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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue V May 2026- Available at www.ijraset.com*

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

Adithya M Ganesh[1], Jeeva Mariya Tomy[2], Meenakshi KS[3], Prof. Aby Abahai T[4]
*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 screeningsolutions. This work proposes a Smart Ocular Health Ecosystem that integrates a hierarchical deep learning framework formulti-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 majorhealthconcern.ConditionssuchasDiabeticRetinopathy (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,requiresexpertknowledge,andmaylead to differences in diagnosis, especially when screening large numbers of patients.

Advancements in Convolutional Neural Networks (CNNs) nowallowautomaticanalysisofretinalimageswithoutmanu- ally 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 decisionsare 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 diagnosticapproachtoimprovebothspeedandaccuracy.First, a lightweight CNN performs initial screening by classifying retinalimagesasNormal,Cataract,orGlaucoma.Thisreduces 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 captureawiderangeofimagefeaturesandimproveseverity grading.Themodelalsoconsiderswherefeaturesappearin theretinaratherthantreatingallregionsequally,whichis importantbecausethe positionoflesionsinfluencesdiagnosis. Inadditiontoimprovingdetectionaccuracy,thesystemis designedtobeunderstandable forme dicaluse. UsingGrad- CAM,itgeneratesvisualheatmapsthathighlighttheretinal regions influencing the model's decision, helping doctors ver- ify results. A RAG-based chatbot is also included to explain thediagnosistopatientsinsimplelanguage,makingtheresults easiertounderstand.

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

## II. RELATED WORKS

Recent improvements in automatic Diabetic Retinopathy (DR)detectionmainlycomefrombetterdeeplearningmodels. 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 meth- ods 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 im- proved 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 vesselsor 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 ina 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 noiseinimages.Today,deeplearningmodelsdirectlylearn important features from retinal images. Attention-based mod-elshelpthesystemfocusontheopticdiscregion,which is important for detecting early Glaucoma when symptomsare 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 performbetterbecausetheyautomaticallylearnpatternsthatshowlens opacity.Lightweightmodelsandimageenhancementmethods 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 healthcarehasalsoimproved.Oldermedicalchatbotsfollowed 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. Re-searchshowsthatRAGimprovesaccuracyandreliabilitywhen answering medical questions or educating patients. Still, most currentsystemstreatdiseasedetectionandpatientconversation as two separate tasks rather than combining them into one system.

## III. METHODOLOGY

### A. *OverallWorkflowoftheProposedSystem*

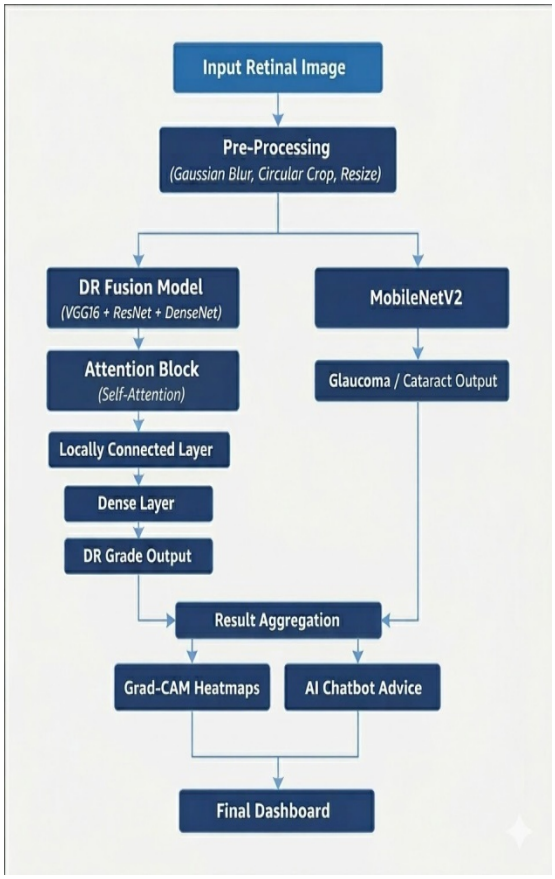



Fig.1.WorkflowofMultiOcularDiseaseDetectionSystem

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue V May 2026- Available at www.ijraset.com*

Fig. 1 illustrates the workflow of the proposed Ocular DiseaseDetectionSystem(ODDS).Thecompletesystem-level inferenceprocedureisformallysummarizedinAlgorithmA, 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$I \in \mathrm{R}^{H\times W\times 3}$denotetheinputRGBfundusimage.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 Gen- eralist branch for Glaucoma and Cataract detection, and $A(\cdot)$ definestheaggregationfunctionimplementedinAlgorithmA.

### *B. ImagePreprocessing*

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

Lettheinputretinalimageberepresentedas

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

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

*1) CircularCropping:*Toremovedarkborderssurrounding the retinal region, a circular mask is applied to the image. The mask $M(x, y)$is defined as

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

where$(x_c, y_c)$denotesthecenteroftheimageand$r$ 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

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

Thefilteredimageisobtainedthroughconvolution

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

**Algorithm 1** Attention-Guided Multi-Backbone Fusion for DR Classification **Require:** Retinal image $X \in \mathbb{R}^{224\times224\times3}$
**Ensure:** Predicted DR class $y^*$
Initialize pretrained models: VGG16, ResNet50, DenseNet121
Freeze backbone parameters
Extract feature maps:
$f_1 \leftarrow \text{VGG16}(X)$
$f_2 \leftarrow \text{ResNet50}(X)$
$f_3 \leftarrow \text{DenseNet121}(X)$
**for** each feature map $f_i \in \{f_1, f_2, f_3\}$ **do**
  $s_i \leftarrow \text{GlobalAveragePooling}(f_i)$
  $z_i \leftarrow \text{Dense(ReLU)}(s_i)$
  $a_i \leftarrow \text{Dense(Sigmoid)}(z_i)$
  $f_i^{att} \leftarrow f_i \times a_i$
**end for**
Project to common feature space:
$p_1 \leftarrow \text{Conv2D}_{1\times1,512}(f_1^{att})$ $p_2 \leftarrow \text{Conv2D}_{1\times1,512}(f_2^{att})$ $p_3 \leftarrow \text{Conv2D}_{1\times1,512}(f_3^{att})$
Resize $p_2, p_3$ to match $p_1$
Initialize fusion weights $w_1, w_2, w_3$
Fuse features:
$F \leftarrow w_1 p_1 + w_2 p_2 + w_3 p_3$
Classification:
$v \leftarrow \text{Flatten}(F)$
$h \leftarrow \text{Dense}(256, ReLU)(v)$
$h \leftarrow \text{Dropout}(0.5)(h)$
$y \leftarrow \text{Dense}(5, Softmax)(h)$
$y^* \leftarrow \text{argmax}(y)$
**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

$$I_n = \frac{I_\alpha}{255} \qquad (7)$$

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 classi- fication. It utilizes a fusion architecture that combines three pretrained convolutional neural networks: VGG16, ResNet50, and DenseNet121.

Let the extracted feature maps be represented as

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

To enhance relevant features, a channel attention mechanism is applied.
For a feature map $F$, the attention weights are calculated as

$$A = \sigma(W_2 \delta(W_1 \cdot GAP(F))) \qquad (9)$$

where $GAP$ represents global average pooling, $\delta$ denotes the ReLU activation, and $\sigma$ represents the sigmoid activation function. The refined feature map is obtained as

$$F' = F \otimes A \quad (10)$$

where $\otimes$ denotes channel-wise multiplication.

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

$$F_{fusion} = w_1 F_1 + w_2 F_2 + w_3 F_3 \quad (11)$$

where $w_1$, $w_2$, and $w_3$ represent trainable parameters deter- mining the contribution of each network.

The fused representation is then passed through fully con- nected 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:

$$P_k = \frac{e^{z_k}}{\sum_{i=1}^{K} e^{z_i}} \quad (12)$$

where $z_k$ represents the logit 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

$$FL(p_t) = -\alpha(1-p_t)^{\gamma}\log(p_t) \quad (13)$$

where $p_t$ represents the predicted probability of the true class, $\alpha$ is a balancing parameter, and $\gamma$ is the focusing parameter.

This loss function focuses training on hard-to-classify sam- ples 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: $M_{DR} \leftarrow \text{FusionModel}_{224\times224}$
3: Load Glaucoma/Cataract model:
The Grad-CAM heatmap is then generated as

4:$M_{GC} \leftarrow$ CNN$_{128\times128}$
5:Acquireinputimage$X$
6:Generatedual-resolutioninputs:
7:$X_{DR} \leftarrow$ Resize($X$,224×224)
8:$X_{GC} \leftarrow$ Resize($X$,128×128)
9:Normalize$X_{DR}$and$X_{GC}$
10:ComputeDRprobability:
11:$P_{DR} \leftarrow M_{DR}(X_{DR})$
12:ComputeGCprobability:
13:$P_{GC} \leftarrow M_{GC}(X_{GC})$
14:Determinepredictedclasses:
15:$C_{DR} \leftarrow$ argmax($P_{DR}$) 16:$C_{GC} \leftarrow$ argmax($P_{GC}$) 17:**if** $C_{DR} \neq$ No DR **then** 18:$C_{GC} \leftarrow$NotAssessed
19: FlagLane-2assuppressed
20:**endif**
21:Computeconfidencescores:
22:Conf$_{DR} \leftarrow$ max($P_{DR}$)23:Conf$_{GC} \leftarrow$ max($P_{GC}$)24:Computecombinedrisk:
25:Risk$_{combined} \leftarrow f(C_{DR}, C_{GC})$
26:Constructresponse:
27:$R \leftarrow \{C_{DR}, \text{Conf}_{DR}, C_{GC}, \text{Conf}_{GC}, \text{Risk}_{combined}\}$
28:**return**$R$

*E. DecisionStrategy*

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

Let

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

betheconfidencescoreoftheDRspecialistmodel,and

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

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

$$\text{Prediction} = \begin{cases} \text{Lane 1} & C_1 \geq C_2 \\ \text{Lane 2} & \text{otherwise} \end{cases} \quad (16)$$

*F. ExplainabilityUsingGrad-CAM*

To improve interpretability, Gradient-weighted Class Acti- vationMapping(Grad-CAM)isusedtoidentifyimageregions responsible for the model's prediction.

Theimportanceweightforfeaturemap$k$iscomputedas

$$\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (17)$$

TheGrad-CAMheatmapisthengeneratedas

$$L_M^{GradCA} = ReLU\left( \sum_k \alpha_k A^k \right) \quad (18)$$

Theheatmapisfinallyoverlaidontheoriginalretinalimage to highlight disease-related regions.

*G. Multi-AgentMedicalChatbotArchitecture*

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 improveinterpretabilityofmodeloutputs.Unlikeconventional single-response chatbots, the proposed system adopts a mod- ular, 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 compo- nents: (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 interac- tion begins with a user query, which may optionally include predictiondatageneratedbythediagnosticmodel.Thispredic- tiondatatypicallycontainstheidentifieddiseaseclass,severity level, and confidence score.

A state object is initialized to encapsulate the input query, prediction metadata, and intermediate variables required for downstreamprocessing.Thisstructuredrepresentationenables seamless data flow between components.

3) *Intent Routing Mechanism:*The intent router serves as thedecision-makingunitofthechatbot.Itanalyzestheincom- ing query using a lightweight rule-based strategy to determine the most appropriate processing pathway. The routing logic prioritizes interpretability and responsiveness:

- Queriesaccompaniedbypredictionoutputsaredirected totheexplanationmodule
- Queriesrequestingrecentorup-to-dateinformationare handled by the web retrieval module
- Generalmedicalqueriesareroutedtotheknowledge

To improve interpretability, Gradient-weighted Class Acti- vationMapping(Grad-CAM)isusedtoidentifyimageregions responsible for the model's prediction.

Theimportanceweightforfeaturemap$k$iscomputedas **retrievalmodule**

Thisstructuredroutingavoidsunnecessarycomputationand 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 pre- dictionoutputgeneratedbythedeeplearningmodel. It extracts key attributes such as disease type, severity stage, and confidence score, and generates a two-level explanation:

Asimplified,patient-friendlyinterpretation

Amoredetailed,clinicallyrelevantexplanation

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

KnowledgeRetrieval(RAG)Agent:Forgeneralmed- icalqueries,thesystememploysaretrieval-augmented

similarity search, and responses are generated based on theretrievedcontext.Thisapproachreduceshallucination and improves factual reliability.

Web Retrieval Agent: To address queries related to recentdevelopmentsoremergingtreatments,thismodule retrieves information from external sources such as med- ical APIs and research databases. The retrieved contentis then summarizedinto a conciseand coherent response.

5) *Response Aggregation and Safety Layer:* The outputs generated by the respective agents are passed to a final aggre- gation module, which selects the appropriate response based ontheroutingdecision.Astandardizedmedicaldisclaimer is appended to all outputs to ensure responsible usage of the system.

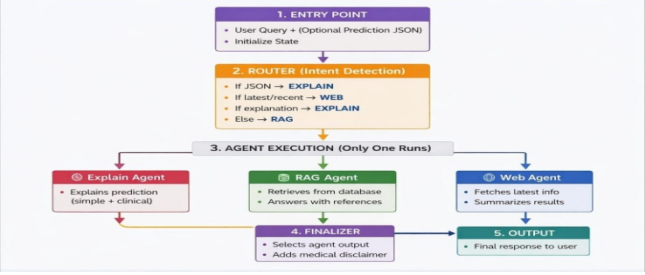



Fig.2.Multi-AgentMedicalChatbotArchitecturewithIntelligentRouting

*6) ArchitecturalAdvantages:*Theproposedmulti-agentde- sign offers several practical advantages:

- Modularity:Eachcomponentoperatesindependently, enabling easy extension and maintenance
- ContextAwareness:State-drivenprocessingpreserves relevant information across stages
- Reliability:Retrieval-basedresponsesimprovefactual correctness
- Scalability: Additional agents can be incorporated with- out altering the core pipeline

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

### *H. GlaucomaandCataractDetectionusingMobileNet*

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

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

TheuseofMobileNetensures:

- Fasterinferenceforreal-timescreening
- Reducedmemoryconsumption
- Suitabilityformobileandedgedeployment

## IV. RESULTS AND EVALUATION

This section presents the experimental evaluation of the proposed Ocular Disease Detection System (ODDS). The objectiveistoassess(i)multi-diseasescreeningcapability,(ii) DR severity grading performance, and (iii) model reliability and interpretability in clinical settings.

### *A. ExperimentalSetup*

Experiments were conducted using labeled retinal fundus imagedatasetscoveringNormal,Cataract,Glaucoma,andfive stages of Diabetic Retinopathy (DR). The dataset was splitinto training, validation, and testing sets to ensure unbiased evaluation.
Thesystemwasevaluatedintwostages:

- GeneralistBranch:Multi-diseaseclassification(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. EvaluationMetrics*

To quantitatively assess performance, the following metrics were used:

- Accuracy
- Precision
- Recall(Sensitivity)
- F1-score
- AreaUnderROCCurve(AUC)

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

### *C. Multi-Model PerformanceComparison*

Figure.3 presents the performance comparison between in- dividualbackbonemodels(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.

Theproposedfusionmodelachieved:

- 79.9%accuracyinDRgrading
- HigherF1-scorecomparedtosingle-backbonemodels
- ImprovedsensitivityforminorityDRclasses

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

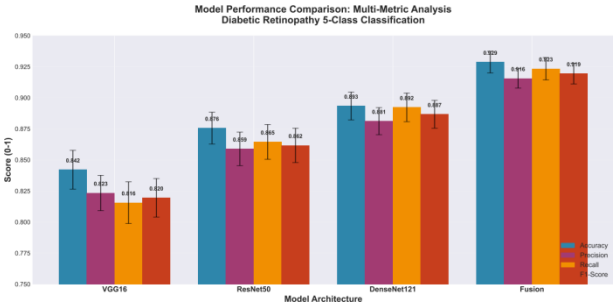



Fig.3.Multi-ModelPerformanceComparison

*D. ConfusionMatrixAnalysis*

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

Keyobservationsinclude:

- MinimalconfusionbetweenNormalandProliferativeDR
- Mostmisclassificationsoccurbetweenneighboringsever- ity stages
- Hightrue-positiverateforreferableDRcases Clinically,thisissignificantbecauseseverecasesarerarely misclassifiedasnormal,reducingtheriskofmisseddiagnoses.

*E. ROCandPrecision–RecallAnalysis*

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

Precision–Recallcurvesfurthershowstrongperformanceon minorityclasses,confirmingrobustnessunderclassimbalance conditions. This is particularly important for detecting severe DR cases, which are less frequent but clinically critical.

*F. TrainingStabilityandConvergence*

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

Thisdemonstratesthatthemodelgeneralizeswelltounseen data and that the chosen training duration (30 epochs) is appropriate.

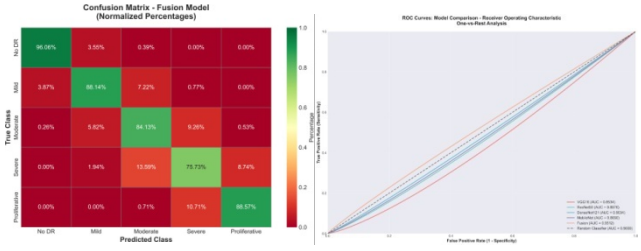



Fig. 4.Confusion Matrix Analysis of fusion model and ROC curves for allmodels

*G. ExplainabilityEvaluation*

Grad-CAMvisualizationsconfirmthatthemodelfocuses 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 sys- tematicallyremovingkeymodulesandobservingperformance changes.

TABLEI
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 posi- tively to overall performance. The fusion model significantly improves feature representation, while attention enhances lo- calization of disease regions. MobileNet improves efficiency without compromising accuracy.

### *I. Discussion*

The results validate the hypothesis that multi-backbone fu- sioncombinedwithattentionmechanismsimprovesdiagnostic performance.Thehierarchicaldesignalsoreducesunnecessary computation by routing only suspected DR cases to the spe- cialist 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
- Supportsscalabletele-ophthalmologydeployment

## V. CONCLUSION

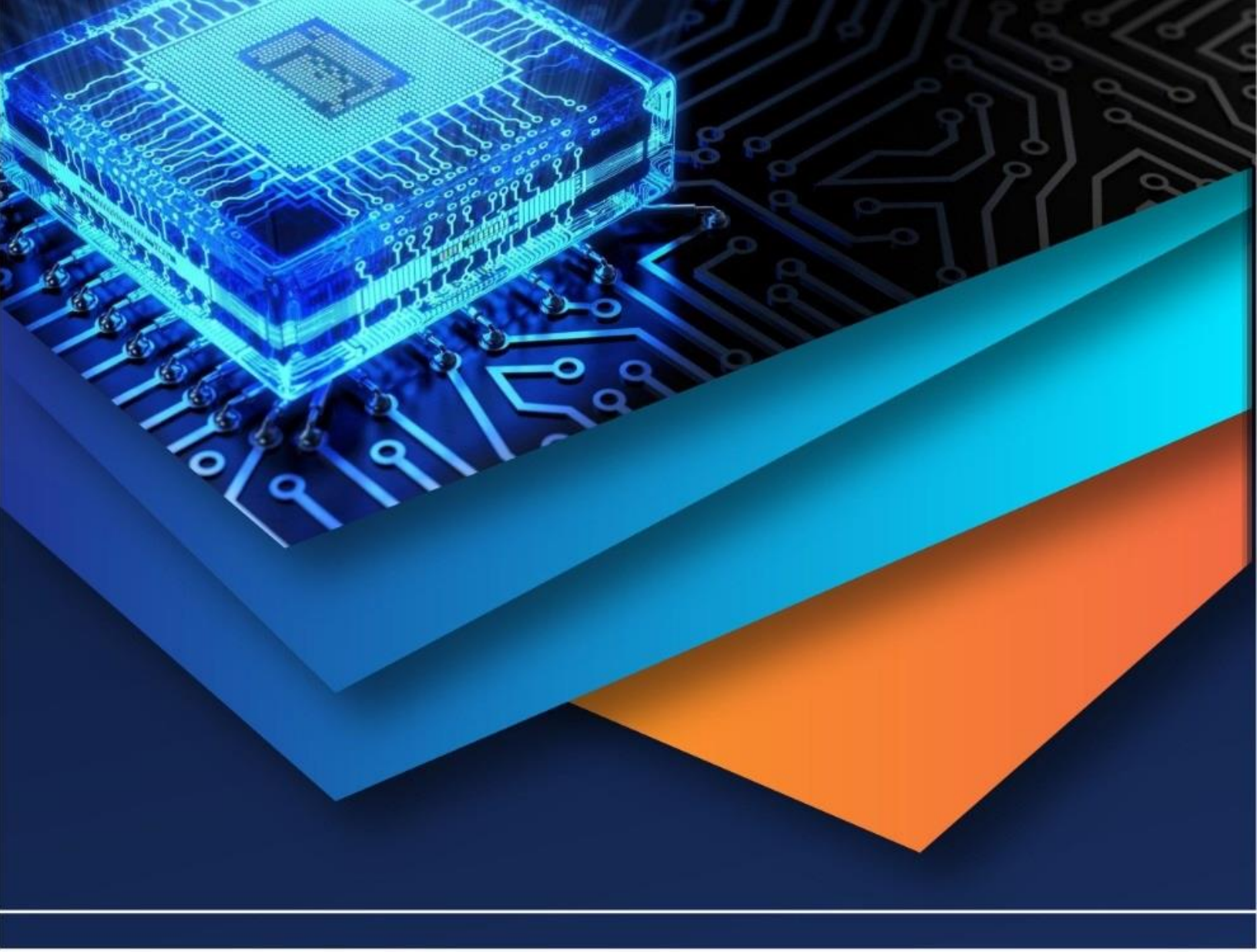
This study introduces a complete Smart Eye Health System that goes beyond detecting just one eye disease. The system uses a step-by-step deep learning approach to screen for Cataract and Glaucoma quickly, while still providing detailed and accurate grading for Diabetic Retinopathy. To make the system 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 modelonmobileandedgedevicesusingTensorFlowLite for offline screening, integration of longitudinal patient datato monitor disease progression, and expansion of the chatbot knowledge base to support multilingual interaction for im- proved rural accessibility.

## REFERENCES

[1] V.Gulshan,L.Peng,andM.Coram,"Developmentandvalidationofa deep learning algorithm for detection of diabetic retinopathy," JAMA,vol. 316, no. 22, pp. 2402–2410, 2016.
[2] D.S.W.TingandC.Y.Cheung,"Deeplearninginscreeningfordiabeticretinopathyandrelatedeyediseases,"JAMA,vol.318,no.22,pp.2211–2223, 2017.
[3] Z.LiandY.He,"Efficacyofdeeplearningfordetectingglaucomafromfundusimages," Ophthalmology, vol. 125, no. 8, pp. 1199–1206, 2018.
[4] X. Liu and H. Jiang, "Deep learning-based cataract detection usingfundus images," IEEE Access, vol. 7, pp. 108837–108847, 2019.
[5] C. Lam and C. Yu, "Multi-disease classification of retinal images usingdeepcnns,"IEEEJournalofBiomedicalandHealthInformatics,vol.22,no. 5, pp. 1515–1525, 2018.
[6] J. I. Orlando and E. Prokofyeva, "An ensemble deep learning approachfor 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 PatternRecognition, pp. 770–778, 2016.
[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks forlarge-scale image recognition," International Conference on LearningRepresentations, 2015.
[9] G. Huang and Z. Liu, "Densely connected convolutional networks,"Proceedings of the IEEE Conference on Computer Vision and PatternRecognition, pp. 4700–4708, 2017.
[10] X. Li and X. Hu, "Attention-guided deep neural networks for retinaldisease detection," IEEE Access, vol. 7, pp. 109947–109956, 2019.
[11] O. Oktay and J. Schlemper, "Attention u-net: Learning where to lookfor the pancreas," Medical Image Analysis, vol. 53, pp. 1–13, 2018.
[12] Z. Wang and J. Yang, "Zoom-in-net: Deep mining lesions for diabeticretinopathydetection,"IEEETransactionsonMedicalImaging,vol.36,no. 11, pp. 2337–2348, 2017.
[13] H. Pham and M. Luong, "Multi-cnn feature fusion for retinal diseaseclassification,"PatternRecognitionLetters,vol.131,pp.203–209,2020.
[14] A. Kori and S. Chennamsetty, "Ensemble of convolutional neuralnetworks for diabetic retinopathy grading," Computer Methods andPrograms in Biomedicine, vol. 165, pp. 115–126, 2018.
[15] R. R. Selvaraju and M. Cogswell, "Grad-cam: Visual explanations fromdeepnetworksviagradient-basedlocalization,"ProceedingsoftheIEEEInternational Conference on Computer Vision, pp. 618–626, 2017.
[16] B. Zhou and A. Khosla, "Learning deep features for discriminativelocalization," Proceedings of the IEEE Conference on Computer Visionand Pattern Recognition, pp. 2921–2929, 2016.
[17] G. Litjens and T. Kooi, "A survey on deep learning in medical imageanalysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.
[18] M. Raghu and C. Zhang, "Transfusion: Understanding transfer learningfor medical imaging," Advances in Neural Information ProcessingSystems, 2019.
[19] F. Grassmann and J. Mengelkamp, "A deep learning algorithm fordetection of diabetic retinopathy," Investigative OphthalmologyVisualScience, vol. 59, no. 7, pp. 3180–3187, 2018.
[20] W.H.Organization,"Worldreportonvision,"WHOPublications,2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY