International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 1, Jan-Feb-2026, ISSN (Online): 2395-566X

Ocular Disease Recognition Using VGG-19 Deep Learning with Multi-Class Classification on Retinal Images

Anuja Shinde

Abstract- — This paper presents a deep learning-based framework for the automated recognition of ocular diseases using retinal fundus imaging data. Leveraging the VGG-19 convolutional neural network (CNN) architecture with transfer learning, the proposed system performs multi-class classification of retinal images to distinguish between seven ocular conditions: Myopia (M), Hypertension (H), Diabetes (D), Cataract (C), Glaucoma (G), Age-related Macular Degeneration (A), and other abnormalities (O). The input images are preprocessed using computer vision techniques including normalization, contrast enhancement, and texture and shape-based feature extraction. Unlike prior binary classification approaches, our system enables simultaneous prediction of multiple diseases within a single retinal image using the Ocular Disease Intelligent Recognition (ODIR) dataset of 10,000 images. Experimental results demonstrate high classification accuracy, with the model achieving competitive precision, recall, and F1-scores. The proposed system has significant implications for clinical ophthalmology, particularly in enabling early, accurate, and scalable eye disease diagnosis in resource-limited environments.

Keywords: Deep Learning, VGG-19, Convolutional Neural Network, Ocular Disease Recognition, Transfer Learning, Retinal Fundus Images, Multi-Class Classification, ODIR Dataset.

I. INTRODUCTION

Eye diseases represent a significant global healthcare burden, affecting millions of individuals and potentially leading to irreversible vision loss and blindness. The World Health Organization (WHO) estimates that over 2.2 billion people worldwide suffer from some form of vision impairment, with conditions such as cataracts, glaucoma, diabetic retinopathy, and age-related macular degeneration (AMD) being among the most prevalent [1]. This challenge is particularly acute in developing nations, where limited access to trained ophthalmologists, diagnostic infrastructure, and healthcare facilities results in a disproportionate share of undiagnosed and untreated eye diseases.

Traditional ophthalmic diagnosis relies on comprehensive clinical examinations—including visual acuity tests, slit-lamp biomicroscopy, and fundus photography—performed by skilled specialists. While these modalities are clinically reliable, they remain labor-intensive, subjective, and inaccessible in underserved geographic regions. The growing shortage of ophthalmologists globally further exacerbates this challenge, underscoring the need for scalable, automated, and cost-effective diagnostic solutions.

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have enabled the development of high-performance, computer-aided diagnostic systems for medical image analysis. CNN-based models have demonstrated remarkable accuracy in detecting conditions such as diabetic retinopathy, glaucoma, and AMD from retinal fundus images, in some cases matching or exceeding specialist-level performance [2]. Transfer learning techniques further enhance model generalizability by leveraging representations learned from large-scale datasets like ImageNet for domain-specific medical imaging tasks.

This paper proposes an Ocular Disease Recognition (ODR) system based on the VGG-19 CNN architecture. The key contribution of our work lies in transitioning from binary to multi-class disease classification, enabling the simultaneous detection of seven distinct ocular conditions in a single fundus image. Using the publicly available ODIR-5K dataset, the system is trained and evaluated for multi-label classification accuracy, offering a clinically meaningful tool for supporting ophthalmic screening, particularly in resource-constrained settings.

Objectives

- To develop a VGG-19-based deep learning model for automated ocular disease classification.

- To implement multi-class classification enabling simultaneous detection of multiple conditions in a single fundus image.
- To apply transfer learning and image preprocessing to enhance model accuracy and generalizability.
- To evaluate model performance using standard metrics including accuracy, precision, recall, and F1-score.
- To contribute a scalable diagnostic tool that can support ophthalmologists and enable early disease detection in underserved communities.

Scope
The scope of this work encompasses the development and evaluation of an automated system for detecting multiple ocular pathologies from retinal fundus images. The system targets conditions including myopia, cataracts, glaucoma, diabetic retinopathy, hypertension-related retinopathy, AMD, and other miscellaneous ocular abnormalities. Beyond single-institution use, the architecture is designed for integration into telemedicine platforms, mobile screening programs, and large-scale population health initiatives, with a focus on continuous model improvement through iterative learning.

II. RELATED WORK

The field of automated ocular disease recognition has witnessed substantial progress over the past decade, driven by advances in deep learning architectures and the availability of large annotated fundus image datasets. This section reviews key contributions from the literature relevant to the present work.

Deep Learning for Retinal Disease Detection
Muchuchuti and Viriri (2023) [1] conducted a comprehensive review of deep learning methods for retinal disease detection and grading, encompassing conditions such as glaucoma, diabetic retinopathy, and AMD. The authors surveyed the application of Deep Convolutional Neural Networks (DCNNs) and vision transformers for Computer-Aided Diagnosis (CAD), highlighting their superior performance over conventional image processing methods. However, the studies reviewed were noted to utilize relatively small datasets (as few as 434 images), and the authors drew attention to the interpretability challenges associated with black-box deep learning models in clinical decision-making contexts.

Smartphone-Based AI Diagnostics
Baig et al. (2023) [2] explored the application of smartphone-based artificial intelligence for ocular disease detection, emphasizing its potential to democratize eye care in remote and underserved populations. The review evaluated diverse AI algorithms including CNNs, Support Vector Machines (SVMs), and decision trees for image-based diagnosis. While acknowledging the accessibility benefits of smartphone deployment, the authors identified significant limitations including constrained diagnostic capabilities, data security concerns, and sensitivity to image acquisition quality.

VGG-19 for Ocular Disease Classification
Khan et al. (2022) [3] applied the VGG-19 architecture to classify multiple ocular diseases using the ODIR dataset, employing transfer learning to improve classification accuracy. The model demonstrated high performance in binary classifications of conditions such as glaucoma, cataract, and myopia, outperforming several baseline CNN architectures. A critical limitation of this approach was the binary classification paradigm, which restricted the system to predicting a single disease per image and thus ignored the clinical reality of comorbid ocular conditions.

AlexNet-Based Ocular Recognition
Tugui and Iftene (2022) [4] developed ocular disease identification algorithms from fundus images using AlexNet, achieving an overall accuracy of 91–93% across multiple pathologies. The model performed well for myopia, cataract, and normal eye classification; however, it exhibited significant difficulty in accurately detecting glaucoma, underscoring the challenge of distinguishing subtle structural changes in the optic nerve head from fundus images alone.

CNN Models for Bangladeshi Eye Diseases
Siddique et al. (2022) [5] trained six pre-trained CNN architectures—including DenseNet121, MobileNet, Xception, InceptionV3, VGG16, and VGG19—for detecting common eye conditions prevalent in Bangladesh (cataract, chalazion, squint). Among the tested models, MobileNet achieved the highest accuracy of 97.49%. The study applied standard preprocessing techniques such as resizing, rotation, zooming, and flipping for data augmentation. Limitations included a narrow disease scope, an unclear dataset size, and the absence of clinical validation.

Coherent CNN for OCT-Based Retinal Detection
Upadhyay (2022) [6] proposed a Coherent CNN (CCNN) built upon a pruned VGG-16 architecture for the detection of retinal diseases from Optical Coherence Tomography (OCT) images, classifying four classes: choroidal neovascularization (CNV), diabetic macular edema (DME), DRUSEN, and normal retina. The CCNN achieved an accuracy of 97.16%, demonstrating the effectiveness of model pruning for optimizing network

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 1, Jan-Feb-2026, ISSN (Online): 2395-566X

connectivity. The study acknowledged that model performance was sensitive to hyperparameter tuning.

Benchmark Multi-Disease Dataset (OIA-ODIR)
Li et al. (2021) [7] introduced the OIA-ODIR benchmark dataset comprising 10,000 fundus images from 5,000 patients, annotated with eight distinct ocular disease categories. Their evaluation of nine deep learning models revealed that increasing model depth alone does not guarantee improved classification accuracy, emphasizing the importance of well-structured feature fusion strategies. The authors also identified challenges including class imbalance, limited geographical diversity of training data, and the absence of a standard benchmark for multi-disease fundus classification.

Machine Learning Survey for Eye Disease Detection
Ramanathan et al. (2021) [8] provided a broad review of machine learning approaches—including CNNs, SVMs, decision trees, and ensemble methods—for eye disease detection. The review highlighted that model performance is heavily dependent on dataset quality, size, and annotation accuracy. The authors also drew attention to interpretability barriers in complex architectures and technical requirements that may limit clinical adoption.

Table I presents a synthesized summary of these works, highlighting key findings and limitations.

TABLE I. Summary of Related Works
Year
Authors
Technique
Key Finding / Limitation
2023
Muchuchuti & Viriri
DCNN, Vision Transformers
Strong accuracy; black-box interpretability concern; small dataset (434 images).
2023
Baig et al.
CNN, SVM, Decision Trees
Accessible smartphone deployment; limited diagnostic scope; data security issues.
2022
Khan et al.
VGG-19, Transfer Learning
High binary classification accuracy; limited to single-disease prediction per image.
2022
Tugui & Iftene
AlexNet
91–93% accuracy; poor glaucoma detection performance.
2022
Siddique et al.
MobileNet, VGG19, DenseNet
MobileNet: 97.49%; limited disease scope; no clinical validation.
2022
Upadhyay
CCNN (pruned VGG-16)
97.16% accuracy on OCT; sensitive to hyperparameter tuning.
2021
Li et al.
VGG-16, ResNet, DenseNet
Introduced ODIR benchmark; class imbalance and geographical diversity limitations.
2021
Ramanathan et al.
CNN, SVM, Ensemble
Comprehensive ML review; data quality dependency; interpretability challenges.

III. PROPOSED METHODOLOGY

The proposed ODR system follows a structured pipeline encompassing dataset acquisition, image preprocessing, deep feature extraction, multi-class model training, and performance evaluation. Figure 1 illustrates the overall proposed system architecture.

Dataset
The Ocular Disease Intelligent Recognition (ODIR) dataset, sourced from Kaggle [9], is used for all experiments. The dataset contains 10,000 color fundus photographs collected from both left and right eyes of 5,000 patients. Each image is labeled with one or more of seven ocular disease categories: Myopia (M), Hypertension (H), Diabetes (D), Cataract (C), Glaucoma (G), Age-related Macular Degeneration (A), and Other abnormalities (O). All images are standardized to 224×224 pixels prior to model input.

TABLE II. ODIR Dataset Characteristics

Dataset
Images
Disease Categories
Resolution
ODIR (Ocular Disease Intelligent Recognition)
10,000
M, H, D, C, G, A, O (7 classes)
224×224 px

Image Preprocessing
Raw fundus images undergo several preprocessing steps to standardize input quality and enhance discriminative features:

- Resizing: All images are uniformly resized to 224×224 pixels to match VGG-19 input requirements.
- Normalization: Pixel intensities are normalized to the [0, 1] range to accelerate convergence.
- Contrast Enhancement: Adaptive Histogram Equalization (CLAHE) is applied to improve the visibility of retinal structures such as blood vessels and lesions.
- Data Augmentation: Training images are augmented using random rotation, horizontal/vertical flipping, and zooming to mitigate overfitting and improve model robustness across varied image orientations.

Feature Extraction
Feature extraction is performed using the convolutional layers of the VGG-19 network. Key retinal features extracted include:

- Vascular Structure Features: Patterns in blood vessel morphology, tortuosity, and branching.
- Shape Descriptors: Geometric characteristics of the optic disc, optic cup, and macula.
- Lesion Features: Presence and distribution of microaneurysms, exudates, hemorrhages, and drusen deposits.

VGG-19 Architecture and Transfer Learning
VGG-19 is a 19-layer deep convolutional neural network proposed by Simonyan and Zisserman (2014), consisting of 16 convolutional layers organized in five convolutional blocks, three fully connected layers, and softmax output. The architecture employs small 3×3 convolutional filters throughout, enabling deep feature hierarchies with manageable parameter counts.

Transfer learning is applied by initializing VGG-19 with weights pretrained on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) dataset. The final fully connected layers are replaced and retrained for the seven-class ODIR classification task. Early convolutional layers are frozen during initial training to preserve low-level feature representations, with fine-tuning applied to higher layers in subsequent epochs.

Multi-Class Classification Strategy
A key contribution of this work is the transition from binary to multi-class classification. Unlike prior approaches [3] that reduce the problem to pairwise binary decisions, our model applies a softmax activation at the output layer across all seven disease classes simultaneously. This enables the system to produce probabilistic confidence scores for each condition and flag multiple comorbid diseases in a single inference pass—more closely reflecting real clinical presentations.

Hardware and Software Configuration

TABLE III. Experimental Setup

Component
Specification
Processor
Intel Core 2 GHz
RAM
8 GB
Storage
180 GB HDD
Operating System
Windows 11
Programming Language
Python
Dataset
ODIR (Kaggle)

IV. RESULTS AND DISCUSSION

The proposed VGG-19-based multi-class classification system was trained and evaluated on the ODIR dataset. Model performance was assessed using four standard evaluation metrics: classification accuracy, precision, recall (sensitivity), and F1-score. Confusion matrices were generated to visualize per-class prediction distributions.

The system demonstrated high accuracy in classifying common ocular conditions including myopia, cataracts, diabetic retinopathy, and AMD. The use of transfer learning from

ImageNet-pretrained weights significantly reduced training time while improving generalization on the relatively small medical image dataset. Data augmentation was effective in reducing overfitting, as evidenced by comparable training and validation performance metrics.

The multi-class output design enabled the model to correctly flag comorbid conditions—for example, simultaneously identifying diabetic retinopathy and hypertensive retinopathy in a single fundus image. This represents a meaningful clinical advance over binary classification systems [3] and addresses a key gap identified in prior literature [7].

Qualitative evaluation of model outputs via GUI screenshots confirmed that the system successfully classifies diabetic retinal scans, correctly identifying abnormalities associated with diabetic retinopathy in example test cases. Post-processing steps including confidence thresholding were employed to refine multi-label output predictions.

Compared to related works, the proposed system offers a more clinically realistic multi-class output while maintaining competitive accuracy. The VGG-19 architecture, though computationally heavier than MobileNet, provides richer feature representations that are particularly advantageous for distinguishing visually similar pathologies.

V. CONCLUSION AND FUTURE WORK

Conclusion

This paper presented a deep learning-based system for automated ocular disease recognition using the VGG-19 convolutional neural network with transfer learning. The proposed system performs multi-class classification of retinal fundus images, distinguishing between seven ocular conditions on the ODIR dataset. By transitioning from binary to multi-class classification, the system more accurately reflects clinical reality and enables simultaneous detection of multiple comorbid diseases in a single inference pass.

Experimental results demonstrated high classification accuracy and strong performance across precision, recall, and F1-score metrics. The system has significant potential to support ophthalmologists in clinical practice, accelerate large-scale screening programs, and improve access to diagnostic services—particularly in regions with limited specialist availability. This work represents a meaningful step toward the deployment of AI-powered diagnostic tools in ophthalmic healthcare.

Future Work

- **Enhanced Architectures:** Exploration of more recent architectures such as EfficientNet, Vision Transformers (ViTs), and ensemble models for improved accuracy and efficiency.
- **Multi-modal Data Integration**: Incorporation of additional imaging modalities (OCT, fluorescein angiography) and clinical metadata (patient history, genetic markers) for richer diagnostic models.
- **Explainability:** Integration of Gradient-weighted Class Activation Mapping (Grad-CAM) and other explainability frameworks to enhance clinical trust and transparency.
- **Real-time Mobile Deployment**: Development of lightweight, optimized models for deployment on mobile or edge devices to enable point-of-care diagnostics.
- **Clinical Validation:** Prospective validation in partnership with ophthalmology clinics and hospitals to assess real-world generalizability and safety.
- **Continuous Learning:** Implementation of online learning pipelines to continuously update the model as new annotated data becomes available.

REFERENCES

1. S. Muchuchuti and S. Viriri, "Retinal Disease Detection Using Deep Learning Techniques: A Comprehensive Review," Journal of Imaging, vol. 9, no. 4, 2023.
2. M. A. Baig, T. Hussain, M. D. Khan, A. Malik, and F. Akram, "Smartphone-Based AI Detection of Ocular Diseases," IEEE Access, 2023.
3. M. S. Khan, N. Tafshir, K. N. Alam, A. R. Dhruba, M. M. Khan, A. A. Albraikan, and F. A. Almalki, "Retracted: Deep Learning for Ocular Disease Recognition: An Inner-Class Balance," BioMed Research International, 2022.
4. I. M. Tugui and A. Iftene, "Ocular Disease Recognition," Procedia Computer Science, vol. 207, pp. 2070–2079, 2022.
5. M. A. A. Siddique, J. Ferdous, M. T. Habib, M. J. Mia, and M. S. Uddin, "Convolutional Neural Network Modeling for Eye Disease Recognition," in Proc. International Conference on Advancement in Electrical and Electronic Engineering, 2022.
6. P. K. Upadhyay, "Coherent Convolution Neural Network Based Retinal Disease Detection Using Optical Coherence Tomographic Images," Expert Systems with Applications, 2022.
7. N. Li, T. Li, C. Hu, K. Wang, and H. Kang, "A Benchmark of Ocular Disease Intelligent Recognition: One Shot for Multi-disease Detection," in Artificial Intelligence in Label-Efficient Medical Image Analysis, Springer, 2021.

8. G. Ramanathan, D. Chakrabarti, A. Patil, S. Rishipathak, and S. Kharche, "Eye Disease Detection Using Machine Learning," in Proc. IEEE International Conference on Intelligent Technologies, 2021.
9. K. Schwartz, "Ocular Disease Recognition ODIR-5K Dataset," Kaggle, 2021. [Online]. Available: https://www.kaggle.com/datasets/kevinschwartz1991/ocular-disease-recognition-odir5k/data