MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3832-3839

A Unified Approach to Multi Disease Prediction using Deep Learning Models and Medical Language Intelligence

Mrs. Deepthi Nair P, Ass. Professor
Department of Computer Science & Engineering Department, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India.
deepthinair@karunya.edu.in
Dr. Subha Hency Jose, Professor
Department of Bio Medical Engineering, Karunya Institute of Technology and Science, Coimbatore, India hency200200@karunya.edu.in
Dr. Anantha Christu Raj P, Professor
Department of Bio Medical Engineering, Karunya Institute of Technology and Science, Coimbatore, India anantha@karunya.edu.in
Mr. Mohamed Arafath A, Student
Department of Computer Science & Engineering Department, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India.
mohamedarafatha23cse@srishakthi.ac.in

ABSTRACT

Retinal diseases such as Diabetic Retinopathy, Glaucoma, and other vision-threatening abnormalities are major causes of preventable blindness when not detected at an early stage. Manual retinal screening requires expert ophthalmologists and is often limited by time, cost, and availability, especially in rural and underserved regions. To address these challenges, this project titled "Retinal Multi-Disease Prediction" presents an intelligent screening and awareness support system that combines Convolutional Neural Networks (CNNs) with Large Language Models (LLMs), including Med-PaLM2, to provide automated retinal disease detection along with patient-friendly explanations. The proposed system accepts a retinal fundus image as input and performs automated analysis using a deep learning–based CNN model built on advanced architectures such as EfficientNet. The CNN extracts high-level visual features from the fundus image and classifies it as Normal, Abnormal, or Dangerous, and can further identify specific retinal conditions such as Diabetic Retinopathy. The image analysis and disease classification are handled exclusively by the CNN model to ensure reliable and focused visual interpretation. Following the image-based prediction, the classification result is passed as textual input only to the LLM and Med-PaLM2, which do not analyze the image directly. Their role is to transform the medical prediction into a clear, patient-friendly explanation using simple language. The generated explanation includes the detected disease, possible causes, recommended next steps, and guidance on when the patient should consult an ophthalmologist. The system also supports bilingual explanations in English and Tamil, improving accessibility and patient understanding. The application features a professional and user-friendly interface, including secure login and sign-up modules, prediction history tracking, dark and light mode themes, and visually appealing background design to enhance usability. This system is not intended to replace medical professionals or provide treatment decisions. Instead, it serves as a screening and awareness support tool aimed at promoting early detection, improving health literacy, and encouraging timely consultation with eye care specialists. By integrating CNN-based medical image analysis with LLM-driven natural language explanation, the proposed system bridges the gap between technical predictions and human understanding, making retinal health screening more accessible, interpretable, and impactful.

Keywords: Convolutional Neural Networks (CNNs), Large Language Models (LLMs), Med-PaLM2, retinal fundus, Diabetic Retinopathy

I. INTRODUCTION

Retinal diseases such as Diabetic Retinopathy, Glaucoma, and other vision-threatening abnormalities are among the leading causes of preventable blindness worldwide. Early detection plays a crucial role in preserving vision; however, conventional retinal screening relies heavily on trained ophthalmologists, specialized equipment, and clinical infrastructure. This dependency often limits timely diagnosis, particularly in rural and underserved regions where access to eye-care specialists is scarce, expensive, or delayed.To overcome these challenges, the project "Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM2)" proposes an intelligent, automated retinal screening and awareness support system. The system integrates the strengths of deep learning–based Convolutional Neural Networks (CNNs) for medical image analysis with the natural language understanding capabilities of Large Language Models (LLMs), including Med-PaLM2. This hybrid approach ensures accurate image-based disease detection while also delivering clear, patient-friendly explanations of the results.In the proposed framework, retinal fundus images are analyzed exclusively by a CNN model built on advanced architectures such as EfficientNet. The CNN extracts high-level visual features and classifies the retinal image into categories such as Normal, Abnormal, or Dangerous, and can further identify specific conditions like Diabetic Retinopathy. This strict separation ensures that medical image interpretation remains reliable and focused within the deep learning model.Once the disease prediction is obtained, the result is provided as textual input to the LLM and Med-PaLM2. These language models do not analyze the image itself; instead, they translate the CNN's prediction into simple, easy-to-understand explanations. The generated output includes details about the detected condition, possible causes, associated risk level (Low, Medium, or High), recommended precautions, and guidance on when to consult an ophthalmologist. To enhance accessibility, the system supports bilingual explanations in both English and Tamil.The application is designed with a professional and user-friendly interface, featuring secure login and sign-up functionality, prediction history tracking, dark and light themes, and an intuitive visual layout. Importantly, the system is intended as a screening and awareness tool rather than a diagnostic or treatment platform. By combining CNN-based retinal image analysis with LLM-driven natural language explanations, the proposed system bridges the gap between technical medical predictions and patient understanding, promoting early detection, health awareness, and timely medical consultation. Asif nawaz et al., 2023, [14], in this research work, a convolutional neural network (CNN) model is proposed for multi-class classification problems with the efficient use of memory consumption. The proposed model has been evaluated on a standard benchmark dataset of Eye Net, having 32 classes of retinal diseases. From experimental evaluation, it has been concluded that the proposed model performs better regarding memory management and accuracy. The overall comparison has been performed based on precision, recall, and accuracy with different numbers of epochs and time consumption by each step. The proposed technique achieved an accuracy of 95% on the Eye-net dataset. Ms. Kavitha. et al., 2024, [15], The study explores hybrid architectures that combine deep learning and transfer learning to assess their capability in achieving a balance between accuracy and minimal memory usage. Evaluation metrics and comparative analyses are presented, shedding light on the performance of various approaches in terms of accuracy, sensitivity, specificity, and computational efficiency. The survey concludes by outlining remaining challenges in achieving efficient multiclass retinal disease detection and proposes future directions for research.

i) Large Language Model (LLM)

A Large Language Model (LLM) is an advanced artificial intelligence model designed to understand, process, and generate human language in a meaningful way. LLMs are trained on large volumes of text data using deep learning techniques, particularly transformer-based architectures, which enable them to learn language patterns, context, and relationships between words. In the proposed Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM2) system, the LLM plays a supportive and explanatory role. It does not analyze retinal images or make medical predictions. Instead, it receives only the textual output generated by the CNN model, such as the detected disease and severity level.

The primary function of the LLM is to convert technical medical predictions into simple, patient-friendly explanations. It provides clear descriptions of the detected condition, possible causes, associated risk levels, and recommended next steps. This helps users better understand their screening results and encourages timely consultation with eye-care professionals. Additionally, the LLM supports bilingual explanation generation, allowing information to be presented in both English and Tamil. By improving clarity and accessibility the LLM bridges the gap between complex AI-based medical analysis and human understanding, making the system more effective as a screening and awareness tool rather than a diagnostic system.

Med-PaLM2 is a medically specialized Large Language Model (LLM) developed to understand and generate high-quality responses in the healthcare domain. It is trained and fine-tuned on a wide range of medical literature, clinical guidelines, and expert-reviewed data, enabling it to provide medically accurate, context-aware, and safe explanations. In the Retinal Multi-Disease Prediction Using CNN and LLM system, Med-PaLM2 is used only for explanation and awareness purposes.

It does not analyze retinal images and does not make disease predictions. All image-based analysis and classification are performed exclusively by the CNN model.

After the CNN predicts the retinal condition and risk level, the result is passed as textual input to Med-PaLM2. Based on this input, Med-PaLM2 generates a clear, patient-friendly explanation that includes:

- A brief description of the detected retinal disease
- Possible causes and risk factors
- General precautions and guidance on consulting an ophthalmologist

Med-PaLM2 is particularly suitable for healthcare applications because it emphasizes medical correctness, clarity, and safety. In this project, it helps bridge the gap between technical AI predictions and human understanding, improving health awareness without replacing professional medical judgment.Overall, the integration of Med-PaLM2 enhances the system's interpretability and usability, making retinal screening results more accessible and meaningful to users from both clinical and non-clinical backgrounds.

II. Related Works

Alyoubi, W. L., Abulkhair, M. F. & Shalash, W. M., 2021, In[1], the researchers proposed a system to classify diabetic retinopathy into five categories using automated diagnostic techniques. They integrated two convolutional neural network models, CNN512 and a YOLOv3-based CNN to handle DR stage classification and lesion localization. The study utilized the APTOS 2019 and DDR datasets. It implemented preprocessing techniques such as CLAHE, noise removal, cropping, and augmentation before training the models. CNN512 was specifically trained for classification tasks, while the YOLOv3 model was used for detecting lesions. This combined approach improved the classification accuracy to 89%, achieving a sensitivity of 89% and a specificity of 97.3%. Çinarer, G., Kilic, K. & Parlar, T.2022, In [2], the authors introduced a deep transfer learning framework for the automatic classification and detection of diabetic retinopathy stages. They utilized the APTOS 2019 Blindness Detection dataset, divided it into training and testing sets and then applied augmentation techniques to the training data. The preprocessing involved five steps including Gaussian Filtering, cropping, rescaling, and the application of CLAHE (Contrast Limited Adaptive Histogram Equalization). The pre-processed images were then fed into three models: ResNet152, DenseNet201 and VGGNet19. Among these models, DenseNet201 achieved the highest test accuracy of 82.7% while ResNet152 attained the highest AUC value of 94.1%.

Methods & Models	Diseases	Performance	Advantages	Research Gaps and Future Work
CNN512 + YOLOv3-based CNN (Alyoubi et al., 2021)	Diabetic Retinopathy (5-class classification + lesion localization)	Accuracy: **89%**Sensitivity: **89%**Specificity: **97.3%**	Joint classification and lesion detection improves diagnostic reliability; effective preprocessing (CLAHE, augmentation)	Computational complexity is high; performance could be improved using attention mechanisms and lightweight architectures; limited generalization testing
ResNet152, DenseNet201, VGG19 (Transfer Learning) (Çinarer et al., 2022)	Diabetic Retinopathy (multi-stage)	Best Accuracy: **82.7% (DenseNet201)**Best AUC: **94.1% (ResNet152)**	Robust feature extraction via transfer learning; systematic preprocessing pipeline	Accuracy is comparatively lower; ensemble or attention-based models could improve results
InceptionV3 + Logistic Regression (PySpark DL Pipeline) (Kotiyal & Pathak, 2022)	Diabetic Retinopathy (Binary)	Accuracy: **95%**AUC: **94.98%**F1-score: **95%**	Efficient big-data processing using Apache Spark; scalable framework	Limited to binary classification; lesion localization and severity grading not addressed
VGG-based CNN (Skouta et al., 2021)	Diabetic Retinopathy (Binary)	Accuracy: **95.5%**	Simple architecture; effective data augmentation; good generalization	Does not support multi-class DR severity classification; lacks explainability
SqueezeNet + AVO Optimization + Blockchain (Uppamma & Bhattacharya, 2023)	Diabetic Retinopathy	Accuracy: **94.2%**Sensitivity: **94.8%**Specificity: **93.4%**	Secure medical data storage via blockchain; optimized feature selection	Blockchain overhead may affect scalability; real-time clinical validation needed
Gray-Level Features + XGBoost Ensemble (Sikder et al., 2021)	Diabetic Retinopathy (Severity Classification)	Accuracy: **94.20%**F-measure: **93.51%**Recall: **92.69%**	Handles class imbalance effectively; interpretable ensemble model	Relies on handcrafted features; deep learning-based feature learning could enhance robustness
AlexNet-based CAD System (OCT Images) (Ghazal et al., 2020)	Non-Proliferative Diabetic Retinopathy	Accuracy: **94%**Recall: **100%**Specificity: **88%**	Optimized for OCT imaging; low computational cost	Limited to OCT modality; fundus image-based validation needed
Hybrid CNN Models (NASNet, EfficientNet, Xception, etc.) (Liu et al., 2020)	Diabetic Retinopathy (5-class)	Best Accuracy: **86.34% (Hybrid-c)**	Multi-model fusion improves robustness; evaluated on multiple datasets	Performance still moderate; training complexity is high; needs explainability
ResNet + GoogleNet + APSO + ML Classifiers (Jabbar et al., 2024)	Diabetic Retinopathy (Binary & Multi-class)	Accuracy: **94%**	Hybrid deep features with optimization improve classification; flexible ML classifiers	Needs real-time deployment analysis; broader dataset validation required
Attention-Guided CNN (AG-CNN) (Moustari et al., 2024)	Diabetic Retinopathy (Severity Classification)	Accuracy: **98.48%**AUC: **0.998**	Attention mechanism enhances lesion-focused learning; state-of-the-art performance	High computational demand; future work could focus on lightweight attention models and clinical deployment

Kotiyal, B. & Pathak, H.,2022 [3] proposed a binary classification framework for detecting diabetic retinopathy using transfer learning. They combined PySpark with deep learning leveraging the capabilities of Big Data tools to process the IDRiD dataset (Indian Diabetic Retinopathy Image Dataset). For preprocessing, the dataset underwent cleaning where dark images were removed followed by cropping to eliminate unwanted spaces and resizing. The classification was performed using a logistic regression (LR) classifier. Authors employed DL Pipelines on

Apache Spark for this purpose, splitting the dataset into 80% for training and 20% for testing. Results indicated that among the three transfer learning models utilized, InceptionV3 demonstrated the best performance achieving an accuracy of 95%, an AUC of 94.98%, and an F1-Score of 95%. Skouta, A., Elmoufidi, A., Jai-Andaloussi, S. & Ochetto, 2021 [4] introduced a binary classification framework for diabetic retinopathy detection using Convolutional Neural Networks (CNNs). They adopted the VGG (Visual Geometry Group) architecture which consists of four blocks, each containing a convolutional layer followed by maximum pooling layers. ReLU was used as the activation function in these layers while the output layer employed the sigmoid function for binary classification. The loss function employed was binary cross-entropy with Adam serving as the optimizer. The dataset was sourced from diabetic-retinopathy-classified via Kaggle with image augmentation applied to enhance diversity and generalization. The proposed model in this study achieved an accuracy of 95.5%.Uppamma, P. & Bhattacharya,2023 In[5], a blockchain-based healthcare framework is proposed for the detection of diabetic retinopathy using deep learning. The research utilizes the publicly available IDRiD dataset from IEEE DataPort. Preprocessing of the dataset involves median filtering followed by lesion segmentation. Hyperparameter tuning is conducted using the Taylor African Vulture Optimization (AVO) algorithm. The most relevant features are then selected and input into the SqueezeNet classifier to predict the occurrence of diabetic retinopathy. The resulting output is securely stored within the blockchain architecture which is accessible to the Electronic Health Record (EHR) manager. Comparative analysis against previous research demonstrated that the proposed model outperformed achieving an accuracy of 94.2%, sensitivity of 94.8%, and specificity of 93.4%. Sikder, N. et al. 2021 In their paper [6], researchers introduced an ensemble learning framework for categorizing the severity of diabetic retinopathy. Their methodology incorporates a gray level intensity algorithm alongside a decision tree-based ensemble learning approach. The APTOS 2019 dataset is utilized in their research. Before classification, the authors implemented several preprocessing techniques including feature extraction and selection. Images were resized and to tackle class imbalance issues, augmentation techniques were applied. Ghazal, M., Ali, S. S., Mahmoud, A. H., Shalaby, A. M. & El-Baz, 2020 In their work [7], researchers introduced a computer-aided diagnostic (CAD) system tailored for detecting non-proliferative diabetic retinopathy through convolutional neural networks (CNNs). Their CNN is optimized for optical coherence tomography (OCT) imaging. They trained the model on an OCT dataset implementing crucial preprocessing steps to extract retina patches for CNN training.

Transfer learning principles and effective feature combination techniques were also employed to improve performance. Utilizing the AlexNet CNN with an input size of 227 × 227, the methodology achieved the highest accuracy with minimal computational complexity. By combining output features from two independently trained CNNs, the system attained 94% accuracy, 100% recall, and 88% specificity.

Liu, H. et al.,2020In [8], the authors introduced three distinct hybrid models for classifying diabetic retinopathy (DR) into five classes: Hybrid-a, Hybrid-f, and Hybrid-c. These hybrids models integrate five base CNN models: NASNet Large, EfficientNetB5, EfficientNetB4, InceptionResNetV2, and Xception. Two loss functions are used to train the base models and then their outputs train the hybrid models. Experiments are conducted on three datasets: APTOS, EyePACS, and DeepDR. Preprocessing involves both initial image enhancement techniques and those applied during training. Among these hybrid models, Hybrid-c achieved the highest accuracy of 86.34%.

Jabbar, A. et al,2024 In [9], the authors introduce an innovative approach for diabetic retinopathy (DR) classification based on deep learning. They propose a hybrid model combining ResNet and GoogleNet for feature extraction, enhanced by adaptive particle swarm optimization (APSO). These features are then classified using machine learning techniques such as SVM, random forest, decision tree, and linear regression. The study uses the EyePACS dataset and employs preprocessing steps like image resizing, green channel extraction, and top-hat/bottom-hat transformations. Their hybrid model achieved an impressive 94% accuracy, surpassing existing binary and multiclass DR detection techniques. Future work includes testing on diverse datasets to ensure broader applicability. Moustari, A. M., Brik, Y., Attallah, B. & Bouaouina, R,2024 In [10], researchers tackle the challenges associated with determining the severity of Diabetic Retinopathy (DR) by presenting a novel convolutional neural network (CNN) architecture known as the attention-guided CNN (AG-CNN) which was characterized by its dual-branch design. They utilize the APTOS 2019 DR Dataset, emphasizing the necessity of initial preprocessing to enhance image quality. The AG-CNN features a global branch for overall attention and a local branch that targets important localized features. Their experimental results indicate that the baseline model, DenseNet-121, achieves an accuracy of 97.46% and an AUC of 0.995, while the AG-CNN significantly improves these metrics, reaching an accuracy of 98.48% and an AUC of 0.998.

III. Methodology

The proposed Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM2) system follows a structured methodology to ensure accurate retinal image analysis and effective communication of results to users. The system workflow is organized into multiple stages, starting from image acquisition to result explanation and presentation.

i) Image Acquisition

The dataset used for the Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM2) system was obtained from the Kaggle platform, which hosts a variety of publicly available medical imaging datasets for research and educational use. The dataset contains color retinal fundus images representing both normal eyes and multiple retinal diseases. These images capture different regions of the retina and reflect variations commonly observed in real clinical settings.

ii) Retinal Diseases Included: The dataset includes retinal images corresponding to the following 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

These diseases were selected to cover both frequently occurring and critical vision-threatening retinal disorders. Including a wide range of conditions allows the model to learn diverse retinal patterns and improves its ability to perform reliable multi-disease screening.

iii) Data Characteristics: The dataset consists of color retinal fundus photographs stored in standard image formats such as JPEG and PNG. The images vary in resolution and acquisition settings, as they originate from different sources.During preprocessing, all images are resized and standardized to ensure uniform input to the CNN model. The dataset includes both macula-centered and optic disc–centered views, enabling effective analysis of retinal structures relevant to different diseases. Image size is 320x314.

iv) Preprocessing. Prior to analysis, the images are preprocessed to improve quality and consistency. Preprocessing steps include image resizing to a fixed resolution, normalization of pixel values, and enhancement techniques to improve contrast and visibility of retinal structures. Data augmentation methods such as rotation, flipping, and scaling are applied during training to increase data diversity and improve model generalization.

v) CNN-Based Retinal Disease Classification.A Convolutional Neural Network (CNN) based on an EfficientNet architecture is employed for retinal image analysis. The CNN automatically extracts discriminative features such as blood vessel patterns, microaneurysms, hemorrhages, and optic disc characteristics from fundus images. Using these features, the model classifies images into Normal, Abnormal, or Dangerous categories and identifies specific retinal disease.

vi) CNN-Based Feature Extraction

A Convolutional Neural Network (CNN) extracts features from the input retinal image using convolution operations.

Let the input retinal image be represented as:

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

where

H = height,

W = width,

C = number of color channels.

The convolution operation for a feature map 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

This operation enables the CNN to learn patterns such as blood vessels, lesions, and structural abnormalities.

vii) Activation Function (ReLU):

After convolution, the Rectified Linear Unit (ReLU) activation function is applied:

ReLU(x)=max(0,x)

This introduces non-linearity and helps the network learn complex retinal patterns.

viii) 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 while maintaining computational efficiency.

ix) Softmax Classification

The final layer uses the Softmax function to classify retinal images into Normal, Abnormal, or Dangerous:

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

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.

x) Loss Function (Categorical Cross-Entropy)

The CNN is trained using categorical 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

This loss function penalizes incorrect predictions and guides model optimization.

xi) Optimization (Adam Optimizer)

The Adam optimizer updates network weights using adaptive learning rates:

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

where:

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

xii) 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

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

These metrics evaluate model performance on unseen retinal images.

xiii) 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.

All image-related analysis and disease prediction are performed exclusively by the CNN to maintain accuracy and reliability.

xiv) Model Training and Evaluation. The CNN model is trained using labeled retinal images with appropriate optimization techniques. The dataset is divided into training, validation, and testing subsets to evaluate model performance. Standard evaluation metrics such as accuracy, precision, recall, and F1-score are used to assess classification effectiveness and ensure robustness on unseen data.

xv) Prediction Output Processing.

Once classification is completed, the CNN generates a structured textual output containing the predicted disease, severity level, and confidence score. This output is used as the only input for the explanation stage, ensuring a clear separation between image analysis and language generation.

xvi) Explanation Generation Using LLM. The textual prediction produced by the CNN is passed to a Large Language Model (LLM), including Med-PaLM 2, for explanation generation. The language model does not process retinal images or influence the prediction. Its function is limited to converting the medical output into clear, patient-friendly explanations. The generated explanation includes a brief description of the detected condition, possible causes, risk level, recommended precautions, and guidance on consulting an ophthalmologist. Explanations are provided in both English and Tamil to improve accessibility.

xvii) LLM-Based Explanation Mapping:

The explanation process can be represented as:

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

where fLLM generates a patient-friendly explanation without accessing image data.

xviii) System Interface and Integration. The complete system is integrated into a user-friendly application interface. The interface includes secure login and registration modules, image upload functionality, display of prediction results, and storage of prediction history. Users can switch between light and dark themes to enhance u.sability and visual comfort.

xix) Ethical Considerations and Usage Scope. The system is designed strictly for screening and awareness purposes. It does not provide medical diagnosis or treatment recommendations. Clear disclaimers are included to encourage users to consult qualified eye-care professionals for clinical decisions.

xx) Final Output Presentation. The final output combines the CNN-based disease prediction with a simplified explanation generated by the LLM. This approach ensures that technical results are presented in an understandable format, supporting early awareness and timely medical consultation.

IV. System Architecture

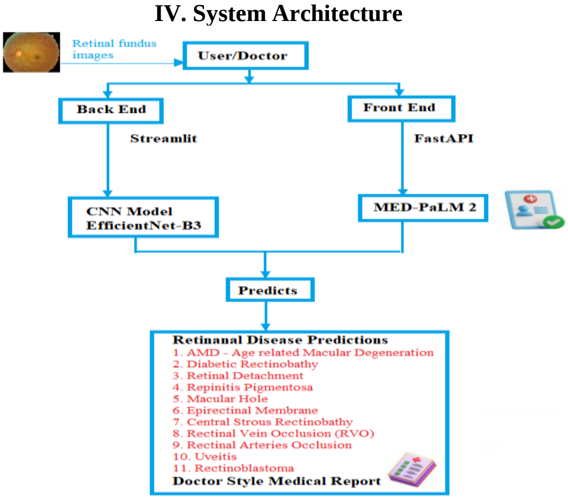


Figure 1: System Architecture

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3832-3839

i) System Architecture working procedure: The diagram illustrates the overall working architecture of the Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM2) system. It shows how retinal fundus images flow through different system components to produce disease predictions and a doctor-style medical report.

1. Input: Retinal Fundus Images

The process begins when a User or Doctor uploads a retinal fundus image. These images capture the internal structures of the eye and are essential for identifying retinal abnormalities.

2. User / Doctor Interface

The uploaded image is handled through the user interface, which connects both the Back End and Front End modules. This layer ensures secure access and smooth interaction between users and the system.

3. Back End – Image Analysis (Streamlit + CNN)

The Back End is implemented using Streamlit, which manages image upload, preprocessing, and interaction with the deep learning model.
The retinal image is passed to a CNN model based on EfficientNet-B3.
This CNN performs all visual analysis, extracting important retinal features and identifying disease-related patterns.
The CNN generates a prediction result indicating the detected retinal disease and its severity.

4. Front End – Explanation Generation (FastAPI + Med-PaLM2)

The Front End, implemented using FastAPI, handles communication with the language model.
The Med-PaLM2 model receives only the textual output from the CNN (not the image).
Its role is to convert the technical prediction into a clear, patient-friendly explanation, similar to how a doctor would explain the condition.

5. Prediction Integration

The outputs from both modules are combined:
The CNN provides the disease prediction.
Med-PaLM2 provides a simplified explanation and guidance.
This integration ensures accuracy in medical analysis while improving understandability for users.

6. Output: Retinal Disease Prediction and Medical Report

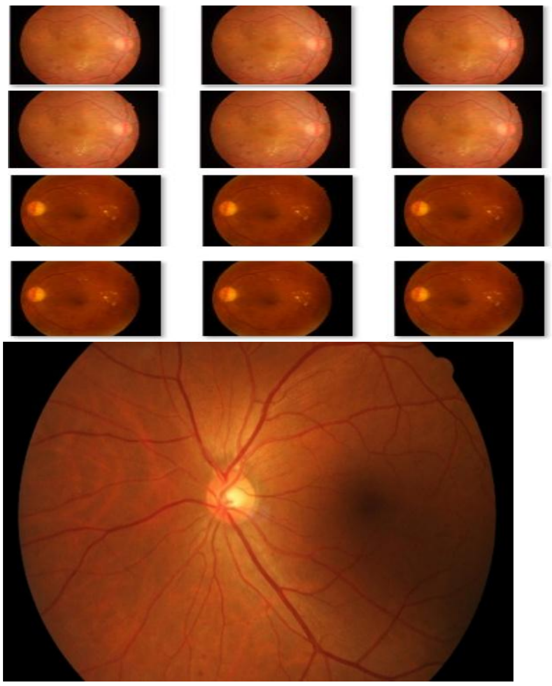
The final output includes:
Detection of one or more retinal diseases such as:

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

A doctor-style medical report explaining the condition, risk level, and recommended next steps.

V. Results And Discussions

Figure 2: Input Images



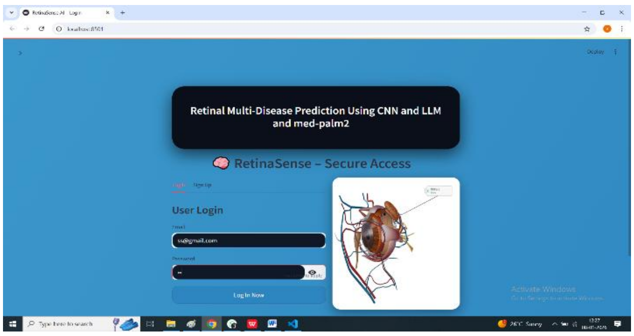


Figure 3: Create Account and Login

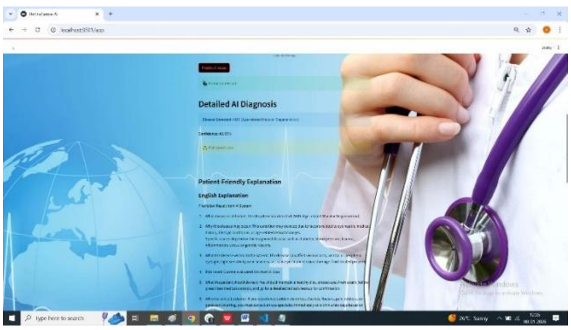


Figure 4: Input Image

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

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3832-3839

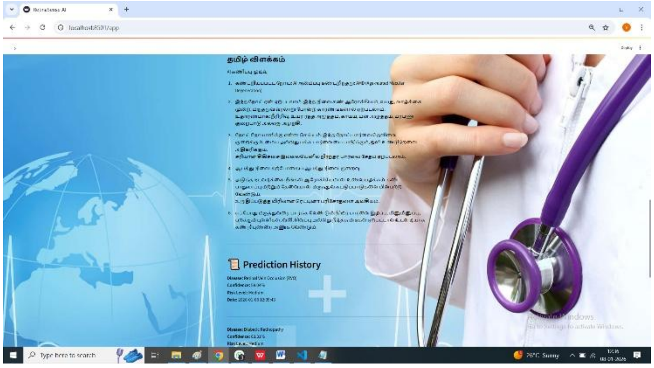


Figure 5: Prediction

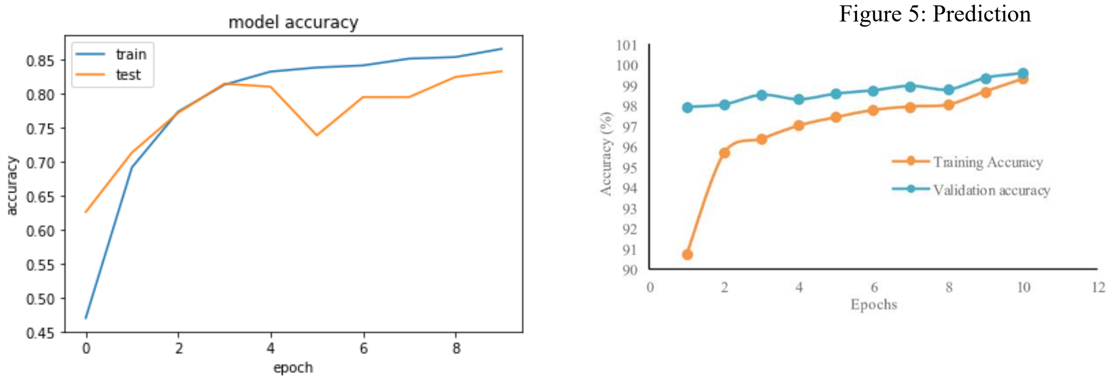


Figure 6: Accuracy of Training and validation accuracy

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 VGG [4]
Accuracy	98.86	95.5
Precision	97.01	95
Recall	97.02	95
F1-score	97.03	95

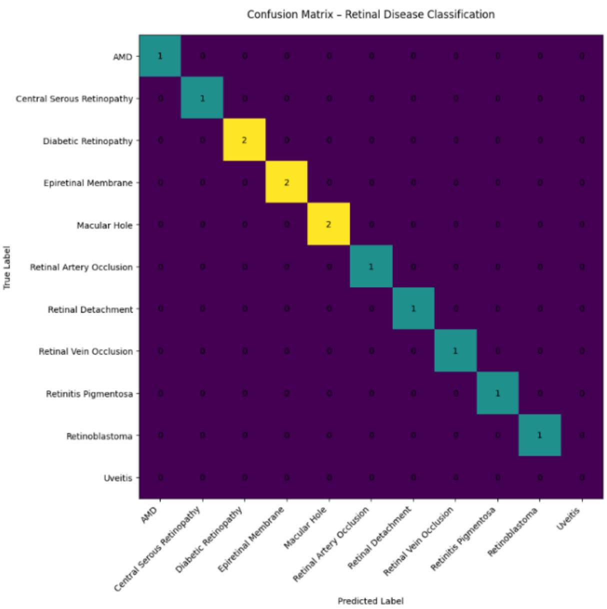


Figure 7: Confusion Matrix for 11 classes

VI. Conclusion and future work

The Retinal Multi-Disease Prediction Using CNN and LLM (Med-PaLM2) system presents an effective and accessible solution for early retinal disease screening and awareness. By combining deep learning–based image analysis with natural language explanation, the system addresses both the technical and communication challenges associated with retinal health assessment. The use of a CNN model based on EfficientNet-B3 enables accurate extraction of retinal features and reliable classification of multiple vision-threatening diseases from fundus images. At the same time, the integration of Large Language Models, including Med-PaLM2, transforms complex medical predictions into clear, patient-friendly explanations, making the results easier to understand for non-technical users. The inclusion of bilingual support further enhances accessibility and usability. The system is designed as a screening and awareness tool, not as a diagnostic or treatment platform. It encourages users to seek timely consultation with qualified ophthalmologists, thereby supporting early intervention and reducing the risk of preventable blindness. Features such as secure authentication, prediction history tracking, and a user-friendly interface contribute to a reliable and practical application. Overall, the proposed system successfully bridges the gap between advanced AI-based retinal analysis and human understanding. By promoting early detection, improving health literacy, and supporting informed decision-making, the system demonstrates strong potential for use in community screening programs and healthcare awareness initiatives.

VII. Reference

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