Journal for Advanced Research and Innovation Studies (JARIS)
An International Peer-Reviewed Multidisciplinary Research Journal
e-ISSN: XXXX-XXXX
Vol. I | Issue I |
June, 2026

# AN EFFICIENT DEEP LEARNING FRAMEWORK FOR RETINAL DISEASE DIAGNOSIS USING MULTI-MODAL AND ROBUST LEARNING TECHNIQUES

**Divya B** [1]
**Dr. Mohankumar T P** [2]

## Abstract

*Retinal disorders, including glaucoma, diabetic retinopathy, and age-related macular degeneration, are among the primary causes of visual impairment and blindness worldwide. Early and accurate diagnosis of these diseases is crucial for preventing irreversible vision loss and improving patient outcomes. Recent advances in Artificial Intelligence (AI), particularly Deep Learning (DL), have significantly enhanced the automated analysis of retinal images, enabling efficient and reliable disease detection. This survey paper presents a comprehensive review of recent developments in deep learning-based retinal disease diagnosis using retinal imaging modalities such as Fundus Imaging and Optical Coherence Tomography (OCT). The study systematically analyses sixteen recent research articles, focusing on image preprocessing techniques, feature extraction methods, Convolutional Neural Network (CNN) architectures, multi-modal learning approaches, and classification strategies employed for retinal disease detection. Furthermore, the review evaluates commonly used datasets, performance metrics, and the strengths and limitations of existing approaches. The analysis reveals several challenges, including class imbalance, noisy annotations, limited dataset diversity, and poor model generalization across different clinical environments. To address these issues, a conceptual framework is proposed that integrates robust preprocessing, advanced CNN-based architectures, and reliable training mechanisms to improve diagnostic accuracy, computational efficiency, and model interpretability. The findings of this survey emphasize the growing potential of AI-assisted retinal disease diagnosis and highlight future research directions aimed at developing clinically applicable, accurate, and trustworthy computer-aided diagnostic systems for ophthalmic healthcare.*

*Keywords: Retinal Disease, Deep Learning, Convolutional Neural Network, Fundus Imaging, Optical Coherence Tomography, Medical Image Analysis, Machine Learning*

## Introduction

Retinal diseases, including glaucoma, diabetic retinopathy (DR), and age-related macular degeneration (AMD), are among the major causes of visual impairment and blindness worldwide. Early and accurate diagnosis of these disorders is essential for preventing irreversible vision loss and improving patient outcomes. Recent advancements in Artificial Intelligence (AI), particularly Deep Learning (DL), have transformed retinal image analysis by enabling automated and efficient disease diagnosis using Fundus and Optical Coherence Tomography (OCT) images.

Several studies have demonstrated the effectiveness of Convolutional Neural Networks (CNNs) for retinal disease classification. Lightweight architectures such as MobileNet provide high accuracy with lower computational complexity, making them

[1] Assistant Professor, Sree Siddaganga College of Arts, Science and Commerce for Women and Research Scholar, Sri Siddhartha Academy of Higher Education (Deemed to be University), Tumkur
[2] Associate Professor, Sri Siddhartha Institute of Technology and Research Supervisor, Sri Siddhartha Academy of Higher Education (Deemed to be University), Tumkur

e-ISSN: XXXX-XXXX
Vol. I | Issue I |
June, 2026

suitable for real-time clinical applications. In addition, multi-modal learning frameworks combining Fundus and OCT images have improved diagnostic performance by utilizing complementary information from different imaging modalities. Advanced models such as RLF-Net, DeepRetinaNet, FANet, and transformer-based approaches have further enhanced disease detection, lesion localization, and model interpretability. Segmentation techniques, domain adaptation, and synthetic data generation methods have also contributed to improving the robustness of retinal diagnostic systems.Despite these advancements, challenges such as limited annotated datasets, class imbalance, noisy labels, high computational requirements, and lack of interpretability still remain.

This survey comprehensively reviews recent deep learning-based retinal disease diagnosis techniques, including preprocessing methods, CNN architectures, multi-modal frameworks, segmentation approaches, and hybrid learning strategies. The survey identifies the limitations of existing studies and highlights future research directions aimed at developing robust, interpretable, computationally efficient, and clinically applicable retinal disease diagnostic systems capable of overcoming current challenges and supporting reliable ophthalmic healthcare.

**Comparative Study**

This survey provides a comprehensive analysis of recent deep learning-based approaches for retinal disease diagnosis. It examines various methodologies, including image preprocessing, CNN architectures, multi-modal learning, and segmentation techniques, along with their strengths and limitations. The study identifies key challenges such as limited dataset diversity, class imbalance, poor generalization, and lack of interpretability, while highlighting future research directions for developing robust, efficient, and clinically applicable retinal diagnostic systems.

**Comparative Analysis**

| **Aspect** | **Existing Approaches** | **Strengths** | **Research Gaps** | **Future Scope** |
| --- | --- | --- | --- | --- |
| Deep Learning Models | CNN, MobileNet, DeepRetinaNet, MDBL-Net | High accuracy in retinal disease classification | High computational cost and dependence on large datasets | Develop lightweight and efficient models |
| Multi-Modal Learning | MultiEYE, OCT-CoDA | Combines Fundus and OCT for improved diagnosis | Limited ability to detect unseen diseases | Extend to multiple modalities and unknown diseases |

| | | | | |
|---|---|---|---|---|
| Image Registration | Hybrid feature-based and area-based methods | Accurate retinal image alignment | Sensitive to image quality and complex deformations | Integrate deep learning-based registration |
| Lesion Detection | RLF-Net, CLAT | Better lesion localization and interpretability | Requires lesion annotations and high-quality images | Develop annotation-efficient models |
| Few-Shot and Meta Learning | Mirror Meta-Learning | Better generalization with limited data | Complex training procedures | Improve simplicity and real-world applicability |
| Segmentation Methods | W-Capsule Network, URVSM | Accurate vessel and optic disc segmentation | Limited validation across diverse datasets | Develop universal segmentation frameworks |
| Synthetic Data Generation | OCTA Synthetic Images | Reduces annotation requirements | Synthetic-to-real domain gap | Generate more realistic and diverse datasets |
| Explainable AI | Transformer and Concept-based Models | Improves interpretability and clinical trust | Computational complexity remains high | Design efficient and explainable AI systems |
| Clinical Deployment | MobileNet-based and Mobile AI Models | Suitable for portable and real-time diagnosis | Limited real-world validation | Develop robust systems for clinical applications |

**Comparative Summery**

Recent advancements in Artificial Intelligence (AI) and Deep Learning (DL) have significantly improved automated retinal disease diagnosis systems. Existing studies have explored various approaches, including CNN-based architectures, multi-modal learning, transformer models, meta-learning, segmentation networks, and synthetic data generation. CNN-based models such as MobileNetV3 and DeepRetinaNet have achieved high classification accuracy [1], [5], while multi-modal frameworks like MultiEYE enhance diagnosis by integrating Fundus and OCT images [2]. Advanced methods such as RLF-Net, FANet, CLAT, and Mirror Meta-Learning further improve lesion detection, robustness, generalization, and interpretability [4], [8], [9], [12]. Segmentation techniques, including W-Capsule Networks and URVSM, also provide accurate retinal structure analysis across multiple modalities [13], [14].

Despite these advancements, several challenges remain. Many models require large annotated datasets and high computational resources, limiting their real-time clinical applicability [5], [12], [16]. Other issues include dependence on single imaging modalities, class imbalance, noisy annotations, limited generalization, and increased complexity in transformer-based and multi-modal frameworks [1], [2], [4], [11], [12]. In addition, synthetic data generation faces challenges due to the gap between synthetic and real-world images [15], while variations in imaging devices and lack of standardized datasets affect model robustness [6], [14].

Therefore, future retinal disease diagnosis systems should focus on lightweight, interpretable, and computationally efficient models that can operate across diverse clinical settings and multiple imaging modalities [5], [9], [12], [16].

## Objectives

- To generate synthetic OCT-like features from retinal fundus images using deep learning techniques.
- To integrate fundus and synthetic OCT features for improved retinal disease analysis.
- To enhance the accuracy and reliability of retinal disease detection.
- To develop a cost-effective, interpretable, and clinically deployable diagnostic system for both urban and rural healthcare environments.

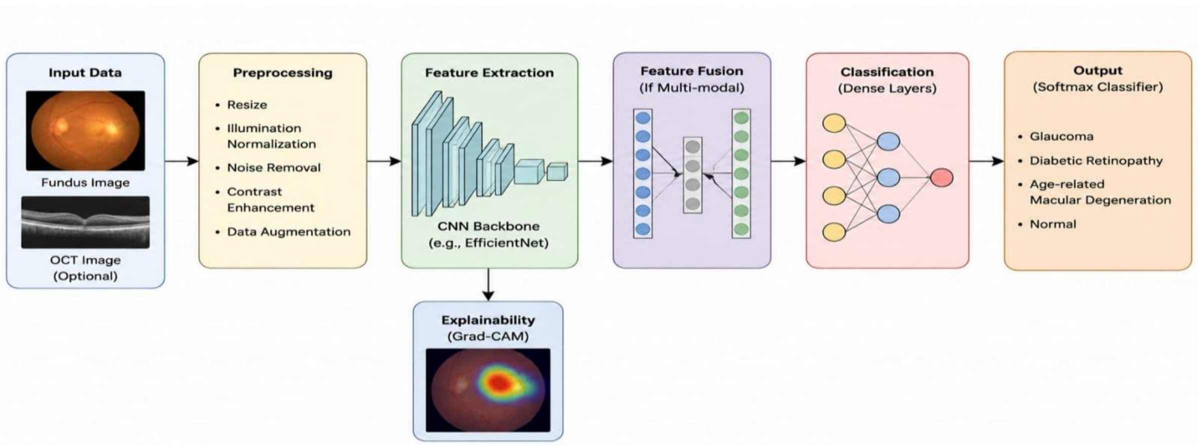
## Problem Statement

Deep learning techniques have shown promising results in diagnosing retinal diseases such as diabetic retinopathy, glaucoma, and age-related macular degeneration. However, many existing approaches rely on single-modality data or require paired Fundus and OCT images, which are often difficult to obtain in real clinical settings. This challenge is more significant in rural and resource-limited healthcare environments where advanced imaging devices may not be available.

Moreover, current models often face challenges such as class imbalance, noisy annotations, limited generalization, and lack of interpretability. Most deep learning systems function as black-box models, reducing clinicians' trust in automated predictions. Therefore, there is a need to develop an efficient, robust, and interpretable retinal disease diagnosis system that can operate effectively with limited data, address real-world challenges, and provide reliable support for clinical decision-making.

## Proposed Future Framework

The proposed framework begins with retinal fundus images as the input, followed by preprocessing steps such as image resizing, noise removal, and contrast enhancement to improve image quality. To overcome the limitation of unavailable OCT images in many healthcare settings, a deep learning-based synthetic OCT generation module is proposed to create OCT-like representations from fundus images. The extracted fundus and synthetic OCT features are then fused to obtain richer image representations for disease analysis. Subsequently, a CNN-based architecture, such as EfficientNet, can be employed for feature extraction and retinal disease classification, including diabetic retinopathy, glaucoma, and age-related macular degeneration. Finally, an Explainable AI technique such as Grad-CAM can be incorporated to provide visual explanations for the predictions, thereby improving interpretability and supporting clinicians in reliable decision-making.




## Conclusion

This study reviewed recent advancements in deep learning-based retinal disease diagnosis and analyzed various techniques used for automated retinal image analysis. The survey showed that CNNs, transformer-based models, multi-modal learning, and lightweight architectures have significantly improved diagnostic accuracy and efficiency. However, challenges such as limited annotated datasets, class imbalance, noisy labels, high computational cost, limited generalization, and lack of interpretability still remain. The reviewed studies indicate that deep learning has strong potential to assist ophthalmologists in early disease detection and clinical decision-making. Future research should focus on developing robust, interpretable, computationally efficient, and clinically deployable retinal disease diagnosis systems.

## References

[1] R. Afshan, A. K. Chakraverti and M. Chhabra, "Glaucoma Detection Approach based on Colour Fundus Pictures," *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2024, pp. 87-92, doi: 10.23919/INDIACom61295.2024.10499019.

[2] L. Wang *et al.*, "MultiEYE: Dataset and Benchmark for OCT-Enhanced Retinal Disease Recognition From Fundus Images," in *IEEE Transactions on Medical Imaging*, vol. 44, no. 4, pp. 1711-1722, April 2025, doi: 10.1109/TMI.2024.3518067.

[3] K. Wisaeng, "Retinal Image Registration Using Partial Intensity Invariant Feature Descriptor and Redundant Keypoints Elimination Techniques," in *IEEE Access*, vol. 12, pp. 83611-83628, 2024, doi: 10.1109/ACCESS.2024.3404242.

[4] X. Chen *et al.*, "Retinal Fusion Network with Contrastive Learning for Imbalanced Multi-Class Retinal Disease Recognition in FFA," in *Big Data Mining and Analytics*, vol. 8, no. 6, pp. 1225-1244, December 2025, doi: 10.26599/BDMA.2025.9020020.

[5] A. K. Sahoo, P. Parida, M. K. Panda, C. Nayak and N. Mohankumar, "DeepRetinaNet: An Automated AI-Based Framework for Retinal Disease Diagnosis," in *IEEE Latin America Transactions*, vol. 23, no. 8, pp. 718-728, Aug. 2025, doi: 10.1109/TLA.2025.11072496.

[6] Y. Turkan *et al*., "Automated Diagnosis of Alzheimer's Disease Using OCT and OCTA: A Systematic Review," in *IEEE Access*, vol. 12, pp. 104031-104051, 2024, doi: 10.1109/ACCESS.2024.3434670.

[7] N. Kalita and S. K. Borgohain, "An Ocular Feature-Based Novel Biomarker Determination for Glaucoma Diagnosis Using Supervised Machine Learning and Fundus Imaging," in *IEEE Sensors Letters*, vol. 8, no. 11, pp. 1-4, Nov. 2024, Art no. 6014504, doi: 10.1109/LSENS.2024.3483990.

[8] X. Guo *et al*., "FANet: Joint Spatial–Frequency Domain Framework for Retinal Disease Classification From OCT Images," in *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1-18, 2025, Art no. 5041618, doi: 10.1109/TIM.2025.3596997.

[9] H. Peng *et al*., "Learn to Learn: A Mirror Meta-Learning Method for Retinal Disease Diagnosis on Fundus Images," in *IEEE Transactions on Artificial Intelligence*, vol. 6, no. 12, pp. 3391-3405, Dec. 2025, doi: 10.1109/TAI.2025.3566082.

[10] d. Young Kim, H. Keun Kim, Y. Jun Lim and M. H. Sunwoo, "Efficient Deep Retinal Fundus Image-Based Network for Alzheimer's Disease Diagnosis Using Mobile Device Applications," in *IEEE Access*, vol. 12, pp. 79166-79176, 2024, doi: 10.1109/ACCESS.2024.3396153.

[11] J. Hou, J. Xu, R. Feng and H. Chen, "QMix: Quality-Aware Learning With Mixed Noise for Robust Retinal Disease Diagnosis," in *IEEE Transactions on Medical Imaging*, vol. 44, no. 8, pp. 3345-3355, Aug. 2025, doi: 10.1109/TMI.2025.3562614.

[12] C. Wen, M. Ye, H. Li, T. Chen and X. Xiao, "Concept-Based Lesion Aware Transformer for Interpretable Retinal Disease Diagnosis," in IEEE Transactions on Medical Imaging, vol. 44, no. 1, pp. 57-68, Jan. 2025, doi: 10.1109/TMI.2024.3429148.

[13] G. Balla, M. F. Hashmi and D. M. Kotambkar, "Selective Routing-Based Optic Disc Segmentation for Glaucoma Diagnosis," in *IEEE Sensors Letters*, vol. 8, no. 5, pp. 1-4, May 2024, Art no. 7002704, doi: 10.1109/LSENS.2024.3384679.

[14] B. Wen *et al*., "Universal Vessel Segmentation for Multi-Modality Retinal Images," in *IEEE Transactions on Image Processing*, vol. 34, pp. 7903-7918, 2025, doi: 10.1109/TIP.2025.3623893.

[15] L. Kreitner *et al*., "Synthetic Optical Coherence Tomography Angiographs for Detailed Retinal Vessel Segmentation Without Human Annotations," in *IEEE Transactions on Medical Imaging*, vol. 43, no. 6, pp. 2061-2073, June 2024, doi: 10.1109/TMI.2024.3354408.

[16] A. Zhang, X. Qian, C. Xu and J. Zhang, "A Novel Artificial-Intelligence-Based Approach for Automatic Assessment of Retinal Disease Images Using Multi-View Deep-Broad Learning Network," in *IEEE Access*, vol. 12, pp. 13248-13259, 2024, doi: 10.1109/ACCESS.2024.3356824.