DIABETES DISEASE PREDICTION USING MACHINE LEARNING ALGORITHMS

[1]SHASHANK K M, [2]SHRUTI M T

[1] Student, Department of MCA, BIET, Davangere

[2]Assistant Professor Department of MCA, BIET, Davangere

ABSTRACT

Diabetes can lead to a condition known as diabetic retinopathy (DR), which affects the eyes by damaging the blood vessels in the retina's light-sensitive tissue. When diabetes is poorly managed, DR becomes the leading cause of blindness among working-age individuals. Current methods for detecting diabetic retinopathy rely on manual examination of retinal images by ophthalmologists. The proposed approach aims to automate this detection process. The model was trained on 35,126 retinal scans, accessible publicly on Kaggle through eyePACS, using a GPU. It achieved an accuracy rate of approximately 81%.

KeyWords:DRdetection,GPU,eyePACS.

I. INTRODUCTION

Deep learning models, characterized by multiple layers of processing, excel in learning data representations across varying levels of abstraction, revolutionizing fields such as drug discovery, genomics, object recognition, and speech analysis. Backpropagation is instrumental in adjusting internal parameters layer by layer to derive these representations. Recurrent neural networks excel with sequential data like text and speech, while deep

edges. Subsequent layers may then combine these patterns into recognizable objects. Notably, deep learning automates the learning of these hierarchical features from data, a departure from manually engineered approaches. This capability is pivotal in addressing long-standing challenges in artificial intelligence.

convolutional networks advance the processing of images, videos, and audio.

Modern society heavily relies on machine learning, influencing social media content filtering, e-commerce recommendations, and everyday consumer technologies like cameras and smartphones. Deep learning methods, which facilitate representation learning through stacked non-linear modules, enable the analysis of raw natural data more effectively than traditional machine learning. By progressively transforming input data into higher-level abstractions, these models can learn complex functions and emphasize features crucial for discrimination while reducing irrelevant variations.

For instance, in image processing, initial layers may detect basic edge orientations, followed by layers identifying patterns composed of these and Micro-aneurysms (HEM) that form in the retina. To avoid blindness, HEM must be diagnosed as soon as possible. DR detection methods have historically made extensive use of texture features like LBP. The utilisation of various texture features for DR is introduced in this paper, with a focus on the Local Ternary Pattern (LTP) and the Local Energy-based Shape Histogram (LESH).

II. LITERATURE SURVEY

Carrera, R. **[1]** Proposed that The most prevalent cause of blindness in the general population and a common eye disease in diabetes individuals is diabetic retinopathy. Patients who have diabetic retinopathy are protected from losing their vision by early identification. Thus, in order to assist individuals in early detection of diabetic retinopathy, this research suggests a computer-assisted diagnosis based on the digital processing of retinal pictures. The major objective is to automatically assess any retinal image for non-proliferative diabetic retinopathy. For this, blood vessels, microaneurysms, and hard exudates are isolated during the initial image processing step in order to extract characteristics that may be used by a support vector machine to determine the retinopathy grade of each retinal picture. 400 retinal photographs from a database were used to test this proposal.

Kardouchi, M. **[2]** prposed that Diabetes mellitus can induce diabetic retinopathy (DR), a medical disorder that can harm the retina of the patient and lead to blood clots. If this illness is not promptly addressed, it can result in a variety of symptoms, from minor vision issues to total blindness. The early symptoms of DR include haemorrhages, hard exudates,

Sven Loncaric [3] proposed that Retinal vessel segmentation and extracting features such as tortuosity, width, length related to those vessels can be used in diagnosis, treatment and screening of many diseases such as retinopathy of prematurity, hypertension and diabetes. Therefore, automatic segmentation of vessels by computers will make the analysis of those diseases easier and will help during the screening, diagnosis and treatment processes. In this study, a solution based on convolutional neural networks (CNN) is proposed for automatic segmentation of retinal vessels. The proposed CNN model is tested on DRIVE dataset and a better performance than literature is achieved.

Bhagwan and Kumar [4]
This study improves the method for detecting diabetic retinopathy by extracting precise information about the location and number of microaneurysms from colour fundus images. For the early detection and treatment of diabetic retinopathy, routine eye examinations are essential. The eye condition known as diabetic retinopathy (DR) is brought on by long-term diabetes mellitus sickness and retinal damage. Microaneurysms (MA) are tiny red spots that form on the retina as a blood vessel's more delicate parts inflate. The primary stage of MA must be identified, and this is the first step towards DR inhibition. Several techniques have been put forth for the identification and diagnosis of DR.
Manjramkar M [5] proposed that This digital document iA condition known as diabetic retinopathy affects the tiny blood vessels that supply the macula and other key areas. The primary cause of visual loss is this gradual illness. Patients with diabetes are susceptible to diabetic retinopathy, a retinal vascular disease. If the level of diabetes is exceedingly high, this damages the retina of the eye and results in vision impairment. There are no warning indications of diabetic retinopathy. In certain circumstances, eyesight can improve or deteriorate over the day. Therefore, the value of automatic macular enema assessment rose. We conducted a survey of the many methods for detecting diabetic retinopathy in this article. A distinctive category of lesions discovered in the retina called diabetic retinopathy consists of

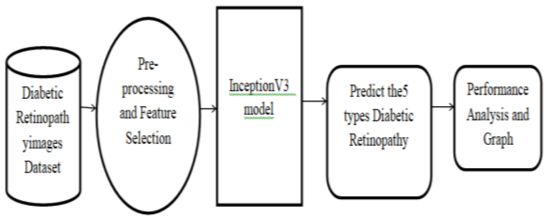

Fig. 1. Proposed Architecture

III. EXISTING MODEL

In their study, Enrique Carrera et al. suggested a method based on SVM to aid in the early diagnosis of diabetic retinopathy. In this method, the blood vessels, microaneurysms, and hard exudates are separated during the preprocessing stage in order to extract features for SVM. The sensitivity of the model tested on the STARE dataset was 94.6%.

The solution suggested by Mohamed Chetoui et al. uses the Local Ternary Pattern (LTP) and Local Energy based Shape Histogram (LESH), which perform better than the LBP derived features. On the 1200 image MESSIDOR database, this model had an accuracy of 90.4%.

To divide blood arteries, Martina Melinscak et al. used a deep convolutional neural network. The model, which had ten layers, had a 94% accuracy rate on the DRIVE dataset, which was made available to the public.

IV.PROPOSED METHODOLOGY

In this system, we provide a deep-learning-based automated method for detecting the stage of diabetic retinopathy using a single image of the human fundus. The suggested approach can be employed as a screening technique for sensitive early diagnosis of diabetic retinopathy. The model that was shown was created in Python using Keras and a TensorFlow backend. Utilising the Inception V3 framework, the suggested system is put into practise.

 The dataset that was used in the proposed study is free to obtain from online resources. All photos were taken with various cameras, different lenses, and various sizes of subjects. Due to the high level of noise in the data, numerous preprocessing processes were used to convert each image into a format appropriate for training.

A universal deep learning model for diagnosing diabetic retinopathy was created based on the experiment, and it could be used to any database.

We offered a straightforward technique that can be used to different medical photos to correct the imbalance of DR datasets.

Achieved a 95% accuracy score, which is higher than the models of the current systems.

V. IMPLIMENTATION

Dataset: We created a system in the first module to obtain the input dataset for training and testing. The data collection is located in the model folder.

There are 2222 photos of diabetic retinopathy in the dataset.

Imbalance Dataset check.

Python will be the language we use for this. In order to build the primary model, partition the training and test data using Sklearn, turn photos into arrays of numbers using PIL, and use other libraries like pandas, numpy, matplotlib, and tensorflow, we must first import the appropriate libraries.

Retrieving the images:
The photos and their labels will be retrieved. The photos should then be resized to (224,224) since they all need to be the same size for recognition. Then create a numpy array from the images.

Splitting the dataset:
Create train and test datasets. 80 percent train data, 20 percent test data.

Building the model:
Convolutional neural networks as an idea.In terms of image recognition, they excel. The convolution operation, which sets CNN apart from conventional neural networks, is the crucial component to comprehend. CNN repeatedly analyses an image once it is entered to look for specific traits. Stride and padding type are the two key factors that can be adjusted for this scanning (convolution). As can be seen in the image below, the initial convolution process results in a collection of new frames, which are displayed in the second column (layer). Each frame comprises details about a specific feature and the presence of the feature in the scanned image. The resulting frame will have higher values in locations where a feature is clearly visible and lower values in locations where there are few or no such characteristics. After that, the procedure

VI. RESULT

Diabetes disease prediction using machine learning involves several critical steps. It starts with the collection of relevant patient data, including demographic information, lifestyle factors, medical history, and clinical test results. Data preprocessing is then performed to handle missing values, normalize data, and select pertinent features. Various machine learning algorithms, such as logistic regression, decision trees, support vector machines, and neural networks, are trained on this data to identify patterns indicative of diabetes. The performance of these models is evaluated using metrics like accuracy, precision, recall, and the area under the ROC curve (AUC) to ensure reliability. The goal is to develop predictive models that can effectively identify individuals at risk of developing diabetes, facilitating early intervention and better disease management.

VII. REFERENCES

[1] "Automated detection of diabeticretinopathy using SVM," 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Cusco, 2017, pp. 1-4. E. V. Carrera, A. Gonzalez, and R. Carrera.

[2] "Diabetic RetinopathyDetection Using Machine Learning and Texture Features," 2018 IEEE Canadian Conference on Electrical & Computer

Engineering (CCECE), Quebec City, QC, 2018, pp. 1-4.

[3] "Retinal vesselsegmentation using deep neural networks," Martina Melinscak, Pavle Prentasic, and Sven Loncaric, 2015, pp 577–582.

[4] S. Kumar and B. Kumar, "Diabetic Retinopathy Detection by ExtractingArea and Number of Microaneurysm from Colour Fundus Image," 5th International Conference on Signal Processing and IntegratedNetworks (SPIN), Noida, 2018, pp. 359-364.

[5] M. Manjramkar, "Survey of Diabetic Retinopathy Screening Methods," in 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, pp. 1-6.

[6] "Diagnosis of Diabetic Retinopathy using Machine Learning Classification Algorithm," Karan Bhatia, Shikhar Arora, and Ravi Tomar.

[7] "Computer aided approachfor proliferative diabetic retinopathy detection in colour retinal images," 2016 International Conference on Next Generation Intelligent Systems(ICNGIS), Kottayam, 2016, pp. 1-6.

[8] S. Kanth, A. Jaiswal, and M. Kakkar, "Identification of different stagesof Diabetic Retinopathy using artificial neural network," 2013 Sixth International Conference on Contemporary Computing (IC3), Noida, 2013, pp. 479-484.

[9] "Automatic Diabetic RetinopathyDetection Using Digital Image Processing," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2018, pp. 0072–0076.

[10] "Feature Selectionof Diabetic Retinopathy Disease Using Particle Swarm Optimisation andNeural Network," 6th International Conference on Cyber and ITService Management (CITSM), Parapat, Indonesia, 2018, pp. 1-4.