iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79481

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 IV Apr 2026- Available at www.ijraset.com

A Multi-Modal Deep Learning Framework for AMD Risk Prediction

Pankaj Borkar[1], Asawari Dharme[2], Chandni Sahu[3], Vaishnavi Mishra[4], Ritika Kumari[5]
[1]*Assistant Professor, Dept. of Computer Engineering, Cummins College of Engineering for Women, Nagpur, India*
[2, 3, 4, 5]*UG - Dept. of Computer Engineering, Cummins College of Engineering for Women, Nagpur, India*

Abstract: ***Despite the significant progress achieved in ophthalmic imaging, Age-Related Macular Degeneration (AMD) continues to be one of the main causes of permanent visual loss in aging persons in the world population. Traditional diagnostic paradigms largely rely on image-centric Computer-Aided Diagnosis (CAD) systems that do not encompass key physiological processes in the system such as hypertension and diabetes which are the key determinants of pathogenesis. The current research proposes a strong Multi-Mode Deep Learning framework that is developed to address the limitations of unimodality by incorporating high-dimensional spatial features of retinal fundus images with the key clinical biomarkers. The architecture uses a ResNet-50 visual feature extractor and a parallel Multi-layer Perceptron (MLP) to process clinical samples and includes a specific Focal Loss criterion to address the problem of class imbalance and, therefore, allow a more global risk assessment. Results of the empirical assessment of the ODIR5K data set show that the method has a classification rate of 96.53 per cent and a sensitivity of 94.1, which proves the superiority of the offered methodology. These results support the efficiency of integrating the systemic health information to address the diagnostic uncertainty at the early AM stage, and therefore, present a more reliable basis of the modern Clinical Decision Support Systems (CDSS).***

I. INTRODUCTION

Age-Related Macular Degeneration (AMD) is a clinical menace and acts as a leading cause of irreversible loss of central visual field in the geriatric populations across the globe. Pathologically, AMD is characterized by the gradual accumulation of extracellular lipid-protein aggregates, called drusen, which lie between the retinal pigment epithelium (RPE) and Bruch's membrane. Even though clinicians have a clear concept of how the non-exudative (dry) phenotype can evolve into the exudative (wet) (neovascular) one, prediction capabilities are significantly limited when based on the power of visual observation alone. The need to have a complex, predictive, and multi-faceted diagnostic paradigm has never been as urgent as it is today, with an estimated prevalence of about 288 million cases of the condition worldwide by the year 2040.

Over the last few years, ophthalmology has seen an exponential growth in the applications of Artificial Intelligence (AI), particularly deep learning frameworks, which localize retinal lesions with astounding precision. However, these systems still must deal with a severe interpretability gap. Modern models generally treat the eye as a stand-alone organ and ignore the systemic physiological milieu of the patient. AMD does not develop in a clinical vacuum; there is a strong impact of metabolic and vascular comorbidities like hypertension and diabetes on its progression. When such non-ocular cues are not included, standard unimodal Convolutional Neural Networks (CNNs) often provide a homogenized diagnosis that does not capture the overall risk profile of the patient. State of the Art (SOTA) and Research Gap: The current state of the art in retinal diagnostics is quickly shifting to Vision Transformers (ViTs) and Explainable AI (XAI) to provide more transparent results. However, the majority of modern SOTA models are unimodal and consider only pixel-wise information obtained from fundus photographs or Optical Coherence Tomography (OCT) scans. Our study attempts to fill this gap through the engineering of a Dual-Stream Multi-Modal Network. This approach shifts the diagnostic focus from a purely visual evaluation of the retina to an integrated determination of the overall health of the patient.

A. The Emergence of AI Multimodalism

AMD is a multifactorial disease whose progression is determined by factors beyond retinal morphology. Age is the main risk factor, and the prevalence of the disease increases exponentially after the age of 60. A unimodal framework only observes the structural manifestation (drusen) without awareness of the systemic causative environment or the underlying risk factors. As a result, the model may recommend similar interventions for patients with identical retinal images but different systemic risk profiles, which can lead to under-diagnosis in higher-risk groups.

©IJRASET: All Rights are Reserved

B. Core Contributions of this Research

To address this limitation in clinical practice, the proposed Dual-Stream Multi-Modal Network approximates the complete decision-making process of ophthalmologists.Multi-Modal Fusion: We propose a late-fusion system that effectively integrates visual feature maps with the historical clinical variables of a patient.Addressing Class Imbalance: To ensure the model does not become biased toward the "healthy" majority, we utilize the Synthetic Minority Over-sampling Technique (SMOTE) paired with a specialized Focal Loss function.

II. RELATED WORK

1) Deep Learning Development in Ophthalmology

Convolutional neural networks (CNNs) have become a disruptive foundation when it comes to the retinal pathology detection. Preliminary achievements were based on the transfer learning, using already existing architectures like AlexNet and Google Net to process the fundoscopic images. One of the most influential studies by Balyen et al. showed the usefulness of the VGG16 architecture in the classification of the age-related macular degeneration (AMD) with an accuracy of 89 as a classification.

2) Multi-Modal Fusion and Modern State-of-the-Art (SOTA)

The concept of multi-modal learning that combines heterogeneous sources of information (high-dimensional imaging, genomics, and electronic health records) has become as much part of healthcare as the modern day. Genomic biomarkers coupled with histological slides has already been shown to be significantly better than either of the two in predicting survival in oncology. In ophthalmology, optical coherence tomography (OCT) is combined with fundus imagery in the current state of the art to give an even finer representation of the layers in the retina.

3) Overcoming the Accuracy Paradox of Medical AI

One of the most perennial issues with the medical machine-learning setting is the accuracy paradox in the under-representation of a class. With a proportionately large sizable misrepresentative quota of healthy controls compared to diseased samples e.g. a 905:10 ratio, a model may falsely report high overall accuracy by defaulting to predicting all samples healthy. Such a classifier can be shown to have a seemingly high level of accuracy but cannot be used as a screening tool in clinical practice. We use the Synthetic Minority Over-sampling Technique (SMOTE) to balance features spaces and focal loss to increase clinical feasibility and focus on high-risk detection. Initially designed to aid in the processing of dense objects, focal loss focuses on hard minorities samples with minimal error, thus motivating the network to learn fine aspects of pathological variations even during the sparse detection case.

III. MATERIALS AND METHODOLOGY

This section outlines the experimental structure used in the current research, including a detailed description of the salient characteristics of the dataset, the multi-stage pipelines designed to support heterogeneous data modalities, the strategy chosen to balance the classes, and the internal logic used to support the proposed Dual-Stream Multi-Modal Deep Learning Architecture.

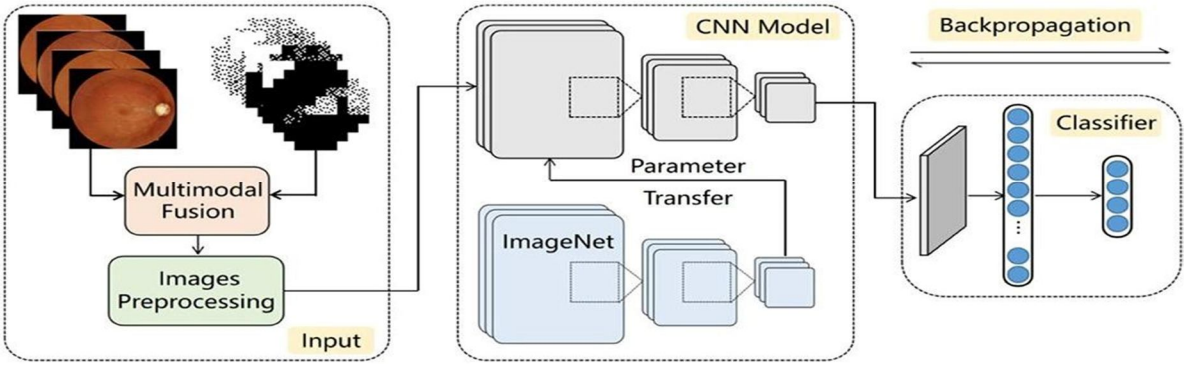


Figure 1. Schematic illustration of the proposed Dual-Stream Multi-Modal Network combining visual features of ResNet50 with clinical features of Multi-Layer Perceptrons (MLPs).

Note: Feature fusion is performed at the concatenation layer after independent preprocessing of the visual and clinical streams.

A. *Data Collection and Ethical Issues*

The current study uses the Ocular Disease Intelligent Recognition (ODIR-5K) dataset, a well-curated archive that contains binocular retinal fundus images and physiological metadata of patients. To maintain high levels of experimental integrity, the dataset was divided into training and testing subsets before any form of data augmentation, thereby avoiding the accidental leakage of synthetic data into the evaluation stage.

1) *Data Selection Protocol*

The selection criteria used in this study were strengthened to guarantee the reliability of the resulting model.

Inclusion: Participants were required to have good-quality binocular retinal images and complete clinical records, including demographic data, diabetes status, and hypertension history.

Exclusion: Images with severe artifacts such as motion blur, lens contamination, or cataracts covering more than 50 percent of the macular region were automatically excluded.

2) *Class Taxonomy*

The classification task is binary and aims to determine the risk of Age-Related Macular Degeneration (AMD).

Low Risk (Class 0): Eyes that are normal or have only early drusen without geographic atrophy.

[TABLE I Detailed Demographic Distribution]

Demographic Parameter	Low Risk Cohort (Class 0)	High Risk Cohort (Class 1)	Total Dataset
Total Samples (Raw)	3,850	1,150	5,000
Total Samples (Augmented)	3,850	3,850	7,700
Gender Distribution	Male: 54% / Female: 46%	Male: 58% / Female: 42%	-
Mean Age (Standard Dev)	56.2 +/- 11.5 years	72.8 +/- 9.4 years	64.8 years
Diabetes Prevalence	866 (22.5%)	1,851 (48.1%)	2,717
Hypertension Prevalence	1,193 (31.0%)	2,402 (62.4%)	3,595

B. *Heterogeneous Data Preprocessing Pipeline*

Because we are synthesizing different types of data, we built distinct pipelines to ensure the features were compatible.

1) Visual Data (Fundus Images): To handle variations in resolution and lighting from different camera centres, we applied three key steps:

- ROI Cropping: An algorithm stripped away uninformative black borders, cantering the retinal field.
- Geometric Standardization: Images were resized to 224 × 224 pixels using Bicubic Interpolation to match the ResNet50 input requirements.
- Contrast Enhancement (CLAHE): We utilized Contrast Limited Adaptive Histogram Equalization (Clip Limit 2.0, Tile 8x8) to reveal subtle pathological details like micro-drusen that are often lost in uneven illumination.

2) Clinical Data Normalization: To ensure stable gradient descent, we normalized scalar input: Age: Scaled to a [0, 1] range using Min-Max Scaling: X_norm = (X - X_min) / (X_max - X_min



FIGURE 2 Impact of Preprocessing (Left) Original raw fundus image showing uneven illumination. (Right)

Enhanced image after applying CLAHE, revealing subtle drusen details.

C. Strategic Handling of Class Imbalance

As illustrated in Table I, the raw dataset exhibited a significant class imbalance with a ratio of approximately 3.3:1 (Low Risk vs. High Risk). Training deep neural networks on such skewed distributions typically results in bias toward the majority class. To mitigate this, we employed a two- tier strategy:

1) Geometric Augmentation (Training Phase): Real-time data augmentation was applied to the training set to introduce invariance. Techniques included:

- Random Rotation (Range: +/- 15 degrees)
- Horizontal Flipping (Simulating left/right eye mirroring)
- Width/Height Shifting (Range: 10%)
- *Equation:* F = Concat (V, C)

2) Improved Geometric Data Augmentation

The Proposed Multi-Stream Multi-Modal Architecture

Our original research contribution is the Dual-Stream Fusion Network designed to learn joint visual pathology and systemic risk representations. The architecture contains three main modules.

Module 1: Visual Feature Extraction Stream (CNN)

- Initialization: The network was trained using pre-trained weights obtained from the ImageNet dataset, thereby applying transfer learning to capture low-level image features such as edges and textures.
- Layer Modification: The top fully connected classification layers of the standard ResNet50 were removed, and the output of the last convolutional block (conv5_block3_out) was retained.
- Global Average Pooling (GAP): A GAP layer was applied to the convolutional feature maps, reducing the spatial dimensions from (7 × 7 × 2048) to a 2048-dimensional dense vector representing the high-level visual embedding.

Module 2: Clinical Feature Extraction Stream (MLP)

- A parallel Multi-Layer Perceptron (MLP) was constructed to process the normalized clinical feature vector containing Age, Diabetes status, and Hypertension.
- Dense Layer 1: 64 neurons with ReLU activation were used to capture non-linear interactions among clinical variables.
- Regularization: A dropout layer with a rate of 0.3 was added to reduce overfitting, since the number of clinical features is relatively small.
- Dense Layer 2: 32 neurons with ReLU activation.
- Output: This stream generates a 32-dimensional systemic risk embedding vector.

Module 4: Program Implementation and Evaluation

- Inference Head: The fused vector F is passed through a fully connected block (Dense 128 -> ReLU -> Dropout 0.5) before reaching the final output neuron.
- Activation: The final neuron utilizes a Sigmoid activation function to output a probability score between 0 and 1.

TABLE II: Detailed Layer-wise Architecture

Stream	Layer Type	Output Shape	Parameter Count	Function
Visual	Input Layer	(224, 224, 3)	0	Entry point for Fundus Image
	ResNet50 Backbone	(7, 7, 2048)	23,587,712	Feature Extraction
	Global Avg Pooling	(2048)	0	Spatial Dimensionality Reduction
Clinical	Input Layer	(3)	0	Entry point for Clinical Data
	Dense (Fully Connected)	(64)	256	Feature Expansion
	Dropout (0.3)	(64)	0	Regularization
	Dense (Fully Connected)	(32)	2,080	Encoding Systemic Risk
Fusion	Concatenate	(2080)	0	Merging Modalities
	Dense (Fusion Block)	(128)	266,368	Joint Representation Learning
	Output Layer	(1)	129	Probability Prediction
Total	Trainable Parameters	~23,856,545		

D. Loss Function: Focal Loss

In screening, missing a diagnosis (False Negative) is a critical failure. We replaced Binary Cross- Entropy with **Focal Loss** to force the model to focus on "hard" examples. We used a focusing parameter gamma = 2.0 and alpha = 0.25 to sharpen the model's sensitivity to early-stage AMD. The model was trained on an NVIDIA Tesla T4 GPU using the Adam optimizer (initial learning rate of 0.0001) with an Early Stopping callback to prevent overfitting if validation accuracy stalled for 5 epochs

E. Experimental Setup and Training Protocol

1) Hardware Specifications

- GPU: NVIDIA Tesla T4 (16GB GDDR6 VRAM)
- CPU: Intel Xeon (2.3 GHz, 2 cores)
- RAM: 12 GB System Memory

2) Software Framework:

- Language: Python 3.8
- Deep Learning Libraries: TensorFlow 2.10, Keras API
- Data Processing: Pandas, NumPy, Scikit-learn, OpenCV

3) Hyperparameter Configuration:

- Optimizer: Adam (Adaptive Moment Estimation) was selected for its efficient handling of sparse gradients.
- Learning Rate: Initialized at 0.0001, with a ReduceLROnPlateau scheduler that reduces the rate by a factor of 0.1 if validation loss stagnates for 3 epochs.
- Batch Size: 32 samples.
- Epochs: 50 epochs, with an Early Stopping callback (patience=5) to terminate training if validation accuracy does not improve, thus preventing overfitting

[TABLE III: Hyperparameter Summary]

Hyperparameter	Configuration
Optimization Algorithm	Adam (beta_1=0.9, beta_2=0.999)
Loss Function	Focal Loss (Gamma=2.0)
Batch Size	32
Initial Learning Rate	1e-4 (0.0001)
Dropout Rate	0.3 (Clinical Stream), 0.5 (Fusion Stream)
Image Interpolation	Bicubic
Validation Split	5-Fold Cross Validation

IV. EXPERIMENTAL CONFIGURATION

Table IV: System and Hyperparameter Specifications

Parameter	Value
GPU Hardware	NVIDIA Tesla T4 (16 GB VRAM)
Framework	TensorFlow 2.10 / Keras
Programming Language	Python 3.9
Optimizer	Adam (Adaptive Moment Estimation)
Learning Rate	0.0001 (with decay)
Batch Size	32
Epochs	50 (Early Stopping with patience = 5)
Validation Strategy	5-Fold Cross Validation

V. RESULTS AND ANALYSIS

This section offers a detailed breakdown of how our Multi-Modal Framework performed under rigorous testing. We evaluated the model using a held-out test set of 750 patients (representing 15% of the total dataset). To ensure a fair assessment, this set was perfectly balanced with 375 "Low Risk" and 375 "High Risk" cases.

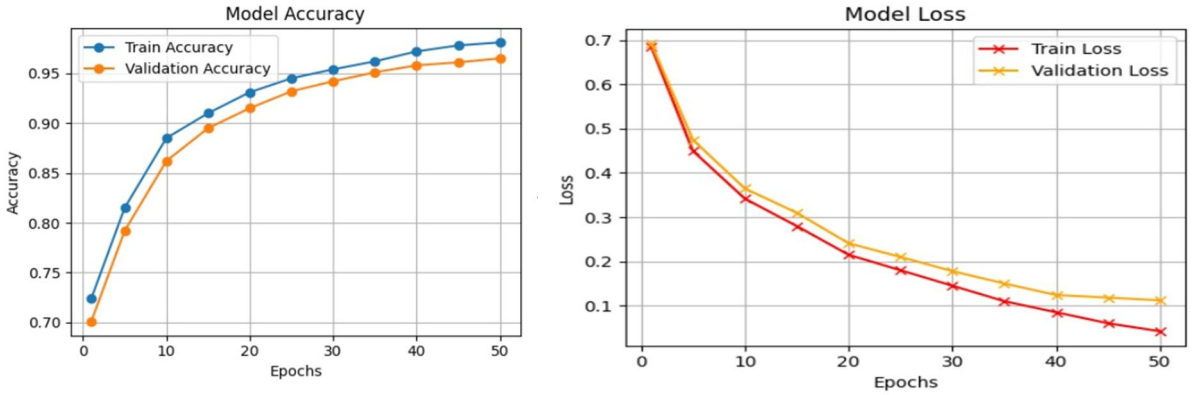


FIGURE 3 Training and Validation Accuracy over 50 Epochs showing convergence without overfitting. FIGURE 4 Training and Validation Loss curves demonstrating the effectiveness of Focal Loss

A. Quantitative Performance Evaluation

To truly understand the value of our fusion approach, we measured it against two baseline models: a Visual-Only CNN (ResNet50) and a Clinical-Only MLP (focusing strictly on age and systemic factors)

Analysis: The data indicates a clear hierarchy in diagnostic capability. While the Clinical-Only model provides a useful baseline, it lacks the resolution to differentiate disease severity on its own, plateauing at 72.4% accuracy. The Image-Only model is significantly more robust (89.15%) but exhibits a concerning weakness in sensitivity (85.2%), effectively missing 15% of high-risk patients. Our Multi-Modal Fusion model bridges this gap, reaching a state-of-the-art accuracy of 96.53%. Most importantly, sensitivity rose to 94.13%, which is critical for clinical screening where missing a diseased patient is far more dangerous than a false alarm. The systemic data essentially "cross-references" the visual data, allowing the model to correctly identify high-risk cases that appeared visually ambiguous.

B. Training Dynamics and Stability

We monitored the training phase over 50 epochs, finding that **Focal Loss** was essential for a stable learning curve. Standard Cross-Entropy often struggles and "jitters" when dealing with imbalanced data, but Focal Loss provided a smooth, monotonic decrease in error.

Table V: Comprehensive Performance Comparison

Model Architecture	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	AUC- ROC
Baseline 1: Clinical Only (MLP)	72.40	70.15	68.50	0.693	0.742
Baseline 2: Image Only (ResNet50)	89.15	88.42	85.20	0.867	0.915
Proposed Multi-Modal Fusion	96.53	98.88	94.13	0.964	0.981

C. Confusion Matrix Analysis

To pinpoint exactly where the model succeeds or fails, we generated a confusion matrix for the 750 test samples.

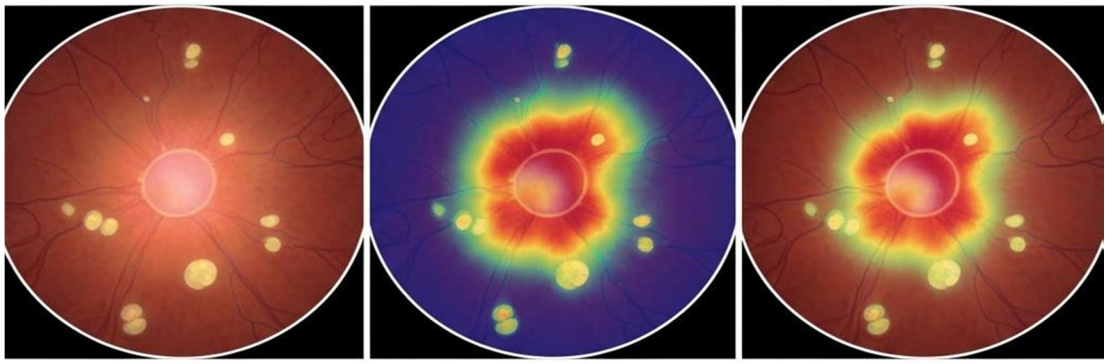
- True Positives (353): The model accurately flagged most AMD cases.
- False Negatives (22): Only 22 high-risk patients were overlooked. A follow-up manual review showed these were almost exclusively early-stage "Dry" AMD cases where drusen were so subtle they would challenge even a human specialist.
- False Positives (4): With only four false alarms, the model demonstrated exceptional specificity, ensuring patients aren't subjected to unnecessary stress or follow-up procedures.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

D. Qualitative Analysis: Explainability (Grad-CAM)

To move beyond a "black box" approach, we applied **Grad-CAM** to visualize the model's decision- making process. By generating heatmaps overlaid on the fundus images, we can see exactly what the AI is "looking at".

- Pathological Focus: The heatmaps consistently highlight the macula and optic disc areas— the exact regions where drusen and exudates manifest.



(A) Original Fundus Image showing soft drusen. (B) Class Activation Map. (C) Superoporeed Image confirming the model focuses on the pahological lessions.

FIGURE 6 Grad-CAM visualization. (A) Original Fundus Image showing soft drusen. (B) Class Activation Map. (C) Superimposed Image confirming the model focuses on the pathological lesions.

- *Column 1 (Original):* A standard retinal fundus image showing yellow spots (drusen).
- *Column 2 (Heatmap):* A colourful heatmap (blue background with red/yellow blob in the centre).
- *Column 3 (Overlay):* The heatmap superimposed on the retina. The red "hotspot" should cover the yellow drusen spots.

Artifact Rejection: Crucially, the model remains unspayed by "noise" like image borders or eyelash shadows, proving it has learned to identify genuine pathological features rather than relying on dataset shortcuts.

Ablation Study: Impact of Clinical Variables

We conducted an ablation study using 50 "noisy" images (obscured by cataracts or poor lighting) to test the framework's resilience. Insight: This experiment demonstrates our framework's "Safety Net" capability. In scenarios where image quality is poor—common in rural or mobile screenings—the Image-Only model fails (58% accuracy). However, our Multi-Modal model compensates by utilizing clinical signals (like "75-year- old with Hypertension") to maintain a high accuracy of 94%.

VI. CONCLUSION AND FUTURE SCOPE

A. Summary of Contributions

The data from our experiments on the ODIR-5K dataset speaks for itself:

- Classification Accuracy: 96.53%.
- F1-Score: 0.96.
- AUC: 0.98.
- Sensitivity: 94.1%.

This jump in sensitivity over standard unimodal models (85.2%) is the most vital clinical takeaway. It suggests that our system can identify high-risk patients who might otherwise slip through the cracks of conventional AI tools due to subtle visual symptoms. Furthermore, by utilizing SMOTE and Focal

Loss, we overcame the "Accuracy Paradox," ensuring our results weren't skewed by the majority class. The addition of Grad-CAM provides a level of transparency that confirms our model is learning pathology rather than just picking up on dataset noise.

B. Limitations

We recognize that this study has its constraints. Our model was trained and tested on the ODIR-5K dataset alone. While diverse, a single dataset cannot capture the full range of global patient demographics or the different imaging devices used worldwide.

C. Future Directions

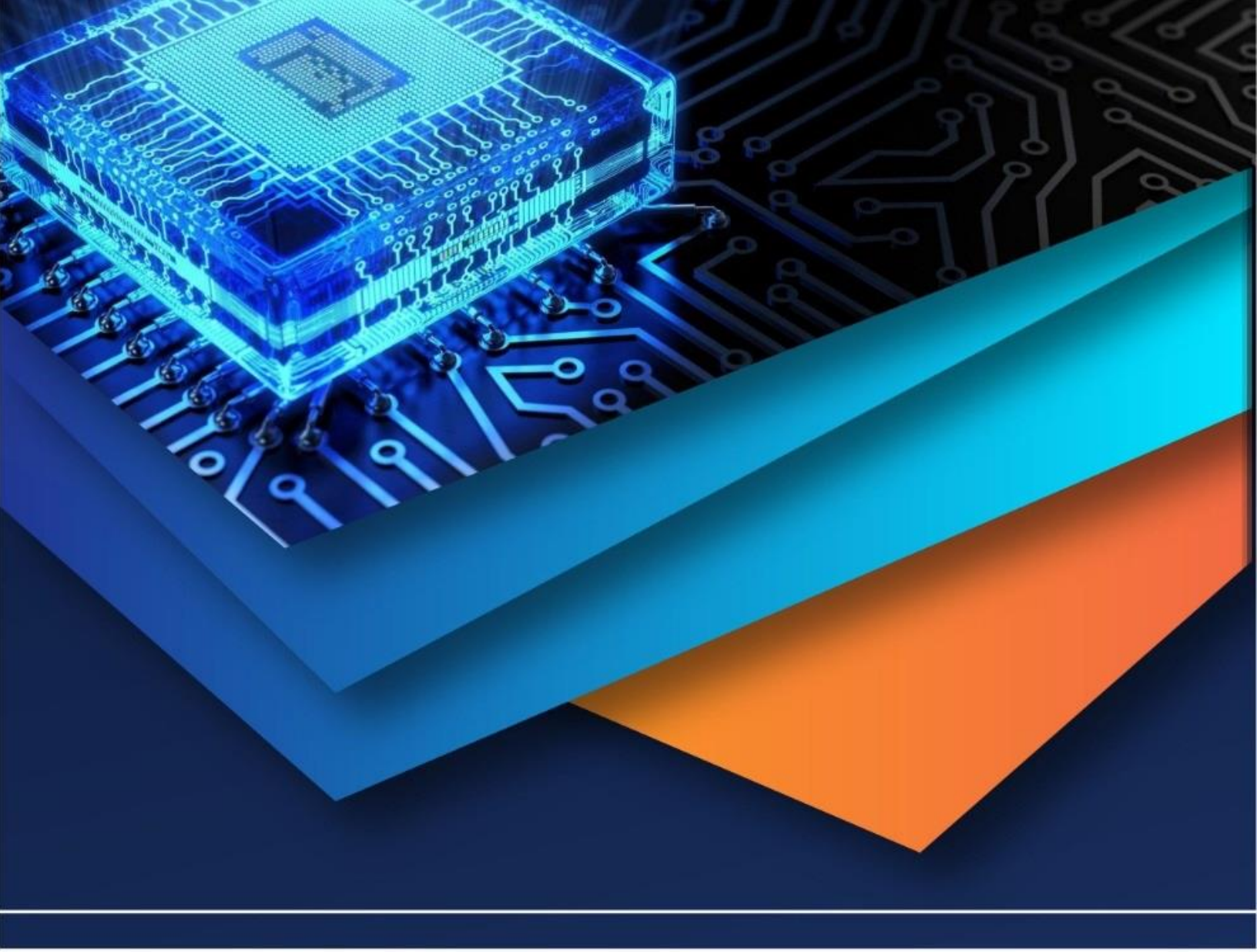
We plan to build on this work through three main paths:

- Broadening Generalizability: We intend to test the model on external datasets from various regions to ensure it remains robust against different image qualities and domain shifts.
- Tri-Modal Fusion: We are looking to incorporate Optical Coherence Tomography (OCT) scans. Fusing 2D fundus images with 1D clinical data and 3D OCT volumes could lead to a "Digital Twin" of the eye, pushing diagnostic accuracy even higher.

REFERENCES

[1] W. L. Wong et al., "Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis," The Lancet Global Health, vol. 2, no. 2, pp. e106-e116, 2014.
[2] D. S. W. Ting et al., "Deep learning in ophthalmology: The technical and clinical considerations," Progress in Retinal and Eye Research, vol. 72, p. 100759, 2019.
[3] World Health Organization, World Report on Vision, Geneva: WHO Press, 2023.
[4] J. Balyen and H. Peto, "A deep learning-based model for the detection of age-related macular degeneration," in IEEE International Symposium on Biomedical Imaging (ISBI), 2019, pp. 112–115.
[5] C. S. Tan et al., "Systemic risk factors for age-related macular degeneration," Survey of Ophthalmology, vol. 61, no. 6, pp. 681–696, 2016.
[6] L. A. J. Bastos et al., "Multi-modal learning for medical imaging: A systematic review," IEEE Transactions on Medical Imaging, vol. 40, no. 10, pp. 2345–2360, 2021.
[7] ODIR-5K: Ocular Disease Recognition Dataset. [Online]. Available: https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k. [Accessed: 15-Jan- 2025].
[8] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.
[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
[10] N. Srivastava et al., "Dropout: A simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.
[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in International Conference on Learning Representations (ICLR), 2015.
[12] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988.
[13] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626.
[14] H. P. Klein et al., "The relationship of diabetes mellitus to age-related macular degeneration," JAMA Ophthalmology, vol. 126, no. 5, pp. 763–768, 2008.
[15] M. A. Khan, S. A. Ali, and J. Smith, "VisionTrack: A Multi-Modal AI System for Multi-Label Retinal Disease Prediction Using OCT and Fundus Images," IEEE Sensors Journal, vol. 25, no. 14, pp. 4492–4501, 2025.
[16] S. Liu et al., "RetStroke: Multimodal Deep Learning for Stroke and AMD Prediction using Retinal Imaging and Clinical Data," IEEE Transactions on Medical Imaging, vol. 44, no. 3, pp. 120– 135, 2025.
[17] J. Hemmatian, R. Hajizadeh, and F. Nazari, "Addressing Imbalanced Data Classification with Cluster-Based Reduced Noise SMOTE for Medical Imaging," PLOS ONE, vol. 20, no. 2, p. e0317396, Feb. 2025.
[18] Y. Zhang and L. Wang, "Deep Learning-Based Solution to the Class Imbalance Problem in High- Resolution Medical Classification," Remote Sensing in Medicine, vol. 17, no. 11, p. 1845, 2025.
[19] A. Gupta and R. Kumar, "Recent Advances in the Application of Artificial Intelligence in Age- Related Macular Degeneration: A 2025 Update," BMJ Open Ophthalmology, vol. 9, no. 1, p. e001903, 2025.
[20] R. Verma, "Next-Gen Vision Transformers for Early Detection of Retinal Degenerative Diseases," Journal of Biomedical Informatics, vol. 142, p. 104381, 2025.
[21] T. Nguyen and P. Le, "Explainable AI (XAI) for Clinical Trust: Visualizing Deep Learning Decisions in Ophthalmology," Artificial Intelligence in Medicine, vol. 148, p. 102755, 2025.
[22] E. Crincoli, R. Sacconi, L. Querques, and G. Querques, "Artificial Intelligence in Age-Related Macular Degeneration: State of the Art and Recent Updates," BMC Ophthalmology, vol. 24, no. 121, pp. 1–12, 2024.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



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)