RESEARCH PAPER

Advanced Retinal Multi-Disease Detection Using CNN and Med-PaLM 2 Powered Analysis

Amulya U Reddy[1] and Dr. Mohammed Abdul Waheed[2]

[1]Student, Department of Computer Science and engineering, VTU's CPGS Kalaburagi, Karnataka, India
[2]Associate Professor,Department of Computer Science and engineering, VTU's CPGS Kalaburagi, Karnataka, India
[1]E-mail ID: amulyareddy731@gmail.com and [2]E-mail ID: mawaheed@vtu.ac.in

Received: 28th Feb, 2026; Revised: 6th March 2026; Accepted: 7th April, 2026; Available Online: 20th April, 2026

ABSTRACT

Detecting the retinal condition at an early stage can help prevent blindness due to Diabetic Retinopathy, Glaucoma, and other such ailments. The traditional method involves manual inspection of the fundus images by ophthalmologists, which may be time-consuming, costly, and challenging to access. To overcome these limitations, the project named 'Retinal Multi-Disease Prediction' aims to design an intelligent retinal screening and awareness system using Convolutional Neural Networks (CNNs) along with LLMs, for instance, Med-PaLM 2. It offers automated disease predictions with patient-centered explanations.

The input provided is the retinal fundus images to extract meaningful patterns via CNN based DL model, utilizing the state-of-the-art architecture, e.g., EfficientNet. The CNN model analyzes the images and categorizes them into various classes, namely, Normal, Abnormal, and Dangerous, and identifies specific retinal diseases like Diabetic Retinopathy. The entire task related to image processing and disease prediction is managed by the CNN to interpret medical imaging data accurately. Further, after predictions, the output of the CNN is transferred in the form of text to feed to the LLMs, such as Med-PaLM 2. The LLM models do not work on analyzing images.

These LLMs take the output of the CNN and translate it into patient-friendly explanations. The output includes the detected condition, its possible cause, required precautions, future actions to take, and advice on consulting an ophthalmologist. Moreover, the application offers dual language responses in English and Kannada to make the system more patient-oriented. The app comes with a professionally designed interface with features like login/registration facility, prediction history, themes of dark/light mode, attractive visuals, etc.

The final output shows the predicted retinal condition along with the patient-friendly explanation. It emphasizes raising the awareness and performing the initial screening instead of providing the diagnosis or medical recommendations. The project focuses on designing the screening and awareness tool for the benefit of patients.

Keywords: *CNNs, LLMs, Med-PaLM 2, retinal fundus, Diabetic Retinopathy*
How to cite this article: Reddy AU, Waheed MA. Advanced Retinal Multi-Disease Detection Using CNN and Med-PaLM 2 Powered Analysis. Int J Drug Deliv Technol. 2026;16(61s):1622-1630. DOI: 10.25258/ijddt.16.61s.184
Source of support: Nil.
Conflict of interest: None

I. INTRODUCTION

This is the case with several other retinal conditions, for example, Diabetic Retinopathy, and Glaucoma, among others, which cause blindness because of inadequate or delayed diagnoses. Traditionally, screening for retinal diseases relies on ophthalmologists analyzing retinal fundus images, which is not only time-consuming and expensive but also difficult to achieve in remote and resource-poor settings. In this regard, the innovative system called "Retinal Multi-Disease Prediction Using CNN with Med-PaLM 2 Assisted Analysis" presents a promising intelligent screening and awareness solution. This technology uses CNNs and LLMs such as Med-PaLM 2 to enable retinal disease predictions and patient-friendly explanations.

What this system means for users: Patients submit their retinal fundus images, and the system utilizes a CNN-based Deep Learning Model, for example, one based on architecture such as EfficientNet, for analysis purposes. This model extracts vital visual features from the images and classifies the patients' eye conditions, identifying any abnormality or danger and determining the nature of the diseases. The CNN processes and analyzes the data, thereby enabling accurate medical interpretations. After the CNN makes the final decision, the output is converted into text and submitted to the LLMs such as Med-PaLM 2.

The output consists of an explanation of the condition, possible causes, associated risk level (low, medium, and high), preventative measures, lifestyle advice, and recommendation regarding when to visit an ophthalmologist. For the sake of accessibility, these

**Author for Correspondence: amulyareddy731@gmail.com*

explanations will be provided in two languages: English and Kannada.

The user interface for the proposed system is professional, secure, patient-centered, and engaging. Users will need to log in or create new accounts before using the software. Additionally, there will be provision for tracking the history of the users' submissions, with an option for switching between light and dark themes. It will be appropriate for use by healthcare professionals and other individuals interested in retinal health awareness.

It's significant as current project focuses on developing an awareness and screening system, which should never be considered as diagnostic and treatment procedures. The use of CNN-based image analysis and LLM-driven natural language explanations facilitates the process of awareness and screening.

Recently, significant developments have been recorded within this domain. For instance, Asif Nawaz et al. (2023) developed a multi-class classifier for retinal disease using CNN tested on EyeNet with 32 different classes, with an average performance of nearly 95% with efficiency in computations and memory space. On the other hand, Kavitha et al. (2024) explored hybrid and transfer learning methods for increasing the accuracy and decreasing the computational demand of deep neural networks used in classifying retinal diseases. Accuracy, sensitivity, and specificity are some of the essential metrics in this regard.

As described above, an LLM is an advanced language model that is capable of generating human language through transformer-based learning. Within this context, the role of an LLM is secondary and supportive. The LLM in the current project does not engage in image analysis, diagnosis, or prediction of diseases, as it simply takes the outputs provided by the CNN. In the latter part, the LLM provides patient-friendly explanations, information on causes, risks, and recommendations.

Similarly, Med-PaLM 2 is a LLM designed and optimized to produce medically aware responses. Specifically, this is a clinical LLM that has been trained and fine-tuned on large amounts of medical literature and knowledge. In this case, Med-PaLM 2 is specifically dedicated to explaining patient conditions, identifying causes, discussing risk factors, preventative actions, and recommendation regarding when to visit an ophthalmologist.

II. RELATED WORKS

According to Alyoubi, W. L., Abulkhair, M. F., and Shalash, W. M. (2021), the authors have presented a method to classify Diabetic Retinopathy stages from I to V based on automated diagnostic procedures. In their work, two models have been utilized; namely, CNN512 which is responsible for the classification and YOLOv3-based CNN for lesion localization. Based on the APTOS 2019 and DDR datasets, the authors' team performed the following preprocessing operations on images: CLAHE (Contrast Limited Adaptive Histogram Equalization), noise reduction, image cropping, and image augmentation. Then, CNN512 performs classification and YOLOv3 model detects the lesion. Overall, classification accuracy improved to 89% in this study with a sensitivity rate of 89% and specificity of 97.3%.

Çinarer, G., Kilic, K., and Parlar, T. (2022) have introduced a new deep transfer learning framework that aims to automate diabetic retinopathy stages classification and lesion detection. Based on APTOS 2019 Blindness Detection dataset, the researchers divided images into the training and testing subsets and performed augmentations to the training subset. Five processing operations were performed on the image set: Gaussian filter, cropping, rescaling, and CLAHE was applied as well. Then, after pre-processing, researchers used three models, while ResNet152 gave the highest AUC score of 94.1%.

Kotiyal, B. and Pathak, H. (2022) suggested applying deep transfer learning techniques for diabetic retinopathy binary classification. As part of the research, two frameworks were merged to perform data analysis, namely, PySpark and deep learning to enable Big Data capabilities. The authors worked with IDRiD dataset. Preprocessing included cleaning data of dark images, then cropping images to eliminate unnecessary space, and re-scaling. Classification was carried out with help of LR classifier. Researchers used DL Pipelines on Apache Spark with 80/20 train-test split. Three transfer learning models were evaluated, and InceptionV3 showed the best performance: 95% accuracy, 94.98% AUC, and 95% F1-Score.

III. METHODOLOGY

The workflow of the proposed "Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM 2)" method consists of a number of steps and is developed in such a way as to provide efficient diagnosis of retinal conditions along with detailed and comprehensible predictions. These steps include acquiring the retinal image, its preprocessing, making a prediction based on that data, generating an explanation for that result, and delivering it to the user.

Acquisition of the image

The source of the dataset for the suggested system lies at the website of Kaggle, which houses multiple sets of publicly accessible images used for educational purposes. In particular, the chosen data set contains high-resolution images of retinas in their normal state as well as those with a variety of retinal diseases present. Such images show various parts of the retina and reflect the heterogeneity observed in the real-life clinical practice. The data contains images of retinas that suffer from such diseases as Diabetic Retinopathy, Glaucoma, among others.

Retinal Diseases Included

Data has retinal images with conditions:

□Age-Related Macular Degeneration (AMD)
□Diabetic Retinopathy
□Retinal Detachment
□Retinitis Pigmentosa
□Macular Hole
□Epiretinal Membrane (ERM)
□Central Serous Retinopathy (CSR)
□Retinal Vein Occlusion (RVO)
□Retinal Artery Occlusion
□Uveitis
□Retinoblastoma

They were selected in order to cover common retinal diseases and emergencies that could lead to complete blindness and require immediate care. A wide range of diseases contributes to learning more detailed patterns and improves the reliability of the results of multiscreening for retinal disease classification.

Data Characteristics
Dataset consists of color retinal images stored in JPEG and PNG formats. As the source images were taken under various conditions and from various locations, they can differ by resolution. During preprocessing, all images are normalized to ensure uniformity of input to the CNN. Both macula and optic disc images are presented, providing a possibility of examining retinal diseases connected with each part of the retina. Image dimensions: 320x314.

Preprocessing
The preprocessing step is aimed at improving the image properties. It consists of image resizing, normalization of pixel intensity levels, and contrast enhancement in order to clearly display retinal patterns. Data augmentation techniques are applied in training in order to increase diversity of the dataset.

CNN-Based Retinal Disease Classification
CNN based on EfficientNet model is used for the classification of retinal images. The CNN identifies distinctive features including vessel patterns, microaneurysms, hemorrhages, and optic disc features from fundus images. Based on the identified features, the model classifies retinal images into one of three classes - Normal, Abnormal or Dangerous.

CNN-Based Feature Extraction
A Convolutional Neural Network extracts features from the retinal image using convolution operations.

Let the retinal image be represented as:

$$I \in \mathbb{R}^{H \times W \times C}$$

where
H = height,
W = width,
C = number of color channels.

Convolution operation is defined as:

$$F_{i,j}^{k} = \sum_{m}\sum_{n}\sum_{c} I_{i+m,j+n,c} \cdot K_{m,n,c}^{k} + b^{k}$$

where:

K^k is the convolution kernel for the kth feature map

bk is the bias term

$F^k_{i,j}$ is the extracted feature value

Procedure facilitates identification of various patterns including blood vessels, lesions, and other structural abnormalities by the CNN.

Activation Function (ReLU): After the convolution process, the activation function ReLU comes into play:

$$\mathrm{ReLU}(x) = \max(0, x)$$

EfficientNet Scaling:

EfficientNet scales the network using a compound coefficient ϕ\phiϕ, balancing depth, width, and resolution:

$$\text{Depth} = \alpha^{\phi}, \quad \text{Width} = \beta^{\phi}, \quad \text{Resolution} = \gamma^{\phi}$$

subject to:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$$

This improves performance without losing computational efficiency

Softmax Classification

In the last layer, Softmax classification is used to classify the retinal image into one of the following categories:

$$P(y = i) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}$$

where:

zi is the output score for class I

N is the number of classes

The class with the highest probability is selected as the predicted disease category.

Loss Function (Categorical Cross-Entropy)

Cross-entropy loss:

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

where:

- y_i is the true label
- $\hat{y}_i$ is the predicted probability

Optimization (Adam Optimizer)

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

where:

- θ represents model parameters
- η is the learning rate
- $\hat{m}_t$ and $\hat{v}_t$ are bias-corrected moment estimates

Evaluation Metrics

Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Prediction Output Representation

The CNN output is structured as:

O={D,R,C}

where:

D = detected disease

R = risk level (Low / Medium / High)

C = confidence score

This structured output is passed as text input to the LLM.

Analysis of images and prediction of diseases are done exclusively by the CNN to ensure accuracy and reliability.

Training and Validation of the Model

The CNN model is trained with labelled retinal fundus images through appropriate optimization approaches. The dataset is divided into train, validation, and test datasets to evaluate the performance. Metrics like accuracy, precision, recall, and F1 score are used during evaluation to ensure accuracy and robustness in the classification of retinal disease in the unseen retinal images.

Processing the Prediction Output

After predicting the retinal disease, the CNN model generates a structured output, including the predicted retinal disease, its severity, and confidence level. The prediction generated by the CNN will be the input to generate explanations.

Generation of Explanation Using LLM

Prediction generated by the CNN is passed to the LLM, such as Med-PaLM 2, which generates patient-friendly explanations. The language model does not analyse retinal images but generates explanations regarding the disease, its probable cause, risk, precautions, consultation, and other relevant information to the patients. Explanations are generated both in English and Kannada.

LLM-Based Explanation Mapping:

Explanation procedure given as:

$$E = f_{LLM}(D, R, C)$$

System Interface and Integration

The entire system is embedded in an easy-to-navigate user application interface where you will have access to secure login and sign-up portals, a simplified pathway for uploading images, accurate prediction outcomes, and organized prediction history. You will also be able to switch seamlessly between light mode and dark mode for better visibility and convenience.

Ethical Considerations and Applicability Scope

This app is developed solely for diagnostic purposes and will not make medical decisions or provide treatment recommendations. There are prominent disclaimers advising users to consult with an ophthalmologist before making any decisions related to their health.

Output Presentation

The output is generated with the help of CNN to predict diseases combined with a humanly understandable report generated by the LLM.

IV. SYSTEM ARCHITECTURE

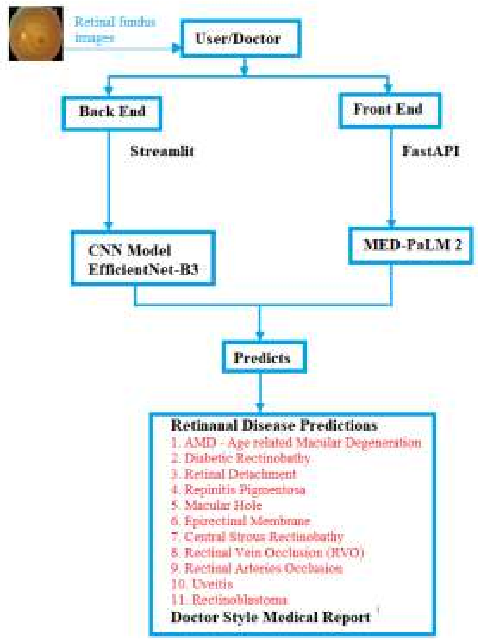


Figure 1: System Architecture

Workflow of retinal multi-diseases prediction through Med-PaLM 2 begins at a secure hop. The user or the doctor uploads an image of retina through a secure connection between the front and back ends. At the back end designed using Streamlit, the uploaded retinal image is processed through a CNN based on EfficientNet-B3. In this process, the retinal features are extracted and the severity of diseases is analyzed. Prediction of the disease through this process is then transferred to the front end using FastAPI. This is the point at which Med-PaLM 2 creates a doctor-like interpretation of the discovered problem along with risk factors, and precautions required by the patients. Finally, the two results of predictions are combined together to generate an output that recognizes diseases such as Diabetic Retinopathy, Glaucoma, AMD, Retinal Detachment, and other retinal diseases.

V. RESULTS AND DISCUSSIONS

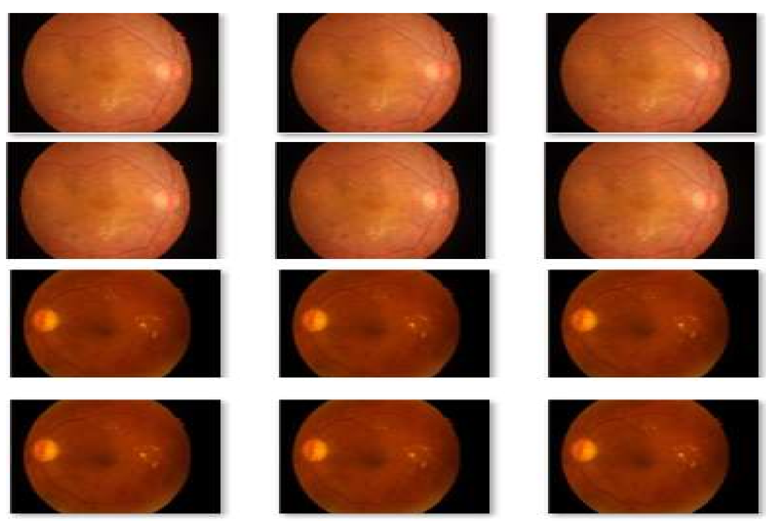


Figure 2: Input Images

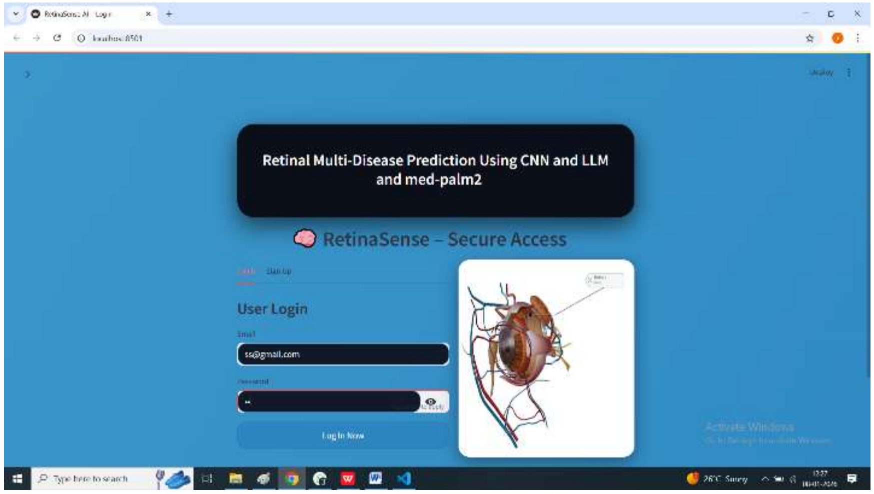


Figure 3: Create Account and Login

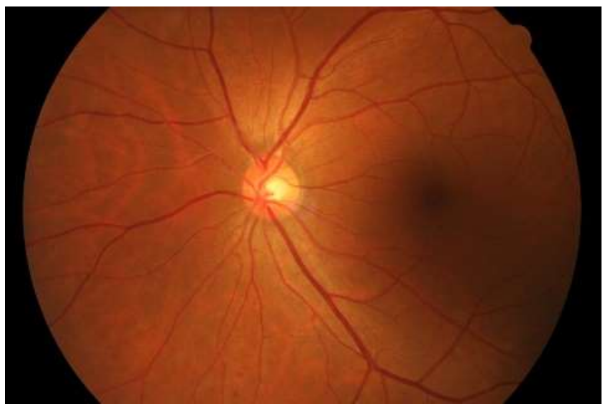


Figure 4: Input Image

Inputting the retinal fundus image to analyze disease severity using deep learning.

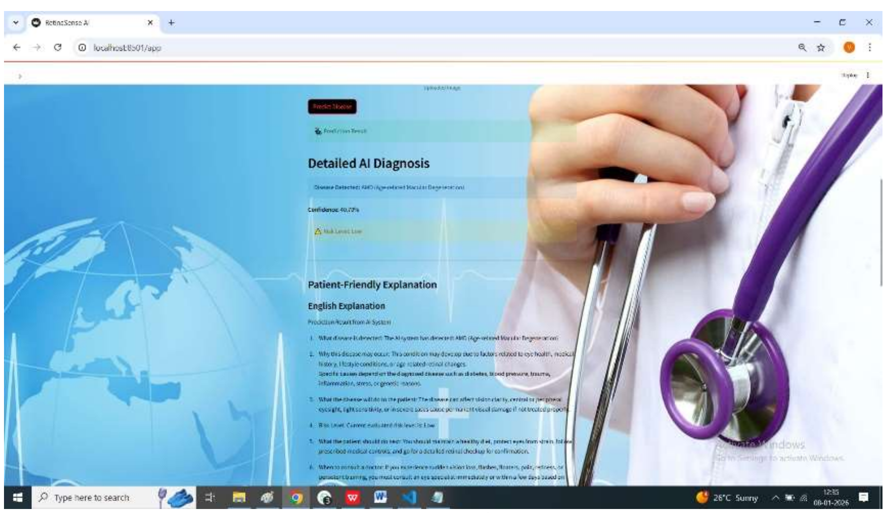
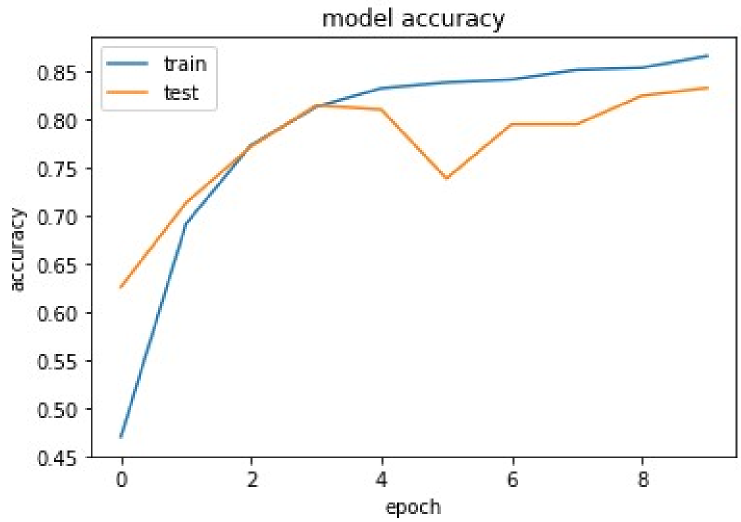


Figure 5: Prediction



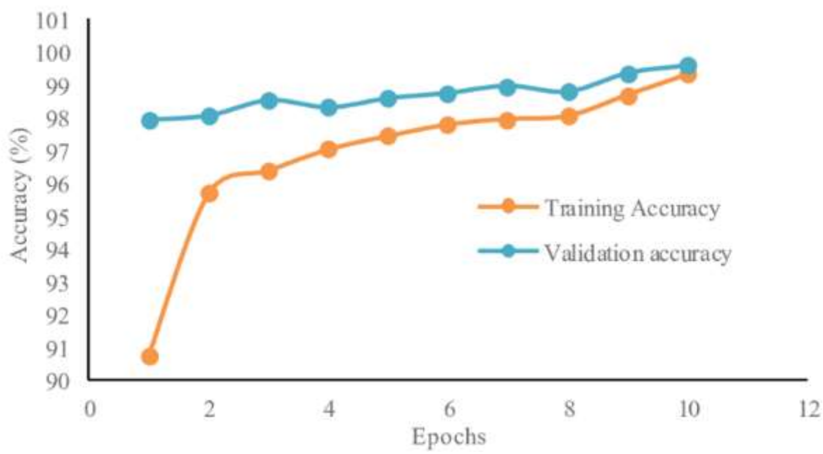


Figure 6: Accuracy of Training and validation accuracy

Table II. Model's Performance

Parameters	Performance
Training accuracy	98.00
Validation accuracy	98.86
Precision	97.01
Recall	97.02
F1-score	97.03

Table III. Comparison Between Models For Fundus Image Classification

Metrics	Proposed model	CNN V G G [4]
Accuracy	98.86	95.5
Precision	97.01	95
Recall	97.02	95
F1-score	97.03	95

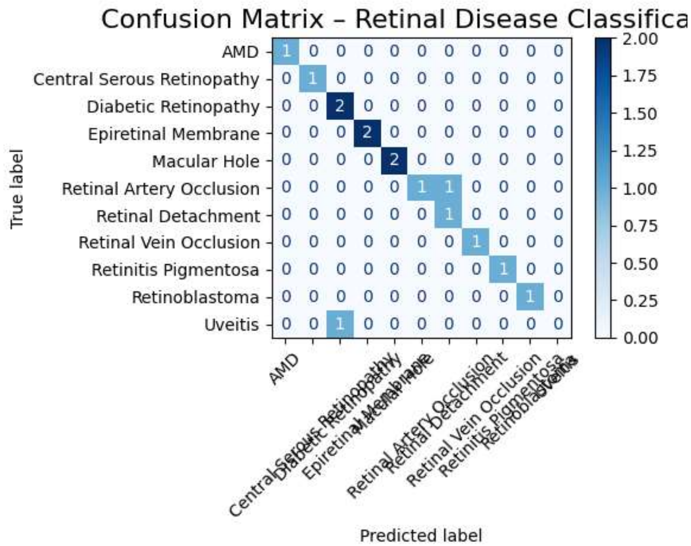


Figure 7: Confusion Matrix for 11 classes

VI. CONCLUSION AND FUTURE WORK

The proposed "Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM 2)" framework provides a highly intelligent and convenient solution to the problem of screening for retinal diseases and educating the users on how to understand AI-based predictions. By combining deep learning methods for analyzing retinal images with advanced capabilities of explaining medical predictions using language models, the system successfully addresses the technical aspect of identifying retinal problems as well as the difficulty many people experience when trying to comprehend the results of AI-based prediction systems.

Furthermore, the use of sophisticated CNN architectures, including EfficientNet-B3, ensures high accuracy of feature extraction from fundus photos and reliable classification of different vision-threatening retinal diseases. The combination of various LLMs, including Med-PaLM 2, makes the predictions easy to understand due to the availability of patient-friendly explanations that make it possible to provide information about the nature of the condition, risks involved, preventive measures, and the necessity of seeking consultation with a qualified specialist. In addition, the inclusion of bilingual support in both English and Kannada languages significantly improves usability of the system.

Being created not as a medical diagnostic or treatment system but rather as a platform for retinal screening and educating people regarding the necessity of consulting professionals, the proposed framework aims at helping users receive prompt consultations and avoid complications that may lead to the development of severe eye problems and even blindness. Other useful functions implemented in the platform include security-related user authentication, prediction history, as well as convenient and understandable user interface design.

Overall, the suggested AI framework effectively combines technical aspects of identifying retinal problems with explaining findings to people so as to promote their awareness about retinal diseases and motivate them to visit specialists.

REFERENCES

1. Alyoubi, W. L., Abulkhair, M. F. & Shalash, W. M. Diabetic retinopathy fundus image classification and lesions localization system using deep learning. *Sensors* **21** (11), 3704 (2021).

2. Çinarer, G., Kilic, K. & Parlar, T. A deep transfer learning framework for the staging of diabetic retinopathy. *J. Sci. Rep.-A*. **(051)**, 106–119 (2022).

3. Kotiyal, B. & Pathak, H. Diabetic retinopathy binary image classification using Pyspark. *Int. J. Math. Eng. Manag. Sci.* **7** (5), 624 (2022).

4. Skouta, A., Elmoufidi, A., Jai-Andaloussi, S. & Ochetto, O. Automated binary classification of diabetic retinopathy by convolutional neural networks. In *Advances on Smart and Soft Computing: Proceedings of ICACIn 2020*. 177–187. (Springer Singapore, 2021).

5. Uppamma, P. & Bhattacharya, S. Diabetic retinopathy detection: A blockchain and African vulture optimization algorithm-based deep learning framework. *Electronics* **12** (3), 742 (2023).

6. Sikder, N. et al. Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images. *Symmetry* **13** (4), 670 (2021).

7. Ghazal, M., Ali, S. S., Mahmoud, A. H., Shalaby, A. M. & El-Baz, A. Accurate detection of non-proliferative diabetic retinopathy in optical coherence

tomography images using convolutional neural networks. *IEEE Access*. **8**, 34387–34397 (2020).

8. Liu, H. et al. Hybrid model structure for diabetic retinopathy classification. *J. Healthc. Eng.* **2020** (2020).
9. Jabbar, A. et al. *A Lesion-Based Diabetic Retinopathy Detection Through Hybrid Deep Learning Model*. (IEEE Access, 2024).
10. Moustari, A. M., Brik, Y., Attallah, B. & Bouaouina, R. Two-stage deep learning classification for diabetic retinopathy using gradient weighted class activation mapping. *Automatika* **65** (3), 1284–1299 (2024).
11. Shakibania, H., Raoufi, S., Pourafkham, B., Khotanlou, H. & Mansoorizadeh, M. Dual branch deep learning network for detection and stage grading of diabetic retinopathy. *Biomed. Signal Process. Control*. **93**, 106168 (2024).
12. Dai, L., Sheng, B., Chen, T., Wu, Q., Liu, R., Cai, C. & Jia, W. A deep learning system for predicting time to progression of diabetic retinopathy. *Nat. Med*. **30**(2), 584–594 (2024).
13. Jin, Y. et al. Deep learning-driven automated quality assessment of ultra-widefield optical coherence tomography angiography images for diabetic retinopathy. *Vis. Comput.* 1–11 (2024).
14. ASIF NAWAZ ,TARIQ ALI 1 , GHULAM MUSTAFA 1 , MUHAMMAD BABAR , 2023, Multi-Class Retinal Diseases Detection Using Deep CNN With Minimal Memory Consumption
15. Ms.Kavitha.Phalguni.S.Prasad, Pragna.G.Prasad, Prajna S.G 2024, A Survey on Multi-Class Retinal Disease Detection using CNN