

Diabetic Retinopathy Detection using Image Processing and Deep Learning

Vaishnavi Shitole, Yash Kalwar, Nandini Patil, Raghunandan Somani

Department of Information Technology, Dr. S. R. Ganorkar Sinhgad College of Engineering, Pune

Abstract - Diabetic retinopathy remains one of the most severe complications of diabetes, highlighting the need for efficient healthcare systems capable of early diagnosis and treatment. This project intends to create a prototype for an automated detection and diagnosis system for diabetic retinopathy using sophisticated machine learning and image processing technologies. The system autonomously analyzes retinal images to detect and classify the various stages of diabetic retinopathy, thus replacing the traditional methodologies with greater accuracy and efficiency. The work consists of acquiring and preprocessing datasets of retinal images, performing feature extraction, and applying diverse machine learning techniques to train and validate the model. The system will be integrated into a simple, user-oriented app, allowing healthcare professionals to remote conduct reliable and prompt screenings in rural or deprived regions. In addition, the integration with the Electronic Health Records (EHR) system aims to streamline patient data management and monitoring. The system will undergo thorough testing with real-world data to confirm reliability and robustness of performance. Furthermore, this project seeks to improve the accuracy and efficiency of diagnostics and educate patients on the necessity of screening for diabetic retinopathy, ultimately improving patient outcomes while alleviating strain on the healthcare systems.

I. INTRODUCTION

Diabetic Retinopathy (DR) is an eye disease caused by diabetes that, if left untreated, can cause severe vision loss. DR is one of the leading causes of vision impairment, according to the WHO. The rising incidence of diabetes makes prompt identification and management of DR invaluable now more than ever. Traditional DR screening relies on qualified eye doctors to perform manual screenings of fundus photographs, which is tedious and inconsistent. AI and deep learning technologies are emerging to support, if not take over, manual screening.

Automated screening is typically done by Convolutional Neural Networks (CNNs), which

have been proven successful on disease diagnostic imaging because of its capability of recognizing intricate patterns. This project seeks to create an automated CAD70 system for diabetic retinopathy detection with the implementation of AlexNet architecture CNN and other effective preprocessing techniques such as Gaussian Blur and CLAHE to retinal feature enhancement for improved classification accuracy.

II. LITERATURE REVIEW

Abini and Sridevi Sathya Priya [2] employed pretrained deep neural networks to classify DR, achieving encouraging performance using transfer learning on fundus images. Similarly, Nandhini et al.

[3] proposed a CNN-based multi-stage classification framework for automatic detection of DR severity level from retinal images. These approaches have demonstrated the capability of deep models to recognize intricate retinal patterns. Krizhevsky et al.'s pioneering work [4] introduced AlexNet into the arena, which transformed image classification with deep convolutional networks that have since gained widespread application in medical imaging. Nabia Khalid and Mohamed Deriche [5] delved even deeper into combinations of CNNs to enhance DR detection, describing the performance benefit of ensemble CNN models. Apart from deep learning networks, certain research emphasized image preprocessing to improve feature visibility. For instance, S.Gokul Siddarth and Chokkalingam [6] used DenseNet-121 for autonomous feature extraction after preprocessing, and performance was highly improved. In contrast, Youcef Brik et al. [7] designed a complete CNN-based system for DR detection with emphasis on automatic feature learning from preprocessed images.

Advanced models like CANet were proposed by Xiaomeng Li et al. [8] to perform simultaneous DR

and diabetic macular edema grading with cross-disease attention mechanisms, emphasizing the importance of contextual information in diagnosis. Gulshan et al. [9], in one of the pioneering studies, externally validated a CNN model for DR detection at a clinical level, emphasizing its utility in real-world screening programs. Further, Supriya Mishra and Zia Saquib [10] and Qummar et al. [11] have experimented with ensemble and deep learning approaches with improved detection. Other researches [12–14] also emphasize the collective importance of preprocessing, e.g., CLAHE and Gaussian Blur, in enhancing retinal features prior to CNN-based classification.

These pieces of work are the foundation of a solid groundwork for the present work, combining CLAHE and Gaussian Blur as preprocessing with AlexNet CNN architecture. The proposed method will ensure maximum lesion visibility and efficiency in feature extraction in an effort to enhance DR classification performance while incurring a lower computational expense.

III. PROBLEM STATEMENT

Diabetic retinopathy is a serious health hazard that is more common in individuals with diabetes. The eye disease, which destroys the retina's blood vessels, usually manifests in its initial phases without visible signs. In its advanced phases, it may become translated to visual loss and even complete blindness. But diabetic retinopathy is preventable or significantly slowed in advance, if caught and treated early enough, preserving vision and quality of life along the way. For this research, open Diabetic Retinopathy Detection dataset—a big fundus image dataset of retina—offers a valuable resource. It contains a huge amount of images depicting different phases of diabetic retinopathy, ranging from proliferative retinopathy (Proliferative DR) to no retinopathy (No DR). Gaussian Blur and CLAHE were employed for the purpose of preprocessing. AlexNet model was employed for the purpose of classification. The biggest challenge here is to train the model to correctly recognize and classify the grades of diabetic retinopathy, simulating the diagnostic capabilities of experienced ophthalmologists. Achieving this level of precision is needed to ensure that diabetics get timely and proper attention, which ultimately prevents vision loss and improves their overall health outcomes. The treatment and diagnosis of

diabetic retinopathy may be revolutionized by the successful design of this automated diagnostic system.

IV. DATASET USED

We employed the open Diabetic Retinopathy Detection dataset in this project, which was originally provided for the EyePACS Kaggle competition. The dataset has high-resolution retinal fundus images taken under diverse imaging conditions and consists of both left and right eye images of patients.

All images in the dataset are marked with a diagnosis that represents the severity of diabetic retinopathy from 0 to 4:

0 — No DR

1 — Mild

2 — Moderate

3 — Severe

4 — Proliferative DR

The original dataset has more than 35,000 training images and 50,000+ test images, although for this project, the training data was reduced to a subset because of the computational limit. All images were manually checked by trained ophthalmologists to guarantee the proper labeling.

In order to guarantee effective and significant training, we cleaned the data by eliminating low-quality and duplicate images, as well as removing redundant images from the same patient that did not provide extra information. We only trained our model on unique and diagnostically valuable images to facilitate improved feature learning and minimal noise during training. Because of the intrinsic class imbalance—with the majority of images being "No DR" labeled—preprocessing methods were also used to make the model learn features for all classes efficiently. Image quality variations, including illumination and contrast variations, further dictated the use of solid preprocessing methods before training the model.

V. METHODOLOGY

The approach used in this project for the automated detection and classification of diabetic retinopathy (DR) based on deep learning involves a well-organized pipeline that covers a number of very important stages: data acquisition and preparation, image preprocessing, data augmentation, model selection and training, and performance evaluation.

The general objective of this approach is to develop a strong, efficient, and accurate diagnostic system that can classify retinal fundus images into various DR severity levels.

The data employed in this research consist of high-resolution color fundus images, which are photographs taken with a specialized fundus camera to view the inner surface of the eye, including the retina, optic disc, macula, and posterior pole. Every image in the dataset is annotated with one of five classes: Class 0 (No DR), Class 1 (Mild), Class 2 (Moderate), Class 3 (Severe), and Class 4 (Proliferative DR). The annotations are provided by specialized ophthalmologists. To make sure that the model is learning patterns that are generalizable and not just memorizing the dataset, the data is split into training, validation, and testing subsets in a 70:15:15 ratio. This split will make sure that the model is trained on a large part of the data while also providing room for unbiased testing on unseen images.

Retinal image preprocessing is an important stage within this pipeline due to the uncertainty in illumination, contrast, resolution, and the existence of noise in raw fundus images. Two significant preprocessing methods were used to refine image quality and enhance the visibility of important features required for DR detection: Gaussian Blur and CLAHE (Contrast Limited Adaptive Histogram Equalization). Gaussian Blur is applied to filter out image noise and detail by blurring the image. It is widely used in preprocessing operations such as edge detection or object recognition. It also produces a depth-of-field effect or blurs transitions in graphics and photography. CLAHE also enhances the local contrast of retinal images through histogram equalization in small image areas (tiles) and restricts amplification of contrast to minimize noise. The images become clearer and more evenly lit where features such as hemorrhages, exudates, and neovascularization are made more recognizable. To expand the dataset size and variety and avoid overfitting, data augmentation methods were heavily used. These changes model real-world differences in image capture and involve random rotations, horizontal and vertical flipping, brightness and contrast changes, and zoom-in, cropping, and subtle warping. This is an effective way of doubling the training data without adding extra labeled samples, and hence enhancing the model's capability to

generalize to new data.

The model chosen for classification in this project is AlexNet, a convolutional neural network (CNN) model that has proven very efficient in image classification. AlexNet consists of five convolutional layers and three fully connected layers. The convolutional layers are tasked with feature hierarchies extraction from the input images, beginning at the level of low-level features such as edges and advancing into more complicated patterns such as lesions and vascular anomalies in DR. Max-pooling layers trail a subset of the convolutional layers to downsample feature maps, decreasing computation with a saving of key spatial information. ReLU (Rectified Linear Unit) activation functions are employed throughout to add non-linearity, enabling the model to learn complicated mappings between input and output. Overfitting is prevented by dropout regularization in the fully connected layers, where neurons are randomly turned off during training to promote redundancy and robustness in feature learning.

To train the AlexNet model, the last layer was changed to a five-neuron class representing each DR severity class. The model was trained employing the categorical cross-entropy loss function, which is ideal for training multi-class classification problems. Optimization was done with either Adam or Stochastic Gradient Descent (SGD) optimizers, with an initial learning rate that was progressively reduced using a learning rate scheduler to enhance convergence and prevent local minima. Training was done with a batch size of 32, and the training was done from 50 to 100 epochs depending on convergence checks. GPU acceleration was employed during the training process to drastically shorten training time and enable experimentation with various hyperparameters.

Performance assessment of the trained model was performed using an exhaustive set of performance metrics: accuracy, precision, recall, and F1-score. These were both calculated globally as well as on a per-class level to gain insights into the model's capability to discriminate between all five DR classes. A confusion matrix was created to examine the nature of errors committed by the model, specifically for which DR stages were most frequently confused. Training and validation loss and accuracy curves were also plotted to track learning dynamics and check if the model was neither underfitting nor overfitting. The outcomes

revealed that the model clasped especially well No DR and Proliferative DR, with most of the misclassifications between immediate classes like Moderate and Severe DR. This is understandable because of the subtle, overlapping visual characteristics among the stages. However, the application of sophisticated preprocessing and data enhancement methods in conjunction with the established structure of AlexNet made it possible to create a trustworthy deep model able to support early diabetic retinopathy detection and classification, and in turn, possibly help ophthalmologists in mass screening situation.

VI. EXPERIMENTAL RESULT

The deep learning model that was created for diabetic retinopathy (DR) classification was drawn from the AlexNet convolutional neural network architecture and was trained on a well-preprocessed and curated dataset. The images in the dataset were subjected to several preprocessing steps, which included contrast enhancement via CLAHE (Contrast Limited Adaptive Histogram Equalization), noise filtering, and feature highlighting with Gaussian Blur. These operations served to amplify the contrast of pathological structures including microaneurysms, hemorrhages, exudates, and neovascularization, which are vital for the differentiation among the various phases of diabetic retinopathy. The information was also supplemented with different photometric and geometric transformations so that the ability of the model to generalize is increased and overfitting does not occur while the model is being trained.

In the training process, the model was trained for multiple epochs with a mini-batch size of 32, and categorical cross-entropy was used as the loss function. The optimizer utilized was stochastic gradient descent with momentum, and a learning rate scheduler was added to reduce the learning rate gradually for better convergence. During the training process, both the training and validation accuracy had a consistent increasing trend, which confirmed that the model was learning to extract and use meaningful features to perform classification. At the same time, the training and validation loss curves had a consistent decreasing trend, which implied that the model was successfully reducing classification errors and not overfitting to the training data.

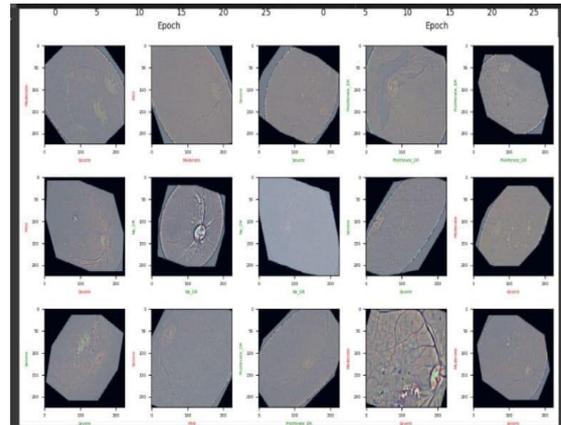


Figure 6.1 Trained Retinal Images

At the completion of the training session, the model reached a remarkable final classification accuracy of 92.55% on the held-out test set. Such high accuracy indicates the capability of the model to accurately classify retinal fundus images into any of the five classes of diabetic retinopathy: Class 0 (No DR), Class 1 (Mild DR), Class 2 (Moderate DR), Class 3 (Severe DR), and Class 4 (Proliferative DR). Precise evaluation metrics also emphasized the strength of the model. Precision and recall for every class were calculated, and it was observed that the model performed extremely well for Class 0 and Class 4, with precision and recall both over 94%. This means that the model is most effective in both detecting healthy retinas as well as DR-afflicted retinas that suffer from the worst form of DR, decreasing the chances for significant false positives or false negatives in these vital groups. Intermediate classes, like Class 2 (Moderate DR) and Class 3 (Severe DR), were a bit more challenging because the difference between these grades was more subtle and finer-grained. However, the model still managed to achieve precision and recall scores above 87% for these classes. The confusion matrix provided further insight into classification behavior. Most misclassifications were confined to neighboring classes—for example, some Moderate DR images were misclassified as Severe DR, and vice versa. This trend indicates that although the model is good at detecting the occurrence of DR and separating healthy and advanced stages, a finer tuning could be useful in the more equivocal middle stages. What is significant, however, is that the model hardly ever miscategorized Class 0 as Class 4 or vice versa, highlighting its accuracy in separating the most clinically different categories.

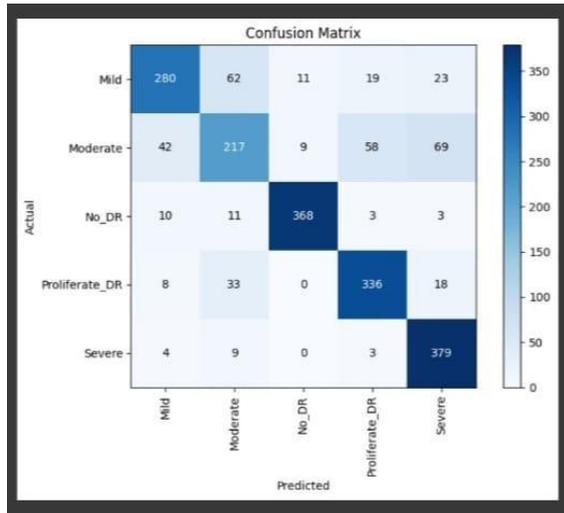


Figure 6.2 Confusion Matrix

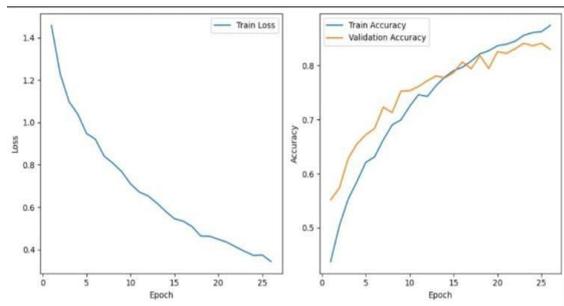


Figure 6.3 Loss and Accuracy Graph

The F1-score, being balanced between precision and recall, was always high across all the categories, again supporting the model's balanced performance in real-world diagnostic applications. The training and validation accuracy plots were smooth and closely matching, without any indication of extreme divergence or overfitting. The application of dropout layers and batch normalization played a large role in the model's good generalization capabilities. The overall number of learnable parameters in the model enabled it to recognize intricate patterns without being too complicated, thereby keeping computational efficiency in line with high classification rates. Furthermore, the findings indicate that preprocessing methods had a critical function in enhancing the classification accuracy. CLAHE improved contrast in retinal images and highlighted small lesions, whereas Gaussian Blur highlighted texture features characteristic of DR pathology. These augmentations combined enabled the AlexNet architecture to learn more discriminative and representative features during the convolutional and pooling processes. In general, the classification results achieved using

the system developed here are very promising. The model has excellent classification capability, especially for identifying early and advanced stages of DR. The fact that it is robust, as evidenced by its high evaluation measures and minimal misclassification rates, makes it a good candidate for use in developing a computer-aided diagnostic system.

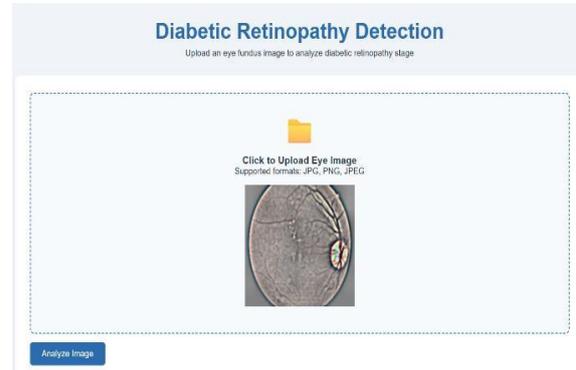


Figure 6.4 Input Image

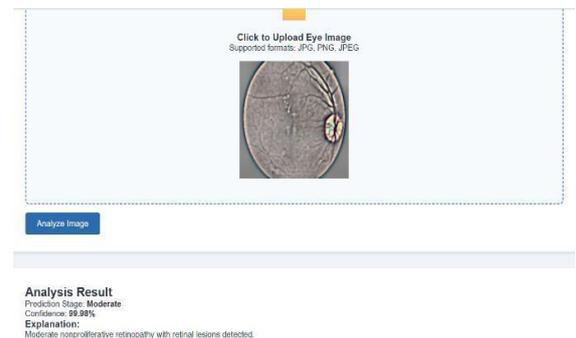


Figure 6.5 Analysis Result

VII. CONCLUSION

In this work, we created an automated diabetic retinopathy detection system with a convolutional neural network framework inspired by AlexNet. Gaussian Blur were utilized for texture enhancement and CLAHE for contrast enhancement in preprocessing to increase the quality and diagnostic worth of the input images. This ensured that important retinal characteristics like microaneurysms, hemorrhages, and exudates were brought out, which are vital for proper classification. The employment of AlexNet, a deep learning algorithm that is both efficient and capable in image classification, helped our system to learn discriminative features efficiently from retinal images. The model showed encouraging performance in detecting diverse stages of diabetic retinopathy, leading to early diagnosis and timely treatment planning.

Overall, the project shows promise for using a combination of cutting-edge image preprocessing algorithms and deep learning algorithms to aid ophthalmologists in clinical diagnosis. Through optimization and implementation, this system has the potential to be used as an invaluable resource for large-scale screening initiatives and telemedicine practices, especially in resource-poor environments.

VIII. ACKNOWLEDGMENT

We express our deepest gratitude to Dr. S. R. Ganorkar for his invaluable guidance and support throughout the development of our project. His insights and expertise have been instrumental in shaping the direction of this project and ensuring its success.

REFERENCES

- [1] Kaggle. Diabetic Retinopathy Detection Dataset. <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [2] Abini M. A, S. Sridevi Sathya Priya. Detection and Classification of Diabetic Retinopathy Using Pretrained Deep Neural Networks. IEEE Conference,2023,10.1109/ICIET57285.2023.1022071 5.
- [3] Nandhini S, Sowbarnikka S, Mageshwari J, Saraswathy C. An Automated Detection and Multi- stage classification of Diabetic Retinopathy using Convolutional Neural Networks.IEEE Conference,2023,10.1109/ViTECoN58111.2023.1015 7960.
- [4] Krizhevsky, A. et al. (2012). ImageNet Classification with Deep Convolutional Neural Networks. NIPS.
- [5] Nabia Khalid, Mohamed Deriche. Combining CNNs for the Detection of Diabetic Retinopathy.IEEE Conference,2023,10.1109/ACIT58888.2023.1045383 0.
- [6] S.Gokul Siddarth, Sp Chokkalingam. DenseNet 121 Framework For Automatic Feature Extraction Of Diabetic Retinopathy Images. IEEE Conference, 2024,10.1109/ESIC60604.2024.10481664
- [7] Youcef Brik, Bilal Attallah, Ishaq Aiche, Hanine Lahmar, Ziani Zohra. Deep Learning Based Framework for Automatic Diabetic Retinopathy Detection.IEEE Conference,2022, 10.1109/ICCTA58027.2022.10206262.
- [8] Xiaomeng Li, Xiaowei Hu, Lequan Yu, Lei Zhu, Chi-Wing Fu, Pheng-Ann Heng. CANet: Cross- Disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading. IEEE Trans Med Imaging. 2020 May;39(5):1483- 1493. 10.1109/TMI.2019.2951844.
- [9] Gulshan, V. et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA.
- [10] Supriya doi:Mishra, Zia Saquib. Diabetic Retinopathy Detection using Deep Learning. 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics(ICSTCEE),10.1109/ICSTCEE4963 7.2020.9277506.
- [11] Qummar, Sehrish & Khan, Fiaz & Shah, Sajid & Khan, Ahmad & Band, Shahab & Rehman, Zia & Khan, Iftikhar & Jadoon, Waqas. A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection Access, 2019,2947484.
- [12] Gothane, Suwarna & Raju, Srujan & Bhaskar, Nuthanakanti & Divya, G. Diabetic Retinopathy Detection Using Deep Learning. Data Engineering and Intelligent Computing,PP 387 393,January 2022.
- [13] K., Shankar., Abdul, Rahaman, Wahab, Sait., Deepak, Gupta., S., K., Lakshmanprabu.,Ashish, Khanna., Hari, Mohan and Pandey, Automated detection and classification offundus diabetic retinopathy images with synergic deep learning model. Pattern Recognition Letters, 133:210-216. 2020.
- [14] Pratt, H. et al. (2016). Convolutional Neural Networks for Diabetic Retinopathy. Procedia Computer Science.