Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Deep Learning for Early Diagnosis of Diabetic Retinopathy, Glaucoma, and Cataract from Fundus Images

Puneet Misra[1], Mohd Usman[2], Uroosha Usman[3]

[1,2,3] University of Lucknow, Lucknow, 226007, India

Abstract

Retinal disorders represent a growing global health crisis, particularly among the working-age population. While manual screening of fundus imagery is the standard, it remains a labour-intensive process dependent on specialized expertise. This study addresses the urgent demand for rapid, automated diagnostic tools by evaluating the efficacy of lightweight Deep learning (DL) architectures for multi-class retinal disease identification.

Utilizing a public dataset of 4,217 images, the study performed a comparative analysis of four Convolutional Neural Network (CNN) architectures: MobileNetV2, VGG16, DenseNet121, and EfficientNetB0. Each model was assigned with classifying images into four categories: Normal, Diabetic Retinopathy, Glaucoma, and Cataract. Experimental results identify MobileNetV2 as the superior model, achieving a peak accuracy of 94.08%, an F1-score of 0.9430, and a mean Average Precision (mAP) of 0.9786.

With a calculated efficiency score of 0.4157, MobileNetV2 demonstrates the ideal balance between high diagnostic precision and low computational demand. These findings highlight that streamlined CNNs are uniquely suited for real-time deployment in resource-constrained environments. Ultimately, this research provides a robust foundation for AI-assisted ophthalmic screening, facilitating earlier clinical intervention and improved patient outcomes globally.

1. Introduction

The global landscape of ocular health is currently facing an unprecedented crisis as the prevalence of retinal diseases has surged to epidemiological proportions, presenting a formidable challenge to international public health infrastructures. This escalation is particularly concerning because retinal pathologies target the delicate neural tissues responsible for visual perception; once these tissues are significantly compromised, the resulting vision loss is often permanent [1]. Global demographic changes, particularly the aging population's explosive growth and the sharp increase in systemic metabolic illnesses, are substantially responsible for the crisis's severity. Among them, diabetes mellitus has become a major cause of ocular morbidity, making diseases like age-related macular degeneration (AMD), diabetic retinopathy (DR), cataracts, and glaucoma worldwide health priorities rather than isolated clinical anomalies.

The most pervasive of these conditions—Diabetic Retinopathy, glaucoma, and cataracts—each present unique physiological hurdles that demand early detection. Diabetic Retinopathy originates from chronic microvascular damage, where sustained hyperglycemia compromises the integrity of retinal blood vessels [2]. In contrast, glaucoma is characterized by progressive degeneration of the optic nerve, a process frequently exacerbated by elevated intraocular pressure that restricts the transmission of visual data to the brain. Meanwhile, cataracts involve the gradual opacification or clouding of the eye's natural crystalline lens, creating a physical barrier to light that results in a profound loss of visual clarity. Despite the high prevalence of these ailments, the current diagnostic paradigm is heavily restricted by its reliance on manual intervention. Traditional screening protocols require highly specialized ophthalmologists to conduct painstaking interpretations of fundus photography and Optical Coherence Tomography (OCT) images [3].

This manual, human-centric workflow is inherently unscalable and cannot keep pace with the exponential growth in patient volume worldwide. To mitigate these limitations, the integration of Artificial Intelligence (AI) and DL into the domain of medical imaging has initiated a fundamental paradigm shift. This transition moves clinical practice away from subjective, qualitative assessments toward objective, automated quantification. Convolutional Neural Networks (CNNs) have established themselves as the industry standard for these complex image classification tasks due to their unique ability to autonomously derive hierarchical feature representations from intricate biological data [4].

These artificial networks learn to recognize subtle biomarkers and structural abnormalities that are essential for precise diagnosis by consuming and evaluating enormous libraries of labeled fundus images [5]. These advanced CNN techniques aid as a vital force multiplier for the medical community, providing high-precision diagnostic support that can be effectively deployed in low-resource or geographically constrained demographic areas. By automating the screening process, these tools ensure that high-quality eye care is accessible to broader populations, ultimately assisting ophthalmologists in preventing irreversible blindness.

2. Related Work

Retinal disorders are among the primary causes of unavoidable blindness worldwide, creating an urgent demand for scalable screening solutions [6]. The traditional approach to detecting these retinal diseases relies heavily on manual interpretation by ophthalmologists or optometrists, a process that has several scalability limitations. In order to overcome this, automated methods—including sophisticated computational techniques—have been developed to more accurately determine the severity of malignancy [7].

A significant milestone in this domain was achieved by Gulshan et al., who developed a deep CNN specifically for the successful detection of diabetic retinopathy (DR) and diabetic macular edema from fundus images. Their applied method demonstrated that DL models could accurately detect early signs of pathology in addition to diagnosing established eye problems [8].

Building on this foundation, subsequent studies focused on identifying a broader range of retinal illnesses. Sarki et al. introduced a DL system designed for the automatic detection of moderate and multi-class diabetic eye diseases using pre-trained CNNs, expanding the scope beyond binary classification [9]. Similarly, a DL framework for the simultaneous diagnosis of several retinal illnesses, including DR, Macular Holes (MH), and Optic Disc Cupping (ODC), was suggested by Ejaz S et al. In their study, statistical criteria including sensitivity, specificity, accuracy, F1-score, recall, and precision were used to assess performance. The findings showed that their created framework performed promisingly, with accuracy rates of up to 89.81% for validation using a 12-layer CNN after data augmentation [10].

More recent research has prioritized efficiency for deployment on mobile devices. A lightweight CNN was suggested by D. K. Qasim et al. for the categorization of two retinal diseases: Macular Hole and Diabetic Retinopathy, in addition to normal eyes. According to their results, the MobileNetV2 architecture outperformed NASNetMobile, achieving a highest accuracy of 90.8% compared to 89.5%, proving its viability for resource-constrained environments [11]. Pushing performance further, Sara Ejaz et al. presented a DL model using a hybridized CNN for the early detection of retinal diseases from fundus images. The proposed CNN model achieved a maximum accuracy of 93.39%, outperforming other classification algorithms in the study and demonstrating the efficacy of hybrid architectures [12].

These architectural advancements are supported by large-scale validation studies. Ting et al. developed a DL system using retinal images from multi-ethnic populations, validating that such systems could identify Diabetic Retinopathy, Glaucoma, and Age-related Macular Degeneration with specialist-level accuracy [13]. Furthermore, the scope of AI has expanded to other imaging modalities; Kermany et al. demonstrated the ability to identify medical diagnoses and treatable diseases using image-based DL on Optical Coherence Tomography (OCT) scans, effectively classifying conditions like Choroidal Neovascularization [14]. Finally, the rapid evolution of this field was captured by Asiri et al., whose survey on DL-based computer-aided diagnosis systems highlights the transition toward end-to-end models that necessitate negligible human interference while maximizing diagnostic accuracy [15].

3. Methodology

3.1 Dataset Selection and Characterization

This research utilizes DL architectures to facilitate the automated detection and categorization of various retinal pathologies. The prime data was sourced from a publicly accessible repository on Kaggle, comprising a total of 4,217 high-resolution fundus images.

The dataset is organized into four distinct categories, representing three specific ocular conditions and a control group:

- Diabetic Retinopathy (DR)
- Cataract
- Glaucoma
- Normal (Healthy Retinas)

To ensure a balanced spreading for model training, each classification contains a relatively uniform distribution of approximately 1,000 images. The specific breakdown and quantitative distribution of these images are detailed in Table 1.

Table 1. Division of images in the dataset across training, validation, and testing sets.

Class Name	Training images	Validation images	Testing images	Total
Normal	859	129	86	1074
Diabetic Retinopathy	878	132	88	1098
Glaucoma	806	120	81	1007
Glaucoma	830	125	83	1038

3.2 Implementation methodology

To facilitate the automated categorization of retinal anomalies, this research implemented a series of lightweight DL frameworks. The dataset was laminated into four distinct diagnostic classes: Normal, Diabetic Retinopathy, Glaucoma, and Cataract.

The investigation leveraged four specific Convolutional Neural Network (CNN) architectures chosen for their computational efficiency:

- MobileNetV2
- VGG16
- DenseNet121
- EfficientNetB0

To optimize convergence and feature extraction, each technique was adjusted using transfer learning with pre-trained ImageNet weights. For the final classification stage, the convolutional bases were coupled with a Multi-Layer Perceptron (MLP) featuring fully connected layers.

Performance Evaluation

The proposed models underwent a rigorous evaluation to assess both predictive power and operational feasibility. Key metrics included Accuracy, F1-score, Mean Average Precision (mAP), and inference latency (processing time per image). The experimental outcomes demonstrate the proposed framework's reliability and scalability, validating its potential for deployment in diverse clinical diagnostic workflows.

4. RESULTS AND DISCUSSION

4.1 Comparative Performance Analysis

The empirical evaluation of the selected DL architectures revealed distinct variations in predictive capability, computational efficiency, and inference latency. The quantitative comparison of all implemented models is summarized in Table 2 and Table 3 shows the efficiency analysis.

Table 2. Statistical analysis of different DL model's performance

Model	Accuracy (%)	F1-Score	mAP	Precision	Recall	Param(m)	Inference Time (ms)
MobileNetV2	94.08	0.9430	0.9786	0.9446	0.9426	2.26	12.266
VGG16	93.20	0.9327	0.9653	0.9334	0.9350	14.72	6.311
DenseNet121	89.05	0.8918	0.9501	0.8971	0.8975	7.04	34.6
EfficientNetB0	43.49	0.3556	0.5648	0.3424	0.4619	4.05	22.115

Table 3. Efficiency comparison of CNN models

Model	Efficiency Score
MobileNetV2	0.4157
VGG16	0.0633
DenseNet121	0.1265
EfficientNetB0	10.73

MobileNetV2 emerged as the optimal architecture for this specific domain. It demonstrated superior overall efficacy, achieving a remarkable efficiency score of 0.4157. Furthermore, it secured the highest classification metrics across the board, with an accuracy of 94.08%, an F1-score of 0.9430, and a mean Average Precision (mAP) of 0.9786. These metrics establish MobileNetV2 as the most robust and reliable model for retinal disease identification in this study.

While MobileNetV2 excelled in accuracy, VGG16 distinguished itself through temporal performance. It recorded the fastest inference time of 6.3110 ms, making it highly suitable for time-sensitive applications. Despite prioritizing speed, it maintained a competitive accuracy of 93.20% and an F1-score of 0.9327.

The DenseNet121 model delivered moderate results, yielding an accuracy of 89.05% and an F1-score of 0.8918, falling slightly behind the leading architectures. Conversely, EfficientNetB0 demonstrated suboptimal performance on this specific dataset, recording the lowest accuracy at 43.49% and an F1-score of 0.3566, indicating potential challenges in feature convergence for this specific task.

To further scrutinize the learning behaviour and diagnostic precision of these models, several visual aids were generated. Figure 1 illustrates the training dynamics, comparing the loss and accuracy curves over successive epochs to visualize convergence stability.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

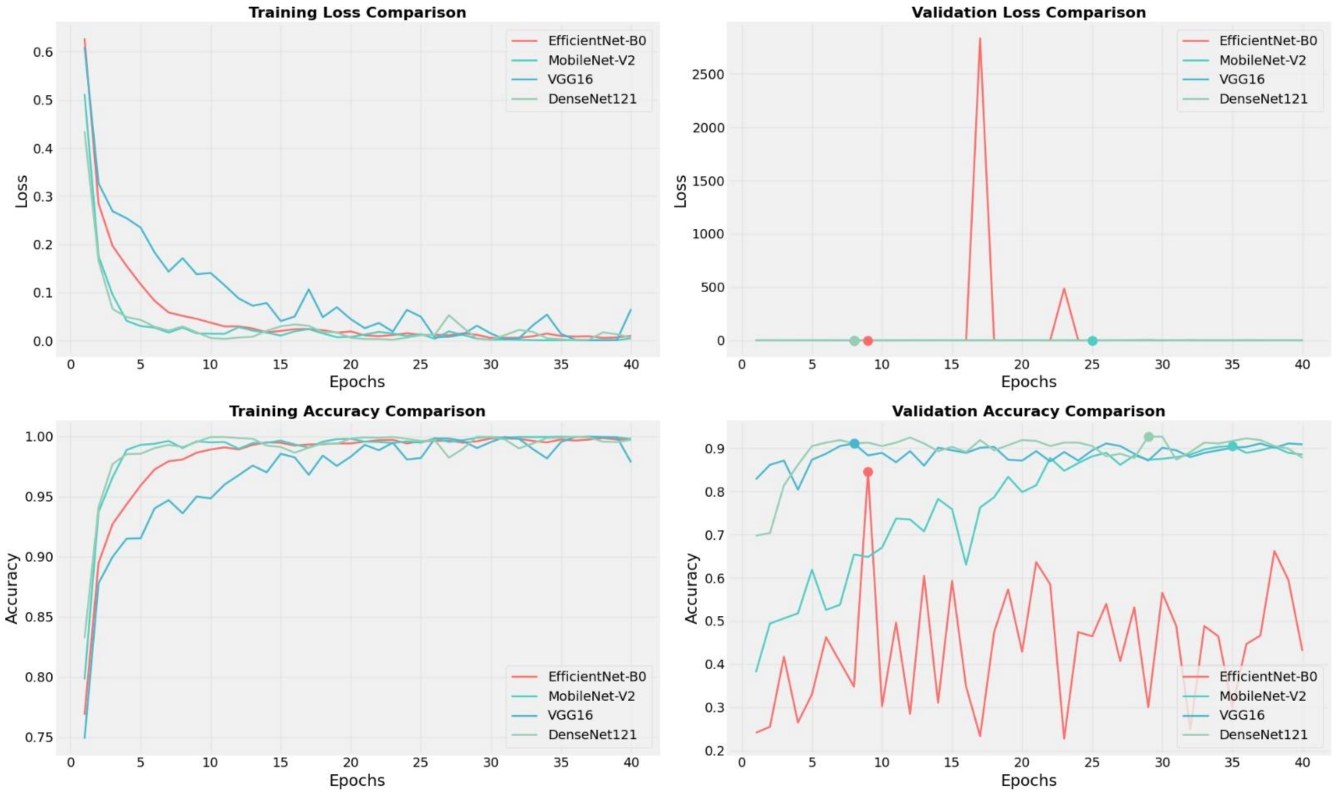

Figure 1. Training histories visualization, including training and validation accuracy and loss over epochs.

Figure 2 presents a comprehensive performance analysis, highlighting trade-offs between precision and recall. Figure 3 depicts the confusion matrices for the models, providing a granular view of misclassification patterns across the four disease categories.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

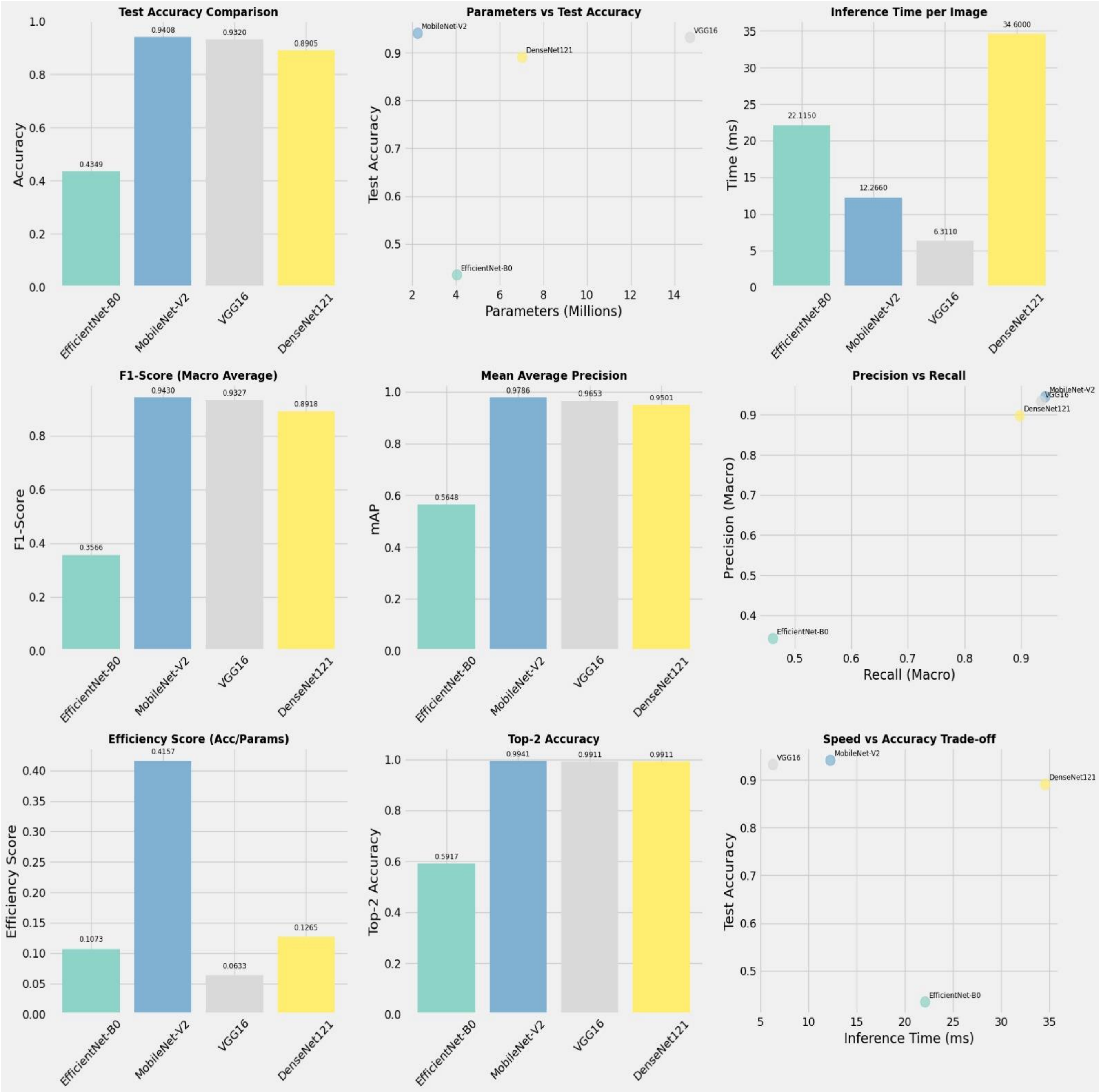

Figure 2. Comparative analysis graphs of DL models- MobileNetV2, VGG16, DenseNet121 and EfficientNetB0.

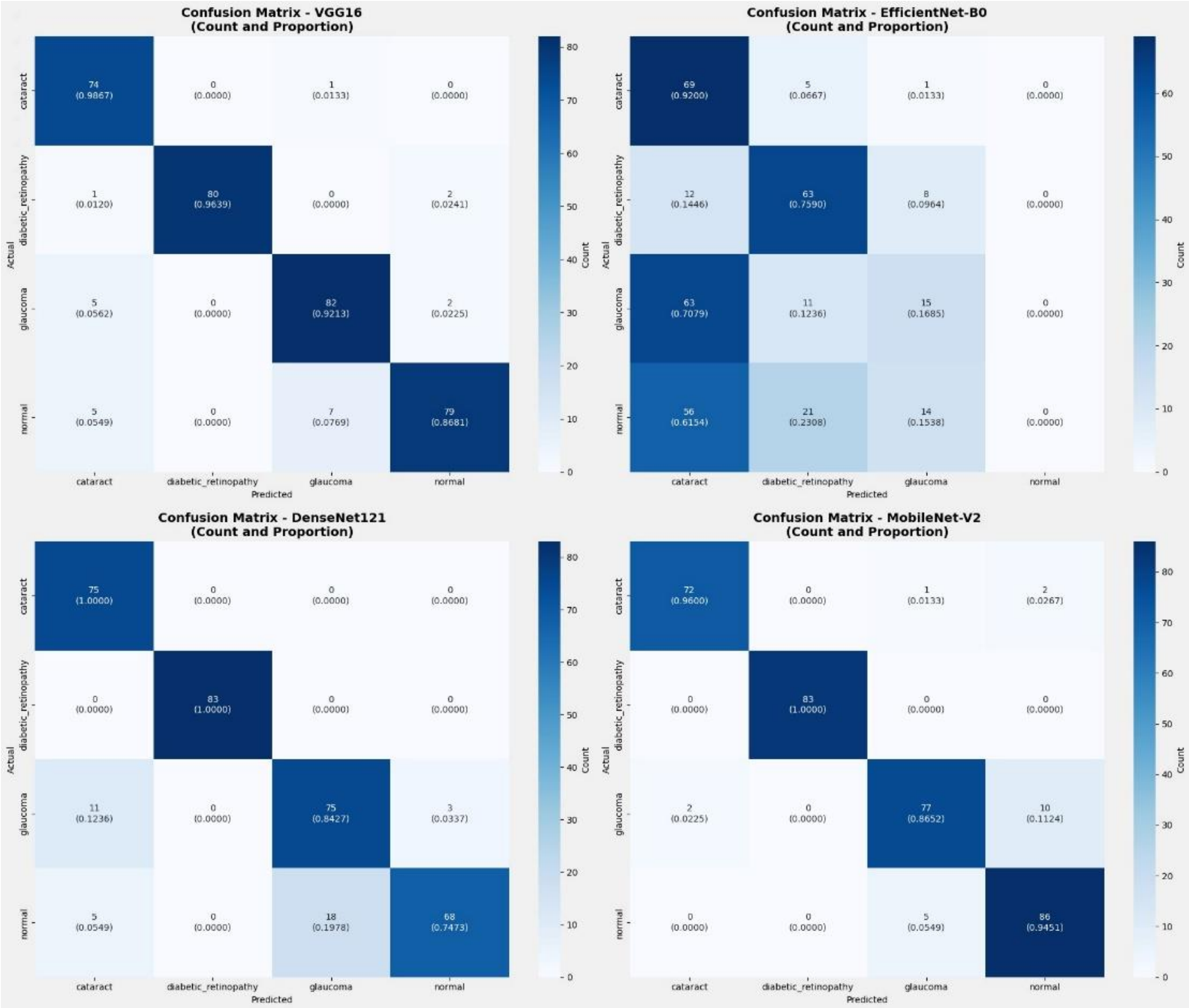

Figure 3. Confusion matrix showing the classification performance of DL models-

Conclusion

The culmination of this research emphasizes the critical role of automated diagnostic systems in addressing the global burden of retinal pathologies, with a primary focus on the classification of Glaucoma, Cataracts, and Diabetic Retinopathy against Normal retinal scans. Through the implementation and rigorous evaluation of several state-of-the-art pre-trained DL architectures—specifically MobileNetV2, VGG16, DenseNet121, and EfficientNetB0—this study identifies distinct performance trade-offs essential for clinical deployment.

MobileNetV2 emerges as the most transformative architecture within this comparative framework, achieving a superior balance between precision and computational demand. With a peak accuracy of 94.08% and an efficiency score of 0.4157, this model demonstrates exceptional operational viability. Such metrics indicate that MobileNetV2 can deliver expert-level diagnostic insights with sub-second latency, making it exclusively suitable for incorporation into resource-constrained mobile devices and remote diagnostic kits. While VGG16 demonstrated a slightly lower accuracy of 93.20%, it achieved the fastest inference time at 6.3110 ms, yielding a robust F1-score of 0.9327. In contrast, DenseNet121 and EfficientNetB0 were found to be less effective in this specific context, with DenseNet121 yielding an accuracy of 89.05% and a corresponding F1-score of 0.8918.

Ultimately, these findings reaffirm that Convolutional Neural Networks (CNNs) remain the most potent and efficient methodology for the automated detection of retinal diseases. The success of the MobileNetV2 framework suggests a clear pathway for deploying high-fidelity screening tools in underserved regions where specialized ophthalmological care is scarce. Future research trajectories should explore the integration of multi-modal data, combining fundus imagery with longitudinal patient history and metabolic profiles. Such a holistic approach promises to further refine diagnostic granularity and deliver more comprehensive provision for clinical decision-making in the fight against preventable blindness.

References

1. Gopalbhai, Patel Vandanabahen, and Jayesh N. Modi. "A Novel approach to multi-level image classification of Retinal Diseases." Journal of Neonatal Surgery 14, no. 4s (2025).
2. JBhulakshmi, Dasari, and Dharmendra Singh Rajput. "A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images." PeerJ Computer Science 10 (2024): e1947.
3. Abdullah, Ahmed Aizaldeen, Ahmed Aldhahab, and Hanaa M. Al Abboodi. "Review of eye diseases detection and classification using deep learning techniques." In BIO Web of Conferences, vol. 97, p. 00012. EDP Sciences, 2024.
4. Ahmed, Faisal. "Hog-cnn: Integrating histogram of oriented gradients with convolutional neural networks for retinal image classification." arXiv preprint arXiv:2507.22274 (2025).
5. Sureja, Nitesh, Vruti Parikh, Ajaysinh Rathod, Priya Patel, Hemant Patel, and Heli Sureja. "Explainable Artificial Intelligence based deep learning for Retinal Disease Detection." Journal of Electronics, Electromedical Engineering, and Medical Informatics 7, no. 2 (2025): 471-483.
6. Albelaihi, Arwa, and Dina M. Ibrahim. "DeepDiabetic: an identification system of diabetic eye diseases using deep neural networks." IEEE Access 12 (2024): 10769-10789.
7. Manhas, Jatinder, Rachit Kumar Gupta, and Partha Pratim Roy. "A review on automated cancer detection in medical images using machine learning and deep learning based computational techniques: Challenges and opportunities." Archives of Computational Methods in Engineering 29, no. 5 (2022): 2893-2933.
8. Gulshan, Varun, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." jama 316, no. 22 (2016): 2402-2410.
9. Sarki, Rubina, Khandakar Ahmed, Hua Wang, and Yanchun Zhang. "Automated detection of mild and multi-class diabetic eye diseases using deep learning." Health Information Science and Systems 8, no. 1 (2020): 32.

10. Ejaz, Sara, Raheel Baig, Zeeshan Ashraf, Mrim M. Alnfiai, Mona Mohammed Alnahari, and Reemiah Muneer Alotaibi. "A deep learning framework for the early detection of multi-retinal diseases." Plos one 19, no. 7 (2024): e0307317.
11. Qasim, Duaa Kareem, Sabah Abdulazeez Jebur, Lafta Raheem Ali, Abdul Jalil M. Khalaf, and Abir Jaafar Hussain. "Lightweight Convolutional Neural Networks for Retinal Disease Classification." arXiv preprint arXiv:2506.03186 (2025).
12. Ejaz, Sara, Hafiz U. Zia, Fiaz Majeed, Umair Shafique, Stefania Carvajal Altamiranda, Vivian Lipari, and Imran Ashraf. "Fundus image classification using feature concatenation for early diagnosis of retinal disease." Digital Health 11 (2025): 20552076251328120.
13. 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, Dec. 2017.
14. D. S. Kermany *et al.*, “Identifying Medical Diagnoses and Treatable Diseases by Image-Based deep learning,” *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018.
15. N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, “Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey,” *Artificial Intelligence in Medicine*, vol. 99, p. 101701, Aug. 2019.