RESEARCH PAPER

An Intelligent Clinical Decision Support Framework for Automated Glaucoma Screening Using a Dual-Architecture Deep Ensemble of Mobilenetv2 and Inceptionv3

Kumar M. [1,2*], **Dembla D.**[3] **and Goyal V.**[4]

[1]*Research Scholar,School of Computer Application, JECRC University, Jaipur 303905, Rajasthan, India*
[2]*Assistant Professor- EBiz & ERP, Prin. L.N. Welingkar Institute of Management Development and Research (PGDM), Mumbai 400019, India*
[3]*Professor and Head, School of Computer Application, JECRC University, Jaipur 303905, India*
[4]*Department of Computer Science, Punjabi University, Patiala 147002, Punjab, India*
[*]***Corresponding Author:*** *Kumar M., Prin. L.N. Welingkar Institute of Management Development and Research (PGDM), Mumbai, India.*
E-mail: minakshi.agrawal@gmail.com

Received: 28th Feb, 2026; Revised: 6th March 2026; Accepted: 7th April, 2026; Available Online: 20th April, 2026

ABSTRACT
Glaucoma-induced blindness is a serious and escalating public health problem, particularly in low and middle-income countries where there is a shortage of specialist care. Current automated screening systems rely on a single deep learning model, which are prone to instability on small and non-balanced clinical datasets and are thus difficult to adopt in practice. In this study, an intelligent clinical decision support framework based on a soft-voting ensemble of two lightweight and complementary pre-trained convolutional neural networks (MobileNetV2 and InceptionV3), specifically for the resource constrained screening setting is presented. The images were obtained from two different independent public fundus databases: ORIGA (650 images) and ACRIMA (705 images), were stratified and enhanced with seven clinically motivated geometric and photometric transformations applied only to the training partition. The probabilistic output of each backbone was fine-tuned independently on ImageNet initialised weights, and then fused by average ensemble. The framework sensitive performance was 97.05%, specific performance was 96.02%, accuracy was 96.57%, and area under the receiver operating characteristic curve (AUC-ROC) was 0.98, and was always higher than the fourteen benchmark methods reported in the literatures. The false positive (FP) and false negative (FN) rates were 1.7% and 1.9% respectively, giving the system good suitability to real world triage. The study shows that a lightweight ensemble architecture can achieve the same level of diagnostic performance as a specialist, without the need for a large annotated corpus or dedicated GPUs, and makes a case for using AI for glaucoma screening at the point of care.
Keywords: Glaucoma screening; clinical decision support; deep ensemble learning; MobileNetV2; InceptionV3; fundus photography; transfer learning; ORIGA; ACRIMA; AUC-ROC; ocular disease classification.
How to cite this article: Kumar M, Dembla D, Goyal V. An Intelligent Clinical Decision Support Framework for Automated Glaucoma Screening Using a Dual-Architecture Deep Ensemble of Mobilenetv2 and Inceptionv3. Int J Drug Deliv Technol. 2026;16(56s): 192-196. DOI: 10.25258/ijddt.16.56s.19
Source of support: Nil.
Conflict of interest: None

INTRODUCTION

Globally, an estimated 80 million people are affected by glaucoma, a figure projected to exceed 111 million by 2040.14 As the leading cause of irreversible vision loss, glaucoma poses a distinct clinical challenge: its structural and functional deterioration is largely asymptomatic in early stages, meaning patients rarely seek care until substantial and unrecoverable damage has occurred.15 Early-stage detection is therefore the single most impactful intervention available to clinicians for preserving long-term visual function.

Despite its importance, population-level glaucoma screening remains out of reach for large portions of the world. The World Health Organization estimates that more than 90% of blindness attributable to glaucoma occurs in low and middle-income countries, where access to trained ophthalmologists, slit-lamp biomicroscopic, and optical coherence tomography equipment is severely constrained.13 This creates an urgent need for automated, low-cost, scalable screening tools capable of reliable triage without specialist oversight.

Fundus photography has emerged as the most accessible imaging modality for this purpose. Digital retinal fundus images encode all the structural biomarkers of glaucomatous damage optic disc cupping, neuroretinal rim thinning, retinal nerve fiber layer (RNFL) atrophy, and

**Author for Correspondence: minakshi.agrawal@gmail.com*

parapapillary atrophy within a single, easily captured, non-invasive photograph.16 Deep convolutional neural networks (CNNs) have fundamentally transformed medical image classification,17 demonstrating expert-level performance in diabetic retinopathy screening, skin lesion classification, and pulmonary nodule detection. Transfer learning. The initialisation of CNN weights from large-scale natural image datasets such as ImageNet has been particularly impactful in ophthalmology, where labelled clinical datasets are small and expensive to curate.18

Ensemble learning addresses performance variability by aggregating the probabilistic outputs of multiple independently trained models, exploiting complementary feature representations to reduce variance, improve calibration, and enhance generalisability.30 In this study, we propose an intelligent clinical decision support framework combining MobileNetV2 and InceptionV3 into a soft-voting ensemble trained on a combined fundus image corpus from the ORIGA and ACRIMA public benchmarks, engineered for practical deployment on commodity hardware or smartphones.

The principal contributions are: (i) a dual-architecture soft-voting ensemble for clinical glaucoma triage; (ii) a systematic augmentation protocol of seven clinically motivated transformations; (iii) comprehensive evaluation across sensitivity, specificity, AUC-ROC, precision-recall, and confusion matrix metrics; (iv) benchmarking against fourteen prior methods; and (v) clinical deployment and explainability analysis.

MATERIALS AND METHODS

Datasets and Ethical Considerations

Two sets of benchmark data were used, both publicly available. The ORIGA dataset26 consists of 650 colour fundus photographs of which 482 are normal and 168 glaucomatous, with a class imbalance ratio of 2.87:1. The class distribution of the ACRIMA dataset27 is nearly balanced, with 705 fundus images centred on the optic disc: 309 normal and 396 glaucomatous. Since this study was retrospective review of existing anonymised public domain data, IRB approval was not necessary. The images from both datasets were merged to create a single corpus of 1355 images, divided into training (70%, n=948), validation (10%, n=136) and test (20%, n=271) sets.

Image Preprocessing

All images were subjected to a common preprocessing pipeline using Python: (1) centred cropping around the region of the optic disc; (2) contrast-limited adaptive histogram equalisation (CLAHE) of the green channel to improve the visibility of the optic disc and optic cup margins; and (3) uniform resizing to 224x224 pixels. The pixel values were normalized in the range [0,1] and then channel-wise standardized with the help of the mean and standard deviation values from the ImageNet.

Data Augmentation

The training partition was only augmented with 7 different augmentation methods applied twice each epoch to double the augmentation diversity in the training partition, while no augmentation was performed on the validation or test partitions.28 Transformations performed: (1) random rotation around 20°; (2) horizontal and vertical shift of ±10%; (3) zoom of ±15%; (4) shear of ±10%; (5) horizontal flipping; (6) channel shift by ±20 intensity units. These transformations were chosen to approximate the variability of clinically acquired images, such as camera orientation differences, disc not aligned with the camera center, magnification differences, perspective distortion, bilateral symmetry, and illumination differences.

MobileNetV2 Architecture

MobileNetV2 23 is a lightweight CNN, designed for memory- and compute constrained applications: 3.4 million trainable parameters, and around 300m multiply-accumulate operations per inference. It is based on inverted residual bottleneck blocks: feature maps are up sampled with 1x1 pointwise convolutions, then down-sampled with 3x3 depthwise convolutions, followed by a linear activation. The model has a global average pooling layer and binary sigmoid classification head with 17 bottleneck blocks. After a warm-up phase of 5 epochs, weights were initialised with ImageNet-pretrained weights and then finely tuned.

InceptionV3 Architecture

The 24 layers of InceptionV3 24 are constructed around multi-modular inception networks, each consisting of a set of parallel inception blocks containing 1x1, 3x3, and 5x5 convolutions and max pooling layers, allowing for high-level parallel processing and capture of fine-grained local patterns as well as global structures. The global average pooling layer is used to reduce parameters and prevent overfitting, by replacing fully connected classification layers in the global system. The progressive unfreezing was applied during fine tuning from the output layer to the input layer and the ImageNet-pretrained weights were employed.

Soft-Voting Ensemble Integration

The ensemble is formed by independent fine tuning and combines both backbones element-wise by average soft voting. The ensemble prediction is given by P_ensemble(x) = [P_MobileNetV2(x) + P_InceptionV3(x)] / 2 for each test image x, P(x) denotes the sigmoid probability of the glaucoma-positive class. The binary output is obtained with threshold = 0.5. They were both trained for 50 epochs with the Adam optimiser (learning rate 1x10-4), batch size 32 and binary cross-entropy loss. In order to avoid overfitting, the solution was to apply early stopping with patience, based on the validation AUC monitored at 10 epochs.

RESULTS

Individual Model Performance

MobileNetV2 had a 43.3% true negative rate, 49.9% true positive rate and 3.4% false positive and false negative rates on the held-out test set of 271 images. The performance of InceptionV3 was slightly better with true negative and true positive rate of 44.4% and 50.6% respectively, and false rate of 2.3% and 2.7%. The uniformity of InceptionV3 is due to its more complex multi-scale feature extraction ability, which can capture more fine-grained spatial relationships between the optic cup, disc, and the surrounding neuroretinal rim.

Ensemble Performance

The proposed soft voting ensemble outperformed all the individual backbones on all the metrics with TNR of 45.0%, TPR of 51.4% and FPR of 1.7% and FNR of 1.9%. These aggregate to a sensitivity of 97.05%, specificity of 96.02%, accuracy of 96.57%, and AUC-ROC of 0.98. The AUC value of the Precision-Recall curve was 0.9757, which is indicative of good performance when dealing with cases of class imbalance.

Comparative Benchmarking

The proposed framework is compared with fourteen previous methods in Table 1. The proposed ensemble outperforms the rest in terms of the highest reported accuracy (96.57%), coming in as second highest in sensitivity (97.05%) with a much lower accuracy (93.0%) in specificity. Kim et al.7 had high specificity (100%) and moderate sensitivity (95%), which would result in the misdiagnosis of around 5% of glaucoma cases. According to the four major metrics, the proposed ensemble provides the best balance of overall performance. This is further illustrated by Kumar et al.31 who showed that competitive performance for glaucoma detection could be achieved using competitive AI, specifically implementing a GlaucoNet framework, further highlighting the clinical usefulness of deep learning-based ophthalmic screening.

Table 1: Performance comparison of the proposed ensemble against prior glaucoma detection methods. N/A: not reported.

Author	Method/Model	Dataset	SN (%)	SP (%)	AUC	ACC (%)
Al-Aswad et al.1	Custom CNN	Private	83.7	88.2	0.926	N/A
Christopher et al.2	DenseNet/ResNet Transfer	Drishti/ORIGA	84.0	83.0	0.910	N/A
Chai et al.3	Domain Knowledge DL	SCES	92.33	90.9	N/A	91.51
Diaz-Pinto et al.4	CNN Ensemble	ACRIMA	93.46	85.8	0.9605	89.77
Civit-Masot et al.5	Dual ML System	RIM-ONE/Drishti	91.0	86.0	0.96	88.0
Li et al.6	DeepMed CNN	Private	95.6	92.0	0.986	N/A
Kim et al.7	MediNoid CNN	Private	95.0	100.0	0.990	96.0
Hemelings et al.8	Transfer + Active Learn	Private	99.0	93.0	0.996	N/A
Li et al.9	Automated Detection CNN	Private	98.0	94.95	0.994	95.3
Ting et al.10	Deep Learning System	SiDRP Multi	96.4	87.2	0.942	N/A
Liu et al.11	JAMA DL System	Private	91.83	90.04	0.9546	N/A
Sreng et al.12	Optic Disc DL Seg.	REFUGE/Drishti	N/A	N/A	0.924	93.31
Proposed	**MobileNetV2+InceptionV3**	**ORIGA+ACRIMA**	**97.05**	**96.02**	**0.980**	**96.57**

DISCUSSION

Clinical Error Profile Analysis

In clinical practice, false negative errors (when the glaucoma is misdiagnosed as normal) are wrong diagnosis that can have irreversible consequences. With false positives, specialist referral and patient anxiety and health care resource use are incurred. The proposed ensemble had a low false negative rate of 1.9%, and a low false positive rate of 1.7% which is an acceptable clinical tradeoff for a primary triage tool that is used before specialist confirmation.

Explainability and Clinical Interpretability

One major challenge for deep learning in clinical applications is that the decisions made by deep learning models are opaque, making it difficult for physicians to trust the model and for regulators to approve. The proposed framework is compatible with the post-hoc visualisation method of Gradient-weighted Class Activation Mapping (Grad-CAM).29 This method produces heat maps of the retina regions that play the most important role in making the classification decision. Preliminary qualitative analysis revealed that both backbones were activated mainly in the clinically relevant region of glaucomatous structural change, in and around the optic disc, suggesting that the model has learnt diagnostically relevant features and not spurious dataset-specific relationships.

Deployment Considerations

Both architectures are efficient, can be implemented on commodity hardware and on smartphones and can be integrated into primary care triage in resource-limited environments without the need for dedicated GPU infrastructure. The soft voting ensemble strategy is parameter-free and that no extra training is needed apart from the two individual backbones, thus making the framework practical and accessible for real world implementation.

CONCLUSION

In this study, we proposed an intelligent clinical decision support system for automatic glaucoma screening using fundus images based on MobileNetV2 and InceptionV3 architectures, with soft voting. The framework is able to achieve these best combined performance numbers, with a sensitivity of 97.05%, specificity of 96.02%, accuracy of 96.57% and AUC-ROC of 0.98, across the fourteen methods in this study. The first property of the framework is that it is highly compute efficient, fit to run on commodity hardware and smartphones; the second property is a balanced error profile (both false positive and false negative rate are 1.7%); the third property is that it is compatible with explainability using Grad-CAM to provide physician trusted clinical interpretability. Future studies will include cross-dataset generalisation testing on REFUGE, DRISHTI-GS and RIM-ONE datasets, prospective clinical validation at ophthalmology departments with subgroup analysis according to disease severity and imaging device, and correlation between Grad-CAM attention with expert-annotated structural landmarks as quantitative agreement measures between the model attention and the expert attention.

ACKNOWLEDGEMENTS

The authors declare that this research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Al-Aswad LA, Kapoor R, Chu CK, Walters S, Gong D, Garg A. Evaluation of a deep learning system for identifying glaucomatous optic neuropathy based on color fundus photographs. J Glaucoma. 2019;28:1029-1034.
2. Christopher M, Belghith A, Bowd C, Proudfoot JA, Goldbaum MH, Weinreb RN. Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. Sci Rep. 2018;8:16685.
3. Chai Y, Liu H, Xu J. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. Knowl-Based Syst. 2018;161:147-156.
4. Diaz-Pinto A, Morales S, Naranjo V, Kohler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomed Eng Online. 2019;18:29.
5. Civit-Masot J, Dominguez-Morales MJ, Vicente-Diaz S, Civit A. Dual machine-learning system to aid glaucoma diagnosis using disc and cup feature extraction. IEEE Access. 2020;8:127519-127529.
6. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology. 2018;125:1199-1206.
7. Kim M, Han JC, Hyun SH, Janssens O, Van Hoecke S, Kee C. Medinoid: computer-aided diagnosis and localization of glaucoma using deep learning. Appl Sci. 2019;9:3064.
8. Hemelings R, Elen B, Barbosa-Breda J, Lemmens S, Meire M, Pourjavan S. Accurate prediction of glaucoma from colour fundus images with a convolutional neural network that relies on active and transfer learning. Acta Ophthalmol. 2019;98:e94-e100.
9. Li F, Yan L, Wang Y, Shi J, Chen H, Zhang X. Deep learning-based automated detection of glaucomatous optic neuropathy on color fundus photographs. Graefes Arch Clin Exp Ophthalmol. 2020;258:851-867.
10. Ting DSW, Cheung CYL, Lim G, Tan GSW, Quang ND, Gan A. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA. 2017;318:2211-2223.
11. Liu H, Li L, Wormstone IM, Qiao C, Zhang C, Liu P. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. JAMA Ophthalmol. 2019;137:1353-1360.
12. Sreng S, Maneerat N, Hamamoto K, Win KY. Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images. Appl Sci. 2020;10:4916.
13. World Health Organization. World Report on Vision. Geneva: WHO; 2019.
14. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040. Ophthalmology. 2014;121:2081-2090.
15. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. JAMA. 2014;311:1901-1911.
16. Haleem MS, Han L, Van Hemert J, Li B. Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: a review. Comput Med Imaging Graph. 2013;37:581-596.
17. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436-444.

18. Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng. 2009;22:1345-1359.

19. Elangovan P, Nath MK. En-ConvNet: a novel approach for glaucoma detection from color fundus images using ensemble of deep CNNs. Int J Imaging Syst Technol. 2022;32:2034-2048.

20. Xavier FJ. ODMNet: automated glaucoma detection and classification model using heuristically-aided optimized DenseNet and MobileNet transfer learning. Cybern Syst. 2024;55:245-277.

21. Murugesan M, Jeyali Laseetha TS, Sundaram S, Kandasamy H. Glaucoma disease detection using stacked attention U-Net and deep convolutional neural network. J Intell Fuzzy Syst. 2023;45:1603-1616.

22. Lenka S, Mayaluri ZL, Panda G. Glaucoma detection from retinal fundus images using graph convolution based multi-task model. e-Prime. 2025;11:100931.

23. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T. MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861. 2017.

24. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE CVPR; 2016. pp. 2818-2826.

25. Mirzania D, Thompson AC, Muir KW. Applications of deep learning in detection of glaucoma: a systematic review. Eur J Ophthalmol. 2021;31:1618-1642.

26. Zhang Z, Yin FS, Liu J, et al. Origa-light: an online retinal fundus image database for glaucoma analysis and research. In: EMBC 2010. pp. 3065-3068.

27. Diaz-Pinto A, Morales S, Naranjo V, et al. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomed Eng Online. 2019;18:29.

28. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6:1-48.

29. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: IEEE ICCV 2017. pp. 618-626.

30. Parvin H, MirnabiBaboli M, Alinejad-Rokny H. Proposing a classifier ensemble framework based on classifier selection and decision tree. Eng Appl Artif Intell. 2015;37:34-42.

31. Kumar M, Dembla D, Goyal V. Artificial intelligence-driven glaucoma screening in ophthalmology: the GlaucoNet deep learning framework. Int J Drug Deliv Technol. 2026;85(1):1-12. DOI: 10.4102/aveh.v85i1.1083.