WEP Warwick Evans Publishing

International Journal of Public Health and Medical Research

ISSN: 3006-7707 (Print), ISSN: 3006-3752 (Online)
Volume 6, Number 7, Year 2026
DOI: https://doi.org/10.62051/ijphmr.v6n7.11



A Review of Deep Learning in Medical Image Recognition for Glaucoma

Zhengrong Li *

Chengdu Medical College, Chengdu, Sichuan, 610500, China
*Corresponding Author: 17336875218@163.com

ABSTRACT

Glaucoma is a leading cause of blindness globally, making early diagnosis and treatment critical for vision preservation. The advancement of deep learning has accelerated progress in medical image recognition, which offers new possibilities for intelligent diagnosis. This paper outlines the research background and significance, detailing techniques such as Convolutional Neural Networks (CNNs), image segmentation, and feature extraction. It reviews achievements in optic disc and cup segmentation, glaucoma classification, and model optimization, specifically discussing techniques such as transfer learning, data augmentation, and attention mechanisms to enhance model robustness, while noting limitations regarding data standardization, model generalization, and interpretability. Currently, computer-aided diagnosis demonstrates high accuracy and clinical potential. Future efforts should focus on refining datasets, developing lightweight models, and advancing interpretability and multi-modal learning to support glaucoma prevention and treatment.

KEYWORDS

Deep Learning; Glaucoma; Medical Image Recognition; Convolutional Neural Networks; Computer-Aided Diagnosis

1. INTRODUCTION

Glaucoma is a primary cause of irreversible blindness worldwide, with an estimated 76 million people affected globally by 2020 [29]. It is characterized by progressive optic nerve damage, thinning of the retinal nerve fiber layer, an abnormal cup-to-disc ratio, and visual field defects. Due to subtle early symptoms, patients often seek medical attention only at advanced stages, missing optimal treatment windows. Traditional diagnosis relies on fundus photography, Optical Coherence Tomography (OCT), and visual field tests. However, the interpretation of these results depends heavily on physician experience, which carries risks of misdiagnosis and missed diagnosis.

In recent years, improvements in medical imaging and computer vision have promoted the application of Artificial Intelligence (AI) in ophthalmology. As a key branch of AI, deep learning utilizes a hierarchical structure similar to the human brain to automatically extract complex information from images. This technology shows significant promise in medical image recognition, helping to improve the accuracy and consistency of glaucoma screening while reducing the workload on clinicians.

This paper reviews the progress of deep learning in glaucoma image recognition. It compares the performance of different models in optic disc/cup segmentation and glaucoma classification. Key factors such as data preprocessing, network architecture design, feature extraction, and model optimization are analyzed. Furthermore, the applicability of CNNs, attention mechanisms, and transfer learning in processing fundus and OCT images is discussed. Current challenges, including

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

data scarcity, weak generalization, and poor interpretability, are identified along with potential strategies. This research not only enriches the theoretical methods of intelligent medical image analysis and promotes multidisciplinary integration but also provides objective and efficient tools for clinical auxiliary diagnosis. These tools are particularly beneficial for screening in resource-limited areas, facilitating large-scale early detection and reducing blindness rates.

2. THEORETICAL OVERVIEW

Deep learning is a significant branch of machine learning rooted in the development of artificial neural networks. Key milestones include the concept of the neuron proposed by McCulloch and Pitts in 1943, the invention of the single-layer perceptron by Rosenblatt in 1958, and the introduction of the backpropagation algorithm by Rumelhart in 1986. The fundamental concept of deep learning involves using deep neural networks with multiple hidden layers to mimic the hierarchical abstraction process of the human brain. This allows for the learning of high-level semantic features directly from raw data. Unlike traditional machine learning methods, deep learning does not require manual feature extractor design. Instead, it employs an end-to-end approach to automatically learn hierarchical feature representations, offering superior expression and generalization capabilities.

The Convolutional Neural Network (CNN) is one of the most successful applications of deep learning in image recognition. By combining convolutional layers, pooling layers, and fully connected layers, CNNs can effectively capture local features and spatial structural information within images. Convolutional layers use learnable kernels for feature extraction, while pooling layers perform dimensionality reduction. This reduces computational complexity and enhances model robustness. In medical image recognition, techniques such as batch normalization, residual connections, and attention mechanisms are integrated to further improve model stability and learning efficiency.

In the context of glaucoma recognition, deep learning is primarily applied to image segmentation and disease classification. Image segmentation involves distinguishing the optic disc and cup from fundus images, a crucial step in diagnosis. The U-Net architecture is widely used in medical image segmentation due to its effective encoder-decoder structure and skip connections. The encoder captures multi-level information, which the decoder then upsamples and fuses to produce detailed results. To further enhance segmentation, some scholars have incorporated attention mechanisms into U-Net, allowing the model to focus automatically on key regions while ignoring irrelevant information.

For disease classification, classic CNN architectures such as ResNet, VGG, and Inception are widely utilized. These networks achieve high-level feature representation by increasing network depth and improving structural design. Transfer learning has also proven highly beneficial for glaucoma classification. Model parameters pre-trained on large-scale natural image datasets can be directly applied, addressing the issue of limited medical image data. This approach accelerates model convergence and improves recognition accuracy. Additionally, methods such as data augmentation and ensemble learning are employed to enhance model robustness and prediction stability.

3. LITERETURE REVIEW

3.1. International Research Status in Medical Image Recognition

Deep learning has emerged as a focal point in medical image recognition, yielding significant achievements in ophthalmic disease detection. Thisara Shyamalee and Dulani Meedeniya (2022) utilized retinal fundus images combined with segmentation and classification techniques for glaucoma detection, laying a foundation for deep learning applications in this field [1]. Building on this, MAMO MICHAEL (2022) proposed a neural network-based approach for medical image segmentation and object recognition, which improved recognition performance by optimizing neural

network applications [11]. Regarding fundamental principles, Viet Khanh Ha et al. (2019) reviewed single-image super-resolution techniques, offering valuable insights for medical image preprocessing [12]. Zhang Pengcheng (2025) summarized the progress and future directions of deep learning in medical image analysis [13], while Qingquan Huang et al. (2025) introduced lightweight models suitable for resource-constrained scenarios, facilitating the development of portable medical devices [14].

Interdisciplinary applications in medical image segmentation and classification also provide critical implications. M. Chitra Devi and M. Ramaswami (2025) constructed an efficient hybrid deep learning network for cardiac angiography, significantly enhancing segmentation and classification efficiency [2]. Similarly, Bei Hui et al. (2021) demonstrated the importance of texture analysis in classifying small hepatocellular carcinomas via CT imaging [3]. Furthermore, Wessam M. Salama and Moustafa H. Aly (2022) applied deep learning to the segmentation and classification of lung CT images for COVID-19 [4]. Feng Huang et al. (2025) proposed a semantic segmentation method fusing multispectral and RGB images for camouflaged objects, which aids in the precise segmentation of complex lesion areas [7]. Extensive research on fundus image analysis in diabetic retinopathy also offers reference values for early glaucoma screening. Gao Weiwei et al. (2018) combined support vector machines with an improved fuzzy C-means algorithm to detect hard exudates [8]. Muhammad Zubair et al. (2014) classified diabetic macular edema using color fundus images and established grading standards [6]. Additionally, S. Karkuzhali (2022) employed deep convolutional neural networks with fovea assessment techniques for macular edema detection, demonstrating the potential of deep learning in fundus disease diagnosis [9].

Advancements in general image classification have also driven improvements in glaucoma diagnosis accuracy. Nilanchal Patel and Brijesh Kumar Kaushal (2010) improved classification precision through principal component image classification and temporal image stacking [5]. Moreover, Deru Li et al. (2023) achieved automatic classification of multi-type pavement damage; their multi-object recognition method provides a reference for identifying multiple lesions in medical images [10]. Finally, AI applications in other diagnostic fields have expanded perspectives for glaucoma recognition. Alexandra Corina Faur et al. (2023) proposed AI for the non-invasive prediction of pancreatic cancer [15], while Madhura Bhide et al. (2023) combined point-of-care ultrasound with deep learning to detect elevated intracranial pressure. Such cross-disease and cross-modal studies collectively promote the intelligent recognition of medical images and the in-depth application of deep learning in glaucoma diagnosis [16].

3.2. Domestic Research Status in Medical Image Recognition

In recent years, substantial progress has been made in medical image recognition within China. Zhang Shuo et al. (2020) provided a comprehensive review of deep learning-based medical image recognition, outlining the application framework and developmental trajectory to support subsequent research [18]. Zhang Qianyu (2018) demonstrated the effectiveness and feasibility of deep learning methods in medical image processing through practical examples [27]. Expanding on this, Zheng Haofu (2024) focused on medical archive images, highlighting the role of deep learning in improving the efficiency of medical information management [22]. Furthermore, Zhu Hongbiao et al. (2023) summarized the application of deep learning in stomatology, underscoring its significant potential in specialized medical image recognition [17]. These studies have collectively propelled the development of deep learning in the medical imaging domain.

Significant contributions have also been made specifically to glaucoma research. Luo Min et al. (2025) reviewed optic disc and cup segmentation using deep learning, elaborating on techniques for feature extraction and lesion identification. This work detailed methods for the automatic measurement of key indicators, such as the cup-to-disc ratio and nerve fiber layer thickness, providing guidance for intelligent diagnosis [23]. From a clinical perspective, Ma Jingyi (2020) explored the role of machine

learning in the early diagnosis of primary open-angle glaucoma. By establishing predictive models, the diagnostic value of machine learning in early screening was confirmed, emphasizing the role of AI in improving early detection rates [28]. Together, these studies have advanced the field from both algorithmic development and clinical application perspectives.

Extensive research on general image recognition in China has laid a solid foundation for medical applications. Zheng Yuanpan et al. (2019) summarized the characteristics and applicable scopes of major models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs), providing guidance for image recognition across various fields [19]. Wang Bin and Fan Donglin (2019) analyzed the advantages of deep learning in processing large-scale, complex remote sensing images [20]. Fu Zhequan et al. (2020) demonstrated the capability of deep learning in solving specific image problems through research on high-resolution range profile recognition for radar targets [25]. In agriculture, He Yushuang [21] et al. (2023), Zhou Huiru and Wu Boming [24] (2021), and Wang Ming and Zhang Qian [26] (2023) reviewed the progress of crop disease recognition.Their work on automatic detection and severity evaluation establishes a benchmark for algorithmic precision in agricultural disease identification, thereby validating the scalability of deep learning models for fine-grained visual tasks applicable to medical lesions.

3.3. Comprehensive Literature Review on Glaucoma Diagnosis

A comprehensive review of deep learning in medical image recognition for glaucoma reveals that robust technical frameworks and extensive application cases have been established. Methodologically, the technology has evolved from simple classification tasks into comprehensive intelligent systems capable of segmentation, feature extraction, and diagnostic classification. The strong feature learning and generalization capabilities of these methods in medical image analysis have been highlighted by Zhang et al. (2020) [18]. Furthermore, valuable technical insights for medical image recognition were provided through an application-focused review by Zheng et al. (2019) [19]. Specifically regarding glaucoma diagnosis, a detailed review of optic disc and cup segmentation was conducted by Luo et al. (2025) [23]. The progress and existing challenges in this area were summarized, offering significant guidance for future research. Additionally, the critical role of intelligent algorithms in assisting clinical decision-making has been demonstrated by Zhang [27] (2018) and Ma [28] (2020), who examined the contributions of deep learning and machine learning from both technical and clinical perspectives.

In terms of technical breadth, deep learning has been widely applied across various medical imaging fields, offering proven methodological insights and transferable frameworks for glaucoma image recognition. For instance, advancements in dental image analysis were reviewed by Zhu et al. (2023) [17] while applications in medical record image analysis were discussed by Zheng (2024) [22]. These studies illustrate the strong cross-domain transferability of deep learning techniques. Zhang (2025) specifically addressed deep learning in medical image processing, noting that diagnostic accuracy is significantly improved by these methods [13]. Moreover, successful applications in other domains—such as the identification of diabetic macular edema, the processing of coronary angiography images, and the automatic detection of various lesions—offer important implications for glaucoma diagnosis. Collectively, these works demonstrate the versatility and superiority of deep learning in analyzing complex medical images, providing a solid theoretical foundation and practical guidance for developing more accurate and efficient intelligent diagnostic systems for glaucoma.

3.4. Model Optimization Strategies and Persistent Challenges

Quantitative analysis of recent literature indicates a marked improvement in diagnostic metrics; for instance, segmentation accuracy has approached near-perfect levels in controlled datasets. Methodologically, the technology has evolved from simple classification tasks to more complex

structural analysis. Significant progress has been made in applying deep learning to medical image recognition for glaucoma; however, methodologies and research directions vary considerably. In optic disc and cup segmentation, hybrid models combining attention U-Net with various convolutional neural network (CNN) encoders have demonstrated superior performance.According to the experimental results reported by Wessam M. Salama et al, (2022) a segmentation Dice coefficient of 97.3% [4] for the optic disc and 90.2% [4] for the optic cup was achieved using an enhanced U-Net combined with a ResNet50 classifier. Furthermore, Inception-v3 based models have exhibited optimal performance in glaucoma classification; for instance, the InceptionCaps model achieved an accuracy of 95.6% [1] on the RIM-ONE v2 dataset.

Regarding feature extraction and classification, hybrid models integrating deep learning with traditional machine learning have garnered significant attention. Deep features extracted by CNNs are often processed by classifiers such as Support Vector Machines (SVM) or Random Forests. While this approach enhances model interpretability to a certain extent, it concurrently increases system complexity. Additionally, multi-modal data fusion has proven effective in improving diagnostic accuracy. By integrating multi-source information—including Optical Coherence Tomography (OCT) images, visual field tests, and intraocular pressure measurements—structural and functional data can be effectively combined, thereby significantly improving the detection rate of early-stage glaucoma.

Various strategies have been employed to optimize model performance. Data preprocessing plays a decisive role in enhancing image quality and model efficacy; techniques such as normalization, filtering, and contrast enhancement are routinely applied. To address the scarcity of medical images, data augmentation methods—including rotation, flipping, scaling, and color transformation—are utilized to expand the training dataset and mitigate overfitting. Furthermore, advanced network architectures incorporating multi-scale feature fusion, residual connections, and dense connections have been adopted to bolster feature extraction capabilities. The application of adaptive learning rate optimization algorithms, such as AdaGrad, RMSProp, and Adam, has accelerated convergence and improved training efficiency.

Despite these advancements, several challenges persist. Poor model interpretability remains a critical barrier to widespread clinical adoption, as the inability of clinicians to comprehend the underlying decision-making process diminishes trust in diagnostic outcomes. Moreover, high computational demands complicate model deployment, restricting the application of deep learning technologies in primary care settings. Finally, the lack of standardized evaluation metrics and datasets across studies hinders direct performance comparisons, presenting issues that require further investigation.

4. CONCLUSION AND FUTURE PERSPECTIVES

Deep learning techniques have demonstrated significant efficacy in the recognition of medical images for glaucoma. Automated detection models based on convolutional neural networks (CNNs) can successfully perform optic disc and cup segmentation as well as glaucoma classification, achieving performance levels that meet or even surpass those of professional ophthalmologists. Extensive research has been conducted in this field, yielding various network architectures, effective feature extraction methods, and advanced model improvement strategies. However, several challenges remain, including the scarcity of sufficiently large datasets, limited model generalization capabilities, and a lack of interpretability.

Future research should prioritize the following areas. First, large-scale, multi-center, and standardized glaucoma image databases must be established to facilitate extensive training and testing. Second, lightweight network architectures and model compression techniques should be developed to reduce computational overhead, thereby enabling the deployment of models on mobile or edge computing devices. Third, model interpretability must be improved by designing visualization tools that demonstrate how decisions are made, which will foster greater clinical trust in AI-assisted diagnoses.

Fourth, multi-source information fusion and cross-modal learning should be investigated to fully leverage complementary data from various examination modalities. Finally, privacy-preserving technologies, such as federated learning, should be utilized to enable collaborative modeling across multiple hospitals while strictly protecting patient privacy.

With continued technological advancements, deep learning is expected to play an increasingly vital role in the early detection, accurate diagnosis, and personalized treatment of glaucoma. Ultimately, these developments will effectively mitigate the incidence of glaucoma-induced blindness and safeguard public visual health.

REFERENCES

[1] Shyamalee, T., & Meedeniya, D. (2022). Glaucoma detection with retinal fundus images using segmentation and classification. Machine Intelligence Research, 19(6), 82-99. https://doi.org/10.1007/s11633-022-1354-zMachine In...

[2] Devi, M. C., & Ramaswami, M. (2025). Efficient hybrid deep learning network model for segmentation and classification of heart angiographic images. Nano Biomedicine and Engineering, (3), 114-129.

[3] Hui, B., Liu, Y., Qiu, J., Cao, L., Ji, L., & He, Z. (2021). Study of texture segmentation and classification for grading small hepatocellular carcinoma based on CT images. Tsinghua Science and Technology, (2), 67-75.

[4] Salama, W. M., & Aly, M. H. (2022). Framework for COVID-19 segmentation and classification based on deep learning of computed tomography lung images. Journal of Electronic Science and Technology, (3), 24-34.

[5] Patel, N., & Kaushal, B. K. (2010). Improvement of user's accuracy through classification of principal component images and stacked temporal images. Geo-Spatial Information Science, (4), 13-18.

[6] Zubair, M., Khan, S. A., & Yasin, U. U. (2014). Classification of diabetic macular edema and its stages using color fundus image. Journal of Electronic Science and Technology, (2), 45-48.

[7] Huang, F., Yang, G., Chen, J., Xu, Y., Su, J., Huang, G., Wang, S., & Liu, W. (2025). Semantic segmentation of camouflage objects via fusing reconstructed multispectral and RGB images. Defence Technology, (8), 328-341.

[8] Gao, W., Shen, J., Wang, M., & Zuo, J. (2018). A novel method for detection of hard exudates from fundus images based on SVM and improved FCM. Journal of Chongqing University (English Edition), (3), 3-12.

[9] Karkuzhali, S. (2022). Identification and prediction of diabetic macular edema on two-dimensional fundus images with raphe assessment and deep convolution neural networks. Biomedical and Environmental Sciences, (12), 70-74.

[10] Li, D., Duan, Z., Hu, X., Zhang, D., & Zhang, Y. (2023). Automated classification and detection of multiple pavement distress images based on deep learning. Journal of Traffic and Transportation Engineering (English Edition), (2), 136-150.

[11] Michael, M. (2022). Research of algorithms for segmentation and object recognition on medical images based on neural networks [Doctoral dissertation]. Hebei University of Technology.

[12] Ha, V. K., Ren, J.-C., Xu, X.-Y., Zhao, S., Xie, G., Masero, V., & Hussain, A. (2019). Deep learning based single image super-resolution: A survey. International Journal of Automation and Computing, (4), 3-16.

[13] Zhang, P. (2025). Special topic on deep learning for medical image processing. Journal of Measurement Science and Instrumentation, (1), 5.

[14] Huang, Q., Zhang, C., Hu, C., Xie, J., Wang, Y., & Zhang, J. (2025). Waterbird image recognition using lightweight deep learning in wetland environment. Avian Research, (4), 268-281.

[15] Faur, A. C., Lazar, D. C., & Ghenciu, L. A. (2023). Artificial intelligence as a noninvasive tool for pancreatic cancer prediction and diagnosis. World Journal of Gastroenterology, (12), 50-62.

[16] Bhide, M., Singh, O., Juneja, D., & Goel, A. (2023). Bedside ultrasonography of optic nerve sheath diameter for detection of raised intracranial pressure in nontraumatic neurocritically ill patients. World Journal of Critical Care Medicine, (1), 13-20.

[17] Zhu, H., Liu, Q., Dai, F., & Song, L. (2023). Research progress of deep learning in stomatological image analysis. Chinese Journal of Practical Stomatology, (4), 110-114.

[18] Zhang, S., Zhang, R., & Zhang, Y. (2020). Review of medical image recognition based on deep learning. Chinese Journal of Health Statistics, (1), 152-158.

[19] Zheng, Y., Li, G., & Li, Y. (2019). Review on the application of deep learning in image recognition. Computer Engineering and Applications, (12), 25-41.

[20] Wang, B., & Fan, D. (2019). Research progress of deep learning in remote sensing image classification and recognition. Bulletin of Surveying and Mapping, (2), 108-111, 145.

[21] He, Y., Wang, Z., Wang, X., Xiao, J., Luo, Y., & Zhang, J. (2023). Research progress of deep learning in crop disease image recognition. Journal of Chinese Agricultural Mechanization, (2), 154-161.

[22] Zheng, F. (2024). Application research of deep learning in medical record image recognition and analysis. Information Systems Engineering, (3), 144-147.

[23] Luo, M., Cao, L., Li, J., He, X., Liu, G., Wen, J., & Huang, X. (2025). Review of optic disc and cup segmentation in glaucoma detection based on deep learning. Computer Engineering and Applications, (9), 65-83.

[24] Zhou, H., & Wu, B. (2021). Research progress on the application of deep learning in crop disease image recognition. Journal of Agricultural Science and Technology, (5), 70-77.

[25] Fu, Z., Li, X., Li, S., Dan, B., & Wang, X. (2020). Review of deep learning in radar target high-resolution range profile recognition. Aero Weaponry, (3), 41-47.

[26] Wang, M., & Zhang, Q. (2023). Research progress of image recognition technology based on deep learning in crop diseases and pests identification in China. China Vegetables, (3), 28-34.

[27] Zhang, Q. (2018). Research and application of deep learning in medical image recognition [Master's thesis]. Taiyuan University of Technology.

[28] Ma, J. (2020). The role of machine learning in early diagnosis of primary open angle glaucoma [Master's thesis]. Harbin Medical University.

[29] Barkana, Y., & Dorairaj, S. (2015). Re: Tham et al.: Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. Ophthalmology, 122(7), e40-e41. https://doi.org/10.1016/j.ophtha.2014.11.030