Retinopathy Detection

Aya Sameh
Systems and biomedical engineering
Cairo university
Cairo, Egypt
aya.ahmed012@eng-std.cu.edu.eg

Ehab Kamal
Systems and biomedical engineering
Cairo university
Cairo, Egypt
ehab.mahmoud00@eng-st.cu.edu.eg

Hanya Ahmad
Systems and biomedical engineering
Cairo university
Cairo, Egypt
hanya.taweel01@eng-st.cu.edu.eg

Mohammed Hashem
Systems and biomedical engineering
Cairo university
Cairo, Egypt
mohamed.abdelwareth02@eng-st.cu.edu.eg

ABSTRACT

Diabetic Retinopathy (DR) is a complication that can arise from diabetes, specifically when high blood glucose levels cause damage to the small vessels in the retina that supply nutrients and oxygen. Initially, diabetic retinopathy was diagnosed through a Fundoscopic Exam, but it is now crucial to have an automated detection method for early detection and prevention of severe vision loss. This study proposes a method for detecting diabetic retinopathy from retinal fundus images using advanced machine learning techniques. The dataset used (1) has 757 color fundus retinal images, out of which 151 images were reserved for testing purposes, acquired at the Department of Ophthalmology of the Hospital de Clínicas, Facultad de Ciencias Médicas, Universidad Nacional de Asunción, Paraguay. Expert ophthalmologists have classified the dataset. The method proposed obtained an accuracy of 70.4% using random forest ensemble classifier for multiclass classification and 88.81% for binary classification.

Keywords—decision tree, random forest, classification, machine learning, diabetic retinopathy.

I. INTRODUCTION

The IDF Diabetes Atlas 10th edition reported in 2021 that 537 million adults (20-79 years) suffer from diabetes.. This number is predicted to rise to 643 million by 2030 and 783 million by 2045. The figure below illustrates the number of diabetic patients around the world. An estimated 26.7% of people with diabetes have Diabetic Retinopathy (DR). This translates to approximately 143 million people with diabetic retinopathy worldwide. (2) .

The early detection of diabetic retinopathy (DR) is crucial, as it is characterized by abnormal and leaky blood vessels in the retina, which can lead to various lesions such as microaneurysms, hemorrhages, cotton wool spots, and hard exudates. Exudates, which are divided into soft and hard exudates, are important indicators of DR in the primary stages.

DR is categorized into two major classes: non-proliferative DR (NPDR) and proliferative DR (PDR), with microaneurysms, hemorrhages, and red lesions being associated with NPDR. While manual screening of the retina can detect DR, it is time-consuming and not efficient for large populations. To address this issue, computer-aided diagnosis (CAD) procedures are being used to develop an automated framework which takes a retinal image as an input, and outputs the classified DR stage as PDR, NPDR, or Normal in the case of multiclass classification, and outputs Normal or PDR in the case of binary classification. This involves pre-processing the input retinal fundus image through grayscale conversion and contrast enhancement, extracting features, generating a feature vector, and finally applying machine learning models for image classification.

II. RELATED WORK

Many studies in the literature have explored automatic detection methods for diabetic retinopathy. These methods can typically be divided into two categories based on the techniques used: traditional machine learning algorithms and deep learning-based approaches. In traditional machine learning methods, external feature extraction techniques using image processing are required, while in deep learning-based approaches, automatic feature extraction is performed.

3.1. retinopathy detection using traditional machine learning algorithms that rely on handcrafted feature extraction

Pawar and Agrawal [3] conducted a study on diabetic retinopathy detection using a dataset of approximately 1040 retinal images, with 200 of them used for testing. They achieved an accuracy of 74.28% and 82.85% for multiclass and binary classification, respectively, using a decision tree classifier. The study involved preprocessing, abnormality detection, and feature extraction.

Odeh, Alkasassbeh, and alauthman [4] used an ensemble machine learning approach for diabetic retinopathy detection, where specific features extracted from the MESSIDOR public fundus dataset were grouped using feature selection algorithms. The system employed several classification algorithms, such as random forest, neural network, and support vector machine, to improve prediction accuracy, with the final predictions produced by a meta-classifier. Feature extraction methods and image analysis were not performed in this study.

Qomariah [5] proposed a diabetic retinopathy detection system that used convolutional neural networks (CNNs) for feature extraction and support vector machine (SVM) for classification. The system aimed to detect exudates, bleeding, and microaneurysms and achieved promising results.

Kahai et al. [6] proposed an automatic diabetic retinopathy detection system that used Bayesian criteria and image processing for feature extraction. The system achieved high sensitivity and good specificity.

3.2. retinopathy detection using deep learning algorithms that rely on automatic feature extraction.

Antal and Hajdu's [7] proposed a system that integrates multiple comparison components used in previous works. These components include image quality assessment, lesion-specific analysis, multi-scale AM/FM-based feature extraction, pre-screening, and anatomical components such as macula and optic disk detection. The system employs an ensemble-based decision-making approach, achieved by training several well-known classifiers alongside energy functions and fusion strategies for ensemble selection. The inclusion of classifiers in the system is based on their ability to yield higher overall accuracy. If a classifier demonstrates superior accuracy, it is included; otherwise, it is excluded. The authors recommended a backward ensemble search methodology that utilizes accuracy and sensitivity energy functions. Initially, all classifiers are considered part of the ensemble, and subsequent elimination tests are performed to determine whether removing a classifier improves accuracy. Antal and Hajdu's ensemble system achieved outstanding results, with an accuracy of 90%, sensitivity of 90%, and specificity of 91% in both disease and no-disease settings. Their approach highlights the significance of combining various components and classifiers to enhance the accuracy of diabetic retinopathy detection.

Gondal et al. [8] proposed a CNN model for the referable Diabetic Retinopathy (RDR). They used two publicly available datasets Kaggle and DiaretDB1, where the Kaggle dataset is used for training and DiaretDB1 is used for testing. They are doing binary classification as normal and mild stages are considered as non-referable DR where the rest of the three stages are used as referable DR. The performance of the CNN model is evaluated based on binary classification resulting in sensitivity 93.6% and specificity 97.6% on DiaretDB1.

Wang et al. [9] proposed a novel architecture that classifies the images as normal/abnormal, referable/ non-referable DR and gets the high AUC on a normal and referable DR task 0.978 and 0.960 respectively and specificity is 0.5. Their proposed method uses three networks: main, attention and crop. The main network uses the Inception model that is trained on ImageNet where the attention network highlights different types of lesions in the images and crop network's crop the high attention image.

III. DATASET AND FEATURES

After the text edit has been completed, the paper is ready The dataset used (1) consists of 7 classes:

1. *No DR*
2. *Mild NPDR*
3. *Moderate NPDR*
4. *Severe NPDR*
5. *Very Severe NPDR*
6. *PDR*
7. *Advanced PDR*

However, to accurately classify the input image as one of those 7 classes, neural networks would be essential, and so we merged classes 2-5 as NPDR, classes 6-7 as PDR, and class 1 as normal.

Data split was done as follows: 80% of images for training, 20% for testing.

Fundus camera images often vary in contrast and have non-uniform lighting, which can make it difficult to detect lesions. To address this issue, it is necessary to enhance the contrast level of the images. An advanced technique called Contrast Limited Adaptive Histogram Equalisation (CLAHE) was used to enhance the contrast of these images. Afterwards, image segmentation is required. Three essential features of DR were segmented:

- *Blood Vessels: Non-proliferative DR can cause narrow blood vessels to be blocked near the optic nerve, which can lead to inadequate blood supply and the release of fluid in the retina. In contrast, proliferative DR can cause the development of new blood vessels in the retina.*
- *Microaneurysms: They appear as small unique patterns in a circular shape around the blood vessels.*
- *Hard Exudates: They are often seen as yellow or white spots in the retina. The presence of hard exudates is indicative of macular edema.*

The below figure shows one sample image from each class.

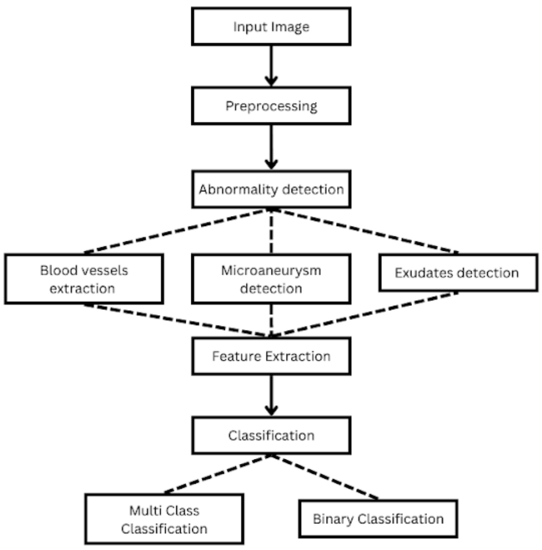
Normal NPDR PDR

IV. METHODS

4.1 Methodology

This paper provides machine learning techniques for disease classification based on the features of the input image. Figure 2 shows the flowchart of the proposed work.

First, preprocessing is performed on the input image for noise reduction and quality enhancement. In the second step, segmentation was implemented to get different lesions from the preprocessed image. In the third step, Gray Level Co-occurrence Matrix (GLCM) was used to generate the feature vector. Finally, machine learning classifiers such as Decision Tree and Random Forest are applied to get the classification results.



4.2 Used Algorithms

4.2.1 Decision Tree Algorithm

Decision Tree is a supervised machine learning algorithm that is used for classification and regression tasks. It is a tree-like model where each internal node represents a feature or attribute, each branch represents a decision rule, and each leaf node represents a class label or a numerical value. Decision Tree algorithm recursively partition the training data based on the feature values to create homogeneous subsets that are pure in terms of the target variable. This partitioning process aims to minimize the impurity or maximize the information gain at each step, leading to optimal splits. The impurity measures commonly used are Gini impurity and entropy.

$H(X) = - \Sigma\ (p(x) * \log_2(p(x)))$ *(Entropy Equation)*

Where:

H(X) represents the entropy of the random variable X.

Σ denotes the sum over all possible values of X.

p(x) is the probability of a specific value x occurring.

4.2.2 Random Forest Algorithm

The random forest algorithm is an ensemble learning technique that combines multiple decision trees to make predictions. In a random forest, each decision tree is trained on a randomly selected subset of the training data, and at each split, a random subset of features is considered. This randomness helps to reduce overfitting and improve the generalization ability of the model. During prediction, the random forest aggregates the predictions from all the individual decision trees. In classification tasks, the final prediction is determined by majority voting, where the class with the most votes is selected. In regression tasks, the predictions from each tree are averaged to obtain the final prediction.

V. EXPERIMENT/RESULTS/DISCUSSION

Intuitively, we thought Random Forest ensemble model would perform best on this problem as classifying images is prone to overfitting, however, we performed grid search with three different classifiers: SVM, Decision Tree, and Random Forest with varying three different hyperparameters for each. Grid search showed that Random Forest classifier works best indeed, with max_depth = 15 and n_estimators = 50 for multiclass classification, and max_depth = 10 and n_estimators = 50 for binary classification.

Those hyperparameters provided an accuracy score of 70.4% for multiclass classification and 88.81% for binary classification. To further illustrate our models' performances, we provided a classification report as shown in the below figure.

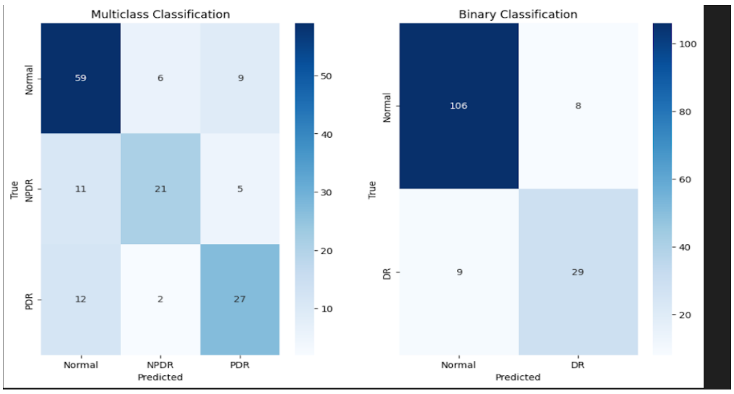
	precision	recall	f1-score	support
DR	0.92	0.93	0.93	114
Normal	0.78	0.76	0.77	38
accuracy			0.89	152
macro avg	0.85	0.85	0.85	152
weighted avg	0.89	0.89	0.89	152

Multiclass classification

	precision	recall	f1-score	support
NPDR	0.75	0.85	0.80	74
Normal	0.76	0.51	0.61	37
PDR	0.70	0.73	0.71	41
accuracy			0.74	152
macro avg	0.74	0.70	0.71	152
weighted avg	0.74	0.74	0.73	152

Binary classification

Moreover, we plotted a heatmap of the confusion matrix as shown in the below figure.



VI. CONCLUSION/FUTURE WORK

Diabetes has emerged as one of the fastest growing illnesses in recent years, with diabetics having a significant risk of developing diabetic *retinopathy (DR) estimated at 30%. If left undiagnosed and untreated, DR can cause vision impairments such as floaters, blurred vision, and ultimately blindness. However, the manual diagnosis of DR from fundus images is complex, time-consuming, and requires highly qualified professionals.*

To address this challenge, we developed a machine learning model that detects DR using decision tree, support vector machine, and random forest classifiers trained on the "Dataset from fundus images for the study of diabetic retinopathy". Our model achieved an accuracy of 70.4% and 88.81% using random forest classifier for multi-classification and binary classification, respectively.

In the future, we aim to explore deep learning techniques, which have shown significant advantages over traditional machine learning methods, especially when it comes to medical images that are complex and require precise analysis. We believe that using efficient deep learning algorithms and better feature extraction techniques could lead to improved accuracy and precision in DR detection. writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

VII. CONTRIBUTION

Our machine learning model was developed collaboratively by a team of four contributors. Aya Sameh was responsible for preprocessing the data and providing a detailed report of the preprocessing steps. Ehab Kamal and Hanya Ahmad worked on feature extraction, with Hanya also contributing to model building and report writing. Mohamed Hashem was involved in model development, Jupiter notebook organization and writing the markdowns, and contributed to feature extraction. Together, the team's diverse expertise and contributions resulted in a robust machine learning model for the task at hand.

VIII. REFERENCES

[1] Castillo Benítez, Veronica Elisa et al. “Dataset from fundus images for the study of diabetic retinopathy.” *Data in brief* vol. 36 107068. 21 Apr. 2021, doi:10.1016/j.dib.2021.107068

[2] International Diabetes Federation. IDF Diabetes Atlas, 10th edn. Brussels, Belgium: International Diabetes Federation, 2021. Available at: IDF Diabetes Atlas 2021 | IDF Diabetes Atlas

[3] M. Pawar, P. (2018). Retinal disease detection using machine learning techniques. HELIX, 8(5), 3932–3937. https://doi.org/10.29042/2018-3932-3937

[4] Odeh, I., Alkasassbeh, M., & Alauthman, M. (2021). Diabetic retinopathy detection using ensemble machine learning. 2021 International Conference on Information Technology (ICIT). https://doi.org/10.1109/icit52682.2021.9491645

[5] Qomariah, D. U., Tjandrasa, H., & Fatichah, C. (2019). Classification of diabetic retinopathy and normal retinal images using CNN and SVM. 2019 12th International Conference on Information & Communication Technology and System (ICTS). https://doi.org/10.1109/icts.2019.8850940

[6] Kahai, P., Namuduri, K. R., & Thompson, H. (2006). A decision support framework for automated screening of diabetic retinopathy. International Journal of Biomedical Imaging, 2006, 1–8. https://doi.org/10.1155/ijbi/2006/45806

[7] Antal, B., & Hajdu, A. (2014). An ensemble-based system for automatic screening of diabetic retinopathy. Knowledge-Based Systems, 60, 20–27. https://doi.org/10.1016/j.knosys.2013.12.023

[8] Gondal, W. M., Kohler, J. M., Grzeszick, R., Fink, G. A., & Hirsch, M. (2017). Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images. 2017 IEEE International Conference on Image Processing (ICIP). https://doi.org/10.1109/icip.2017.8296646

[9] Wang, Z., Yin, Y., Shi, J., Fang, W., Li, H., & Wang, X. (2017). Zoom-in-net: Deep mining lesions for diabetic retinopathy detection. Medical Image Computing and Computer Assisted Intervention − MICCAI 2017, 267–275. https://doi.org/10.1007/978-3-319-66179-7_31