Predictive Modeling Of Diabetic Retinopathy Disease By Use Of Machine Learning

Sandeep Prakash[1], Dr Pankaj Prajapati[2]

[1]*M.Tech Scholar, Industrial Automation & Robotics , Ambalika Institute of Management & Technology Lucknow, India*

[2]*Head Of Department, Industrial Automation & Robotics Ambalika Institute of Management & Technology Lucknow, India*

[1]*Email-thebestsandeep1@gmail.com*   [2]*Email-comp.officialmails@gmail.com*

Abstract*: According to a report by the World Health Organization (WHO), one of the leading causes of death by the end of 2030 will be diabetes, which is a serious disease. Timely treatment of this disease can prevent serious complications, including death. The number of people getting infected with diabetes is millions. The risk of getting this infection is common now a days and is more prevalent in women than men. Diagnosis process for diabetes is quite tedious. Diabetes retinopathy is a disorder that is caused by uncontrolled diabetes and can cause complete blindness if left untreated. Therefore, if detected early its treatment can prevent the unfavourable effects of diabetic retinopathy. The actual diagnosis of diabetes retinopathy by eye doctors takes a lot of time and patients need to suffer more during this time. Thus the latest achievements in science and technology makes it easier to predict the disease. The aim is to diagnose whether a person is diabetic or not using a phase-based machine learning method. This paper reviews, classifies and compares algorithms with previously suggested strategies to develop better and more efficient algorithms.*
Keywords: Diabetic Retinopathy, WHO, Classification based Learning techniques, Algorithms.

1. INTRODUCTION

Diabetes is a disease that affects the metabolic processes when a person has high blood sugar levels, either because the body does not produce enough insulin, or because the cells are not able to effectively use the insulin produced. High blood glucose in sugar is associated with damage to the retinal blood vessels, leading to diabetic retinopathy (DR). Diabetic retinopathy can cause blood vessels inside the retina to leak fluid or bleeding (bleeding), which can lead to blurred or paralyzed vision. In its later stages, new abnormal blood vessels enlarge the surface of the retina, which can cause scarring and cell loss in the retina. Diabetesic retinopathy is one of the most common complications of diabetes. It is a serious and widespread disease. The risk of this disease increases with age and as a result, older adults and adults with diabetes are at risk for Diabetic Retinopathy. People with DR whose eyes are at risk can be treated with laser, to prevent visual impairment or blindness. But at present there is no cure that can restore the vision that has already been lost. Detecting DR is a time-consuming and personalized procedure that requires a trained physician to examine and evaluate digital color fundus retina images. Hence early detection of Diabetic retinopathy is really important[1].
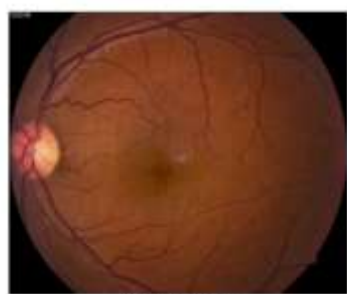
The Figures shown below represent the different stages or types of Diabetic Retinopathy severity of the disease if uncured timely.
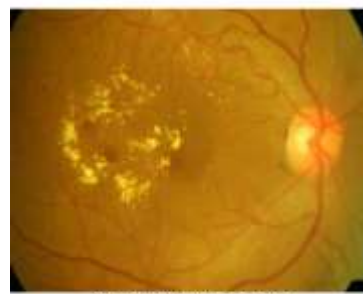
Figure A: Mild NPDR
Figure B: Moderate NPDR
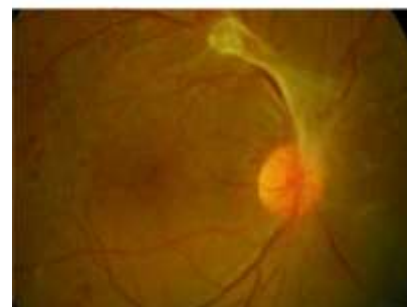Figure C: Severe NPDR
Figure D: PDR

(A): Mild NPDR

(B): Moderate NPDR with DME

(C): Severe NPDR

(D): PDR

A. SCOPE OF THE PRESENT WORK

In this paper, the principles of DR include the categories, causes, symptoms, and risks involved in it as well as the basic concepts of the human eye and body associated with DR. This study aims to provide guidance to the researcher to suggest a new way to diagnose DR.

According to this study, the main cause of DR is an abnormal increase in glucose levels. The first symptoms of DR are small capillary dilation known as micro aneurysms. It states that the development of DR causes neovascularization, macular edema and exudates and the area of cotton wool which in the later stage causes retinal detachment.

The four categories of DR are classified as:

1. Central Developmental Retinopathy: The first stage in which micro aneurysms occur.

2. Common Rheumatic Retinopathy: As the disease progresses, fewer blood vessels supply the blocked retina.

3. Severe Retinopathy Non-proliferative Retinopathy: Many blood vessels are blocked, which reduces blood flow to the retina. These areas send signals to the body to stimulate new blood vessels to be nourished.

4. Growing Retinopathy: This is an advanced stage, the symptoms that the retina sends to nourish itself cause the growth of new blood vessels. This condition is called proliferative retinopathy.

These new blood vessels are unusual and soft. These blood vessels alone do not hunt for symptoms or loss of vision. With small and weak walls, they can leak blood, causing severe vision loss and even blindness.The retina image is an important diagnostic tool and helps computer-assisted eye doctors to diagnose, treat, and diagnose various epidemic eye diseases including DR. They say there are no signs of DR in the early stages and any warning signs. In the first stage of Non proliferative diabetic retinopathy (NPDR), diagnosis can only be made with a fundus image. Reduced fluorescein angiography detection was performed to detect the posterior part of the eye. In the second stage, neovascularization develops in the back of the eye, which can lead to bloating, bleeding, and blurred vision, as the new blood vessels are weakened. It is not difficult when it first happens, it will leave a few spots in the vision. On a funduscopic examination, the doctor will notice cotton wool spots, bleeding flames, and dot bleeding.

2. LITERATURE REVIEW

A lot of work has been done in this field. Before starting the main task we go through the same paper to find out about the whole system such as what things we need to consider in order to detect diabetic retinopathy. AkaraS. , Matthew N. Dailey suggested "A mechanical study method for automatic exudate detection in retinal images from diabetic patients" [l]. In their paper they presented a series of character selection tests and classification of exudates using K- Nearest Neighbor (KNN) and vector machine class (SVM) classification separators. Rajendra Acharya U., E. U-Y. K. Ng, Kwan-Hoong Ng and Jasjit S. Suri introduced algorithms to automatically detect diabetic retinopathy using digital fundus images [2] in which they developed an algorithm used to exclude certain features from images. digital fundus. In addition, Varun G. and Lily P. have used in-depth study to diagnose diabetic retinopathy [3].

In the research paper "Diagnosis of Diabetic Retinopathy using Machine Learning" S. Gupta and K. AM have attempted to detect micro-aneurysms of the retina and extract retinal compounds from images [4]. After prior analysis, morphological operations are performed to determine the feature and the features are extracted such as GLCM and splat for classification. They achieved 87% sensitivity and clarity respectively with 86% accuracy.

Tiago T.G. in his paper "Machine Learning with Diabetic Retinopathy Debrecen Dataset" he used R language to predict diabetic retinopathy [5]. Use a database where features are extracted from photos of a diabetic patient. In his work he used eight division algorithms and also showed some comparisons. He earned 78% accuracy in his work.

Those are some of the related papers in our article where we took information and ideas to improve the new version. In our work we will use machine learning classification algorithms to differentiate diabetic retinopathy.

Automated Detection of Diabetic Retinopathy using Fluorescein Angiography Photographs [6]

Convolutional neural networks (CNNs) and audio-cutting techniques were used to identify the presence and severity of Diabetic Retinopathy in Fluorescein Angiography images. Data provided by EyePacs includes images of fundus with varying degrees of DR difficulty prescribed by doctors. The various CNN building models used were:

1. AlexNet: The first pre-trained model used by AlexNet [7]. The model is used for loading pre-trained weights, and also trains only the last layer that is fully connected to 4 classes out of 5 out of 1000. Loading a pre-trained model and re-training the final layer significantly improved the results produced by the foundation, and produced the first one. official results. A training accuracy of 72.9 percent was achieved, so they were able to clearly pass their data. Interestingly, as they continued to overflow (0.1 loss), the accuracy of the verification remained unfounded. AlexNet accuracy results for different classes are summarised below:

Classes	Accuracy
2	0.6695
3	0.5705
5	0.4073

2. GoogLeNet: The second pre-trained model used by GoogLeNet [8]. Pre-trained weights are loaded into the network, and retrain the final layer to predict 5 classes rather than 1000. GoogLeNet seemed to do a little (12%) better than AlexNet in almost every case where both tried. Like AlexNet, it has managed to achieve much higher training accuracy than verification accuracy (74.2%   vs. 41.7%), indicating that it was overloaded with training   data.

Below summarizes the accuracy results for all numbers of different classes:

Classes Accuracy

Classes	Accuracy
2	0.7105
3	0.5821
5	0.4168

3. Foundation: As a foundation, they have built a convolutional neural network from the beginning that acts as a control. The model is trained using random search hyperparameter. The basic structure is - The model was started using Xavier's launch program, and was updated using Adam.

The foundation has done very badly. In the case of Class 2, Class 3, and 5, it did better than random guesses in verifying and testing sets. When tested on a training set, however, it was able to perform significantly better than random guesses; it can be concluded that the network was able to at least read the parameters of the decision

	Accuracy	Recall	Precision
2nd class	0.541	0.502	0.489
3rd class	0.353	0.387	0.301
5th class	0.227	0.201	0.235

After analyzing the misplaced images, they saw a few things that were happening and the images were not properly sorted. First, the images in which the appearance of the black space were most often misinterpreted. To fix this, they tried to crop images where there was only an eye but not all images with blackspace, which caused them a few problems. Second, eye color was very different. Since colors are no longer evenly distributed between classes, models can learn that certain colors are compatible with certain classes. Black and negative images also cause a problem for model predictions. An immediate random prediction was made.

Machine learning approach to automatic exudate detection in retinal images from diabetic patients [9]

In this paper, they present a series of experiments on feature selection and classification using non-sense Bayes class dividers and vector machine (SVM) formats. They first matched the empty Bayes model in a training set that included 15 features drawn from both positive and negative images. Then make SVM in the best training set of naïve Bayes and repeatedly add features previously released to the filter. For each combination of features, they perform a grid search to determine the best combination of hyper-parameters (tolerance for training errors) and (breadth of base seradial function).They compare the best categories of mindless Bayes and SVM with the first category of nearby neighbours (NN) using the best feature sets from both dividers. Therefore, the senseless Bayes and SVM categories performed better than the NN classifier.

Prior to this, there have been several attempts to resolve the issue. A few are based on regional growth and development, special features and morphological reconstruction strategies. These techniques are very sensitive to image variability.

But the above work is mostly done on extended student photos as student extensions take time and are not appropriate for patients. So in this paper they found useful results for exudate in extinct fundus images.

Feature output provided the following features: Pixel intensity after pre-processing, typical deviation of the pre-considered density values in the pixel window, pixel color, number of edge pixels in the area around the pixel, intermediate pixel thickness. pixel set, pixel group size, measured by pixels, medium pixel density in pixel collection area, ratio between pixel cluster size and optic disc, distance between pixel cluster and optic disc, six differences in Gaus filter such as DoG1, DoG2, DoG3, DoG4, DoG5 and DoG6.

The Naïve Bayes theory is applied to all 15 aspects and the resulting separator has a total PR of 64.67%. Then a few features are removed to determine the amount of additional PR and this is done until the PR stops improving.

In the SVM algorithm excellent performance was achieved using 10 features with a high PR of 72.67%.

The NN category with Euclidean and Mahalanobis scores received a PR of 65.15 and 64.99%, respectively.

Both mindless Bayes and SVM separators sometimes miss out on exudates and find them wrongly as exudate image artifacts or retina structures that share certain features with exudates. It takes time as many experiments to select features need to be done.

Detection and Classification of Diabetic Retinopathy using Retinal Images [10]

The basis for the classification of different stages of diabetic retinopathy is the detection and measurement of blood vessels and bleeding present in the retina image. The retinal arteries are divided using the differences between the blood vessels and the surrounding back. Bleeding individuals were detected using density.analysis and binding box techniques. The classification of different eye diseases was performed using the Random Forests procedure based on location and circulatory system and bleeding. The objectives of this work are: vascular detection, bleeding detection, differentiation of common diabetic, moderate non-proliferative diabetic retinopathy (NPDR) and severe NPDR.

The advanced Parameters without advanced parameters - Range Forests (RF) are used for classification. Six features - the location and perimeter of each of the R, G, B sections of blood vessels and haemorrhages were excluded. The location is the number of white pixels (blood vessel and blood) present within the vessels and the boundary is determined by the number of pixels present on the vessel boundary. These extruded elements were used as input into the RF separator to separate the three categories of retina images. Ordinary cases were classified with 90% accuracy while Medium and robust NPDR cases were 87.5% accurate.

Detection of Diabetic Retinopathy Using K- Means Clustering and Self-Organizing Map [11]

The retinal detachment feature and isolation is a common procedure for diagnosing and diagnosing diabetic retinopathy. In this paper, they used the K-means algorithm used to extract the features and the separator used to separate the Editing Map (SOM).

First turn the images into a grey image and then use the K-means compilation. The release of the grey scale image after merging has six grey intensities. Each strength represents one set. The optic disc and rear are well integrated and clearly defined because K means integration.

The centroid values downloaded from the combination of K methods are then added to the neural network synthetic network filter map and the simplified values are points on the weight map. Neurons will then stop to surround these points to better explain them. This is how a self-organizing map trains its neurons. The moderator distinguishes cluster centroids from the common class and the category of diabetic retinopathy.

In this work, red, blue, green (RGB) layers of colored fundus image are integrated into the K merger algorithm that subtracts the centroid values for each collection in each RGB layer per image. In this paper, SOM is able to show a 3D heavy map showing the significant differences between normal images and diabetic retinopathy. But this model only differentiates whether DR exists or does not exist, it does not classify DR into different categories.

Features released are: blood vessels, Microaneurysms and strong exudates.

The 8 features used by the classier are: Normal Red Degree, Normal Degree Degree, Normal Degree Degree Blue, Concentration of Blood vessels, Potential number of microaneurysms, Real number of microaneurysms, Density of strong exudates, -entropy of the green part.

There are 2 main effects on this discovery of the DRNP and the division of the DRNP range. To find out, they used 301 retina images, 152 0th grade and 149th grade 3 retina. They trained the SVM section with all the features of these images and tested it with a 10-fold verification process. Performance was further enhanced by selecting the most appropriate features and parameters for SVM. To divide the distances they used 400 retina images. The SVM class of several classes (one to one) was trained in all aspects and then tested using 10-fold cross-sectional validation.

Their recommendation was tested on a database of 400 retina images labelled for grade 4 nonproliferative diabetic retinopathy. As a result, a high sensitivity of 94.6% and a predictive volume of 93.8% were obtained.

Detection of Diabetic Retinopathy using Image Processing and Machine Learning [12]

In this paper, the diagnosis of diabetic retinopathy in the image of the fundus is made by analyzing the images and methods of machine learning. The Probabilistic Neural Network (PNN) and Vector support equipment (SVM) two models were taken to detect diabetic retinopathy in the fundus model and their effects were analyzed and compared.

A powerful microscope called an ophthalmoscope or a fundus camera is connected to a digital camera and captures an internal image that includes the retina, optic discs, macula and blood vessels. macula. Images have a resolution of 1280 x 1024 or 700 x 605 Pixels. The image captured on the fundus camera is in 24-bit JPEG format. The gray scale conversion program converts an RGB image captured on a Fundus camera into a gray image.

Adaptive Histogram Equalization integrates several histograms and redistributes light values across the image and enhances image contrast values. Following the variable histogram measurement, the image in the image is removed using a modified filter response (MFR). The final step of the process involves combining with the Fuzzy c-method that makes the image's blood vessels more clearly visible and helps to measure the severity of the disease and the automatic diagnosis of diabetic retinopathy. After image processing, features such as Radius, Diameter, Area, Arc length, Center Angle and Half area are calculated for each image. The extracted elements are transmitted to machine learning algorithms such as Support Vector Machine and the Probabilistic Neural Network to categorize image into Normal, Non prolific Diabetic retinopathy (NPDR) and Prolific Diabetic Retinopathy (PDR).

The retina image has a rounded shape, which is why circular features are removed. After previous processing steps, the circular features of retina images such as Radius, Diameter, Area, Arc length, Center Angle and Half Area are calculated.

Results for PNN:

Classes	Training data	Testing data	Correctly classified	% Classification
Normal	10	10	6	60
NPDR&PDR	20	15	12	80
Average				70

Sensitivity 81.42%, **Specificity** 100%

3.PROPOSED SYSTEM

Problem Statement

"To detect diabetic retinopathy using machine learning."

Problem Elaboration

Diabetic retinopathy(DR) is a disease which causes blindness in people having diabetes. Currently, to detect DR, medical staff has to thoroughly examine images of the retina manually taken by the technique of Fundus photography. This is time consuming. We proposed a model to detect DR using machine learning techniques such as Neural networks to make the detection process automated as well as accurate.

Proposed Methodology

Machine learning consist of a number of stages to detect retinopathy in the fundus images that includes converting image to suitable input format, denoising and various preprocessing techniques. It also includes training a model with a training set and validating with a different testing set.

Method proposed in this project can be listed in two steps: Image Preprocessing, and Supervised learning and Feature Extraction.

First, the images are preprocessed. They are converted from RGB to grayscale. Proper resizing of image is also done. As the images are heterogeneous they compressed into a suitable size and format. Layer separation will also performed. For making intensity variations uniform histogram equalization to the image can be applied. Morphological operation will be done to remove the noise present in the background of retinal image. We then plan to use Convolutional neural network(CNN) architecture for feature extraction and prediction of the class of DR. A CNN is able to capture the temporal and spatial dependencies in images and fits better due to the decrease in parameters used and weight reusability. It has the ability to train to understand the complexity of the image more efficiently.

4. CONCLUSIONS

This paper summarizes our study and review of few literatures related to the detection of Diabetic retinopathy.

A number of studies use neural networks and image processing for detection using different architectures. It is certain that using machine learning techniques will give us good results along with good accuracy for prediction. In this report, we explored the potential usage of the CNN in retinal image classification. Due to the tedious manual methods by medical personnel, an automated system can reduce the labor involved in diagnosing large quantities of retinal images significantly.

5.REFERENCES

[1] Varun G., Lily P., Mark C., "Development and validation of a deep learning Algorithm for Detection of Diabetic Retinopathy", December 2016.

[2] Tiago T.G. "Machine Learning on the Diabetic Retinopathy Debrecen Dataset", knowledge-Based System60, 20-27. Published on June 25, 2016.

[3] Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care. 2012;35:556--64

[4] Boser B. E, Guyon I. M. and Vapnik V. N. (1992). "A training algorithm for optimal margin classiers".Proceedings of the 5th Annual Workshop on Computational Learning Theory COLT'92, 152 Pittsburgh, PA, USA. ACM Press, July 1992. On Page(s): 144-152

[5] Leske MC, Wu SY, Nemesure B, Hennis A Barbados Eye Studies Group. Causes of visual loss and their risk factors: An incidence summary from the Barbados Eye Studies. Rev PanamSaludPublica.2010;27:259-67

[6] Marco Alban, Tanner Gilligan, "Automated Detection of Diabetic Retinopa-thy using Fluorescein Angiography Photographs.

[7] Suvajit Dutta, Bonthala CS Manideep, Syed Muzamil Basha, Ronnie D. Caytiles1 and N. Ch. S. N. Iyengar, "Classification of Diabetic Retinopathy Images by Using Deep Learning Models"

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 25, pages 1097– 1105. Curran Associates, Inc., 2012.

[9] Akara Sopharak a* , Matthew N. Dailey b ,Bunyarit Uyyanonvara a Sarah Barman c ,Tom Williamson d , Khine Thet Nwe b and Yin Aye Moe b, "Machine learning approach to automatic exudate detection in retinal images from diabetic patients"

[10] Kanika Verma, Prakash Deep and A. G. Ramakrishnan, Senior Member, IEEE, "Research Paper 4:Detection and Classification of Diabetic Retinopa-thy using Retinal Images".

[11] Wong Li Yun and Muthu Rama Krishnan Mookiah Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489 , "Detection of Diabetic Retinopathy Using K-Means Clustering and Self-Organizing Map "

[12] Enrique V. Carrera , Ricardo Carrera, Andres Gonzalez, "Automated detection of diabetic retinopathy using SVM "

[13] Salman Sayed, Dr. Vandana Inamdar , Sangram Kapre , "Detection of Diabetic Retinopathy using Image Processing and Machine Learning".

[14] Shouyuan Yang, Zanjie Song and Jong Hyuk Park "High capacity CDMA Watermarking Scheme based on orthogonal Pseudo random subspace projection". International Conference on Multimedia and Ubiquitous Engineering, June 2011.

[15] Siwar Rekik, Driss Guerchi,Habib Hamam & Sid-Ahmed Selouani "Audio Steganography Coding Using the Discrete Wavelet Transforms".International Journal of Computer Science and Security (IJCSS), Volume (6) : Issue (1) : 2012.

[16] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions, 2014. Engineering,June2011

[17] Lionel Fillatre "Adaptive Steganalysis of Least Significant Bit Replacement in Grayscale Natural Images" IEEE Transactions on Signal Processing ,Vol. 60, No. 2, February 2012

[18] R.R.Ahirwal, Deep chand Ahirwal and Jpgendar jain "A High Capacitive and Confidentiality based Image Steganography using Private Stego key" International coference on Information Science and applications, Feb 2010.

[19] Rikzy M. Naguraha "Implementation of Direct sequence Spread Spectrum on Audio Data" International Conference on Informatics Engineering, June 2011.

[20] Siwar Rekik, Driss Guerchi,Habib Hamam & Sid-Ahmed Selouani "Audio Steganography Coding Using the Discrete Wavelet Transforms".International Journal of Computer Science and Security (IJCSS), Volume (6) : Issue (1) : 2012