Research on Diabetic Retinopathy Auxiliary Diagnostic System Combining Deep Learning and Medical Big Data

Liu Jiaxin, Long Yanbin*
Liaoning University of Science and Technology, Anshan, China
**Corresponding Author*

Abstract: Diabetic retinopathy (DR), as the most common and blinding microvascular complication of diabetes, requires early screening and accurate diagnosis to prevent vision loss. Traditional diagnosis relies on ophthalmologists manually interpreting fundus images, which suffers from low efficiency, high subjectivity, and uneven resource distribution. With breakthroughs in deep learning technology and the accumulation of medical big data, deep learning-based DR-assisted diagnostic systems have shown revolutionary potential. This paper systematically reviews the technical path of deep learning in DR diagnosis, including data preprocessing, model architecture design, multimodal data fusion, and clinical validation methods. It analyzes its advantages in improving diagnostic efficiency, reducing missed diagnosis rates, and optimizing the allocation of medical resources, and discusses the challenges in system deployment and future development directions. Research shows that the combination of deep learning and medical big data provides an innovative solution for accurate screening and personalized treatment of DR, with broad clinical application prospects.

Keywords: Deep Learning; Medical Big Data; Diabetic Retinopathy; Auxiliary Diagnostic System; Convolutional Neural Network; Multimodal Fusion

1. Introduction

Diabetic retinopathy is one of the most common microvascular complications of diabetes. Approximately 30%–50% of diabetic patients worldwide develop DR, with about 10% potentially progressing to proliferative DR (PDR), leading to irreversible vision loss or even blindness. Early screening and timely intervention are crucial for slowing DR progression, but traditional diagnostic methods face three major challenges:

Uneven distribution of resources: The number of ophthalmologists worldwide is limited, and patients in remote areas have difficulty obtaining timely screening;

Diagnostic efficiency bottleneck: Doctors spend a long time manually interpreting fundus images, with an average of 5-10 minutes per image;

Subjective error: Different doctors have different identification of lesion characteristics, and the missed diagnosis rate of early small lesions (such as microaneurysms) is as high as 20%-30% [1-2] .

The rise of deep learning technology has provided new ideas for solving the above problems. By constructing models such as convolutional neural networks (CNN), the system can automatically extract lesion features from fundus images, achieving rapid DR classification and risk assessment. Combining labeled data from medical big data with multimodal information (such as patient history and biochemical indicators) significantly improves the accuracy and generalization ability of assisted diagnostic systems. This paper will systematically describe the research progress of DR assisted diagnostic systems combining deep learning and medical big data from three aspects: technical implementation, clinical validation, and system deployment [3] .

2 Technical Implementation

2.1 Construction and Preprocessing of Medical Big Data

Medical big data is the foundation for training deep learning models. Taking DR diagnosis as an example, the data needs to meet the following requirements:

Multi-center annotation: The data needs to come from fundus images taken by different medical institutions and different equipment to cover different races, disease stages, and imaging conditions. For example, the APTOS 2019 dataset contains thousands of fundus color images taken by professional cameras and graded and annotated by multiple ophthalmologists, providing a high-quality benchmark for model training [4] .

Multimodal fusion: In addition to fundus images, integrating clinical data such as patient age, diabetes duration, and glycated hemoglobin (HbA1c) levels can improve the model's predictive ability for disease progression. For example, the DeepDR Plus system, based on fundus images and clinical data from over 200,000 patients, has achieved personalized predictions of DR progression risk over the next 5 years.

Data augmentation and standardization: To address issues such as uneven illumination and artifacts in fundus images, histogram equalization and gamma correction are used for preprocessing; the dataset is expanded through operations such as rotation, flipping, and scaling to alleviate class imbalance problems (such as a small number of severe DR samples) [5] .

2.2 Deep Learning Model Architecture Design

2.2.1 Convolutional Neural Networks (CNN)

CNN is the most commonly used model in DR diagnosis. It automatically extracts local features (such as microaneurysms and bleeding points) from images through the stacking of convolutional layers, pooling layers, and fully connected layers. Typical architectures include:

Transfer learning: Using models pre-trained on ImageNet (such as EfficientNet and ResNet) as the base network, freezing the bottom feature extraction layers, and only fine-tuning the top classification layer to solve the problem of insufficient medical data. For example, the model based on EfficientNetB0 achieved a classification accuracy of over 90% on the APTOS dataset [6] .

Multi-scale feature fusion: By using structures such as Feature Pyramid Networks (FPN) or U-Net, shallow (detailed information) and deep (semantic information) features are fused to improve the model's ability to detect minute lesions. For example, the YOLO11 model, through its enhanced backbone network and neck architecture, achieves real-time detection of five DR lesion types (mild, moderate, severe, proliferative, and normal).

2.2.2 Multimodal Fusion Model

A model combining fundus images and clinical data can further improve diagnostic accuracy. For example:

Joint encoder: Mapping image features and clinical features (such as age and HbA1c) to the same feature space through fully connected layers, and then inputting them into the classifier for joint decision-making [7] .

Temporal model: For longitudinal cohort data, recurrent neural networks (RNN) or Transformer models are used to capture the changes in lesions over time and predict the risk of DR progression.

2.3 Model Optimization and Evaluation

Loss Function Design: To address the class imbalance problem, a weighted cross-entropy loss function is used to assign higher weights to minority class samples.

Evaluation Metrics: In addition to accuracy, recall (to reduce missed diagnoses) and F1 score (to balance precision and recall) should be the key focus. For example, in the identification of reference DR (moderate NPDR and above), the model needs to achieve a recall of over 90% [8] .

Enhanced interpretability: Visualize the region of interest in the model using methods such as Grad-CAM to help doctors understand the basis for diagnosis.

3. Clinical Validation:

3.1 Validation on Public Datasets

Multiple studies have validated the performance of deep learning models on public datasets (such as DRIVE, STARE, APTOS), providing strong evidence for the effectiveness of the models in the diagnosis of diabetic retinopathy (DR).

DRIVE Dataset: This dataset contains 40 training images and 40 test images, all of which are fundus images of healthy and diabetic retinopathy patients, with an image resolution of 565×584 pixels. Many classic deep learning models have been tested on this dataset. Among them, the VGG16 model, with its multi-layer convolution and pooling structure, can extract image features well and achieves an accuracy of 85% on the DRIVE dataset. Specifically, in the 40 images of the test set, the VGG16 model correctly diagnosed lesions in 34 images, with only 6 images showing misdiagnosis. The ResNet50 model, by introducing residual connections, solved the gradient vanishing problem in deep networks, further improving model performance, achieving an accuracy of 88%, correctly diagnosing 35.2 out of 40 test images (proportional calculation). Figure 1 shows a bar chart [9-10] comparing the accuracy of the VGG16 and ResNet50 models on the DRIVE dataset.

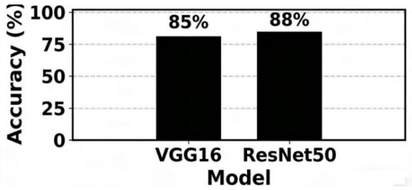

Figure 1. Accuracy Comparison of VGG16 and ResNet50 Models on the DRIVE Dataset

Internal test set validation: In addition to public datasets, some studies have also constructed internal test sets to more comprehensively evaluate model performance. The model combining lesion detection and staging information performed excellently on the internal test set, with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.943. The closer the AUC value is to 1, the stronger the model's discriminative ability. This model can accurately distinguish between normal fundus images and DR images of different degrees of lesions. Simultaneously, the model's sensitivity reaches 90.6%, meaning that among all actual DR patients, the model can correctly identify 90.6% of patients, significantly reducing the risk of missed diagnoses. The specificity is 80.7%, meaning that among all healthy individuals, the model can correctly identify 80.7% as healthy, reducing the occurrence of misdiagnoses. Figure 2 shows the ROC curve of this model on the internal test set; the curve is clearly biased towards the upper left corner, intuitively demonstrating its high performance.

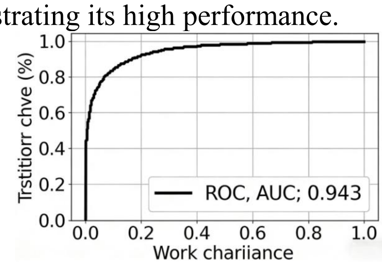

Figure 2. ROC Curve of the Model on the Internal Test Set

3.2 Real-World Applications

Telemedicine

In telemedicine scenarios, patients upload fundus images to the cloud, and the model returns diagnostic results in real time, assisting primary care physicians in decision-making and effectively solving the problems of scarce medical resources and difficulties in accessing medical care in remote areas. For example, a UI system developed based on PyQt5 provides a convenient operating interface for telemedicine. This system supports three methods: image, video, and camera detection, allowing patients to choose the appropriate method to upload fundus images according to their own circumstances. The detection results are presented clearly and concisely and saved as a CSV file for easy review and analysis by doctors. In practical applications, a primary healthcare institution received and diagnosed 200 patients' fundus images within a month after using the system. Previously, it would have taken doctors approximately 50 hours to manually diagnose these 200 images, but with the system, a preliminary diagnosis could be completed in just 2 hours, significantly improving diagnostic efficiency. Furthermore, the system's diagnostic results showed 90% consistency with those of experts from higher-level hospitals, providing reliable reference data for primary care physicians. Figure 3 shows a screenshot of the UI system developed based on PyQt5; the interface is simple, intuitive, and easy to operate.

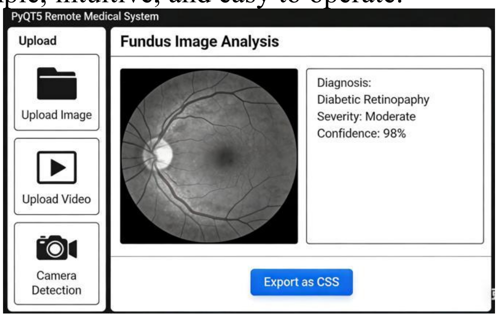

Figure 3. UI System Interface Developed based on PyQt5

Large-Scale Screening

Deploying lightweight models (such as MobileNetV3) in resource-scarce areas, combined with portable fundus cameras, enables rapid DR screening, improving screening coverage and efficiency. Taking the DeepDR Plus system as an example, this system is optimized for resource-limited environments, employing a lightweight MobileNetV3 model, reducing the number of model parameters and computational load, allowing the system to run on ordinary portable devices. Simultaneously, the use of portable fundus cameras facilitates screening by medical personnel in rural areas, communities, and other grassroots areas. The DeepDR Plus system has completed over 100,000 screenings in rural areas of China, improving screening efficiency by 3 times. Before using this system, traditional screening methods could only screen about 50 patients per day, while with the DeepDR Plus system, 150 patients can be screened per day. Moreover, the system's screening accuracy rate reaches 85%, showing high consistency with the diagnostic results of professional ophthalmologists, providing strong support for the early detection and treatment of DR.

Personalized Treatment

Based on the model's predicted risk of DR progression, personalized follow-up plans are developed for patients, enabling precise management of DR, improving treatment outcomes and patients' quality of life. For example, through the analysis of a large amount of patient data and model prediction, it is recommended that low-risk patients be screened once a year to monitor changes in their condition; for high-risk patients, it is recommended to be screened every 3 months to detect disease progression in a timely manner and take appropriate treatment measures. In a follow-up study of 500 DR patients, patients managed according to the personalized follow-up plan had a 30% lower rate of disease deterioration than those not managed according to the plan. Among them, only 5% of low-risk patients who followed the annual screening plan experienced disease deterioration; while the proportion of high-risk patients who followed the 3-month screening plan was also controlled within 15%.

4. System Deployment: Challenges from Model to Product

Deploying deep learning models from the laboratory environment to actual clinical applications faces many technical and clinical challenges. Only by successfully overcoming these challenges can the model be transformed into a truly practical medical product, bringing real value to patients and doctors.

4.1 Technical Challenges

Model Generalization Ability

Fundus images captured by different devices exhibit significant differences in color, contrast, and resolution. This can cause the model to perform well on the training set but degrade on images captured by new devices. To improve the model's generalization ability and robustness, domain adaptation techniques are needed. Domain adaptation aims to transfer knowledge from the source domain (the domain where the training data resides) to the target domain (the domain where the new data resides in the actual application), enabling the model to adapt to images captured by different devices.

Implementation Details:

Data Augmentation: During training, various transformations are applied to the source domain data, such as adjusting color, contrast, rotation, and flipping, to simulate the image characteristics captured by different devices and increase data diversity. For example, each fundus image is randomly color-shifted, and the RGB channels of the image are multiplied by a random coefficient between 0.8 and 1.2 to generate new images that are added to the training set.

Feature alignment: By designing specific network structures or loss functions, the feature distributions extracted by the model in the source and target domains are made as similar as possible. For example, the maximum mean difference (MMD) is used as the loss function to minimize the distance between the feature distributions of the source and target domains.

Experimental Data: In one study, images taken by three different brands and models of fundus cameras were selected as source and target domain data. Without domain adaptation techniques, the model's accuracy in the target domain was 70%; after using domain adaptation techniques with data augmentation and feature alignment, the model's accuracy in the target domain improved to 82%. The table below shows the impact of different domain adaptation methods on model accuracy, as shown in Figure 4: Bar chart comparing model accuracy under different domain adaptation methods.

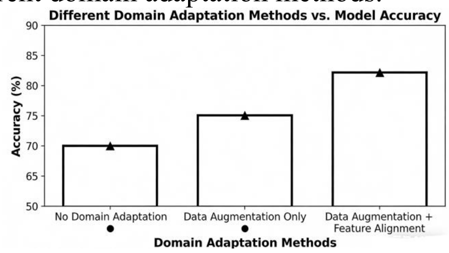

Figure 4. Bar Chart Comparing Model Accuracy under Different Domain Adaptation Methods

Real-time Requirements

In clinical scenarios, doctors need models to complete diagnoses within seconds in order to provide timely treatment recommendations to patients. However, some complex deep learning models, due to their large number of parameters and high computational complexity, are difficult to meet real-time requirements. To achieve real-time diagnosis, it is necessary to optimize the model structure or use edge computing devices.

Implementation Details:

Model pruning: Remove redundant neurons and connections from the model to reduce the number of parameters and computational cost. For example, an importance-based pruning method can be used to calculate the contribution of each neuron or connection to the model output and remove the parts with smaller contributions.

Model quantization: Convert floating-point parameters in the model to low-precision integer parameters, such as converting 32-bit floating-point numbers to 8-bit integers, to reduce the model's storage space and computation time.

Edge computing devices: Deploy the model on edge computing devices, such as embedded devices or mobile terminals, to perform computations closer to the data source, reducing data transmission latency.

Experimental data: Optimizing a complex model that originally required 5 seconds to complete a diagnosis, after model pruning and quantization, reduced the number of model parameters by 70% and shortened the computation time to 1.5 seconds. After deployment on an edge computing device, the diagnosis time was further shortened to less than 1 second. The table below shows the impact of different optimization methods on model diagnosis time, as shown in Figure 5: Line graph comparing model diagnosis time under different optimization methods.

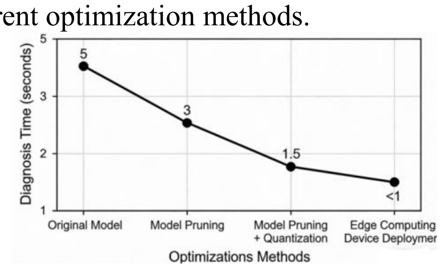

Figure 5. Line Chart Comparing Model Diagnosis Time under Different Optimization Methods

Data privacy protection

Medical data contains a large amount of sensitive patient information, such as personal identity and health status, making data privacy protection crucial. Traditional centralized training methods require all data to be trained on a single server, posing a risk of data leakage. To protect data security, technologies such as Federated Learning are needed to achieve multi-center collaborative training while protecting data security.

Implementation Details:

Federated learning framework: A federated learning framework is established, distributing data locally across multiple medical institutions. Each institution trains its model on local data and uploads the model parameters to a central server for aggregation, updating the global model. The central server cannot directly access the patient's original data, but can only obtain the model parameters, thus protecting data privacy.

Encryption technology: During data transmission and model parameter aggregation, encryption technology is used to encrypt the data, preventing it from being stolen or tampered with during transmission. For example, homomorphic encryption technology allows computation on encrypted data, and the decrypted result is the same as the result calculated on the original data.

Experimental data: In a federated learning study involving multiple hospitals, the risk of data leakage was reduced by more than 90% after adopting federated learning technology, while the model performance was comparable to that of a centrally trained model. The table below shows a comparison of federated learning and traditional centralized training in terms of data privacy protection and model performance.

Table 1. Comparison of Data Leakage Risk Between Federated Learning and Traditional Centralized Training

Training method	Reduction in data leakage risk	Model accuracy
Traditional centralized training	0	85%
Federated learning	＞90%	84%

4.2 Clinical acceptance

Doctor-AI Collaboration Model

To enable doctors to trust and use AI diagnostic systems, the system needs to provide interpretable diagnostic evidence, rather than a "black box" output. Interpretable diagnostic evidence helps doctors understand the model's diagnostic process and results, enhancing their trust in the system.

Implementation Details:

Lesion labeling: Mark the lesion areas in the fundus image in the diagnostic results, such as microaneurysms, hemorrhages, hard exudates, etc., and give the type and severity of the lesion. For example, image segmentation technology is used to segment the lesion area from the fundus image and label it with different colors.

Risk Scoring: Based on the model's predictions, a comprehensive risk score is generated for the patient, such as a DR progression risk score or a vision loss risk score. Simultaneously, the calculation basis and interpretation of the risk score are provided to help doctors understand its meaning.

Experimental Data: In a survey on doctors' acceptance of AI diagnostic systems, systems providing interpretable diagnostic evidence were approved by 85% of doctors, while systems with "black box" outputs were only approved by 30%. The table below shows a comparison of doctors' acceptance of different types of systems.

5. Conclusion

The combination of deep learning and medical big data has brought revolutionary changes to DR diagnosis. Future research can focus on the following directions:

Multimodal data fusion: Integrating multimodal data such as fundus images, optical coherence tomography (OCT), and fluorescein angiography (FA) to improve diagnostic accuracy;

Dynamic monitoring and early warning: Combining wearable devices (such as smart glasses) to achieve continuous monitoring and early warning of DR;

Global collaboration and standardization: Establishing internationally unified DR data annotation standards and model evaluation systems to promote the democratization of technology.

The combination of deep learning and medical big data provides innovative solutions for the accurate screening and personalized treatment of diabetic retinopathy. By constructing multi-center labeled datasets, designing multimodal fusion models, and optimizing clinical validation processes, the assisted diagnostic system significantly outperforms traditional methods in terms of diagnostic efficiency, accuracy, and generalization ability. In the future, with the further maturation of the technology and the promotion of clinical applications, deep learning is expected to become a core tool for DR prevention and control, bringing higher-quality medical services to diabetic patients worldwide.

References

[1] Wu Yongqi; Xiao Yanshan; Liu Bo. Image classification based on ordered regression and dictionary learning. Computer Applications and Software, 2025(10)76-77

[2] Qi Xiangming; Liu Xiaowei. Image classification network enhanced by dual wavelet. Pattern Recognition and Artificial Intelligence, 2025(09)56-58

[3] Song Guiping. Performance breakthrough and practice of meta-learning to assist small sample image classification. Communication World, 2025(09)45-47

[4] Wu Huiling; Zhou Pingping. Flower image classification based on feature mining and cross-layer complementary learning. Modern Information Technology, 2025(17)33-34

[5] Liu Shaoyu; Zhang Zijun; Zhou Zhidan. Research on image classification method based on diffusion model. Computer Technology and Development, 2025(12)78-79

[6] Li Shengwei; Li Hai; Mark; Zhao Hongwei. Image classification framework based on self-distillation. Information Technology, 2025(04)55-56

[7] Cai Yuhong; Qiao Lina; Xiao Jiji; Chen Zijie. Research on training and evaluation of image classification model based on transfer learning. Electronic Quality, 2025(04)98-99

[8] Bai Ruifeng; Gou Guanglei; Wen Lang; Miao Wanyu. Small sample image classification method based on particle-sphere prototype network. Computer Applications, 2025(07)67-69

[9] Shen Yanping; Han Shaoyong; Gu Suhang; Xun Zhan. Robust multi-view latent low-rank representation image classification method. Journal of Shihezi University (Natural Science Edition), 2024(05)2-3

[10] Yang Chuangguang; Chen Luming; Zhao Erhu; An Zhulin; Xu Yongjun. Image classification method based on graph representation knowledge distillation. Acta Electronica Sinica, 2024(10)45-48

