Engineering and Scientific International Journal (ESIJ)
Volume 13, Issue 2, April – June 2026
ISSN 2394-7187(Online)
ISSN 2394 - 7179 (Print)

An Intelligent Deep Learning Framework for Ocular Disease Diagnostics: The RetinaVision Approach

Bonasi Poojitha*[1], Dr.S.Usharani[2]
[1]*Student, Department of Computer Applications, Viswam Engineering College, Andhra Pradesh, India*
[2]*Professor & Head, Department of Computer Applications, Viswam Engineering College, Andhra Pradesh, India*
[1]bonasipoojitha@gmail.com, [2]anjanasundar80@gmail.com

Abstract —Vision impairment and blindness caused by ocular diseases remain a major global health challenge, affecting over 2.2 billion people worldwide. Many conditions such as diabetic retinopathy, glaucoma, cataract, age-related macular degeneration, hypertension-related retinopathy, and myopia are progressive but preventable with early diagnosis. However, limited access to ophthalmologists, high consultation costs, and time-intensive manual diagnosis hinder timely detection, especially in developing regions. RetinaVision AI addresses these challenges through a deep learning-based automated system for retinal disease detection using fundus images. The system is built on the EfficientNet-B3 architecture, known for its high accuracy and computational efficiency. It employs two models: a multi-disease classifier that detects eight retinal conditions, and a specialized model for diabetic retinopathy severity grading across four levels.To enhance interpretability, the system integrates Grad-CAM, which generates heatmaps highlighting important regions in the retinal image, ensuring clinically relevant decision-making. RetinaVision AI is deployed as a user-friendly web application using Streamlit, enabling easy access through a browser. Overall, the system demonstrates the effective use of deep learning, transfer learning, and explainable AI for scalable and reliable ophthalmological diagnostics.

***Keywords* —** *Ocular Disease Detection; Deep Learning; EfficientNet, Fundus Imaging; Grad-CAM, Diabetic Retinopathy; Explainable AI.*

1. Introduction

The human eye is a complex organ, with the retina playing a vital role in vision by converting light into neural signals sent to the brain. Fundus imaging, which captures detailed images of the retina, is widely used to diagnose ocular and systemic diseases. Retinal abnormalities provide critical insights into conditions such as diabetic retinopathy, glaucoma, age-related macular degeneration, cataract, hypertensive retinopathy, and myopia, all of which are major contributors to global vision impairment. Diabetic retinopathy is a leading cause of preventable blindness, affecting a large proportion of diabetic patients. Its progression from mild retinal changes to severe vision-threatening stages highlights the importance of early detection and timely treatment. Similarly, glaucoma causes irreversible damage to the optic nerve and often goes undetected in early stages. Other conditions like macular degeneration and cataract further contribute to vision loss, making large-scale screening essential. However, traditional diagnosis relies on expert ophthalmologists, making it time-consuming and inaccessible in many regions. Recent advances in artificial intelligence, particularly deep learning, have enabled automated analysis of fundus images with accuracy comparable to medical experts. These systems can support early detection, reduce diagnostic burden, and improve accessibility in underserved areas.Retina Vision AI builds on these advancements by developing a deep learning-based diagnostic system using EfficientNet-B3. It incorporates transfer learning to achieve high accuracy with limited data and integrates Grad-CAM visualizations to enhance interpretability. The system is deployed as a web application using Streamlit, ensuring easy accessibility and usability for clinical and research purposes.

1.1 Project Objectives

The primary objective of RetinaVision AI is to develop and deploy an automated system for detecting multiple ocular diseases from fundus images with high accuracy, interpretability, and accessibility.

- Multi-disease classification: Develop a model to detect eight retinal conditions, supporting multi-label classification for co-existing diseases.
- Diabetic retinopathy grading: Implement a dedicated model to classify severity into four stages for improved clinical decision-making.
- Model interpretability: Integrate Grad-CAM to generate heatmaps highlighting regions influencing predictions.
- Web-based deployment: Create an interactive application enabling real-time diagnosis through a simple interface.
- Reproducibility: Provide clear documentation for system replication, validation, and future enhancements.

Table 1: Detected Ocular Diseases and Description

Disease Name	Description
Diabetic Retinopathy (DR)	Damage to retinal blood vessels due to diabetes

DOI: **10.30726/esij/v13.i2.2026.1320029**

ISSN 2394-7187(Online)
ISSN 2394 - 7179 (Print)

Glaucoma	Optic nerve damage caused by increased intraocular pressure
Cataract	Clouding of the eye lens leading to blurred vision
Age-related Macular Degeneration	Deterioration of central vision due to macula damage
Hypertensive Retinopathy	Retinal damage caused by high blood pressure
Myopia	Near-sightedness affecting distant vision
Normal	No detectable retinal abnormalities
Others	Miscellaneous retinal conditions not categorized above

2. Literature Survey

Diabetic retinopathy (DR) is one of the leading causes of blindness worldwide, making early detection crucial for preventing vision loss. According to the World Health Organization (2023), the global burden of visual impairment continues to rise, emphasizing the need for automated and scalable diagnostic solutions. With the advancement of deep learning, significant progress has been made in medical image analysis. Foundational work by LeCun et al. (2015) established deep learning as a powerful approach for hierarchical feature extraction, enabling breakthroughs in computer vision tasks. Large-scale datasets such as ImageNet (Deng et al., 2009) have played a critical role in training deep neural networks and improving generalization. Early convolutional neural network (CNN) architectures like VGGNet (Simonyan & Zisserman, 2014) demonstrated the effectiveness of deep architectures using small convolution filters. Later, ResNet (He et al., 2016) introduced residual learning, which addressed the vanishing gradient problem and enabled training of very deep networks.

Building upon these ideas, EfficientNet (Tan & Le, 2019) proposed a compound scaling method that balances network depth, width, and resolution, achieving superior performance with fewer parameters. In the medical domain, Gulshan et al. (2016) developed one of the first deep learning models for detecting diabetic retinopathy from retinal fundus images, achieving performance comparable to ophthalmologists. Similarly, Ting et al. (2017) extended this work by validating a deep learning system across multiethnic populations, demonstrating robustness and generalizability. Abramoff et al. (2018) further advanced the field by introducing an autonomous AI-based diagnostic system tested in real clinical settings. Recent studies have also explored alternative architectures such as Vision Transformers (Dosovitskiy et al., 2020), which model global image relationships and have shown promising results in image classification tasks. Additionally, review studies like Li et al. (2021) summarize the growing applications of deep learning in fundus image analysis, highlighting improvements in detection accuracy and efficiency. Interpretability remains a key concern in medical AI. Techniques like Grad-CAM (Selvaraju et al., 2017) and Class Activation Mapping (Zhou et al., 2016) provide visual explanations by highlighting important regions in retinal images, increasing trust and transparency in model predictions. Handling imbalanced datasets is another critical challenge in medical diagnosis (Aneeshkmar, 2016). The SMOTE technique (Chawla et al., 2002) is widely used to address class imbalance by generating synthetic samples for minority classes, improving model performance.Practical implementation of these models relies on powerful tools and frameworks. PyTorch (Paszke et al., 2019) provides flexibility for building and training deep learning models, while OpenCV (Bradski, 2000) and NumPy (Van der Walt et al., 2011) support efficient image processing and numerical computation. Platforms like Streamlit (2024) enable deployment of user-friendly web applications for real-time diagnosis.Datasets such as the APTOS 2019 Blindness Detection dataset have been instrumental in benchmarking models for diabetic retinopathy classification, fostering innovation and comparison among different approaches.

3. System Proposal

3.1 Existing System

Existing approaches for automated ocular disease detection include traditional image processing methods, deep learning models, and commercial AI-based screening systems. Traditional methods rely on handcrafted feature extraction to detect retinal abnormalities such as hemorrhages, exudates, and optic disc changes. These techniques use classical computer vision methods like filtering, segmentation, and machine learning classifiers. However, they are sensitive to image quality variations and require manual tuning. Deep learning approaches, especially CNN-based models, have significantly improved diagnostic accuracy. Studies by Varun Gulshan et al. (2016) and Daniel Ting et al. (2017) demonstrated performance comparable to specialists. However, these systems require large datasets and high computational resources. Commercial systems such as IDx-DR, EyeArt, and RetinalScreen are clinically validated but often focus on single diseases, lack interpretability, and involve high costs .Recent research explores multi-disease detection using advanced architectures, but most systems are not publicly accessible or user-friendly. Interpretability tools like Grad-CAM exist but are rarely integrated into complete applications.

3.1.1 Advantages of Existing System

- Clinically validated systems: Commercial platforms like IDx-DR provide proven accuracy through regulatory approval and clinical trials.

DOI: **10.30726/esij/v13.i2.2026.1320029**

ISSN 2394-7187(Online)
ISSN 2394 - 7179 (Print)

- High diagnostic performance: Deep learning models trained on large datasets achieve accuracy comparable to medical experts.
- Better generalization: Large-scale training data improves performance across diverse populations and imaging conditions.
- Regulatory compliance: Approved systems meet strict safety, sensitivity, and specificity standards.
- Algorithm transparency (traditional methods): Classical image processing techniques are fully explainable and easier to audit.
- Reproducibility: Traditional approaches provide consistent and predictable outputs due to fixed algorithms.

3.2 Proposed System

RetinaVision AI is an open-source, multi-disease ocular diagnostic system that combines deep learning, interpretability, and easy deployment through a web interface. It enables simultaneous detection of multiple eye diseases, severity grading for diabetic retinopathy (DR), and visual explanations using Grad-CAM.The system is built on the EfficientNet-B3 architecture, chosen for its high accuracy and computational efficiency. Compared to earlier models like VGG and ResNet, it provides better performance with fewer parameters, making it suitable for medical image analysis.A multi-label classification approach using sigmoid activation allows detection of multiple diseases simultaneously, reflecting real clinical scenarios. The system identifies eight conditions: Normal, Diabetic Retinopathy, Glaucoma, Cataract, Age-related Macular Degeneration, Hypertension, Myopia, and Others.
For diabetic retinopathy, a separate severity grading model (also based on EfficientNet-B3) classifies cases into Normal, Mild, Moderate, and Severe. This two-stage approach improves accuracy by focusing specifically on DR severity after initial screening.To improve interpretability, Grad-CAM is used to generate heatmaps highlighting important regions in retinal images. These visual explanations help users understand model predictions.The system is deployed using Streamlit, providing a user-friendly web interface with features like image upload, prediction display, severity analysis, and visualization outputs. It supports both GPU and CPU execution for flexible deployment.

3.2.1 Disadvantages of the Proposed System

- Lack of clinical validation: The system is not approved as a medical device and requires clinical trials and regulatory clearance.
- Limited training data: Uses smaller public datasets compared to large proprietary datasets, affecting generalization.
- Performance ariability: May not perform consistently across different imaging devices or patient populations.
- Threshold dependency: Multi-label classification relies on manually set thresholds that may need further optimization.
- Computational overhead: Grad-CAM visualization increases processing time compared to standard inference.
- Limited interpretability scope: Grad-CAM provides partial explanations and may not fully represent model decision-making.

4. Implementation

4.1 Modules

RetinaVision AI is organized into five functionally distinct implementation modules: the Model Loading and Initialization Module, the Multi-Disease Screening Module, the Diabetic Retinopathy Severity Grading Module, the Gradient-weighted Class Activation Mapping Module, and the Streamlit Web Application Interface Module. Each module encapsulates a well-defined set of responsibilities that collectively implement the complete diagnostic pipeline from raw fundus image input to comprehensive visual diagnostic output.

Table 2: Model Architecture and Configuration

Parameter	Value/Description
Base Model	EfficientNet-B3
Input Image Size	300 × 300 pixels
Number of Classes	8 (multi-disease), 4 (DR severity)
Activation Function	Sigmoid (multi-label), Softmax (DR grading)
Framework Used	PyTorch
Hardware Support	CPU and GPU (CUDA)
Preprocessing	Resizing, normalization, tensor conversion

4.2 Modules Description

4.2.1 Model Loading and Initialization Module

This module loads and initializes the trained deep learning models for inference in the Streamlit application. To avoid repeated loading and improve performance, models are cached using st.cache_resource.
Two models based on EfficientNet-B3 are loaded:

- Multi-disease classification model (8 classes)
- Diabetic Retinopathy (DR) severity model (4 classes)

The system automatically selects GPU (CUDA) if available, otherwise CPU. Both models are set to evaluation mode.

DOI: **10.30726/esij/v13.i2.2026.1320029**

ISSN 2394-7187(Online)
ISSN 2394 - 7179 (Print)

Input images are preprocessed using resizing (300×300) and tensor conversion for compatibility with the model.

4.2.2 Multi-Disease Screening Module

This module performs primary disease detection from uploaded retinal images. The image is preprocessed and passed through the model to generate predictions. A sigmoid activation function produces independent probabilities for each disease, enabling multi-label classification. A threshold (0.40, with 0.45 for DR) is applied to identify detected conditions. Results include:

- List of detected diseases
- Probability scores for all classes

If no disease is detected, the system labels the image as Normal.

4.2.3 Diabetic Retinopathy Severity Grading Module

This module activates when DR probability exceeds a trigger threshold (0.20).
A second EfficientNet-B3 model classifies DR into:

- Normal
- Mild
- Moderate
- Severe

Softmax activation is used since severity levels are mutually exclusive. The predicted class and probability scores are displayed, helping identify borderline cases.

4.2.4 Grad-CAM Visualization Module

This module generates visual explanations using Grad-CAM.It identifies important regions in the retinal image that influenced the model's prediction by analyzing gradients and feature maps from the final convolutional layer.The output is a heatmap highlighting critical areas, which is overlaid on the original image using color mapping. This improves transparency and interpretability of predictions

4.2.5 Streamlit Web Application Interface Module

This module provides the user interface using Streamlit.,Key features include:

- Image upload (JPEG/PNG)
- Side-by-side display of input image and results
- Disease predictions and probabilities
- DR severity output (if applicable)
- Grad-CAM visualization

The interface uses a clean two-column layout and automatically updates when a new image is uploaded. It supports both GPU and CPU execution and provides a simple, interactive diagnostic dashboard.

5. Results and Discussions

Table 3: System Output and Features

Feature	Description
Disease Prediction	Detects multiple retinal diseases from fundus image
Probability Scores	Displays confidence level for each disease
DR Severity Classification	Classifies DR into Normal, Mild, Moderate, Severe
Grad-CAM Heatmap	Highlights important regions influencing prediction
Image Upload	Accepts retinal images (JPEG/PNG format)
Real-time Processing	Provides instant results via Streamlit interface

The RetinaVision AI system was successfully developed using the EfficientNet-B3 model and deployed through a Streamlit interface. The system was tested on retinal fundus images and demonstrated effective performance in detecting multiple ocular diseases such as Diabetic Retinopathy (DR), Glaucoma, Cataract, and others.

The use of a multi-label classification approach with sigmoid activation allowed the model to identify more than one disease in a single image, which closely reflects real clinical scenarios.The Diabetic Retinopathy severity grading module further enhanced the system by classifying DR cases into four stages: Normal, Mild, Moderate, and Severe.

This two-stage approach improved both efficiency and accuracy by activating severity analysis only when DR was detected. The Grad-CAM visualization module generated heatmaps that highlighted important regions of the retinal images, providing interpretability and helping users understand the model's decision-making process.The system performed efficiently in both GPU and CPU environments, with faster predictions observed when GPU support was available.

The Streamlit interface ensured smooth user interaction by providing real-time predictions, probability scores, and visual outputs in a clean and accessible format.Despite these strengths, certain limitations were observed. The performance of the model depends on the quality and diversity of the training dataset, which may affect generalization to real-world clinical data.

Additionally, threshold-based classification requires careful tuning to balance sensitivity and specificity.Overall, the results indicate that RetinaVision AI is an effective and user-friendly system for automated ocular disease detection, offering a strong foundation for further improvements and potential clinical applications.

DOI: **10.30726/esij/v13.i2.2026.1320029**

6. Conclusion and Future Enhancement

6.1 Conclusion

RetinaVision AI demonstrates an effective application of deep learning for automated ocular disease detection using retinal fundus images. Built on EfficientNet-B3, the system integrates multi-disease screening, diabetic retinopathy (DR) severity grading, and interpretability through Grad-CAM, all deployed via a Streamlit interface.

The multi-label classification approach enables detection of multiple coexisting eye diseases, reflecting real clinical conditions. The two-stage pipeline for DR—screening followed by severity grading—improves diagnostic relevance and efficiency. Grad-CAM visualizations enhance transparency by highlighting important image regions, supporting user trust and understanding. The web-based interface ensures accessibility and ease of use, making the system suitable for academic and preliminary clinical support applications.Despite its strengths, RetinaVision AI remains an academic prototype and requires clinical validation before real-world deployment. Its performance is influenced by dataset size and diversity, and threshold-based classification may need further optimization for different clinical settings.

6.2 Future Enhancement

- Dataset Expansion: Incorporating larger and more diverse datasets will improve model generalization across populations and imaging conditions.
- Class Imbalance Handling: Techniques like SMOTE, data augmentation, and weighted loss functions can enhance detection of rare diseases.
- Advanced Architectures: Integration of transformer-based models such as Vision Transformers could improve performance by capturing global image features.
- Improved Interpretability: Adding methods like SHAP and LIME can provide deeper and more reliable explanations of model decisions.
- Multi-Modal Support: Extending the system to include OCT imaging can broaden diagnostic capability.
- Mobile Application Development: A mobile version can enable remote screening in low-resource settings.
- Clinical Validation: Rigorous testing against ophthalmologists is necessary to ensure safety, accuracy, and regulatory approval.
- Overall, Retina Vision AI provides a strong foundation for future advancements in AI-driven ophthalmic diagnostics.

References

[1] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," Proc. ICML, vol. 97, pp. 6105–6114, 2019.

[2] V. Gulshan et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," JAMA, vol. 316, no. 22, pp. 2402–2410, 2016.

[3] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," Proc. ICCV, pp. 618–626, 2017.

[4] D. S. W. Ting et al., "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," JAMA, vol. 318, no. 22, pp. 2211–2223, 2017.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. CVPR, pp. 770–778, 2016.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[7] J. Deng et al., "ImageNet: A large-scale hierarchical image database," Proc. CVPR, pp. 248–255, 2009.

[8] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," Adv. Neural Inf. Process. Syst., vol. 32, pp. 8026–8037, 2019.

[9] Aneeshkumar A. S., "Relevance of Data Mining for Presumption of Anarchy," International Journal of Linguistics and computational Applications (IJLCA), IRDP Group of Journals, vol.3, issue 4, pp. 72-74, 2016.

[10] G. Bradski, "The OpenCV library," Dr. Dobb's J. Softw. Tools, vol. 25, no. 11, pp. 120–125, 2000.

[11] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy array: A structure for efficient numerical computation," Comput. Sci. Eng., vol. 13, no. 2, pp. 22–30, 2011.

[12] World Health Organization, World Report on Vision. Geneva, Switzerland: WHO Press, 2023.

[13] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, pp. 115–118, 2017.

[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, 2015.

[15] Streamlit Inc., "Streamlit documentation," 2024. [Online]. Available: https://docs.streamlit.io

[16] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," arXiv preprint arXiv:2010.11929, 2020.

[17] B. Zhou et al., "Learning deep features for discriminative localization," Proc. CVPR, pp. 2921–2929, 2016.

[18] N. V. Chawla et al., "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., vol. 16, pp. 321–357, 2002.

[19] Kaggle, "APTOS 2019 blindness detection dataset," 2019. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection

[20] T. Li et al., "Applications of deep learning in fundus images: A review," Med. Image Anal., vol. 69, p. 101971, 2021.

[21] M. D. Abramoff et al., "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices," npj Digital Medicine, vol. 1, p. 39, 2018.

DOI: **10.30726/esij/v13.i2.2026.1320029**