Baghdad Science Journal

Volume 22 | Issue 8 Article 29

8-27-2025

An Efficient Glaucoma Classification for Fundus Images by Pre-trained Deep CNNs and Accuracy-Weighted Voting Classifier Fusion

Salam Jabbar Edan
Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran, salam.jabbar@tabrizu.ac.ir

Mohammad-Reza Feizi-Derakhshi
Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran AND College of Engineering, Uruk University, Baghdad, Iraq, mfeizi@tabrizu.ac.ir

How to Cite this Article

Edan, Salam Jabbar and Feizi-Derakhshi, Mohammad-Reza (2025) "An Efficient Glaucoma Classification for Fundus Images by Pre-trained Deep CNNs and Accuracy-Weighted Voting Classifier Fusion," *Baghdad Science Journal*: Vol. 22: Iss. 8, Article 29.
DOI: https://doi.org/10.21123/2411-7986.5040

BAGHDAD SCIENCE JOURNAL 2025;22(8):2813–2823

Scan the QR to view the full-text article on the journal website

RESEARCH ARTICLE

An Efficient Glaucoma Classification for Fundus Images by Pre-Trained Deep CNNs and Accuracy-Weighted Voting Classifier Fusion

Salam Jabbar Edan [1], **Mohammad-Reza Feizi-Derakhshi** [1,2,*]

[1] Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
[2] College of Engineering, Uruk University, Baghdad, Iraq

ABSTRACT

Glaucoma, a chronic eye disease characterized by progressive optic nerve damage, remains a significant public health concern worldwide. Early and accurate diagnosis is crucial for timely intervention and preventing irreversible vision loss. In this paper, developed an approach for glaucoma classification utilizing pre-trained convolutional neural networks (CNNs), specifically ResNet-101, ResNet-18, and DarkNet-19. Also employed accuracy-based weighted voting as a classifier fusion technique. Leveraging the capabilities of these CNN architectures, extract high-level features from retinal fundus images and apply transfer learning to adapt the networks for glaucoma classification. Leveraging the capabilities of these CNN architectures, extract high-level features from retinal fundus images and apply transfer learning to adapt the networks for glaucoma classification. To enhance classification performance, a fusion strategy based on accuracy-weighted voting is introduced, where the contribution of each classifier is weighted by its individual accuracy on the validation set. Experimental evaluations conducted on three glaucoma datasets—ACRIMA, RIMONE-v2, and Drishti-GS—using an 8:2 ratio for training and testing yielded promising results: 99.3% accuracy for the ACRIMA dataset, 95.6% accuracy for the RIMONE-v2 dataset, and 90% accuracy for the Drishti-GS dataset. These results confirm the efficiency of the approach proposed in the current study, particularly with regard to the accuracy of glaucoma diagnosis across different datasets. By enabling intervention after early detection, this approach could have a significant impact on glaucoma management, which in turn improves patient outcomes.

Keywords: Classifier fusion, Convolutional neural network, Deep learning, Glaucoma classification, Transfer learning

Introduction

Glaucoma is a widespread eye condition caused by optic nerve damage that links the brain with the eye. Chronic vision loss could be the result if it is left untreated. Worldwide, it ranks as the second most prominent cause of blindness. It is often called the “silent thief of sight” due to its undetectable nature in the early stages. [1]

In recent years, there has been a rising number of glaucoma patients reported globally, as indicated by the World Health Organization (WHO). [2] It is anticipated that the prevalence of glaucoma will increase from 64 million to 76 million by 2020 and reach 111.8 million by 2040. Africa and Asia are likely to experience a more significant impact compared to other regions due to a shortage of trained ophthalmologists for diagnosis. [3,4] Therefore, this illness is considered as an important public health issue, and early detection is essential to prevent blindness.

Presently, there is a significant advancement in Computer-Aided Diagnosis systems based on medical images, aimed at providing doctors with secondary diagnoses or medical decisions support. [5] Therefore, authors have produced extensive efforts to achieve remarkable and robust throughputs concerning

Received 23 May 2024; revised 24 September 2024; accepted 16 September 2024.
Available online 27 August 2025

* Corresponding author.
E-mail addresses: salam.jabbar@tabrizu.ac.ir (S. J. Edan), mfeizi@tabrizu.ac.ir (M.-R. Feizi-Derakhshi).

https://doi.org/10.21123/2411-7986.5040

image recognition, object detection, and classification. The efficiency of the proposed solutions is affected by diverse important factors such as quality solutions, computational efficiency, feature complexity, methods or techniques used, and domain-specific constraints. Recently, Deep learning has played a significant role in the extraction features mainly through its CNN technique which extracts features from image pixels by a convolutional operator. CNN is considered highly effective for image detection and recognition due to its need for minimum rate for preprocessing compared to competitors' algorithms. The CNN architecture has several layers to process the images input such as convolutional, pooling, activation, and FC layers on the other side, it includes various layers in the training and testing processes for image classification.[6] Additionally, clustering techniques have become increasingly important, aiding in the identification of patterns and the organization of data into meaningful groups, enhancing the diagnostic process and improving the accuracy of automated systems.[7]

In the medical diagnostics sector, deep CNNs have been deemed robust tools for analyzing image which process complex features from images automatically to make it a powerful pattern recognition approach that points out to a disease.[8] To overcome limitations arising from a scarcity of input samples and computational resources for running deep learning methods, transfer learning by pre-trained CNNs such that ResNet-101, ResNet-18, and DarkNet-19 are frequently utilized in developing of deep learning frameworks. This technique leverages the weights and parameters learned from prior large labeled datasets and adapts them to the new task.[2]

On the other side, several attempts' studies have been conducted to find a solution for such issue by training different classifiers such as classifier fusion to avoid the limited capability for individual-optimized deep CNN which is not adopted for various datasets. The fusion classifier produced a promising solution in the individual-optimized deep CNN using a set of diverse deep CNNs rather than a single architecture. Thus, presenting a fusion approach for different deep CNN models is needed to gain the capability of single classification.[2] Various studies have indicated that traditional classification methods are unsatisfactory due to the challenges associated with the feature extraction process.[9] Therefore, the main objective of this study is to develop a method based on pre-trained CNNs and accuracy-based weighted voting as a classifier fusion technique for better classifies fundus images into two groups: "Glaucoma" or "Non-Glaucoma". Three different pre-trained deep CNNs models such ResNet-101, ResNet-18, and DarkNet-19 are fused in parallel. the output scores are optimized using accuracy-based weighted voting as a classifier to find the best deep CNNs combination. Also, this proposed method achieved high accuracy compared with the other studies in glaucoma classification.

Related work

In reference,[3,4] there is an extensive review of the deep learning techniques used for detecting glaucoma through retinal fundus images. This part presents some of these techniques:

Velpula and Sharma,[10] developed a stage classification model based on automated glaucoma using CNN (convolutional neural network) pre-trained deep models and classifier fusion. They devised classifier fusion to combine the outputs of all CNN models through a maximum voting-based method.

Velpula and Sharma[11] suggested a model for automatically detecting glaucoma in fundus images, employing three deep CNNs (convolutional neural networks): ResNet101, NasNet-Large, and NasNet-Mobile. They evaluated the model using five publicly accessible fundus image datasets: RIMONE-v2, ACRIMA, FTVD, Drishti-GS, and the HVD (Harvard Dataset). Akbar et al.[12] introduced a new approach combining DarkNet and DenseNet to differentiate between glaucoma-affected and standard fundus images. The proposed approach was trained and evaluated using three datasets: RIM 1, high-resolution fundus (HRF), and ACRIMA. To classify healthy eyes, 658 images were implemented, and 612 images were processed for glaucoma-affected eyes.

Kumar and Gupta[13] produced several of models of transfer learning that facilitate eye disease prediction. In this study, various approaches were implemented for deep transfer learning, including AlexNet 2, Basic CNN, deep CNN, Inception V3, Xception, DenseNet121, and ResNet 50. As confirmed by the simulation results, ResNet50 demonstrated 98.9% validation accuracy, out-functioning all other methods. Moreover, the Xception performed significantly, recording 98.4% accuracy. A precise and robust algorithm was proposed by Ajitha et al.[14] They used a CNN (convolutional neural network) to function as an automatic diagnosis of glaucoma. They processed 1113 fundus images, 660 of them were regular and the rest were glaucomatous images from four databases, which were used for the diagnosis. To extract vital features, A 13-layer convolutional neural network was trained effectively on the collected dataset. During testing, they were subsequently classified into either normal or glaucomatous classes.

In addition to dataset technique which was introduced by Ovreiu et al.[15] This set contained fundus

color images of early-stage glaucoma. In the study, a pre-trained ResNet50 network was employed dealt with the ImageNet dataset. The findings suggest that using deep learning algorithms could enable the development of a cost-efficient screening tool for the early and economical detection of glaucoma. Joshi et al.[16] proposed an automated and rapid method based on deep learning for detecting glaucoma. Fundus images of varying resolutions from three-eye datasets are augmented by rotation and scaling at various degrees. The automated configuration can handle fundus image processing in 0.2 seconds. The method was evaluated on 1220 fundus images and demonstrated an accuracy average of 93.698%.

Materials and methods

Datasets

Three public datasets (ACRIMA, RIM-ONE v2, and Drishti-GS) were utilized for training and testing various deep convolutional neural networks (CNNs). Details of these datasets are provided in Table 1. Also, Fig. 1 displays a selection of sample images from the datasets utilized in this study, giving visual examples of the fundus images used for glaucoma detection. The ACRIMA dataset contains 705 fundus images, comprising 396 images showing signs of glaucoma and 309 images depicting normal conditions. Each image in the ACRIMA database underwent annotation by two glaucoma specialists, each possessing eight years of experience. These images were sourced from patients with glaucoma and normal conditions, with their prior consent obtained in compliance with the ethical principles established in the 1964 Declaration of Helsinki. The majority of fundus images in this dataset were captured from eyes that had been dilated beforehand and were focused on the optic disc. Additional clinical information was not considered when labeling the images.[17]

The RIM-ONE v2 dataset was released by the Miguel Servet and Canaria University Hospitals and the San Carlos Clinic, which contains a total of 455 fundus images, comprising 200 images showing signs of glaucoma and 255 images depicting normal conditions that underwent manual segmentation by a medical specialist.[18]

The Drishti-GS dataset was initially assembled at the Aravind Eye Hospital in Madurai and contains 101 fundus images, comprising 70 images showing signs of glaucoma and 31 images depicting normal conditions. The ground truth data was acquired from four experts with clinical experience spanning 3, 5, 9, and 20 years.[19]

Table 1. Datasets details.

Dataset	Normal	Glaucoma	Total
ACRIMA[17]	309	396	705
RIM-ONE v2[18]	255	200	455
Drishti-GS[19]	31	70	101

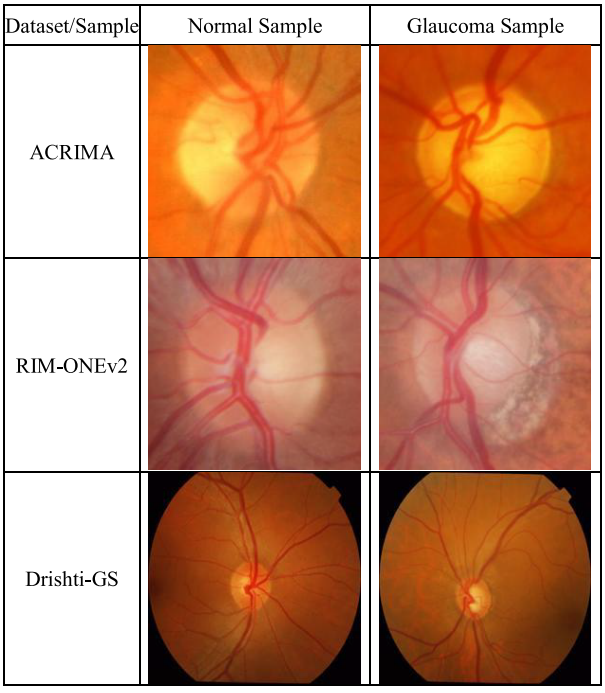

Fig. 1. Sample fundus images from the datasets employed in this study.

Proposed methodology

The automated diagnosis of glaucoma using fundus images, as depicted in Fig. 2, comprises four sequential steps: Data collection and preprocessing, Data augmentation, Deep transfer learning, and Classifier fusion. The following subsections offer detailed explanations for each of these steps:

1) **Data Collection and Preprocessing:** Three distinct collections of retinal images are assembled from multiple sources, as discussed in subsection (Datasets) of this section. In the preprocessing stage, adjustments are made to the dimensions of the datasets to align with the input requirements of the CNN models. Each dataset is resized to match the input size required by the respective CNN models. For instance, the input image sizes for ResNet-101, ResNet-18, and Darknet-19 CNN models are 224×224, 224×224, and 256×256, respectively. Subsequently, each dataset is randomly split into training and testing sets at a ratio 8:2 before training and testing the CNNs.

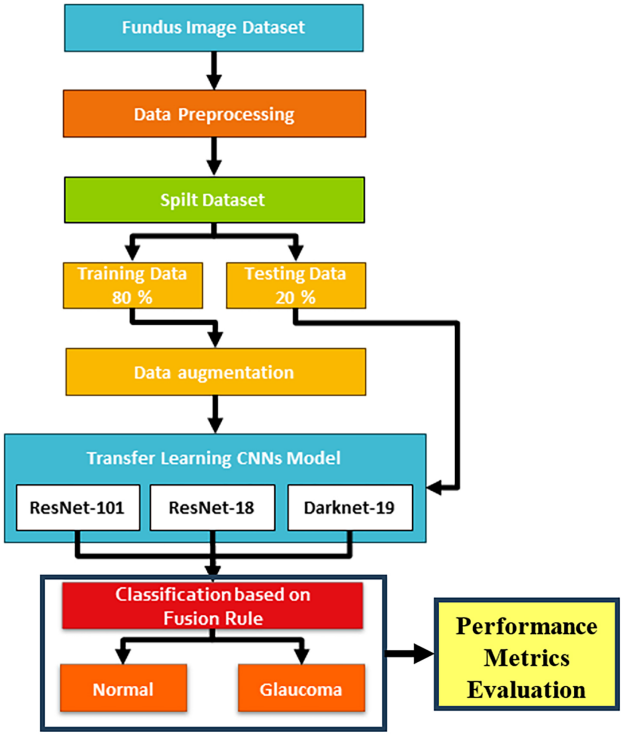

Fig. 2. The proposed model for glaucoma classification.

Table 2. Augmentation parameters used in training.

Function	Value
Rand Scale	[0.5 1]
Rand Rotation	[−5 to 5]
Rand X Reflection	1
Rand Y Reflection	1
Rand X Translation	[−3 to 3]
Rand Y Translation	[−3 to 3]

2) **Data Augmentation:** Data augmentation involves artificially increasing the size of a dataset by introducing slightly modified or transformed copies of existing data. By applying data augmentation to the training data, enhance the robustness and generalization capability of the model, thereby mitigating the risk of overfitting.[20,21] The augmentation techniques utilized during training include scaling, rotation, reflection, and translation, as listed in Table 2. Sample augmented images from the ACRIMA dataset are depicted in Fig. 3.
3) **Deep Transfer Learning:** Convolutional Neural Networks (CNNs), a class of deep neural networks specifically designed for image recognition tasks, are pivotal in the transfer learning phase. Convolutional Neural Networks consist of several layers of pooling and convolutional processes. These processes enable the CNN to learn hierarchical

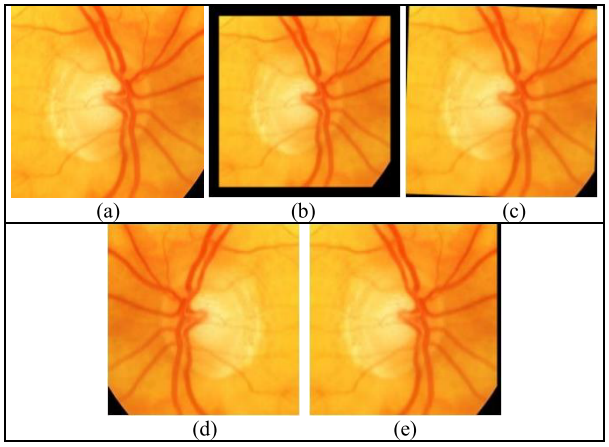

Fig. 3. Sample augmented images from the ACRIMA dataset: (a) Original image, (b) Scaled image with factor [0.5, 1], (c) Rotated image with angle [–5°, 5°], (d) Reflected image, (e) Translated image with displacement [–3, 3].

feature representations automatically from raw pixel data. These networks proved outstanding performance in various computer vision tasks such as segmentation, object detection, and classification.[22–24] Phase transfer learning aims to implement the proficiency of multiple CNNs, which were trained already on ImageNet or any other large-scale datasets, to use their feature extraction and generalization capabilities. By adjusting these pre-trained CNNs on specific glaucoma dataset such as the set proposed in the current study, they are enabled to adapt their learned features to the nuances of our target task. This approach does not only decrease the training time and computational cost, but also enhances the performance of the model by increasing the experience encoded in the trained CNNs. In the present study, three deep pre-trained CNNs are implemented on ImageNet to diagnose glaucoma using retinal fundus images: ResNet-18, Darknet-19, and ResNet-101. The following paragraphs illustrate the basic structure of each CNN.

ResNet-101 and ResNet-18[24–28] are convolutional neural networks, with ResNet-101 having 100 layers and ResNet-18 having 18 layers. These networks play a significant role in computer vision tasks such as image classification, object detection, and image segmentation. Initially, these networks were trained using over a million images from the ImageNet database, classifying images into 1000 object categories. The input to these networks is an RGB image of size 224-by-224. In this study, the ResNet-101 and ResNet-18 architectures were modified for glaucoma classification. After modification, the fully connected (FC) layer

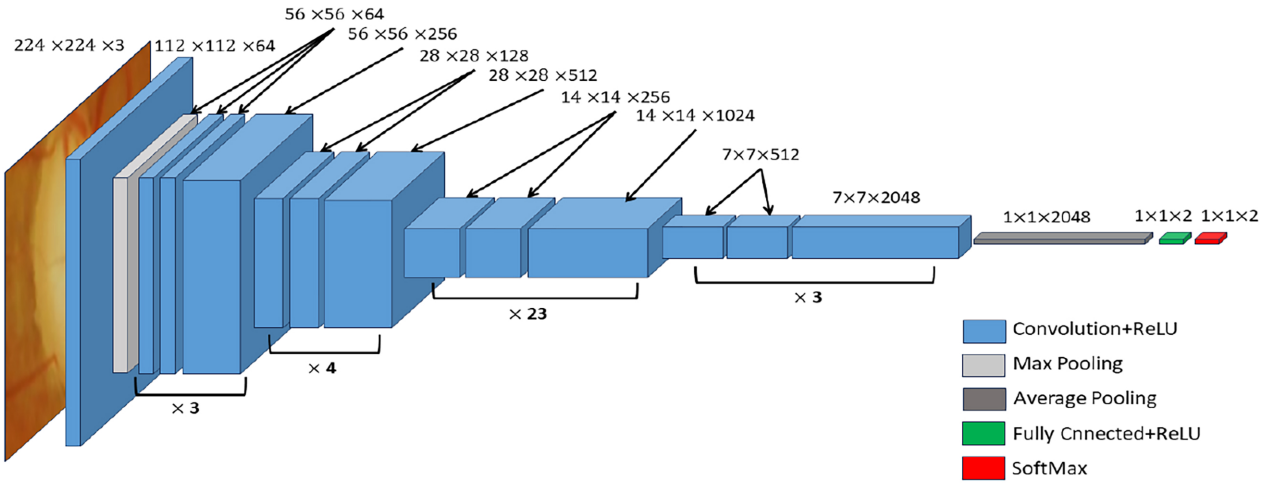

Fig. 4. The proposed architecture of ResNet-101 used for glaucoma classification.

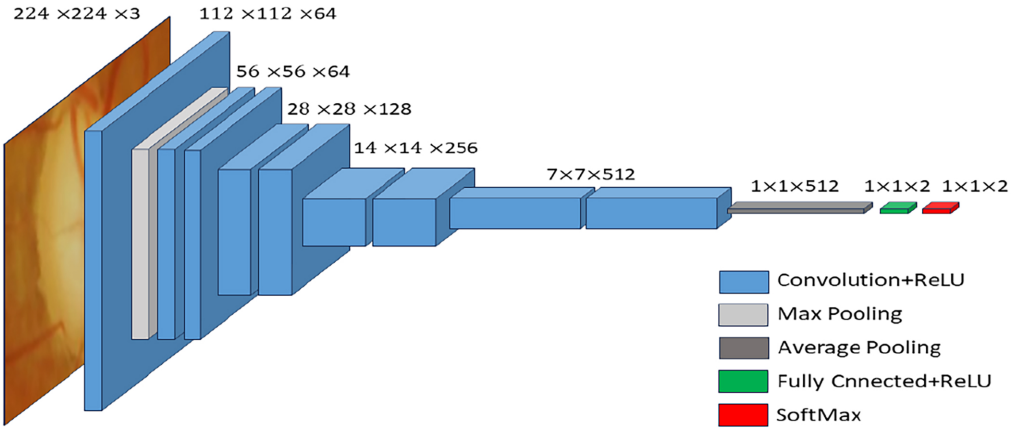

Fig. 5. The proposed architecture of ResNet-18 used for glaucoma classification.

was replaced with a new FC layer comprising only two classes. The architectures of the modified models are illustrated in Figs. 4 and 5. Darknet-19,[29–31] developed by Joseph Redmon, is a novel classification model acting as a backbone network for feature extraction. It comprises 19 convolutional layers and five max-pooling layers, requiring fewer operations for image processing while maintaining high accuracy. Similar to ResNet models, Darknet-19 was trained using over a million images from the ImageNet database, which classifies images into 1000 object categories.[32] In this study, the Darknet-19 architecture was modified for glaucoma classification. After modification, the FC layer was replaced with a new FC layer comprising only two classes. The architecture of the modified model is illustrated in Fig. 6.

4) **Classifier Fusion:** Classifier fusion involves combining different classifiers to address the same problem. In this study, all the mentioned architectures are integrated into a single classifier. Various forms of classifier fusion exist, differing in their structure and fusion operation type. For this study, the Accuracy-Based Weighted Voting (AWV) fusion method is employed. The steps of the AWV fusion method are as follows:

Step1: Calculates the Accuracy for the convolutional neural networks that namely ResNet-10, ResNet-18, and Darknet-19 respectively.

Step2: Calculate weights assigned to each convolutional neural networks model that are determined based on their individual accuracy or performance scores according to the Eq. (1):[33]

$$w_{i,j} = \frac{ACC_{i,j}}{\sum_{i=1}^{N} ACC_{i,j}} \quad (1)$$

Where $w_{i,j}$ denotes the weight assigned to the j_{th} class of the dataset in the i_{th} CNN classifier, $ACC_{i,j}$

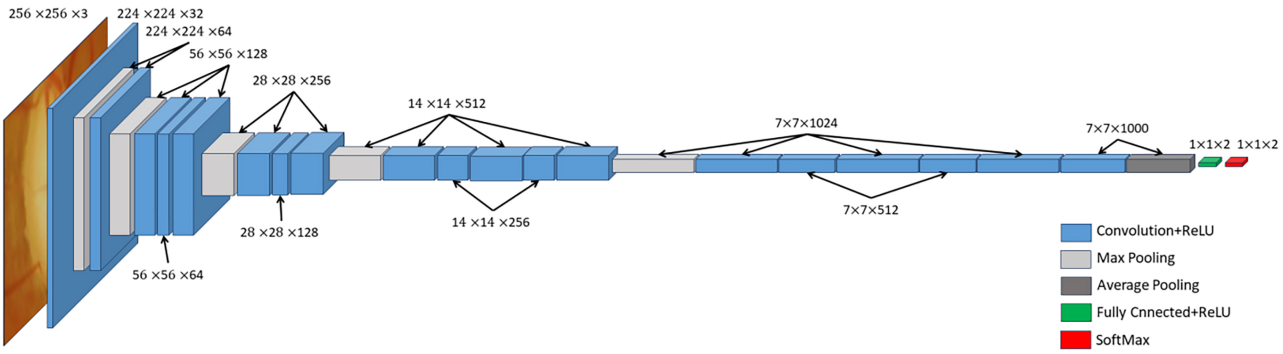

Fig. 6. The proposed architecture of Darknet-19 used for glaucoma classification.

denotes the accuracy of the j_{th} class in the i_{th} CNN classifier, and N is the number of CNN classifiers.

Step3: Calculate the AWV by the Eq. (2)[33] as follows:

$$AWV_{i,j} = \frac{\sum_{i=1}^{N} w_{i,j} * P_{i,j}}{N} \quad (2)$$

Where $P_{i,j}$ represents the probability computed by the j_{th} class in the i_{th} CNN classifier, and N is the number of CNN classifiers.

Step4: To determine the maximum accuracy values of the calculated AWSs and assigned labeling, the final decision can be calculated by Eq. (3) as follows:

$$Output = Label\ (max\ (AWV_{i,j})) \quad (3)$$

Where the max () function selects the highest value and Label () is defined to assign the class number corresponding to this highest value as the final prediction for the gesture class.

Results and discussion

The experiments were conducted in the MATLAB R2023a environment on a computer system equipped with an Intel Core i5 processor, 8 GB of RAM, and a processing speed of 3 GHz. During the training process, initialized epoch 50, used a mini-batch size 10, and executed the environment on the CPU. Additionally, employed the ADAM optimizer with a learning rate of 1e-5 (0.00001) for all networks. The hyperparameters were carefully tuned to optimize outcomes in our experiments.

The primary performance metrics for classification tasks typically include Accuracy, Sensitivity (Recall), Specificity, Precision, F-measure, and G-mean. In this study, all of these metrics were utilized to evaluate the performance of the proposed system. The calculation of these metrics can be done employing the Eqs. (4) to (9)[10,11,34]:

$$Accuracy\ (Acc) = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Sensitivity\ (Sen) = Recall\ (Rec) = \frac{TP}{TP + FN} \quad (5)$$

$$Specificity\ (Spe) = \frac{TN}{TN + FP} \quad (6)$$

$$Precision\ (Pre) = \frac{TP}{TP + FP} \quad (7)$$

$$F\text{-}measure\ (Fm) = \frac{2 \times (Pre) \times (Rec)}{(Pre + Rec)} \quad (8)$$

$$G\text{-}mean\ (Gm) = \sqrt{(Sen \times Spe)} \quad (9)$$

Here, TP, TN, FP, and FN represent the True Positive, True Negative, False Positive, and False Negative values, respectively. Tables 3 to 5 present the performance of the ACRIMA, RIM-ONE v2, and Drishti-GS datasets, respectively. Among the models, ResNet-101 consistently demonstrates the highest accuracy.

Table 3. Performance metrics for the ACRIMA dataset.

CNN model	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
Resnet-101	0.9858	0.9839	0.9873	0.9839	0.9839	0.9856
Resnet-18	0.9433	0.8971	0.9863	0.9839	0.9385	0.9406
Darknet-19	0.9716	1.0000	0.9518	0.9355	0.9667	0.9756
Fusion	**0.9929**	**1.0000**	**0.9875**	**0.9839**	**0.9919**	**0.9937**

Table 4. Performance metrics for the RIM-ONE v2 dataset.

CNN model	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
Resnet-101	0.9341	0.9412	0.9250	0.9412	0.9412	0.9331
Resnet-18	0.8352	0.7812	0.9630	0.9804	0.8696	0.8674
Darknet-19	0.9121	0.9057	0.9211	0.9412	0.9231	0.9133
Fusion	**0.9560**	**0.9434**	**0.9737**	**0.9804**	**0.9615**	**0.9584**

Table 5. Performance metrics for the Drishti-GS dataset.

CNN model	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
Resnet-101	0.8500	0.8000	0.8667	0.6667	0.7273	0.8327
Resnet-18	0.8000	0.6667	0.8571	0.6667	0.6667	0.7559
Darknet-19	0.8500	0.8000	0.8667	0.6667	0.7273	0.8327
Fusion	**0.9000**	**1.0000**	**0.8750**	**0.6667**	**0.8000**	**0.9354**

For the ACRIMA dataset, ResNet-101 achieves accuracies of 98.58%, 94.33%, and 97.16% for ResNet-18 and Darknet-19, respectively. Notably, employing a classifier fusion of all three models results in an accuracy of 99.29%. Similarly, for the RIM-ONE v2 dataset, the accuracy increases from 93.41% to 95.6% with classifier fusion. Likewise, for the Drishti-GS dataset, the accuracy improves from 85% to 90%.

Fig. 7 illustrates the confusion matrices for each dataset obtained through classifier fusion. In Fig. 7(a), the confusion matrix of the ACRIMA dataset shows a 100% accuracy rate for the "Normal" class and 98.8% for the "Glaucoma" class, resulting in an average accuracy of 99.3%. Fig. 7(b) presents the confusion matrix of the RIM-ONE v2 dataset, with a 94.3% accuracy rate for the "Normal" class and 97.4% for the "Glaucoma" class, resulting in an average accuracy of 95.6%. In Fig. 7(c), the confusion matrix of the Drishti-GS dataset displays a 75% accuracy rate for the "Normal" class and 100% for the "Glaucoma" class, resulting in an average accuracy of 90%.

Based on the results outlined previously, the proposed automated glaucoma diagnosis system demonstrates significant benefits and efficiency. Unlike conventional approaches that rely on manual feature extraction strategies, automatic glaucoma feature extraction is employed in this study, leading to superior classification accuracy across various publicly available datasets. By assessing six separate metrics, including Accuracy, Sensitivity, Specificity, Precision, F-measure, and G-mean, a comprehensive evaluation is ensured for the system's performance, demonstrating consistent performance in glaucomatous classification.

Furthermore, comprehensive experiments were conducted to assess the performance of three deep CNNs (ResNet-101, ResNet-18, and Darknet-19). Among these architectures, ResNet-101 consistently outperforms the others. It's worth noting that the Drishti-GS dataset presents a more significant challenge for each type of network. In contrast, the ACRIMA dataset yields satisfactory results, likely due to its suitability for deep learning classification tasks as a newly developed dataset.

Additionally, our proposed ensemble framework surpasses previously suggested methods, as demonstrated in Table 6. These findings could promote the processes of diagnosis and classification of glaucoma. The results emphasize the efficiency and accuracy of our approach to glaucoma diagnosis.

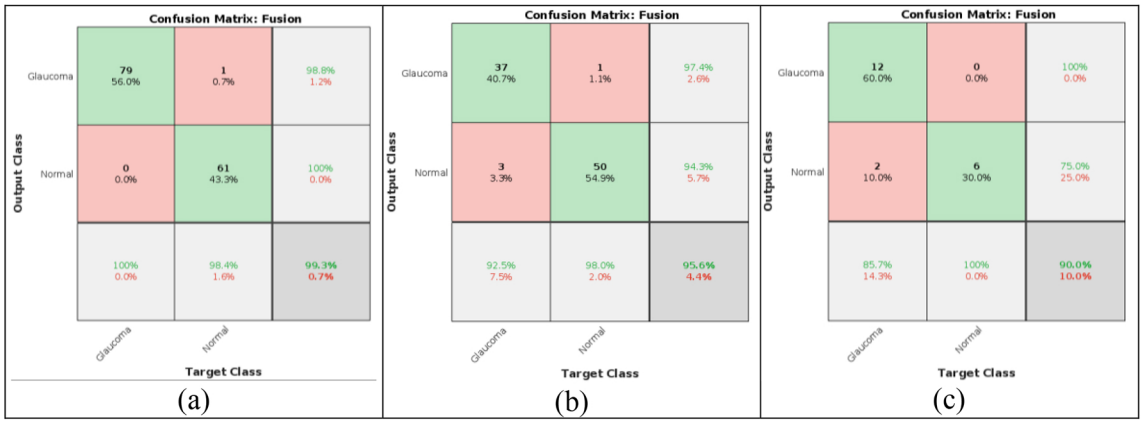

Fig. 7. Confusion matrices of all datasets by classifier fusion: (a) ACRIMA, (b) RIM-ONE v2, (c) Drishti-GS.

Table 6. Comparison of our proposed method with existing state-of-the-art methods.

Author	Year	Method	Dataset	Accuracy (Training and Test Data)	Accuracy (Test Data)
Velpula and Sharma[8]	2023	Transfer learning and fusion	ACRIMA Drishti-GS RIMONE HVD	99.57 90.55 94.95 85.43	–
Velpula and Sharma[9]	2023	Resnet-101 Nasnet_mobile Nasnet_large	ACRIMA RIMONE-v2 Drishti-GS FTVD HVD	99.43 92.75 82.18 78.15 90.23	–
Akbar et al.[10]	2022	Transfer learning and fusion	HRF RIM-1 ACRIMA	– 	99.7 89.3 99
Kumar and Gupta.[11]	2022	Transfer learning models	Private data	–	98.9
Ajitha et al.[12]	2021	Customized CNN	HRF, ORIGA, and Drishti-GS1	–	93.86
Ovreiu et al.[13]	2020	Resnet-50	ACRIMA, ORIGA and HRF	–	96.98
Joshi et al.[14]	2020	Darknet-53	Drishti GS, Refuge and Local Hospital	–	93.69
Our proposed method	**–**	**Pre-trained CNNs and Classifier Fusion**	**ACRIMA** **RIMONE-v2** **Drishti-GS**	**99.8** **99.12** **98.01**	**99.3** **95.6** **90**

Conclusion

To sum up, the efficiency of pre-trained CNNs were examined in the current study, namely, ResNet-18, Darknet-19, and ResNet-101 in the classification and diagnosis of glaucoma using the ACRIMA, Drishti-GS and RIM-ONE datasets. As a fusion technique, accuracy-based weighted voting was used and significant advancements in glaucoma classification accuracy were achieved. On the other hand, remarkable performance was shown by the experimental results, with accuracies of 90%, 95.6%, and 99.3% for Drishti-GS, RIM-ONE v2 and ACRIMA datasets, respectively. The results reveal the suggest promoting pre-trained CNNs and classifier fusion methodologies to improve glaucoma classification and diagnosis. In addition, to ensure the efficiency and generalizability of the findings of the present study across diverse clinical scenarios, multiple datasets can be used with an 8:2 standardized split ratio to verify safe training and testing. This paper supports the growing literature work that looks forward to enhancing the advanced techniques of machine learning in the field of ophthalmology. This will ultimately leads to improve the efficiency and accuracy of tools for early diagnosis and management of glaucoma. Future work is recommended to integrate other datasets and adjustment strategies to improve the classification performance further and increase the accuracy of the proposed approach in clinical utility.

Authors' declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images, that are not ours, have been included with the necessary permission for republication, which is attached to the manuscript.
- No animal studies are present in the manuscript.
- Ethical Clearance: The project was approved by the local ethical committee at University of Tabriz.

Authors' contribution statement

S.J. E. and M.R.F. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript. M.R.F. supervised the work.

Data availability

The datasets analyzed during the current study are publicly available in the following repositories: ACRIMA: Available at Figshare repository, https://figshare.com/s/c2d31f850af14c5b5232 RIM-ONE r2: Available at https://medimrg.webs.ull.es Drishti-GS: Available at https://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php.

References

1. Chayan TI, Islam A, Rahman E, Reza MT, Apon TS, Alam MGR. Explainable AI based glaucoma detection using transfer learning and LIME. Proc IEEE Asia Pac Conf Comput Sci Data Eng. 2022:1–6, Gold Coast, Australia. http://doi.org/10.1109/CSDE56538.2022.10089310.
2. Rehman AU, Taj IA, Sajid M, Karimov KS. An ensemble framework based on Deep CNNs architecture for glaucoma classification using fundus photography. Math Biosci Eng. 2021 Jun 16;18(5):5321–46. http://doi.org/10.3934/mbe.2021270.
3. Guergueb T, Akhloufi MA. A review of deep learning techniques for glaucoma detection. SN Comput Sci. 2023;4:274. http://doi.org/10.1007/s42979-023-01734-z.
4. Ashtari-Majlan M, Dehshibi MM, Masip D. Deep learning and computer vision for glaucoma detection: A review. arXiv preprint arXiv:2307.16528,2023. https://doi.org/10.48550/arXiv.2307.16528.
5. Kadhim KA, Mohamed F, Najjar FH, Salman GA. Early diagnose alzheimer's disease by convolution neural network-based histogram features extracting and canny edge. Baghdad Sci J. 2024;21(2(SI)):0643–0643. http://doi.org/10.21123/bsj.2024.9740.
6. Khan A, Pin K, Aziz A, Han JW, Nam Y. Optical coherence tomography image classification using hybrid deep learning and ant colony optimization. Sensors. 2023;23(15):6706. http://doi.org/10.3390/s23156706.
7. Bakhshi M, Feizi-Derakhshi MR, Zafarani E. Review and comparison between clustering algorithms with duplicate entities detection purpose. Int J Comp Sci Emerg Tech. 2012;3:108–14.
8. Da Silva CA, Miranda PB, Cordeiro FR. A new grammar for creating convolutional neural networks applied to medical image classification. IEEE 34th SIBGRAPI Conf Graph Patterns Images. 2021:97–104. Gramado, Rio Grande do Sul, Brazil. http://doi.org/10.1109/SIBGRAPI54419.2021.00022.
9. Shin Y, Balasingham I. Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification. IEEE Eng Med Biol Soc Annu Conf. 2017:3277–80, Jeju, Korea. http://doi.org/10.1109/EMBC.2017.8037556.
10. Velpula VK, Sharma LD. Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion. Front Physiol. 2023 Jun 13;14:1175881. http://doi.org/10.3389/fphys.2023.1175881.
11. Velpula VK, Sharma LD. Automatic glaucoma detection from fundus images using deep convolutional neural networks and exploring networks behaviour using visualization techniques. SN Comput Sci. 2023;4(5):487. http://doi.org/10.1007/s42979-023-01945-4.
12. Akbar S, Hassan SA, Shoukat A, Alyami J, Bahaj SA. Detection of microscopic glaucoma through fundus images using deep transfer learning approach. Microsc Res Tech. 2022;85(6):2259–76. http://doi.org/10.1002/jemt.24083.
13. Kumar Y, Gupta S. Deep transfer learning approaches to predict glaucoma, cataract, choroidal neovascularization, diabetic macular edema, drusen and healthy eyes: an experimental review. Arch Comput Methods Eng. 2023;30(1):521–41. https://doi.org/10.1007/s11831-022-09807-7.
14. Ajitha S, Akkara JD, Judy MV. Identification of glaucoma from fundus images using deep learning techniques. Indian J Ophthalmol. 2021;69(10):2702–9. https://doi.org/10.4103/ijo.IJO_92_21.
15. Ovreiu, S, Cristescu I, Balta F, Sultana A, Ovreiu E. Early detection of glaucoma using residual networks. 13th Int Conf Commun. 2020:161–164. Bucharest, Romania. https://doi.org/10.1109/COMM48946.2020.9141990.
16. Joshi RC, Dutta MK, Sikora P, Kiac M. Efficient convolutional neural network based optic disc analysis using digital fundus images. 43rd Int Conf Telecommun Signal Process. 2020:533–6. Milan, Italy. https://doi.org/10.1109/TSP49548.2020.9163560.
17. Diaz-Pinto A, Morales S, Naranjo V, Köhler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomed Eng Online. 2019;18:1–19. https://doi.org/10.1186/s12938-019-0649-y.
18. Fumero F, Alayón S, Sanchez JL, Sigut J, Gonzalez-Hernandez M. RIM-ONE: An open retinal image database for optic nerve evaluation. 24th Int Symp Comput Based Med Syst. 2011:1–6. Bristol, UK. https://doi.org/10.1109/CBMS.2011.5999143.
19. Sivaswamy J, Krishnadas SR, Joshi GD, Jain M, Tabish AUS. Drishti-gs: Retinal image dataset for optic nerve head (ONH) segmentation. IEEE 11th Int Symp Biomed Imaging. 2014:53–56. Beijing, China. https://doi.org/10.1109/ISBI.2014.6867807.
20. Zeng C, Lu H, Chen K, Wang R, Tao J. Synthetic minority with cutmix for imbalanced image classification. Proc SAI Intell Syst Conf. 2022:543–556. Cham, Switzerland. https://doi.org/10.1007/978-3-031-16078-3_37.
21. Burkov A. Machine Learning Engineering. Vol. 1. Montreal, QC, Canada: True Positive Incorporated; 2020.
22. Farhadi Z, Bevrani H, Feizi-Derakhshi MR. Combining regularization and dropout techniques for deep convolutional neural network. Global Energy Conf. 2022:335–339. Batman, Turkey. https://doi.org/10.1109/GEC55014.2022.9986657.
23. Iraji MS, Feizi-Derakhshi MR, Tanha J. COVID-19 detection using deep convolutional neural networks and binary differential algorithm-based feature selection from X-ray images. Complexity. 2021:1–10. https://doi.org/10.1155/2021/9973277.
24. Atalay E, Özalp O, Devecioğlu ÖC, Erdoğan H, İnce T, Yıldırım N. Investigation of the role of convolutional neural network architectures in the diagnosis of glaucoma using color fundus photography. Turk J Ophthalmol. 2022;52(3):193. https://doi.org/10.4274/tjo.galenos.2021.29726.
25. Tong Y, Lu W, Deng QQ, Chen C, Shen Y. Automated identification of retinopathy of prematurity by image-based deep learning. Eye Vis (Lond). 2020;7:1–12. https://doi.org/10.1186/s40662-020-00206-2.
26. Rao Y, He L, Zhu J. A residual convolutional neural network for pan-sharpening. Int Workshop Remote Sens Intell Process. 2017:1–4. Shanghai, China. https://doi.org/10.1109/RSIP.2017.7958807.
27. Targ S, Almeida D, Lyman K. Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029. 2016. https://doi.org/10.48550/arXiv.1603.08029.
28. Chen JX, Shen YC, Peng SL, Chen YW, Fang HY, Lan JL, *et al.* Pattern classification of interstitial lung diseases from computed tomography images using a ResNet-based network with a split-transform-merge strategy and split attention. Phys Eng Sci Med. 2024;47:755–767. https://doi.org/10.1007/s13246-024-01404-1.
29. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021;8:1–74. https://doi.org/10.1186/s40537-021-00444-8.

30. Lai Z, Deng H. Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. Comput Intell Neurosci. 2018; 2018:2061516. https://doi.org/10.1155/2018/2061516.
31. Elharrouss O, Akbari Y, Almaadeed N, Al-Maadeed S. Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. arXiv preprint arXiv:2206.08016, 2022. https://doi.org/10.48550/arXiv.2206.08016.
32. Fei L. Imagenet. 2021.
33. Ma Y, Xu T, Han S, Kim K. Ensemble learning of multiple deep CNNs using accuracy-based weighted voting for ASL recognition. Appl Sci. 2022;12(22):11766. https://doi.org/10.3390/app122211766.
34. Mijwil MM, Aljanabi M. A comparative analysis of machine learning algorithms for classification of diabetes utilizing confusion matrix analysis. Baghdad Sci J. 2024;21(5):1712–1728. https://doi.org/10.21123/bsj.2023.9010.

تصنيف فعال لمرض الجلوكوما لصور قاع العين بواسطة الشبكات العصبية التلافيفية العميقة المدربة مسبقًا ومصنف التصويت الموزون بالدقة

سلام جبار عيدان[1]، محمد رضا فيضى درخشى[2،1]

[1] *قسم هندسة الحاسوب، كلية الهندسة الكهربائية و هندسة الحاسوب، جامعة تبريز، تبريز، إيران.*
[2] *كلية الهندسة ، جامعة أوروك، بغداد، العراق.*

المستخلص

الجلوكوما هو مرض مزمن في العين يؤدي الى تلف تدريجي في العصب البصري ، ومايزال يمثل مشكلة صحية عامة كبيرة في جميع أنحاء العالم. التشخيص المبكر والدقيق أمر مهم جدا للتدخل في الوقت المناسب ومنع فقدان البصر الذي لا رجعة فيه. في هذه الورقة، قمنا بتطوير نهج لتصنيف الجلوكوما باستخدام الشبكات العصبية التلافيفية المدربة مسبقًا (CNNs)، وتحديدًا ResNet-101، وResNet-18، وDarkNet-19 . كما استخدمنا التصويت المرجح بناءً على الدقة كطريقة دمج بين المصنفات. من خلال الاستفادة من قدرات هذه الهياكل العصبية التلافيفية، نستخرج ميزات عالية المستوى من صور قاع العين ونطبق التعلم التكيفي لتكييف الشبكات لتصنيف الجلوكوما. لتعزيز أداء التصنيف، تم تقديم استراتيجية دمج تعتمد على التصويت المرجح بالدقة، حيث يتم وزن مساهمة كل مصنف بناءً على دقته الفردية على مجموعة التحقق. وقد أظهرت التقييمات التجريبية التي أجريت على ثلاث مجموعات بيانات للجلوكوما - ACRIMA وRIMONE-v2 و Drishti-GS - باستخدام نسبة 8:2 للتدريب والاختبار نتائج واعدة: دقة بنسبة 99.3% لمجموعة بيانات ACRIMA، ودقة بنسبة 95.6% لمجموعة بيانات RIMONE-v2، ودقة بنسبة 90% لمجموعة بيانات Drishti-GS. تؤكد هذه النتائج كفاءة النهج المقترح في الدراسة الحالية، لا سيما فيما يتعلق بدقة تشخيص الجلوكوما عبر مجموعات البيانات المختلفة. من خلال تمكين التدخل بعد الاكتشاف المبكر، يمكن أن يكون لهذا النهج تأثيرا كبيرا للسيطرة على الجلوكوما، مما يحسن بدوره نتائج المرضى.

الكلمات المفتاحية: دمج المصنفات، الشبكة العصبية التلافيفية، التعلم العميق، تصنيف الجلوكوما، نقل التعلم.