Identifying Diabetic Retinopathy through image processing in the field of Machine Learning

Mr. R.Ezhilarasan[1] , Mr.S.Sridharan[2] , Mr. S. Sivatharanish[3] , Mr. I. Janeshpraveen[4] , Mr. M.Suryaprakash[5] , Dr .R. Umamaheswari[6]

[1]Assistant Professor, Department of CSE, Gnanamani College of Technology, Namakkal, Tamilnadu, India [2, 3, 4, 5] UG Students, Department of CSE, Gnanamani College of Technology, Namakkal, Tamilnadu, India [6] Professor and Head Of Department /CSE ,Gnanamani college of Technology,Namakkal, Tamilnadu, India

ABSTRACT - Diabetic retinopathy (DR) is a condition affecting the eyes, resulting from damage to the retina due to prolonged diabetic mellitus. Detecting and diagnosing DR typically involves three stages: preprocessing of color fundus images, extracting diagnostic features, and classifying DR severity. Utilizing various image processing algorithms, researchers can determine the thickness of blood vessels, a crucial indicator of diabetes severity.In the realm of healthcare, predictive analytics plays a vital role, enabling practitioners to make informed decisions based on extensive data. While predictive analytics in healthcare presents challenges, it offers the potential for timely interventions and treatment decisions. This paper delves into predictive analytics within healthcare, employing six distinct machine learning algorithms. To conduct experiments, a dataset comprising patient medical records is utilized, and these algorithms are applied to analyze the data.Furthermore, the challenges and future directions of ML-based DR detection in improving screening efficiency and enhancing patient outcomes are explored. The integration of ML algorithms into clinical practice holds the potential to revolutionize DR screening by providing accurate, timely, and cost- effective solutions for early diagnosis and intervention.

Index Terms-Diabetes Retinopathy, Machine Learning, Deep Learning, (KNN) K-Nearest Neighbors, (SVM) Support Vector Machine and Detection,Classification

I.INTRODUCTION

Healthcare industry contains very large and sensitive data and needs to be handled very carefully. Diabetes Mellitus is one of the growing extremely fatal diseases all over the world. Medical professionals want a reliable prediction system to diagnose Diabetes. Different machine learning techniques are useful for examining the data from diverse perspectives and synopsizing it into valuable information. The accessibility and availability of huge amounts of data will be able to provide us useful knowledge if certain data mining techniques are applied to it. The main goal is to determine new patterns and then to interpret these patterns to deliver significant and useful information for the users.Diabetes contributes to heart disease, kidney disease, nerve damage, and blindness.Mining the diabetes data in an efficient way is a crucial concern. The data mining techniques and methods will be discovered to find the appropriate approaches and techniques for efficient classification of Diabetes dataset and in extracting valuable patterns. In this study, we aim to apply the bootstrapping resembling technique to enhance the accuracy and then applying Naïve Bayes, Decision Trees and (KNN) and compare their performance. With an increase in the computing and processing power of computers and the development of image processing techniques in recent years, the idea of using a computer to analyze medical images and automatically diagnose diseases has attracted the attention of many medical and computer specialists. This idea of analyzing retinal images and diagnosing ocular and vascular diseases is very efficient and inexpensive. Retinal vascular structures contain a great deal of medical information. Diabetes, hypertension, cardiovasculadiseases, and some others can only be detected through the examination of retinal fundus images. Monitoring is one of the acknowledged methods in examining patients. Annual screening of the eye fundus of diabetic patients is one of the essential parts oftheir treatment. However, the analysis of retinal fundus images demands an ophthalmologist expert.An automated detection of DR, classifies each patient into one of these classes through the examination of their retinal fundus images. Generally, there are five different DR categories: No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. An automated detection of DR, classifies each patient into one of these classes through the examination of their retinal fundus images. medical research with higher detection performance.

II. LITERATURE SURVEY

Graham is the winner achieving the accuracy score of 0.84957. He is the first preprocessed the retina images to remove the illumination difference and used a convolutional neural network, SparseConvNet and random forest for classification by augmenting the retina images to increase the number of images in the training set [17]. [Farrikh Alzami, 2019] described a system for diabetic retinopathy grade classification based on fractal analysis and random forest using MESSIDOR dataset. Their system segmented the images, then computed the fractal dimensions as features. They failed to distinguish mild diabetic retinopathy to severe diabetic retinopathy.

[Qomariah 2019] proffered an automated system for classification of Diabetic Retinopathy and normal retinal images using concurrent neural network (CNN) and support vector machine (SVM). Features comprised of exudates, haemorrhage and microaneurysms. The author partitioned the proposed system into 2 parts: the first part composed with feature extraction based on neural networks and the second part performed classification using SVM. [18] proposed using Remidio Fundus on Phone (FOP) device to capture high- quality retina images compared with thetraditional fundus devices. The FOP is a high- quality portable fundus camera that is capable of capturing wide retinal color photography covering macula, nasal to the optic disc, superior-temporal andinferior-temporal quadrants. The EyeArt system also offers image processing and machine learning techniques such as image gradeability, image enhancement, image restoration, interest region detection, and descriptor computation. Compared with manually graded results by two ophthalmologists, deep artificial neural network method shows high sensitivity and specificity for retina images captured by FOP device.

FOP proves the concept of smartphone-based designs and shows the technological and economic feasibility of the portable retinal imaging systems. However, due to their fewer controllable parameters such asaperture and inexpensive lenses, smartphone-based systems have a lower image quality compared to the fundus camera and FOP. Therefore, the existing algorithms could not be applied directly to the retinal images capturedwith smartphone-based retinal imaging systems because the quality of captured retina images plays an important role in the accuracy of the deep learning techniques. They have used CNN methods for image classification and trained their algorithm with Inception-v3 architecture [23]with about 128175 images using EyePACS and 8Messidor-2 datasets. They have mentioned the variety of DR problems such as referable Diabetic Retinopathy (rDR), vision- threatening Diabetic Retinopathy (vtDR), and referable Diabetic Macular Edema (rDME). Based on the results, they suggest that such a CNN network gives the best results when the network trained with 60000 images in terms of sensitivity and specificity. For the sensitivity, they reached about 97%. In order to apply these results to the clinical environment, more research is needed. They obtained an accuracy of 97.5% and 96.1% choosing two different threshold values. Abramoff et al. [24] developed Iowa detection program for detecting rDR and they have used their own DR database and publicly available Messidor-2 dataset for training and testing, respectively.

Based on the test results on the Mesidor-2 dataset, they achieved 96.8% sensitivity and 59.4% specificity for detecting rDR. After that, they improved their results and reported 96.8% sensitivity and 87% specificity for detecting rDR using machine learning methods. [25] Gargeya et al. [26] used a customized CNN technique to detect DR. They trained their system with 75137 fundus images from their own dataset and tested with Messidor-2 and E-Optha datasets. They classified images into two categories, one with the healthy eyes, the other with any DR stage, in other words, mild and worse DR. They acquired 94% sensitivity and 98% specificity from their own dataset. Also, they tested their model with

Messidor-2 dataset, and they achieved 93% sensitivity and 87% specificity. Philip et al.

[27] developed a DR assessment system based on healthy and disease conditions, also known as mild DR and worse. They trained their algorithm with 1067 images and tested with 14406 images. The performance of their algorithm was 86.2% and 76.8% with sensitivity and specificity, respectively. 9Carson et al. [28] introduced a CNN based deep learning techniques to detect DR using various classification models including but not limited to 2-ary, 3-ary, and 4-ary. Pretrained AlexNet and GoogLeNet models were investigated and transfer learning approaches were applied using Kaggle EyePACS and Messidor-1 dataset. They suggested using image processing techniques to increase validation accuracy especially for the detection of mild DR such as image normalization and contrast adjustments using histogram equalization. They augmented the retina images to increase the number of images in the training set and prevent overfitting. They received 95% sensitivity in the validation set and its accuracy in the testing set 74.5%, 68.75%, and 51.25% for 2- ary, 3- ary, and 4-ary models, respectively. Pires et al. proposed a solution for detecting rDR using data-driven approaches [29]. They used transfer learning techniques by applying to CNN. They applied on their training stage to data augmentation, multi-resolution, feature extraction, per patient analysis, and testing their solution with cross dataset logicby using Kaggle EyePACS dataset as a training, Messidor-2 dataset for testing. Based on the results, they obtained a 98.2% Receiver Operating Characteristic (ROC) curve for predicting rDR. Since these methods in the literature focus on fundus images, they cannot be easily applied to smartphone-based images. For this reason, we need to investigate and create our own synthetic DR dataset using the FoV approach.

PROBLEM STATEMENT

The problem statement aims to address the development of automated systems for diabetic retinopathy detection using machine learning and image processing techniques. This involves designing algorithms capable of accurately analyzing retinal images to identify signs of DR, thus enabling early intervention and preventing vision loss. Key challenges include optimizing algorithm performance for diverse image datasets, ensuring robustness against variations in image quality, and integrating the developed systems into existing healthcare infrastructure for widespread adoption. The successful implementation of automated DR detection systems has the potential to revolutionize diabetic eye care by improving screening efficiency, reducing healthcare costs, and ultimately preserving vision for millions of individuals worldwide.

ALGORITHM

Support Vector Machine (SVM) is a supervised Machine Learning model (a dataset has been labeled). It means if we have a dataset a try to run SVM on it, we will getexperiments of various multi-label learning strategies were performed to substantiate the effectiveness and utility of the planned framework . The sickness Prediction in health care system is Associate in nursing data system that gives data and customized info to users in enhancing health and care outcomes. Hence we use SVM algorithm to handle multi structured health care data for more effective and accurate outcomes.often pretty good results. This is because it is based on the strong and beautiful mathematical background.This algorithm is used for both classification (SVC) and for regression (SVR) also. In this post we will mainly focus on the classification and look at the main idea behind SVM. often pretty good results. This is because it is based on the strong and beautiful mathematical background.This algorithm is used for both classification (SVC) and for regression (SVR) also.

KNN -k Nearest Neighbour

The k-Nearest Neighbors (KNN) algorithm is a straightforward and effective approach in machine learning, especially suited for classification tasks and managing extensive datasets. It operates by assuming similarity between new data and existing data points, placing the new data into the category most resembling the available categories. KNN is versatile and can be used for both classification and regression tasks, although it's primarily utilized for classification. One of its key advantages is its non-parametric nature, meaning it doesn't assume any underlying data distribution. KNN is often referred to as a lazy learner because it stores the dataset and performs actions on it during classification rather than learning immediately from the training set. Additionally, KNN is known for its simplicity in implementation, robustness to noisy training data, and potential effectiveness with large training datasets.

Advantages:

- ✔ It is simple to implement.
- ✔ It is robust to the noisy training data
- ✔ It can be more effective if the training data is large

Working

The process of detecting diabetic retinopathy using the k-Nearest Neighbors (KNN) algorithm starts with gathering a dataset containing images of retinas, specifically focusing on diabetic patients and healthy individuals. These images are then preprocessed to enhance quality, remove noise, and ensure consistency. Subsequently, feature extraction is performed to identify relevant characteristics such as blood vessel patterns and abnormalities unique to diabetic retinopathy.The next step involves labeling the images based on whether they exhibit signs of diabetic retinopathy, enabling supervised learning. The dataset is divided into training and validation sets for model development and evaluation purposes. The KNN model is trained using the training set to recognize patterns associated with diabetic

performance metrics such as accuracy, precision, recall, and F1 score. This evaluation helps determine the model's effectiveness in detecting diabetic retinopathy. Subsequently, the model can be tested on new, unseen images to validate its real-world applicability.To create a functional model, implement these steps

C. Hardware requirements:

They using a programming language or machine learning framework like Python with libraries such as scikit-learn. Load and preprocess image data, extract relevant features, split the dataset, train the KNN

model, evaluate its performance, and finally test it on new images to detect diabetic retinopathy. Fine-tune the model as needed for optimal results, ensuring a robust and effective diabetic retinopathy detection system..

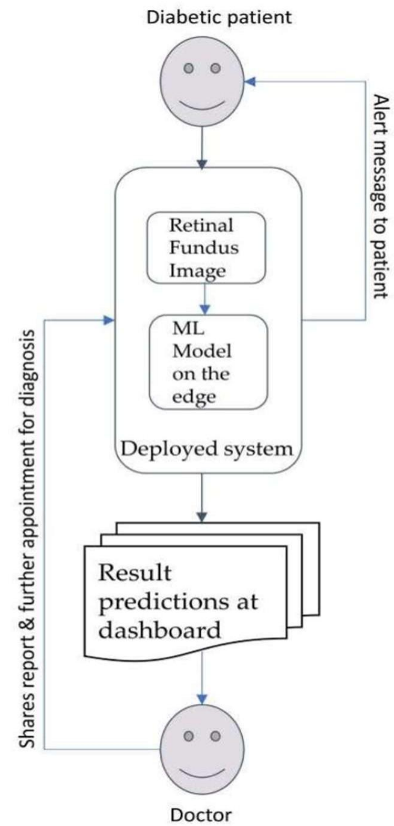

Fig.2 Working Model of The Application System

CONCLUSION

With the use of this system, diabetic patients will be able to identify diabetic retinopathy more easily and efficiently. Users don't have to be concerned about the confidentiality or security of the photos they submit. Additionally, since all of the data is kept up to date in a database and is retrievable or editable as needed, the user doesn't have to worry about it being saved. The administrator occasionally keeps an eye on, secures, and maintains the portal. Our model receives fundus images for the left and right eyes as inputs and converts them into blocks that resemble Siamese. The fully linked layer receives the data from the two eyes, and the model ultimately outputs the diagnosis results for each eye separately.

FUTURE ENCHANCEMENT

The current state of research on diabetic retinopathy is outlined in this project, along with potential directions for future study. Improved metabolic management will result from new understandings of the biology of diabetes and diabetic retinopathy. Analyses of structure and function are providing new information on diabetic retinopathy. Intraocular medication therapy yields better visual results. When combined, these actions will improve methods for identifying and measuring vision loss as well as for creating patient-specific therapies that help diabetics maintain their vision.

REFERENCES

[1.] A. D. Fleming, K. A. Goatman, and J. A. Olson, " The role of haemorrhage and exudate detection in automated grading of diabetic retinopathy," British Journal of Ophthalmology, vol 94, no. 6, pp. 706- 711, 2010 .
[2.] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.
[3.] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.
[4.] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State
[5.] Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.
[6.] C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant

[7.] Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
[8.] Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis and Implementation of 89C51 Controller
[9.] Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools & Technology. 2022; 12(2): 34–41p.
[10.] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.
[11.] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756
[12.] G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:744-749
[13.] C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
[14.] M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
[15.] M Suganthi, N Ramesh, CT Sivakumar, K Vidhya, "Physiochemical Analysis of Ground Water used for Domestic needs in the Area
[16.] of Perundurai in Erode District", International Research Journal of Multidisciplinary Technovation, pp: 630-635, 2019
[17.] Ahmad Fadzil, et. Al, " Analysis of Foveal Avascular Zone in Colour Fundus Images for Grading of Diabetic Retinopathy Severity" , 32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina, August 31 - September 4, 2010, pp. 5632-5635 .
[18.] Gang Luo, OpasChutatape, Huiqi Lei, Shankar. M. Krishnan, "Abnormality Detection in Automated Mass Screening System of Diabetic Retinopathy", available online at "http://luoinnovation.tech.officelive.com
[19.] /l uo research/cbms2001.pdf" (accessed on March 3, 2013).
[20.] Keerthi Ram, Gopal Datt Joshi, and JayanthiSivaswam, "A Successive ClutterRejectionBased Approach for Early Detection of Diabetic Retinopathy", IEEE Transactions On Biomedical Engineering, Vol. 58, No. 3, March 2011, pp.664-673.National Eye Institute, National Institutes of Health, " Diabetic Retinopathy: What you should know," Booklet, NIH Publication, no: 06-2171, 2003.
[21.] Omar, M. Hanafi, S. Mashohor, N. Mahfudz, and M. Muna'im, "Automaticdiabetic retinopathy detection and classification system," in 2017 7th IEEE International Conference on System Engineering and Technology (ICSET), pp. 162– 166, IEEE, 2017 .
[22.] R. Sivakumar, G. Ravindran, M. Muthayya, S. Lakshminarayanan, and C. U. Velmurughendran, " Diabetic Retinopathy Analysis" , Journal of Biomedicine and Biotechnology, 2005 pp. 20– 27. Z.
[23.] S. R. Nirmala, M. K. Nath, and S. Dandapat, " Retinal Image Analysis: A Review," International Journal of Computer & Communication Technology (IJCCT), vol-2, pp. 11- 15, 2011.
[24.] Shijian Lu, et al., " Automatic Fundus Image Classification for Computer-Aided Diagnosis" , 31st Annual International Conference of the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6, 2009, pp.1453-1456.
[25.] XialongZeng,Haiquan Chen, Yuan Lao, Wenbin Ye," Automated Diabetic Retinopathy Detection on Binocular Vol 09 Issue01, Jan 2020 ISSN 2456– 5083 Page 8 Siamese-Like Convolution Neural Network" ,February 2019.
[26.] Y. Haniza, A. Hamzah, and M. Norrima, "Edge sharpening for diabetic retinopathy detection," IEEE Conference Publications, 2010