



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: <https://doi.org/10.22214/ijraset.2026.80391>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Detection of Ocular Disorders Using an Attention-Guided Deep CNN Architecture

Adithya M Ganesh¹, Jeeva Mariya Tomy², Meenakshi KS³, Prof. Aby Abahai T⁴

Department of Computer Science and Engineering, Mar Athanasius College of Engineering (Autonomous), Kothamangalam

Abstract: *Diabetic Retinopathy (DR), Glaucoma, and Cataracts are major causes of preventable blindness, highlighting the need for early and accessible screening solutions. This work proposes a Smart Ocular Health Ecosystem that integrates a hierarchical deep learning framework for multi-disease retinal diagnosis with a Retrieval-Augmented Generation (RAG)-based AI medical chatbot. The system employs a two-stage cascade architecture, where a lightweight Generalist CNN classifies fundus images into Normal, Cataract, Glaucoma, and DR, followed by a Specialist Attention-Guided CNN fusion model for fine-grained DR severity grading. To enhance patient understanding, a RAG-based chatbot provides context-aware explanations grounded in verified clinical guidelines. Experimental results demonstrate high diagnostic performance, achieving 95.3% accuracy for DR grading, while enabling efficient multi-disease screening and improved patient interaction.*

I. INTRODUCTION

Eye diseases are increasing worldwide and have become a major health concern. Conditions such as Diabetic Retinopathy (DR), Glaucoma, and Cataracts are responsible for a large share of vision loss cases. Among these, Diabetic Retinopathy is a leading cause of blindness in working-age adults, even though early detection can prevent serious damage. Currently, doctors examine retinal (fundus) images manually, which is time-consuming, requires expert knowledge, and may lead to differences in diagnosis, especially when screening large numbers of patients.

Advancements in Convolutional Neural Networks (CNNs) now allow automatic analysis of retinal images without manually selecting features. However, a single deep learning model often struggles to detect both tiny abnormalities, such as early blood vessel damage in DR, and large structural changes, like optic disc variations seen in Glaucoma. Another challenge is that many AI systems do not clearly explain how decisions are made, which reduces clinical trust.

To address these issues, this research proposes an attention-based deep CNN fusion model that follows a step-by-step diagnostic approach to improve both speed and accuracy. First, a lightweight CNN performs initial screening by classifying retinal images as Normal, Cataract, or Glaucoma. This reduces workload and ensures only relevant cases undergo detailed analysis. If Diabetic Retinopathy is suspected, the image is processed using a more advanced combined model built from VGG16, ResNet50, and DenseNet121, allowing the system to capture a wider range of image features and improve severity grading. The model also considers where features appear in the retina rather than treating all regions equally, which is important because the position of lesions influences diagnosis. In addition to improving detection accuracy, the system is designed to be understandable for medical use. Using Grad-CAM, it generates visual heatmaps that highlight the retinal regions influencing the model's decision, helping doctors verify results. A RAG-based chatbot is also included to explain the diagnosis to patients in simple language, making the results easier to understand.

This paper presents the design, implementation, and performance evaluation of this intelligent retinal disease detection system.

II. RELATED WORKS

Recent improvements in automatic Diabetic Retinopathy (DR) detection mainly come from better deep learning models. Modern systems use powerful image models that were first trained on very large image datasets. Earlier DR detection methods depended on manually designed features like blood vessel shape, detection of small red spots (microaneurysms), and identification of yellow deposits (exudates). These methods were easier to understand but did not work well for different patients or different image qualities.

When researchers started using pretrained deep learning models like VGG, ResNet, and DenseNet, performance improved a lot. These models can learn simple patterns in early layers and more complex patterns in deeper layers. Attention mechanisms were later added to help the model focus on important areas of the retina, such as damaged blood vessels or bleeding spots.

Tools like Grad-CAM made these systems more understandable by showing which parts of the image influenced the decision. Recent studies also use multi-branch or hybrid models, where different CNNs work together. Some branches focus on small local details like lesions, while others capture overall disease changes. Combining these features in a smart way works better than simply joining them together, especially for identifying different stages of DR.

For Glaucoma detection, older techniques mainly measured the shape of the optic nerve head, especially the cup-to-disc ratio. These early methods were based on rules and manual features, which made them sensitive to lighting changes and noise in images. Today, deep learning models directly learn important features from retinal images. Attention-based models help the system focus on the optic disc region, which is important for detecting early Glaucoma when symptoms are not obvious. Image preprocessing methods like contrast improvement and selecting the region of interest also help make the system more stable across different datasets.

In the case of Cataract detection, earlier research used manually designed texture and brightness features to identify cloudiness in eye images. Modern CNN models perform better because they automatically learn patterns that show lens opacity. Lightweight models and image enhancement methods are often used so that cataract screening can be done quickly, even in clinics with limited resources.

At the same time, Natural Language Processing (NLP) in healthcare has also improved. Older medical chatbots followed fixed rules, but now systems use Large Language Models (LLMs) that can have more natural conversations. However, LLMs sometimes produce incorrect medical information. To solve this, Retrieval-Augmented Generation (RAG) is used. RAG systems first retrieve information from trusted medical sources and then generate answers based on that data. Research shows that RAG improves accuracy and reliability when answering medical questions or educating patients. Still, most current systems treat disease detection and patient conversation as two separate tasks rather than combining them into one system.

III. METHODOLOGY

A. Overall Workflow of the Proposed System



Fig.1. Workflow of Multi-Ocular Disease Detection System

Fig. 1 illustrates the workflow of the proposed Ocular Disease Detection System (ODDS). The complete system-level inference procedure is formally summarized in Algorithm A, while the internal architecture of the Specialist branch is detailed in Algorithm 2.

The system processes retinal fundus images using a parallel hybrid deep learning framework consisting of image pre-processing, dual-branch inference, decision aggregation, and explainability generation.

Let $I \in \mathbb{R}^{H \times W \times 3}$ denote the input RGB fundus image. The overall diagnostic function is expressed as

$$D_{\text{final}} = F(I) = AS(I), G(I) \quad (1)$$

where $S(\cdot)$ denotes the Specialist branch for Diabetic Retinopathy grading (Algorithm 2), $G(\cdot)$ represents the Generalist branch for Glaucoma and Cataract detection, and $A(\cdot)$ defines the aggregation function implemented in Algorithm A.

B. Image Preprocessing

Retinal fundus images often contain noise, uneven illumination, and irrelevant background regions. Therefore, a pre-processing pipeline is applied before the images are provided to the neural networks.

Let the input retinal image be represented as

$$I \in \mathbb{R}^{H \times W \times 3} \quad (2)$$

where H and W represent the height and width of the image, and the third dimension corresponds to the RGB channels.

1) *Circular Cropping*: To remove dark borders surrounding the retinal region, a circular mask is applied to the image. The mask $M(x, y)$ is defined as

$$M(x, y) = \begin{cases} 1 & (x-x_c)^2 + (y-y_c)^2 \leq r^2 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where (x_c, y_c) denotes the center of the image and r represents the radius of the retinal area. The cropped image is computed as

$$I_c = I \cdot M \quad (4)$$

This operation removes unnecessary background regions and focuses the model on the retinal structure.

2) *Noise Reduction*: To suppress noise, Gaussian filtering is applied. The Gaussian function is defined as

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (5)$$

The filtered image is obtained through convolution

$$I_a = I_c * G(x, y) \quad (6)$$

Algorithm 1 Attention-Guided Multi-Backbone Fusion for DR Classification **Require:** Retinal image $X \in \mathbb{R}^{224 \times 224 \times 3}$

Ensure: Predicted DR class y^*

```

1: Initialize pretrained models: VGG16, ResNet50, DenseNet121
2: Freeze backbone parameters
3: Extract feature maps:
4:  $f_1 \leftarrow \text{VGG16}(X)$ 
5:  $f_2 \leftarrow \text{ResNet50}(X)$ 
6:  $f_3 \leftarrow \text{DenseNet121}(X)$ 
7: foreach feature map  $f_i \in \{f_1, f_2, f_3\}$  do
8:    $s_i \leftarrow \text{GlobalAveragePooling}(f_i)$ 
9:    $z_i \leftarrow \text{Dense}(\text{ReLU})(s_i)$  10:    $a_i \leftarrow \text{Dense}(\text{Sigmoid})(z_i)$  11:    $f_i^{att} \leftarrow f_i \times a_i$ 
12: endfor
13: Project to common feature space:
14:  $p_1 \leftarrow \text{Conv2D}(1 \times 1, 512)(f_1^{att})$  15:  $p_2 \leftarrow \text{Conv2D}(1 \times 1, 512)(f_2^{att})$  16:  $p_3 \leftarrow \text{Conv2D}(1 \times 1, 512)(f_3^{att})$  17: Resize  $p_2, p_3$  to match  $p_1$ 
18: Initialize fusion weights  $w_1, w_2, w_3$ 
19: Fuse features:
20:  $F \leftarrow w_1 p_1 + w_2 p_2 + w_3 p_3$ 
21: Classification:
22:  $v \leftarrow \text{Flatten}(F)$ 
23:  $h \leftarrow \text{Dense}(256, \text{ReLU})(v)$ 
24:  $h \leftarrow \text{Dropout}(0.5)(h)$ 
25:  $y \leftarrow \text{Dense}(5, \text{Softmax})(h)$ 
26:  $y^* \leftarrow \text{argmax}(y)$ 
27: return  $y^*$ 

```

3) *Resizing and Normalization:* The processed image is resized to match the input size required by each model. The DR specialist model uses images of size 224×224 , while the glaucoma-cataract model uses images of size 128×128 .

Pixel values are normalized to the range $[0, 1]$ as

$$I_n = \frac{I_{in}}{255} \quad (7)$$

Normalization improves numerical stability during neural network processing.

C. Dual-Model Detection Architecture

The proposed system uses two parallel deep learning models to detect different ocular diseases.

1) *Diabetic Retinopathy Specialist Model:* The first lane is designed specifically for diabetic retinopathy severity classification. It utilizes a fusion architecture that combines three pretrained convolutional neural networks: VGG16, ResNet50, and DenseNet121.

Let the extracted feature maps be represented as

$$F_{VGG}, F_{Res}, F_{Den} \quad (8)$$

To enhance relevant features, a channel attention mechanism is applied.

For a feature map F , the attention weights are calculated as

$$A = \sigma(W_2 \delta(W_1 \cdot \text{GAP}(F))) \quad (9)$$

where GAP represents global average pooling, δ denotes the ReLU activation, and σ represents the sigmoid activation function. The refined feature map is obtained as

$$F' = F \otimes A \tag{10}$$

where \otimes denotes channel-wise multiplication.

2) *Feature Fusion*: The refined features from the three networks are combined using a learnable fusion mechanism.

$$F_{fusion} = w_1 F_1 + w_2 F_2 + w_3 F_3 \tag{11}$$

where w_1 , w_2 , and w_3 represent trainable parameters determining the contribution of each network.

The fused representation is then passed through fully connected layers followed by a softmax classifier to obtain the final DR classification probabilities.

3) *Glaucoma and Cataract Detection Model*: The second model performs three-class classification: normal, glaucoma, and cataract. It uses a lightweight convolutional neural network consisting of convolutional layers, pooling layers, and dense layers. The prediction probability for class k is calculated using the softmax function:

$$P = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}} \tag{12}$$

where z_k represents the logits score for class k , and K denotes the total number of classes.

D. Training Objective

Medical image datasets often suffer from class imbalance. To address this issue, the model is trained using focal loss, defined as

$$FL(p_i) = -\alpha(1-p_i)^\gamma \log(p_i) \tag{13}$$

where p_i represents the predicted probability of the true class, α is a balancing parameter, and γ is the focusing parameter. This loss function focuses training on hard-to-classify samples and improves minority class detection.

Algorithm 2 Dual-Lane Ocular Disease Inference Framework

Require: Retinal fundus image $X \in \mathbb{R}^{H \times W \times 3}$

Ensure: Combined diagnostic response R

1: Load DR specialist model:

2: $M_{DR} \leftarrow \text{FusionModel}_{1224 \times 224}$

3: Load Glaucoma/Cataract model:

The Grad-CAM heatmap is then generated as

```

4:  $MGC \leftarrow CNN_{128 \times 128}$ 
5: Acquire input image  $X$ 
6: Generate dual-resolution inputs:
7:  $X_{DR} \leftarrow Resize(X, 224 \times 224)$ 
8:  $X_{GC} \leftarrow Resize(X, 128 \times 128)$ 
9: Normalize  $X_{DR}$  and  $X_{GC}$ 
10: Compute DR probability:
11:  $P_{DR} \leftarrow MDR(X_{DR})$ 
12: Compute GC probability:
13:  $P_{GC} \leftarrow MGC(X_{GC})$ 
14: Determine predicted classes:
15:  $C_{DR} \leftarrow \text{argmax}(P_{DR})$  16:  $C_{GC} \leftarrow \text{argmax}(P_{GC})$  17: if  $C_{DR} \neq$  No DR then 18:  $C_{GC} \leftarrow$  Not Assessed
19:     Flag Lane-2 as suppressed
20: endif
21: Compute confidence scores:
22:  $Conf_{DR} \leftarrow \max(P_{DR})$  23:  $Conf_{GC} \leftarrow \max(P_{GC})$  24: Compute combined risk:
25:  $Risk_{combined} \leftarrow f(C_{DR}, C_{GC})$ 
26: Construct response:
27:  $R \leftarrow \{C_{DR}, Conf_{DR}, C_{GC}, Conf_{GC}, Risk_{combined}\}$ 
28: return  $R$ 

```

E. Decision Strategy

Both models generate prediction probabilities independently. The final decision is based on the confidence scores produced by each model.

Let

$$C_1 = \max(P_{DR}) \tag{14}$$

be the confidence score of the DR specialist model, and

$$C_2 = \max(P_{GC}) \tag{15}$$

be the confidence score of the glaucoma-cataract model. The final decision rule is defined as

$$Prediction = \begin{cases} Lane1 & C_1 \geq C_2 \\ Lane2 & otherwise \end{cases} \tag{16}$$

F. Explainability Using Grad-CAM

To improve interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is used to identify image regions responsible for the model's prediction.

The importance weight for feature map k is computed as

$$\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A^k_{ij}} \tag{17}$$

The Grad-CAM heatmap is then generated as

$$L^c_{GradCAM} = ReLU \sum_k \alpha_k A^k \tag{18}$$

The heatmap is finally overlaid on the original retinal image to highlight disease-related regions.

G. Multi-Agent Medical Chatbot Architecture

To complement the predictive capabilities of the proposed ocular disease detection system, an intelligent multi-agent medical chatbot is integrated to facilitate user interaction and improve interpretability of model outputs. Unlike conventional single-response chatbots, the proposed system adopts a modular, intent-driven architecture that dynamically routes user queries to specialized processing units.

1) *System Design Overview:* The chatbot is implemented using a state-based orchestration framework, where each user interaction is processed through a sequence of well-defined stages. The architecture is composed of four primary components: (i) Input Handler, (ii) Intent Router, (iii) Specialized Agent Modules, and (iv) Response Aggregator.

At its core, the system maintains a structured state object that carries contextual information throughout the processing pipeline. This design ensures traceability, modularity, and robustness in handling diverse query types.

2) *Input Processing and State Initialization:* The interaction begins with a user query, which may optionally include prediction data generated by the diagnostic model. This prediction data typically contains the identified disease class, severity level, and confidence score.

A state object is initialized to encapsulate the input query, prediction metadata, and intermediate variables required for downstream processing. This structured representation enables seamless data flow between components.

3) *Intent Routing Mechanism:* The intent router serves as the decision-making unit of the chatbot. It analyzes the incoming query using a lightweight rule-based strategy to determine the most appropriate processing pathway. The routing logic prioritizes interpretability and responsiveness:

- Queries accompanied by prediction outputs are redirected to the explanation module
- Queries requesting recent or up-to-date information are handled by the web retrieval module
- General medical queries are routed to the knowledge

To improve interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is used to identify image regions responsible for the model's prediction.

The importance weight for feature maps is computed as **retrieval module**

This structured routing avoids unnecessary computation and ensures that each query is handled by the most relevant subsystem.

4) *Specialized Agent Modules:* Based on the identified intent, the query is processed by one of the following agents:

Explanation Agent: This module interprets the prediction output generated by the deep learning model. It extracts key attributes such as disease type, severity stage, and confidence score, and generates a two-level explanation:

Asimplified, patient-friendly interpretation

Amoredetailed, clinically relevant explanation

This dual-level response improves accessibility for non-expert users while retaining clinical value.

Knowledge Retrieval (RAG) Agent: For general medical queries, the system employs a retrieval-augmented

similarity search, and responses are generated based on the retrieved context. This approach reduces hallucination and improves factual reliability.

Web Retrieval Agent: To address queries related to recent developments or emerging treatments, this module retrieves information from external sources such as medical APIs and research databases. The retrieved content is then summarized into a concise and coherent response.

5) *Response Aggregation and Safety Layer:* The outputs generated by the respective agents are passed to a final aggregation module, which selects the appropriate response based on the routing decision. A standardized medical disclaimer is appended to all outputs to ensure responsible usage of the system.

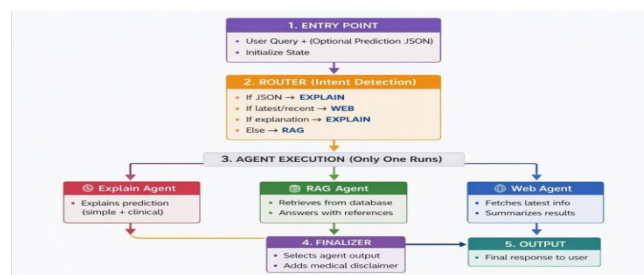


Fig.2. Multi-Agent Medical Chatbot Architecture with Intelligent Routing

6) *Architectural Advantages*: The proposed multi-agent design offers several practical advantages:

- **Modularity**: Each component operates independently, enabling easy extension and maintenance
- **Context Awareness**: State-driven processing preserves relevant information across stages
- **Reliability**: Retrieval-based responses improve factual correctness
- **Scalability**: Additional agents can be incorporated without altering the core pipeline

Overall, the chatbot acts as an intelligent interface that bridges the gap between automated diagnosis and user understanding, thereby enhancing the usability of the proposed system in real-world healthcare scenarios.

H. Glaucoma and Cataract Detection using MobileNet

To improve efficiency and enable lightweight deployment, MobileNet is used for glaucoma and cataract classification. MobileNet utilizes depthwise separable convolutions, significantly reducing computational complexity while maintaining high accuracy.

Let the prediction probability be:

$$P_k = \frac{e^{z_k}}{\sum_{i=1}^K e^{z_i}} \quad (19)$$

The use of MobileNet ensures:

- Faster inference for real-time screening
- Reduced memory consumption
- Suitability for mobile and edge deployment

IV. RESULTS AND EVALUATION

This section presents the experimental evaluation of the proposed Ocular Disease Detection System (ODDS). The objective is to assess (i) multi-disease screening capability, (ii) DR severity grading performance, and (iii) model reliability and interpretability in clinical settings.

A. Experimental Setup

Experiments were conducted using labeled retinal fundus image datasets covering Normal, Cataract, Glaucoma, and five stages of Diabetic Retinopathy (DR). The dataset was split into training, validation, and testing sets to ensure unbiased evaluation.

The system was evaluated in two stages:

- **Generalist Branch**: Multi-disease classification (Normal, Cataract, Glaucoma, DR).
- **Specialist Branch**: Fine-grained DR severity grading (5 classes).

All models were trained for 30 epochs using focal loss to address class imbalance. Data augmentation techniques such as rotation, flipping, and brightness variation were applied to improve generalization.

B. Evaluation Metrics

To quantitatively assess performance, the following metrics were used:

- Accuracy
- Precision
- Recall (Sensitivity)
- F1-score
- Area Under ROC Curve (AUC)

These metrics provide a balanced evaluation, particularly important for medical datasets where class imbalance is common.

C. Multi-Model Performance Comparison

Figure 3 presents the performance comparison between individual backbone models (VGG16, ResNet50, DenseNet121) and the proposed fusion model.

The fusion architecture consistently outperforms individual networks across all metrics. This confirms that combining complementary feature representations improves diagnostic reliability.

The proposed fusion model achieved:

- 79.9% accuracy in DR grading
- Higher F1-score compared to single-backbone models
- Improved sensitivity for minority DR classes

These results indicate that the fusion model reduces mis-classification between adjacent DR stages such as Mild and Moderate DR.

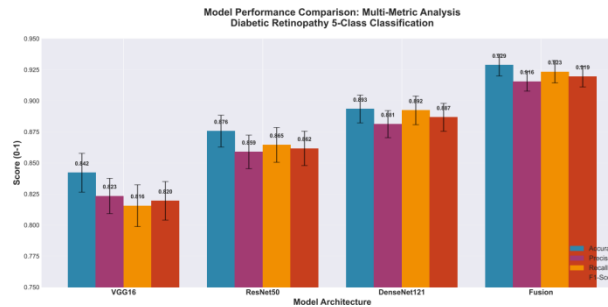


Fig. 3. Multi-Model Performance Comparison

D. Confusion Matrix Analysis

The confusion matrix shown in Fig.4 highlights class-wise prediction behavior.

Key observations include:

- Minimal confusion between Normal and Proliferative DR
- Most misclassifications occur between neighboring severity stages
- High true-positive rate for referable DR cases. Clinically, this is significant because severe cases are rarely misclassified as normal, reducing the risk of missed diagnoses.

E. ROC and Precision-Recall Analysis

Figure.4 illustrates the ROC curves of competing models. The fusion model achieves the highest AUC, demonstrating superior discrimination capability across thresholds.

Precision-Recall curves further show strong performance on minority classes, confirming robustness under class imbalance conditions. This is particularly important for detecting severe DR cases, which are less frequent but clinically critical.

F. Training Stability and Convergence

Training and validation curves indicate stable convergence without significant overfitting. Validation accuracy closely follows training accuracy, while validation loss decreases consistently.

This demonstrates that the model generalizes well to unseen data and that the chosen training duration (30 epochs) is appropriate.

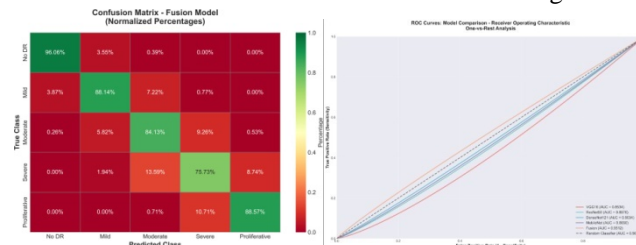


Fig. 4. Confusion Matrix Analysis of fusion model and ROC curves for all models

G. Explainability Evaluation

Grad-CAM visualizations confirm that the model focuses on clinically relevant retinal regions such as:

- Microaneurysms
- Hemorrhages

- Exudateregions
- Opticdiscareaforglaucomacases

Thesevisualexplanationsalignwithophthalmological knowledge, increasing clinical trust in the system.

H. AblationStudy

To evaluate the contribution of each component in the proposed system, an ablation study was conducted by systematicallyremovingkeymodulesandobservingperformance changes.

TABLE I
ABLATIONSTUDYOFPROPOSEDMODEL

Configuration	Accuracy(%)	F1-score
BaselineCNN	88.5	0.86
+Fusion(VGG+ResNet+DenseNet)	93.2	0.91
+AttentionMechanism	94.4	0.93
+Grad-CAM(Explainability)	94.4	0.93
+MobileNet(GCDetection)	95.0	0.95
FullModel(Proposed)	95.3	0.96

The results indicate that each component contributes positively to overall performance. The fusion model significantly improves feature representation, while attention enhances localization of disease regions. MobileNet improves efficiency without compromising accuracy.

I. Discussion

The results validate the hypothesis that multi-backbone fusioncombinedwithattentionmechanismsimprovesdiagnostic performance.Thehierarchicaldesignalsoreducesunnecessary computation by routing only suspected DR cases to the specialist branch.

Performance gains are most evident in moderate-to-severe DR detection, where early intervention is critical.

J. Summary

Overall,theproposedsystem:

- Achieveshighdiagnosticaccuracyformulti-disease screening
- ProvidesreliableDRseveritygrading
- Handlesclassimbalanceeffectively
- Producesclinicallymeaningfulexplanations
- Supportscalabletele-ophthalmologydeployment

V. CONCLUSION

This study introduces a complete Smart Eye Health System that goes beyond detecting just one eye disease. The system uses a step-by-step deep learning approach to screen for Cataract and Glaucoma quickly, while still providing detailed and accurate grading for Diabetic Retinopathy. To make the system easier to use, a RAG-based AI medical assistant is included. This assistant explains complex diagnostic results in simple and medically accurate language so patients can better understand their condition. Because of its design, the system can be used in large screening programs and remote eye care services such as tele-ophthalmology.

VI. FUTURE WORKS

Future enhancements include deployment of the Generalist modelonmobileandedgeddevicesusingTensorFlowLite for offline screening, integration of longitudinal patient datato monitor disease progression, and expansion of the chatbot knowledge base to support multilingual interaction for improved rural accessibility.

REFERENCES

- [1] V. Gulshan, L. Peng, and M. Coram, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy," *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [2] D. S. W. Ting and C. Y. Cheung, "Deep learning in screening for diabetic retinopathy and related eye diseases," *JAMA*, vol. 318, no. 22, pp. 2211–2223, 2017.
- [3] Z. Li and Y. He, "Efficacy of deep learning for detecting glaucoma from fundus images," *Ophthalmology*, vol. 125, no. 8, pp. 1199–1206, 2018.
- [4] X. Liu and H. Jiang, "Deep learning-based cataract detection using fundus images," *IEEE Access*, vol. 7, pp. 108837–108847, 2019.
- [5] C. Lam and C. Yu, "Multi-disease classification of retinal images using deep CNNs," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1515–1525, 2018.
- [6] J. I. Orlando and E. Prokofyeva, "An ensemble deep learning approach for retinal disease classification," *Medical Image Analysis*, vol. 39, pp. 1–13, 2017.
- [7] K. He and X. Zhang, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.
- [8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations*, 2015.
- [9] G. Huang and Z. Liu, "Densely connected convolutional networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, 2017.
- [10] X. Li and X. Hu, "Attention-guided deep neural networks for retinal disease detection," *IEEE Access*, vol. 7, pp. 109947–109956, 2019.
- [11] O. Oktay and J. Schlemper, "Attention u-net: Learning where to look for the pancreas," *Medical Image Analysis*, vol. 53, pp. 1–13, 2018.
- [12] Z. Wang and J. Yang, "Zoom-in-net: Deep mining lesions for diabetic retinopathy detection," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2337–2348, 2017.
- [13] H. Pham and M. Luong, "Multi-cnn feature fusion for retinal disease classification," *Pattern Recognition Letters*, vol. 131, pp. 203–209, 2020.
- [14] A. Kori and S. Chennamsetty, "Ensemble of convolutional neural networks for diabetic retinopathy grading," *Computer Methods and Programs in Biomedicine*, vol. 165, pp. 115–126, 2018.
- [15] R. R. Selvaraju and M. Cogswell, "Grad-cam: Visual explanations from deep networks via gradient-based localization," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 618–626, 2017.
- [16] B. Zhou and A. Khosla, "Learning deep features for discriminative localization," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2921–2929, 2016.
- [17] G. Litjens and T. Kooi, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [18] M. Raghu and C. Zhang, "Transfusion: Understanding transfer learning for medical imaging," *Advances in Neural Information Processing Systems*, 2019.
- [19] F. Grassmann and J. Mengelkamp, "A deep learning algorithm for detection of diabetic retinopathy," *Investigative Ophthalmology Visual Science*, vol. 59, no. 7, pp. 3180–3187, 2018.
- [20] W. H. Organization, "World report on vision," *WHO Publications*, 2023.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)