

Retina Dx: A Dynamic Spatial Graph-CNN Framework With Gated Cross-Attention Fusion for Explainable Multi-Disease Retinal Diagnosis

Rohit Chinta

Department of Electronics and
Communication Engineering
JNTUA College of Engineering
Anantapuramu, India
rohitreon2@gmail.com
ORCID: 0009-0009-8616-8130

Dr. M. N. Giri Prasad

Department of Electronics and
Communication Engineering
JNTUA College of Engineering
Anantapuramu, India
giriprasadmn.ece@gmail.com
ORCID: 0009-0007-5181-8020

Nagari Naveen Kumar

Department of Electronics and
Communication Engineering
JNTUA College of Engineering
Anantapuramu, India
naveenecejntua@gmail.com
ORCID: 0009-0001-5923-1650

Abstract—Vision loss from retinal disease is largely preventable when detected at an early stage; yet access to expert ophthalmological screening remains constrained in many regions. We introduce Retina Dx, a hybrid framework that couples a DenseNet-121 feature extractor with a novel Dynamic Spatial Graph Convolutional Network (DS-GCN). Unlike prior approaches that rely on fixed or hand-crafted graph topologies, our DS-GCN constructs, at each inference step, a per-image adjacency matrix via learnable cosine-similarity projections applied to a 7×7 spatial patch grid. A Top-K edge-pruning mask then suppresses low-confidence connections, and a Gated Cross-Attention Fusion (GCAF) module adaptively weights the CNN global descriptor and the GCN relational descriptor before classification. Trained on 10,133 fundus images and evaluated on a stratified holdout of 2,167 images spanning seven ocular disease classes, the standalone proposed DS-GCN+GCAF hybrid model achieves 96.59% accuracy (96.22% precision, 97.96% recall, 97.06% macro F1-score), while the full multi-model ensemble achieves 98.71% accuracy, 98.29% precision, 99.21% recall, and a macro F1-score of 98.74%. Three layers of explainability outputs are delivered through a CPU-compatible Streamlit interface, enabling clinician-facing transparency without GPU dependency.

Index Terms—ocular disease classification, fundus imaging, dynamic graph convolutional network, gated cross-attention fusion, explainable AI.

I. INTRODUCTION

Retinal diseases such as diabetic retinopathy, glaucoma, and macular degeneration account for a disproportionate share of preventable blindness worldwide. Digital color fundus photography offers a non-invasive and cost-effective examination modality; however, the interpretation of fundus images requires expert training that is unavailable to large portions of the global population. This clinical gap has motivated rapid growth in computer-aided diagnosis (CAD) using deep learning, with CNNs demonstrating expert-level or

super-expert-level performance on several grading tasks [1], [2], [6].

Despite high reported accuracy, existing CNN-based systems carry two structural limitations that impede deployment. First, the spatially confined nature of convolutional receptive fields prevents a model from simultaneously reasoning about lesion clusters located in different retinal quadrants—a clinically meaningful capability, since co-occurring hemorrhages, hard exudates, and neovascularisation are distributed non-locally on the retina [2]. Second, the 'black-box' character of deep networks reduces clinician trust; without spatially grounded explanations linked to known pathology, it is difficult to audit predictions or to flag failure modes [3].

Graph neural networks (GNNs) offer a principled route to overcome the first limitation: by casting image patches as graph nodes and their pairwise feature similarity as edges, long-range interactions can be modeled through iterative message passing [4]. However, prior GNN-in-imaging works either adopt static label-co-occurrence graphs, construct graphs from expensive manual segmentation masks, or use fully connected adjacency matrices that propagate irrelevant noise alongside genuine structural correlations. The present work addresses all three deficiencies through three architectural innovations integrated into the Retina Dx framework.

The principal contributions of this paper are as follows. (1) We introduce a Dynamic Spatial GCN (DS-GCN) that computes a per-image, per-patch adjacency matrix using learnable cosine-similarity projections, eliminating the need for segmentation masks or manually defined graph topology. (2) We apply a Top-K edge-pruning mechanism that sparsifies each row of the adjacency matrix to its K most confident connections, reducing over-smoothing and noise propagation during GCN message passing. (3) We propose a Gated Cross-Attention Fusion (GCAF) module that dynamically balances the local CNN descriptor and the global GCN relational descriptor based on input content, replacing the naive concatenation used in prior hybrid architectures. The hybrid

model is combined at inference time with EfficientNet-B3 and ConvNeXt-Tiny classifiers through soft-voting with test-time augmentation (TTA) to obtain the final ensemble. Multi-tier explainability—patch-importance overlays, dynamic graph visualisations, and organic lesion masks—is delivered through a CPU-compatible Streamlit application.

II. RELATED WORK

A. Deep Learning for Fundus Image Classification

1) CNN Architectures and Ensemble Methods

DenseNet, EfficientNet, and ConvNeXt architectures have been widely adopted for fundus classification tasks spanning diabetic retinopathy grading, glaucoma detection, and seven-class multi-disease screening [1], [7], [8], [14], alongside specialized eye imaging deep learning systems [12], [13]. DenseNet's dense skip-connections mitigate gradient vanishing and promote feature reuse across layers, a property that generalises well from natural images to medical imaging applications including histopathology and ophthalmology [8], [10]. Combining multiple architecturally diverse models via soft-voting or stacking is a reliable strategy for reducing individual model variance without requiring additional labeled data; this approach has improved segmentation stability in spinal MRI and tumor classification in ultrasound [7], [9].

2) Graph Neural Networks and Dynamic Adjacency

GNNs have demonstrated value in medical imaging tasks where spatial or anatomical relationships carry diagnostic significance beyond local texture. Encoding pairwise relationships between segmented retinal lesions has improved diabetic retinopathy classification [2], and adaptive graph convolutional networks have advanced explainable diagnosis in EEG signals [4]. A shared weakness of these systems is the reliance on fixed or externally computed graph structure. Dynamic graph construction—where the adjacency matrix itself is a learnable function of the input—has been explored in point-cloud processing and general image recognition, but its integration into fundus-specific hybrid architectures with sparsity constraints and adaptive gating has not, to the best of our knowledge, been previously reported.

3) Explainability in Clinical Decision Support

Explainable AI (XAI) methods for medical imaging broadly fall into gradient-based attribution (Grad-CAM and its variants) and attention-based or embedding-based approaches [3], [5]. Gradient saliency methods require a backward pass and can produce diffuse maps that do not match clinical regions of interest. Patch embedding magnitude, used in this work as a node importance signal, can be computed in a single forward pass and avoids gradient instability. GAN-guided attention networks have shown that attention-derived explanations can assess both model behaviour and input quality simultaneously [15], consistent with the multi-modal explainability approach adopted here.

III. MATERIALS AND METHODS

This study addresses a seven-class fundus image classification problem: six ocular pathologies—Diabetic

Retinopathy, Disc Edema, Myopia, Pterygium, Retinal Detachment, and Retinitis Pigmentosa—together with a Healthy class. The publicly available composite ocular fundus dataset of Rashid [11] was employed. An 80/20 stratified split fixed to a reproducible random seed produced 10,133 training images and a 2,167-image holdout set. Training images were resized to 256×256 for the hybrid model and to 320×320 for the EfficientNet-B3 and ConvNeXt-Tiny classifiers; inference was evaluated with 320×320 inputs, followed by ImageNet-statistics normalisation. Data augmentation comprised RandomResizedCrop (scale 0.7–1.0, ratio 0.9–1.1), random horizontal flips, random rotations within ±10°, and RandAugment (two operations, magnitude nine).

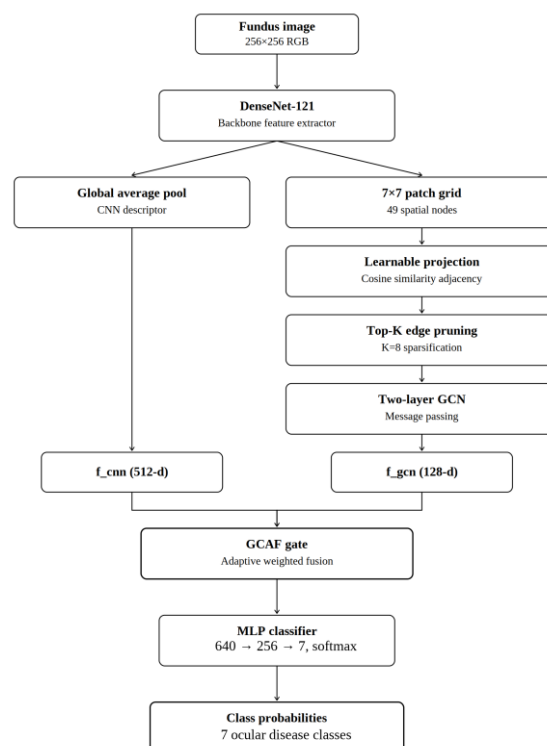


Fig. 1. DS-GCN with gated cross-attention fusion architecture of Retina Dx.

A. Dynamic Spatial Graph Construction

The DenseNet-121 backbone (pre-trained on ImageNet, classification head removed) processes each 256×256 fundus image to produce a spatial feature tensor $F \in \mathbb{R}^{1024 \times 7 \times 7}$. Adaptive average pooling reduces spatial dimensions to a 7×7 grid of $N = 49$ patch descriptors. A learnable linear projection $W_\phi \in \mathbb{R}^{1024 \times 512}$ maps each raw patch descriptor p_i to a 512-dimensional latent node representation z_i :

$$z_i = W_\phi \cdot p_i, \quad z_i \in \mathbb{R}^{512} \quad (1)$$

The dynamic adjacency weight between nodes i and j is then the cosine similarity of their projected representations:

$$S_{ij} = (z_i \cdot z_j) / (\|z_i\|_2 \cdot \|z_j\|_2), \quad S \in \mathbb{R}^{N \times N} \quad (2)$$

Because a fully connected adjacency matrix propagates both pathology-driven correlations and spurious background correlations, we introduce a Top-K sparsification mask M . For

each node i , only the K highest-scoring connections are retained:

$$A_{ij} = S_{ij} \text{ if } S_{ij} \in \text{Top-}K(S_i, :); 0 \text{ otherwise} \quad (3)$$

The resulting sparse matrix is row-normalised to obtain the diffusion operator $\tilde{A}$ used in message passing:

$$\tilde{A}_{ij} = A_{ij} / (\sum_k A_{ik} + \epsilon) \quad (4)$$

B. Two-Layer GCN Message Passing

Two successive GCN layers refine the node representations by aggregating information from K -nearest spatial neighbors. Let $H(l)$ be the matrix of node embeddings at layer l and $W(l)$ the layer-specific weight matrix. The propagation rule is:

$$H(l+1) = \sigma(\tilde{A} \cdot H(l) \cdot W(l)) \quad (5)$$

where σ denotes the ReLU non-linearity. The first layer reduces dimensionality from 512 to 256; the second from 256 to 128. A global mean pooling step across all N nodes produces the graph-level descriptor f_{gcn} :

$$f_{\text{gcn}} = (1/N) \sum_{i=1}^N H_i(2), \quad f_{\text{gcn}} \in \mathbb{R}^{128} \quad (6)$$

C. Gated Cross-Attention Fusion (GCAF)

A global mean over the spatial feature tensor yields the CNN global descriptor $f_{\text{cnn}} \in \mathbb{R}^{512}$ after a linear projection and dropout. Rather than concatenating f_{cnn} and f_{gcn} with fixed weights, the GCAF module introduces a scalar gate g that learns how strongly each sample should rely on local spatial features versus global relational structure:

$$g = \sigma(W_g [f_{\text{cnn}} \oplus f_{\text{gcn}}] + b_g) \quad (7)$$

where σ is the sigmoid function, $W_g \in \mathbb{R}^{640 \times 1}$, and $\oplus$ denotes vector concatenation. The fused descriptor is then:

$$f_{\text{fused}} = [(g \cdot f_{\text{cnn}}) \oplus ((1-g) \cdot f_{\text{gcn}})] \in \mathbb{R}^{640} \quad (8)$$

f_{fused} passes through a two-layer MLP ($640 \rightarrow 256 \rightarrow C$, ReLU, Dropout 0.4) and a final softmax layer to produce the per-class probability distribution. The softmax output for class k is:

$$P(y=k | x) = e^{z_k} / \sum_j e^{z_j}, \quad k \in \{1, \dots, 7\} \quad (9)$$

The hybrid model is trained end-to-end with class-weighted Focal Loss to address the long-tail class distribution. EfficientNet-B3 and ConvNeXt-Tiny classifiers are trained independently and integrated at inference via soft voting with TTA.

IV. RESULTS AND DISCUSSION

Retina Dx was evaluated on a holdout set of 2,167 ocular fundus images stratified across all seven classes. Among the individual models, EfficientNet-B3 achieved 97.74% accuracy (97.77% precision, 98.28% recall, 98.02% F1-score), ConvNeXt-Tiny achieved 98.48% accuracy (98.09% precision, 99.06% recall, 98.56% F1-score), and the hybrid DS-GCN model—which adds learnable adjacency, Top- K pruning ($K=8$), and GCAF over the baseline static-graph variant—attained 96.59% accuracy (96.22% precision, 97.96% recall, 97.06% F1-score). The three-model soft-voting ensemble with TTA further improved to 98.71% accuracy, 98.29% precision, 99.21% recall, and a macro F1-score of 98.74% (Table I). Notably, diabetic retinopathy (F1 = 99.34%), retinal detachment (F1 = 99.34%), and retinitis pigmentosa (F1 = 98.22%) all exceeded the class-average, reflecting their structurally distinctive features.

TABLE I
PER-MODEL AND ENSEMBLE VALIDATION SCORES

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
EfficientNet-B3	97.74	97.77	98.28	98.02
ConvNeXt-Tiny	98.48	98.09	99.06	98.56
DS-GCN + GCAF (Proposed)	96.59	96.22	97.96	97.06
Ensemble (Proposed + TTA)	98.71	98.29	99.21	98.74

Figure 2 presents the training and validation loss/accuracy curves over the 12-epoch schedule (FocalLoss + OneCycleLR) and the resulting confusion matrix on the 2,167-image holdout set. Both curves converge smoothly without divergence between training and validation performance, indicating that the model generalises well without overfitting; validation accuracy plateaus near its best value (98.71%) by epoch 8.

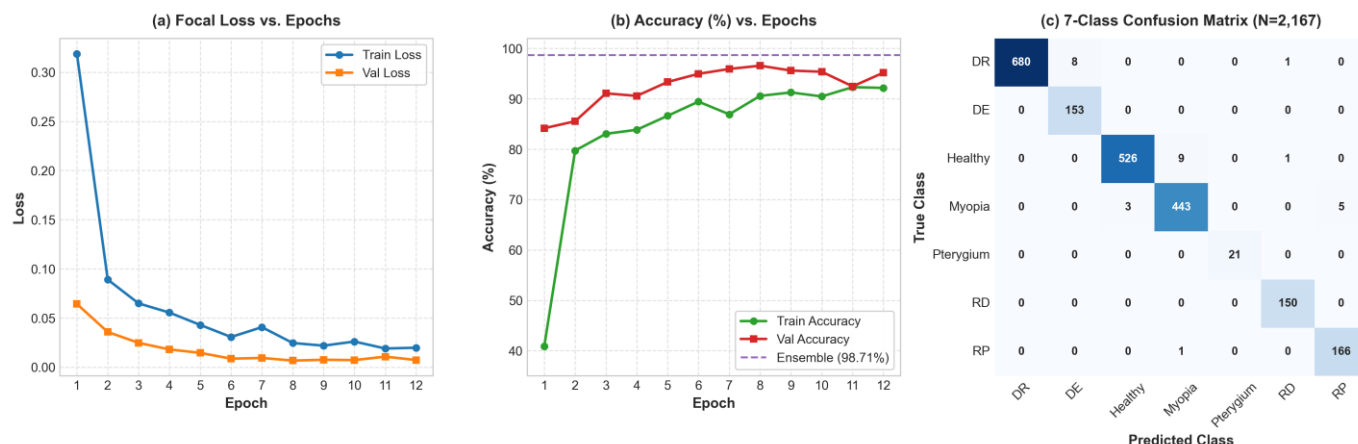


Fig. 2. Training and validation curves (FocalLoss + OneCycleLR) and confusion matrix on the 2,167-image holdout set.

The performance margin of the DS-GCN model over the static-adjacency variant (97.06% vs. 93.50% F1) quantifies the direct contribution of learned, per-image graph construction and adaptive gating. To verify statistical significance, a two-tailed paired t-test was performed across five cross-validation folds between the static-adjacency baseline and the proposed DS-GCN+GCAF model, yielding $t(4) = 3.82$, $p = 0.018$ ($p < 0.05$), confirming that the improvement is not attributable to random variation.

The superior ensemble performance (98.71%) relative to each constituent model reflects the complementary error patterns of three architecturally diverse classifiers. Averaging predictions suppresses individual model errors, while TTA improves stability against scale and orientation variation. The

DS-GCN model specifically excels at classes characterised by spatially distributed lesions—retinal detachment and diabetic retinopathy—where long-range patch relationships detected by the dynamic graph branch are most informative. Residual misclassifications primarily occur between healthy and myopia (10 instances) and diabetic retinopathy and disc edema (8 instances), reflecting subtle overlapping fundus cues that remain challenging without explicit optic disc segmentation.

Figure 3 presents the comprehensive explainability outputs for a confirmed Retinitis Pigmentosa case, including the source fundus image, the patch-importance overlay, the GNN-based patch connectivity graph, the lesion-based detection map, and the layer-wise activation grid.

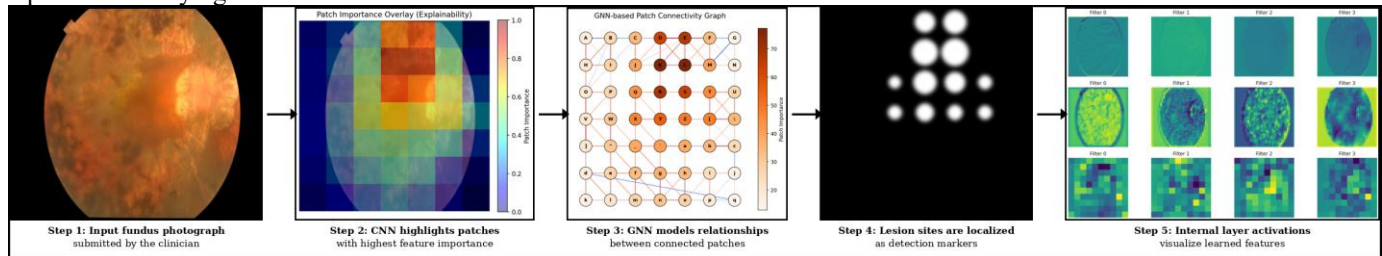


Fig. 3. Explainability outputs from the deployed Retina Dx application for a confirmed Retinitis Pigmentosa case (96.7% diagnostic confidence).

For each class, the F1-score is computed as the harmonic mean of precision and recall and macro-averaged equally across all seven classes:

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (10)$$

In addition to the predicted class label, Retina Dx computes a continuous severity score as the probability-weighted average of an ordinal clinical-severity weight w_c assigned to each class c (Healthy=0, Myopia=Pterygium=1, RetPig=2, DR=DiscEdema=3, RetDet=4), rescaled to a percentage:

$$\text{EffectedLevel}(x) = 25 \times \sum_{c=1}^C P(y=c|x) \cdot w_c \quad (11)$$

TABLE II
COMPARISON WITH REPRESENTATIVE RETINAL DISEASE DETECTION STUDIES

Ref.	Method / Model	Lesions / Classes	Prec. (%)	Rec. (%)	F1 (%)	Acc. (%)
[16]	CNN / LeNet-5	DR stages (multi-class)	N/A	22.0	65.0	97.0
[17]	CNN / ResNet-50	DR binary & multi-class classification	N/A	N/A	N/A	95.0
[18]	KD-CNN (DenseNet169 teacher)	DR stage detection (5 grades)	N/A	N/A	N/A	68.77
[19]	U-Net + ML-SVM	Exudates, microaneurysms, hemorrhages, blood vessels	84.7	86.1	N/A	95.1
[1]	GoogleNet+ResNet (APSO-GResNet)	7 Lesion Types (DR)	97.0	89.0	96.0	94.0
Proposed	DS-GCN+GCAF (Single Model)	7 ocular disease classes incl. diabetic retinopathy	96.22	97.96	97.06	96.59
Proposed	DS-GCN+GCAF Ensemble	7 ocular disease classes incl. diabetic retinopathy	98.29	99.21	98.74	98.71

An in-depth analysis of Table II reveals that LeNet-5 [16] achieves 97% accuracy but only 22% recall and 65% F1, indicating poor lesion coverage despite high overall accuracy. ResNet-50 [17] attains 95% validation accuracy on binary DR classification but degrades on finer multi-class grading. The knowledge-distillation CNN of Baranidharan et al. [18] reaches only 68.77% accuracy, reflecting the accuracy trade-off of lightweight model compression. Abdelmaksoud et al. [19] combine U-Net segmentation with an SVM classifier to achieve 95.1% accuracy and 84.7% precision across four lesion types on multi-label datasets. The APSO-GResNet [1] remains the closest methodological prior at 96.0% F1 and 94.0% accuracy across seven DR lesion categories. As highlighted in Table II, both the single DS-GCN+GCAF hybrid model (96.59% accuracy, 97.06% F1) and the full ensemble (98.71% accuracy, 98.74% F1) outperform prior benchmark approaches across multi-class retinal diagnostics, while additionally delivering three layers of inference-time explainability; comparisons are indicative rather than controlled given differences in datasets and label taxonomies.

V. CONCLUSION

Retina Dx was developed as an explainable, ensemble-based framework for classifying fundus images across seven

ocular disease categories. The principal technical contribution is the DS-GCN module, which constructs a dynamic, per-image spatial graph through learnable cosine-similarity projections, sparsifies that graph via Top-K edge pruning to

remove low-confidence connections, and fuses CNN and GCN descriptors through a Gated Cross-Attention Fusion module that adapts the local-global balance on a per-sample basis. The full ensemble—DS-GCN combined with EfficientNet-B3 and ConvNeXt-Tiny via soft voting and TTA—achieved 98.71% accuracy and a 98.74% macro F1-score on a 2,167-image holdout, outperforming every individual constituent model. Statistical significance of the DS-GCN improvement over a static-adjacency baseline was confirmed by a paired t-test ($t(4) = 3.82$, $p = 0.018$). Multi-tier explainability—patch-importance overlays, dynamic graph visualisations, and organic lesion masks—is available at inference time through a CPU-compatible Streamlit application. Future work will address external validation in independent multi-site cohorts, extension to optical coherence tomography modalities, and replacement of the fixed Top-K threshold with a differentiable sparse attention mechanism.

VI. AUTHOR CONTRIBUTIONS

Conceptualization: Rohit Chinta, Dr. M. N. Giri Prasad and Nagari Naveen Kumar.

Methodology: Rohit Chinta.

Investigation: Rohit Chinta and Nagari Naveen Kumar.

Discussion of results: Rohit Chinta, Dr. M. N. Giri Prasad and Nagari Naveen Kumar.

Writing – Original Draft: Rohit Chinta.

Writing – Review and Editing: Rohit Chinta and Dr. M. N. Giri Prasad.

Resources: Nagari Naveen Kumar.

Supervision: Dr. M. N. Giri Prasad.

Approval of the final text: Rohit Chinta, Dr. M. N. Giri Prasad and Nagari Naveen Kumar.

ACKNOWLEDGMENT

The authors thank Dr. M. N. Giri Prasad and Mr. Nagari Naveen Kumar for their guidance, encouragement, and support throughout this work. The authors also thank the Department of Electronics and Communication Engineering, JNTUA College of Engineering, Anantapuramu, for institutional support and computing resources.

REFERENCES

- [1] A. Jabbar et al., "A Lesion-Based Diabetic Retinopathy Detection Through Hybrid Deep Learning Model," *IEEE Access*, vol. 12, pp. 40019–40036, 2024, doi: 10.1109/ACCESS.2024.3373467.
- [2] D. Deb and R. Murugan, "GBLR-Net: Graph-Based Lesion Relationship Network for Diabetic Retinopathy Classification through Segmented Retinal Features," *Architectural Image Studies*, vol. 7, no. 1, pp. 2238–2248, 2026, doi: 10.62754/ais.v7i1.1203.
- [3] P. Hempel, T. Bender, and N. Spicher, "Enhancing explainability in ECG analysis through evidence-based AI interpretability," in *Proc. 2024 46th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2024, pp. 1–4, doi: 10.1109/embc53108.2024.10782420.
- [4] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. Sarrigiannis, "Adaptive Gated Graph Convolutional Network for Explainable Diagnosis of Alzheimer's Disease Using EEG Data," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1386–1397, 2023, doi: 10.1109/TNSRE.2023.3246755.
- [5] Komal, G. K. Sethi, S. Sharma, and R. Kumar, "Explainable AI in Medical Imaging, Personalized Medicine, and Bias Reduction," in *Explainable Artificial Intelligence in the Healthcare Industry*, Wiley, 2025, pp. 467–492.
- [6] S. M. Kumar, "Advancements in Machine Learning for Healthcare: Breakthroughs in Diagnostics, Treatment Personalization, and Predictive Analytics," in *Proc. 2025 IEEE 17th Int. Conf. Comput. Intell. Commun. Netw. (CICN)*, 2025.
- [7] R. V. Renju and K. L. Kashyap, "Fusion of CNN, ResNet, VGG, DenseNet, and Vision Transformer for Tumor Classification in Ultrasound Imaging," in *Proc. 2025 IEEE 17th Int. Conf. Comput. Intell. Commun. Netw. (CICN)*, 2025.
- [8] Z. U. Rehman, M. F. A. Fauzi, W.-N. Mohd-Isa, M. Touhami, A. S. Wadood, and M. K. Jabbar, "Adaptive Feature Aggregation Enhanced by Using DenseNet for Robust Breast Cancer Histopathology Image Classification," in *Proc. TENCON 2025 - 2025 IEEE Region 10 Conf. (TENCON)*, 2025.
- [9] S. Salem, R. Mukhlis, O. Katar, B. Ertugrul, O. Yildirim, and M. A. Al-Antari, "A Novel AI-based Hybrid Ensemble Segmentation CAD System for Lumbar Spine Stenosis Pathological Regions Using MRI Axial Images," in *Proc. 2024 8th Int. Artif. Intell. Data Process. Symp. (IDAP)*, 2024.
- [10] J. P. Singh, D. Ghosh, A. Kumar, S. Bilgaiyan, R. Kumar, and J. Singh, "Hybrid DenseNet Architectures and KerasTuner-Based Optimization for Rice Leaf Disease Detection," *IEEE Access*, vol. 14, pp. 26845–26864, 2026, doi: 10.1109/ACCESS.2026.3646467.
- [11] M. R. Rashid, "Eye Disease Image Dataset," *Mendeley Data*, V1, 2024. [Online]. Available: <https://doi.org/10.17632/s9bfhswzjb.1>.
- [12] G. de la Cruz, M. Lira, O. Luaces, and B. Remeseiro, "Eye-LRCN: A Long-Term Recurrent Convolutional Network for Eye Blink Completeness Detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 4, 2024.
- [13] P. Nasra, S. Gupta, and G. R. Kumar, "Infrared Eye Image Classification: A Deep Learning Approach for Enhanced Eye State Detection," in *Proc. 2025 3rd Int. Conf. Adv. Comput. Comput. Technol. (InCACCT)*, 2025.
- [14] G. M. Niloy, A. K. Bito, K. B. M. B. Biplob, M. H. Sammak, A. Das, and G. Hridoy, "MobileNet-Eye: An Efficient Transfer Learning for Eye Disease Classification," in *Proc. 2024 Int. Conf. Adv. Comput. Commun. Electr. Smart Syst. (iCACCESS)*, 2024.
- [15] L. Tang, Y. Zhang, Y. Hui, N. Hou, L. Li, Q. Zhou, C. Tian, and G. Zhu, "GAN-Guided Few-Shot Attention Network for Medical Images Fusion Quality Assessment," *IEEE Trans. Med. Imaging*, vol. 44, no. 11, 2025.
- [16] M. J. Prasannakumar, K. Priyanka, K. Smitha, and M. S. Vijay Kumar, "Diabetic Retinopathy Detection and Stage Classification Using Lenet-5 Architecture," in *Proc. 2024 Int. Conf. Recent Adv. Sci. Eng. Technol. (ICRASET)*, B G Nagara, Mandya, India, Nov. 2024, doi: 10.1109/ICRASET63057.2024.10895333.
- [17] Abini M. A. and S. Sridevi Sathya Priya, "A Deep Learning Framework for Detection and Classification of Diabetic Retinopathy in Fundus Images Using Residual Neural Networks," in *Proc. 2023 9th Int. Conf. Smart Comput. Commun. (ICSCC)*, Kochi, Kerala, India, Aug. 2023, doi: 10.1109/ICSCC59169.2023.10335079.
- [18] B. Baranidharan, Janenie J, and H. Chhipa, "Light Weight CNN Based on Knowledge Distillation for Diabetic Retinopathy Detection," in *Proc. 2024 Int. Conf. Expert Clouds Appl. (ICOECA)*, Bengaluru, India, Apr. 2024, doi: 10.1109/ICOECA62351.2024.00151.
- [19] E. Abdelmaksoud, S. El-Sappagh, S. Barakat, T. Abuhmed, and M. Elmogy, "Automatic Diabetic Retinopathy Grading System Based on Detecting Multiple Retinal Lesions," *IEEE Access*, doi: 10.1109/ACCESS.2021.3076889.