intraretinal micro-vascular abnormalities. PDR is the advanced stage of DR which leads to neovascularisation, a natural formation of new blood vessels in the form of functional micro-vascular networks that grow on the inside surface of the retina[3]. Diabetic Retinopathy is a progressive disease, detection of disease can determine its severity, before a patient loses the vision[1]. Detection of diabetic retinopathy has been done manually, but manual detection requires the ability of the experts and needs a longer time, also sometimes there is misdiagnosis so that the automation techniques can be applied[1].Automated grading of diabetic retinopathy has potential benefits such as increasing efficiency, reproducibility, and coverage of screening programs; reducing barriers to ac- cess; improving patient outcomes by providing early detection and treatment[8].For an automated system to be clinically viable, it should be able to classify retinal images based on clinically used severity scales[7]. Nadeem et al. categorized the features of diabetic retinopathy, i.e. micro-aneurysms, haem- orrhages, exudates, and blood vessels, and group algorithms using computer-aided diagnosis systems into 4 categories, i.e. optic disc localization and segmentation exudates, blood ves- sel segmentation, geometric and haemodynamic features and diabetic retinopathy detection and classification[1]. Versatile application in the medical field makes use of digital image processing which has become very popular[4]. To maximize the clinical utility of automated grading,, an algorithm to detect referable diabetic retinopathy is needed[8]. Machine learning has been leveraged for variety of classification tasks including automated classification of diabetic retinopathy[8].Recently the deep learning algorithm Convolutional Neural Networks (CNN) has significant changes in computer vision and im-age classification, several studies were conducted to classify diabetic retinopathy[1].CNN architecture can perform feature extraction from the input images and classification at once but the classification process using CNN with fine-tuning requires longer computation time[1] . Therefore this study proposes the feature extraction and classification method using CNN. Colour and Edge Directivity Descriptor(CEDD) is a multiple feature extraction algorithm [9] and incorporates colour and texture information in a histogram[10]. While Absolute dis- tance is an algorithm used for classification purposes. Deep learning is a machine learning technique that learns the most predictive features directly from the images given a large data set of labelled examples[8]. Hence we use the data set from Messidor for this study. Our main aim includes comparing all these pairs of machine learning algorithms : CNN-InceptionNet and CEDD,EHD-Absolute distance and find out which pair gives the maximum accuracy for detection of diabetic retinopathy. Hence we have made two modules forthe detection of same.

II. System Overview

In this study we have proposed two systems for detection of diabetic retinopathy. In the first system DR is predicted by using CEDD, EHD and absolute distance algorithms. First the .csv file is loaded which contains all the labelled textual data of the images from Messidor such as-the image serial number, image name and retinopathy grade. The dataset is then selected and loaded into the system. The model is then trained and features are extracted using CEDD and EHD Algorithms. After that a feature file is created which is then loaded. Again dataset is selected for testing and the images are then compared and classified. By using absolute distance algorithm for classification a minimum value is returned which allocates the DR grades to the image. In the second system wehave used CNN (InceptionNet model) for the prediction of DR. The dataset is used from Kaggle and it is loaded into the system. First the training of model takes place which includes feature extraction, testing, classification of the images, then the testing is done and the result of DR is displayed.

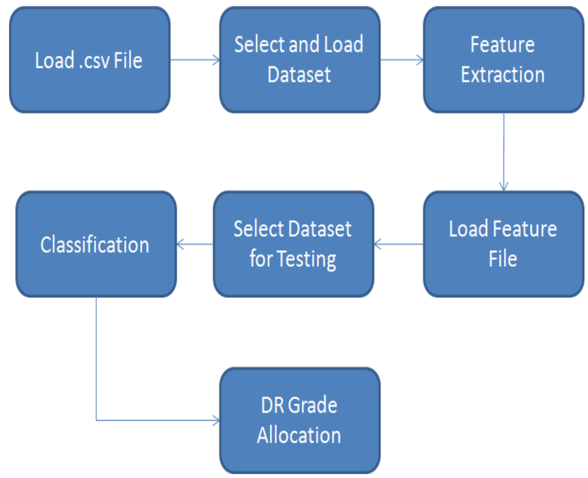


Fig. 1. System overflow using CEDD, EHD and Absolute Distance Algo- rithms

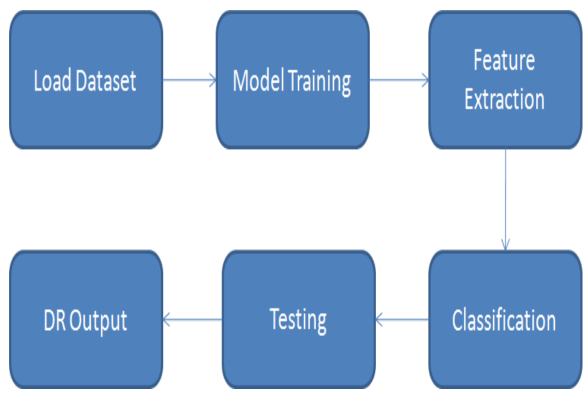


Fig. 2. System overflow using CNN (InceptionNet

III. Datasets and Methods

A. Dataset

For the first module we have used Messidor dataset whichis publicly available. It consists of 1200 images and which are taken from 3 different ophthalmology departments. We have taken 80 images from each folder of Base11-14 for training and we have used 94 images from the folder of Base11-24 for testing. For the second module we have used 861 images from Kaggle dataset for training as well as testing. The severity of NPDR is based on the appearance, spread and size or area of exudates, micro-aneurysms and haemorrhages as shown in Fig. 3[1]. Exudates are the bright areas with yellowish appearance

Fig. 3. Image with DR

which are caused because of the ruptured blood vessel that contains lipid. While haemorrhages are caused by ruptured micro-aneurysms in the blood vessels. The images are rated as per the presence of diabetic retinopathy[4] on the scale of 0-3, shown below. 0- No DR 1- Mild 2- Moderate 3- Severe

B. CEDD and EHD based feature extraction

1. Colour Edge and Directivity Descriptor

This algorithm incorporates colour and texture information ina histogram[10].CEDD size is limited to 54 bytes per image, making this descriptor suitable for large image datasets. One of the most important attribute of the CEDD is the low computational power needed for its extraction, in comparison with the needs of the most MPEG-7 descriptors[10]. Fig. 4 shows the block diagram of CEDD algorithm. Firstly, a

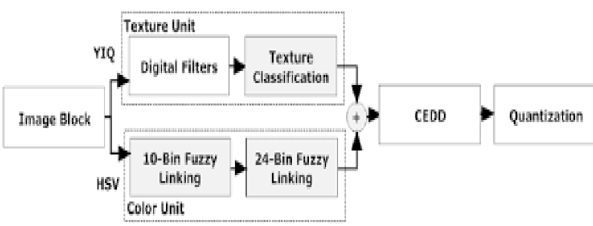
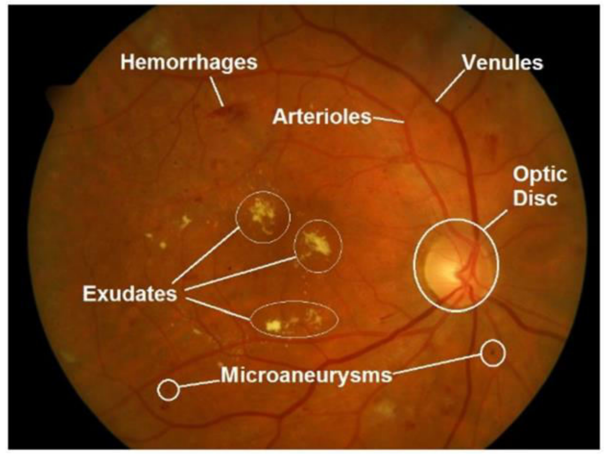


Fig. 4. Block Diagram of CEDD

histogram linking technique was proposed in HSV colour space[12]. A set of fuzzy rules undertake the extraction of colour information[12]. Due to the fuzzy system a 24 bins histogram is formed with 3 channels of HSV inputs, in which the bin represents a preset colour. With the use of 10 bins histogram the process is started and ten colour, grey typesare defined. To get the further 24-bins histogram for hue,there were 7 fuzzy areas been divided: (0) Red, (1) Orange, (2) Yellow, (3) Green, (4) Cyan, (5) Blue and (6) Magenta and (7) Blue[12]. The 24-bins histogram is achieved when Dark Colour and Light Colour for each hue area is combined with white, gray.The second step of CEDD is the textureextraction using 5 digital filters, proposed in the MPEG-7 Edge Histogram Descriptor which are able to detect edges in vertical, horizontal, 45-degree, 135-degree and non-directional edges[12]. Added the 5 edge descriptor with the original information, each region contains 6 fixed texture regions and totally a 144-bins histogram[12].



2. Edge Histogram Descriptor

Edge is an important texture feature of images which contains the outline and texture information[11]. In this each imageis divided into sub-images where each sub-block represents one pixel. 6-bins are extracted based on the sub-images. Five types of edges are selected and five corresponding detection operators are recommended in MPEG-7 for the algorithm EHD[11]. The five types are vertical edge, horizontal edge, 45-degreee diagonal edge, 135-degree diagonal edge and non-directional edge, as shown in Fig. 5 and Fig. 6. The main purpose of EHD is used in MPEG-7 to find the relevant images efficiently[13]. At first the image is converted to gray level and the image is divided into sub-block to find edge of each block. Edges are extracted by the direction and non-directional edge feature[13]. The division of image is based on a fixed size of image blocks. Each sub-block image size is 2x2 matrix, this algorithm perform to masking for

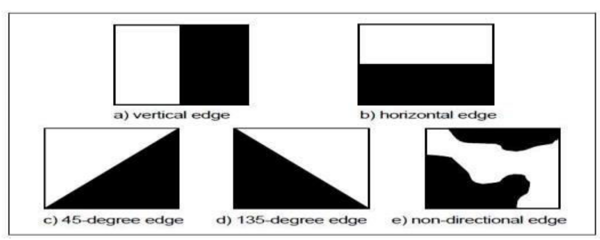


Fig. 5. 5 Edge Types

1	-1	1	1	$\sqrt{2}$	0	0	$\sqrt{2}$	2	-2
1	-1	-1	-1	0	$-\sqrt{2}$	$-\sqrt{2}$	0	-2	2

Fig. 6. Edge Detectors

each image block and finally EDH is calculated for a query image[13].

C. Absolute Distance Algorithm

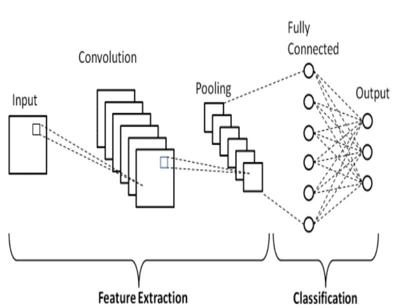
Distance metrics play prominent role in machine learning. A distance measure is an objective score that summarizes the relative difference between two objects in a problem domain. The process of classification of features is done by this algorithm. In image processing sum of absolute distance is the measure of similarity between image blocks. It is calculated by taking the absolute distance of each pixel in the original block and corresponding pixel in the block used for comparison. The differences are summed to create a simple metric of block similarity, the L1 norm of the difference image or Manhattan distance between two images. The distance d, between two vectors p, q in an n-dimensional real vector space with fixed Cartesian coordinate system is the sum of lengths of projections of the line segments between the points onto the coordinate axes is given by,

$$d(p, q) = ||p - q|| = \sum |p - q| \qquad (1)$$

where (p, q) are vectors. The sum of absolute differences provide a simple way to automate the searching of objects inside image, but may be unreliable due to the effects of contextual factors such as changes in lighting, colour, viewing direction, size, or shape. The absolute is used with object recognition methods-edge detection, to improve the reliability of results.

D. Convolutional Neural Networks

CNNs are class of Deep Neural Networks that can rec-ognize and classify particular features from the images and



can be widely used for analyzing visual images. The term 'Convolutional' in CNN indicates the mathematical function of convolution which is a special kind of linear operation in which two functions are multiplied to produce a third function which shows how the shape of one function is modified by the other. Simply, two images can be represented as matrices and can be multiplied to give an output that is used to extract features from the image. The CNN architecture consists of two main parts- A convolutional tool which separates and identifies the various features of the image for analysis which is known as feature extraction; A fully connected layer that uses the output from the convolutional process predicts the class of image on the basis of feature extracted in previous stages. Fig. 7 shows the architecture of CNN. The convolutional

Fig. 7. CNN Architecture

layer is the first layer which is used to extract the various features from the input images. The input image and a filter of particular size are multiplied-MxM. After sliding the filter over the input image their dot product is taken between filter and the parts of input image with respect to size of the filter (MxM). The output is given which is known as Feature Map and it gives us the information about the various attributes of image such as edges and corners. Then the pooling layer used which decreases the size of the convolved feature map to decrease the computational costs. Next the fully connected layer is used which consists of weights and biases along with neurons, and it connects the neurons between two layers. To overcome the problem of overfitting, a dropout layer is used in which few neurons are dropped from the neural network during the training process and it results in the reduced size of the model. Finally, the activation function is used to learn and approximate any kind of continuous and complex relationship between variables of the network. It decided which informa- tion should be forwarded and which should not at the end of network. Here, the CNN takes tensors of shape (image height, image width, color channels) and ignoring the batch size.

E. InceptionNet

The inception module was described and used in the GoogLeNet model by Christian Szegedy[14]. In order to avoid patch-alignment issues, current incarnations of the Inception architecture are restricted to filter sizes 1×1, 3×3 and 5×5; this decision was based more on convenience rather than necessity. Additionally, since pooling operations have been essential for the success of current convolutional networks, it suggests that adding an alternative parallel.pooling path in each such stage should have additional beneficial effect, too[14].InceptionNet is a very simple and powerful architectural unit which allows the model to learn not only parallel filters of the same size, but also parallel filters of differing sizes, making it possible to learn at multiple scales. Systematic sizing of filters for parallel convolutional layers is not used as the model is highly optimized. We can parameterize the module so that we can setout the number of filters to use in each of the 1×1, 3×3, and 5×5 filters. An improvement to the module was made to reduce the amount of computation required. 1x1 convolutional layers were added to decrease the number of filters before the 3×3 and 5×5 convolutional layers, and to increase the number of filters after the pooling layer. This leads to the second idea of the Inception architecture: judiciously reducing dimension wherever the computational requirements would increase too much otherwise i.e. 1×1 convolutions are used to compute reductions before the expensive 3×3 and 5×5 convolutions[14]. Besides being used as reductions, they also include the use of rectified linear activation making them dual-purpose[14]. Inception has an input image and three channels (Red, Green, Blue)[1]. With convolutional layers and pooling layers in block

1 and convolution layers in block two followed by three inception blocks for reducing parameter, then the pooling layerand classification is shown [1] in Fig 8.

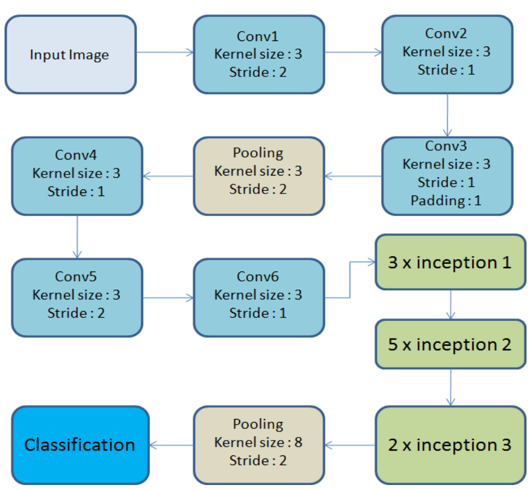


Fig. 8. InceptionNet Architecture

IV. EXPERIMENTAL RESULTS

We can find many dissimilar aspects of image processing, unchangeable pattern recognition and image encoding pose estimation for which moments can be applied[4]. Statistical moment helps in describing the image content (or distribution) with respect to axis[4].The statistical features extracted here are :

1. Mean () : Mean value gives the contribution of individual pixel intensity of a complete image[4]. The foundation ofall statistical measure is the mean. In the process of filteringin image processing using mean it is known as spatial filtering and which helps to reduce the noise[4]. The Fig. 9 shows the mean scalar graph of our CNN module. Scalar values visualize classification accuracy. The X-axis shows the training steps (or epochs) and the Y-axis corresponds to the distribution i.e. the values represent number of occurrences of the corresponding weight value in the layer .The orange line represents the training accuracy and the blue line, validation accuracy.

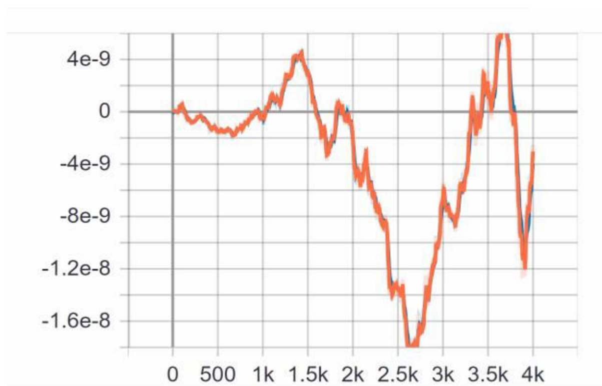


Fig. 9. Mean Scalar Graph

2. Standard Deviation :

Measuring the variability gives the standard deviation which is most widely used. It describes the amount of variation that exists from an average in terms of image processing[4]. A lesser value of standard deviation indicates that the data tends to be very close to the mean[4]. The Fig. 10 below shows the standard deviation of CNN module. The learning rate for 4000 training steps is 0.04.

3. Cross Entropy

Cross entropy is commonly used as a loss function. It is a measure which is generally used for calculating the difference between two probability distributions (shown in Fig. 11).The red and blue lines on the graph indicate training and testing loss respectively. For 4000 training steps the training loss of the model is 0.15 and the testing loss is 0.40.

4. Accuracy

Model accuracy is the measurement which is used to determine that which model is best at identifying relationships and patterns between variables in a dataset based on the input, or training data. The Fig. 12 below shows the accuracy of our

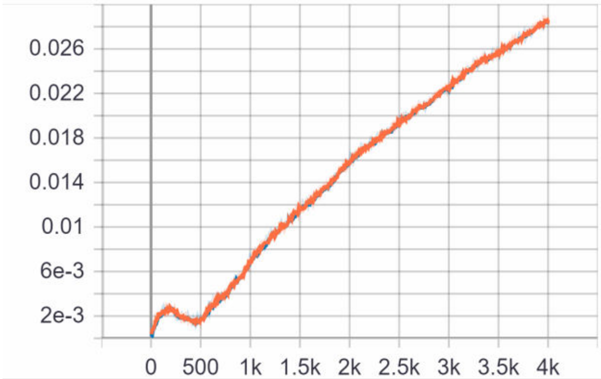


Fig. 10. Standard Deviation

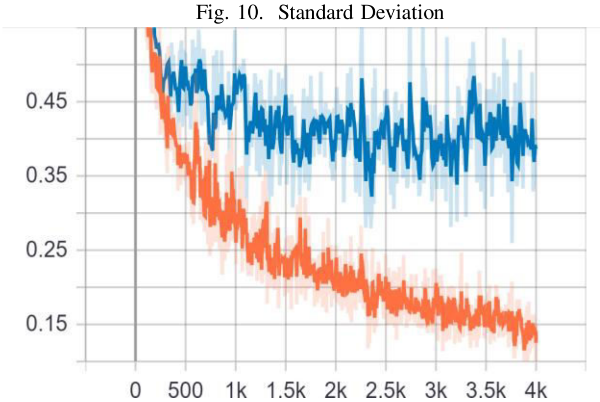


Fig. 11. Standard Deviation

CNN module. For 4000 steps the training accuracy is 0.96while the validation accuracy is 0.78.

V. FINAL RESULTS

The first system was developed using CEDD, EHD and absolute distance algorithms in which we created a dataset which contained testing images from base 11-24 of Messidor dataset. The dataset contained 94 images which gave an accuracy of 71%. Hence we conclude that the overall accuracy

using these algorithms 71% for detecting diabetic retinopathy. The second module was developed using CNN(Inception Net model). We used the dataset from Kaggle containing around 900 images. Our CNN based Inception model gave Model training accuracy of 96% and the final Test accuracy is 86%.

TABLE I RESULTS

Algorithm	Accuracy
CEDD, EHD and Absolute distance	71%
CNN (InceptionNet) model training accuracy	96%
CNN (InceptionNet) final testing accuracy	86%

VI. CONCLUSION AND FUTURE WORK

The early detection of diabetic retinopathy can considerably help a person to recover from this disease. The clinical diagno- sis process is very costly and time consuming therefore in this paper we proposed a plan for detection of diabetic retinopathy using different feature extraction and classification algorithms like EHD, CEDD, absolute distance, CNN(InceptionNet) on retinal fundus images from adults having diabetes. From the experiments performed, CNN (InceptionNet) gave the maxi- mum accuracy which is 96% and hence it can be concluded that this pair is the best fit for detection of diabetic retinopathy.In future, further research can be made by using differentalgorithms and bigger amount of data.

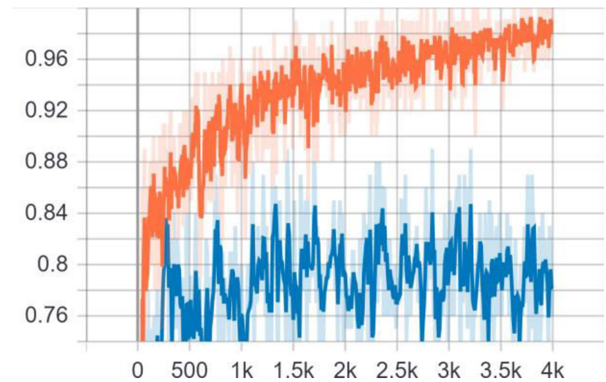


Fig. 12. Standard Deviation

REFERENCES

[1] D. U. N. Qomariah, H. Tjandrasa and C. Fatichah, ”Classification of Diabetic Retinopathy and Normal Retinal Images using CNN and SVM,” 2019 12th International Conference on Information Communication Technology and System (ICTS), Surabaya, Indonesia, 2019, pp. 152- 157, doi: 10.1109/ICTS.2019.8850940.

[2] Q. Chen, X. Sun, N. Zhang, Y. Cao and B. Liu, ”Mini Lesions Detection on Diabetic Retinopathy Images via Large Scale CNN Features,” 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), Portland, OR, USA, 2019, pp. 348-352, doi: 10.1109/IC- TAI.2019.00056.

[3] S. Qummar et al., ”A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection,” in IEEE Access, vol. 7, pp. 150530-150539, 2019, doi: 10.1109/ACCESS.2019.2947484.

[4] K. S. Lotlekar and S. D. Desai, ”Multilevel Classification Model for Diabetic Retinopathy,” 2018 International Conference on ComputationalTechniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, 2018, pp. 326-331, doi: 10.1109/CTEMS.2018.8769207.

[5] M. A. Habib Raj, M. A. Mamun and M. F. Faruk, ”CNN Based Diabetic Retinopathy Status Prediction Using Fundus Images,” 2020 IEEE Region 10 Symposium (TENSYMP), 2020, pp. 190-193, doi:10.1109/TENSYMP50017.2020.9230974.

[6] Javeria Amin, Muhammad Sharif, Mussarat Yasmin, ”A Review on Re- cent Developments for Detection of Diabetic Retinopathy”, Scientifica, vol. 2016, Article ID 6838976, 20 pages, 2016.

[7] Sahlsten, J., Jaskari, J., Kivinen, J. et al. Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema

Grading. Sci Rep9, 10750 (2019).

[8] Gulshan, Varun et al. "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs." JAMA 316 22 (2016): 2402-2410 .

[9] Y. Zhou, L. Li, T. Zhao and H. Zhang, "Region-based high-level seman- tics extraction with CEDD," 2010 2nd IEEE InternationalConference on Network Infrastructure and Digital Content, 2010, pp. 404-408.

[10] Chatzichristofis S.A., Boutalis Y.S. (2008) CEDD: Color and Edge Directivity Descriptor: A Compact Descriptor for Image Indexing and Retrieval. In: Gasteratos A., Vincze M., Tsotsos J.K. (eds) Computer Vision Systems. ICVS 2008. Lecture Notes in Computer Science, vol 5008. Springer, Berlin, Heidelberg.

[11] He Zhang and Xiuhua Jiang, "An improved algorithm based on texture feature extraction for image retrieval," 2015 12th International Confer- ence on Fuzzy Systems and Knowledge Discovery (FSKD), 2015, pp. 1270-1274, doi: 10.1109/FSKD.2015.7382125.

[12] Y. Zhou, L. Li, T. Zhao and H. Zhang, "Region-based high-level seman- tics extraction with CEDD," 2010 2nd IEEE International Conferenceon Network Infrastructure and Digital Content, 2010, pp. 404-408, doi: 10.1109/ICNIDC.2010.5657800.

[13] A. Nazir, R. Ashraf, T. Hamdani and N. Ali, "Content based image retrieval system by using HSV color histogram, discrete wavelet trans- form and edge histogram descriptor," 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), 2018, pp. 1-6, doi: 10.1109/ICOMET.2018.8346343.

[14] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015,pp. 1-9, doi: 10.1109/CVPR.2015.7298594.