DOI: 10.54254/2755-2721/97/20241465

Deep Learning AI for DR Microaneurysm Screening Based on Fundus Images

Yunfei Qi

School of Faculty of Data Science, City University of Macau, Macau, China

D22090102812@cityu.edu.mo

Abstract. Since 2014, Artificial Intelligence (AI) image detection has developed rapidly in the medical field, while the global incidence of diabetes has increased rapidly. As an early sign of diabetic retinopathy (DR), the detection of diabetic microaneurysms is particularly important for the prevention and treatment of diabetes. Therefore, the artificial intelligence exhibition of DR microaneurysm screening based on fundus images shows great potential in this regard. This paper reviews the research progress of artificial intelligence technology for DR microangioma screening based on fundus images in recent years, and focuses on the improvement measures taken for different depth learning methods. These improvement measures aim to optimize model performance, improve training efficiency and enhance the generalization ability of the model. Some model experimental results are compared and analyzed in order to get the best model in a specific dataset or result. Finally, the current challenges are discussed and suggestions for the development of this field are put forward.

Keywords: diabetic retinopathy, microaneurysm, deep learning.

1. Introduction

In recent years, with the rapid development of artificial intelligence technology, its application in the field of medical image analysis has become increasingly widespread. Diabetes, as a chronic disease that is widely prevalent worldwide, has become one of the main causes of vision loss among working-age people due to its complication, diabetic retinopathy. According to the World Health Organization, the number of people with diabetes has increased dramatically worldwide, and it is expected to reach nearly 600 million by 2035, which undoubtedly increases the urgency of preventing and controlling diabetic retinopathy. As an early landmark lesion of diabetic retinopathy, the early detection and accurate diagnosis of microaneurysms are crucial for timely intervention and treatment, as the paper as delaying the progression of the disease. However, the diameter of microaneurysms is usually small, ranging from tens to hundreds of microns, and their morphology is variable, making it difficult to directly observe and diagnose them with the naked eye. Traditional manual screening methods based on funduscopy are not only time-consuming and laborious, but also have a high rate of misdiagnosis and missed diagnosis. Therefore, how to quickly and accurately achieve automated screening of microaneurysms has become a key issue that needs to be addressed urgently. In this context, artificial intelligence technology provides a new solution for the screening of microaneurysms with its powerful data processing and analysis capabilities.

At present, deep learning algorithm have made remarkable achievements in the field of micro aneurysm (MA) detection. On this basis, many micro aneurysm detection methods based on deep learning have emerged. In particular, convolutional neural network (CNN) and its variants are used to extract features from retinal images, and then to classify and locate microaneurysms. For example, Xinpeng Zhang et al. proposed a kind of SODNet based on deconvolution neural network (DNN) to restore the erased feature map details in the convolution layer by embedding the deconvolution operation [1]. In addition, other methods such as Ramin Almasi et al. used CNN and transfer learning to detect microaneurysms (MA) in OCT images by stacking and generalizing four pre trained CNN models after fine-tuning [2]. Gao Weiwei et al. proposed an improved Faster R-CNN micro aneurysm automatic detection algorithm based on multi feature scale fusion. By integrating the feature maps of different layers and performing feature fusion, the feature utilization of small targets was improved [3].

This paper aims to review the research progress of artificial intelligence technology for DR microaneurysm screening based on fundus images in recent years. Through combing and analyzing the relevant literature, this paper will focus on the improvement of the artificial intelligence deep learning model used in the detection and segmentation of micro aneurysms by different methods. At the same time, by comparing the accuracy, sensitivity, F1 and other experimental results of different methods, the paper can get a relatively better method. Finally, the paper proposes suggestions for the development of artificial intelligence for DR microaneurysm screening based on fundus images.

2. Micro aneurysm detection method based on deep learning

2.1. Deconvolution neural network (DNN)

In 2019, Xinpeng Zhang et al. proposed a kind of SODNet based on deconvolution neural network (DNN) to restore the erased feature map details in the convolution layer by embedding the deconvolution operation [1].

The architecture of SODNet consists of three convolution layers (Conv), three deconvolution layers (Deconv) and two full connection layers (FC). The deconvolution layer is a key innovation in SODNet. The deconvolution layer achieves this by transposing the size of the feature map through the upsampling operation, and restores some details lost in the process of convolution and pooling, which is equivalent to the reverse process of the convolution operation. Through deconvolution layer, SODNet can combine low-level and high-level features to improve the detection ability of small targets. The full connection layer receives the output from the deconvolution layer, and further processes these features through activation functions (such as ReLU) and dropout and other technologies to reduce over fitting. Finally, these features are passed to the softmax layer for classification to distinguish whether the input image contains MA objects.

2.2. Convolutional Neural Network (CNN)

2.2.1. Optical coherence tomography (OCT). In 2019, Ramin Almasi et al. proposed a method based on convolutional neural network and optical coherence tomography (OCT) to identify microaneurysms in images in their research [2].

This paper uses four pre trained CNN models (VGG16, VGG19, Xception, and Inception V3) as basic learners. These models are pre trained on ImageNet dataset. Fine tune these models to adapt to the four classification tasks of OCT strips, including removing the top layer, adding a new full connection layer, batch normalization layer, dropout layer, and softmax layer. Use training set and verification set to train basic learners, reduce overfitting by freezing some layers, and gradually train more layers. Cross entropy loss function and SGDM optimizer are used for training, and early stopping method and saving the best model are used to improve the training effect. In the Stacking Ensemble stage, basic learners first make independent classifications, and then meta learners synthesize these classification results to produce final classification decisions. Finally, the model output includes the prediction results of normal, abnormal, vascular and other labels.

2.2.2. Feature embedding. In 2021, Muhammad Mateen et al. proposed a method based on the pre training CNN model VGG-19 and Inception v3, which fused the features extracted from the two models through feature embedding technology to achieve automatic detection of microaneurysms in fundus images [4].

VGG-19 used in this paper is a classic depth convolution neural network, which has a strong ability of image feature extraction. In the feature extraction phase, the full connection layer of VGG-19 is removed, and only the convolution layer and pooling layer are retained for extracting depth features from fundus image data blocks. Inception-v3 model is also an efficient depth neural network, with convolution kernels of various sizes, which can capture image features of different scales. Similar to VGG-19, the full connection layer is removed and only the convolution layer is used for feature extraction. Feature embedding is one of the core steps of this method, which aims to effectively fuse the features extracted from VGG-19 and Inception v3 models to improve the classification performance. The gradient descent optimizer and softmax function are used to calculate the cross entropy loss of each feature, and the weight and offset are updated through backpropagation to finally obtain a fused feature vector for subsequent classification tasks. The fused feature vectors are sent to the softmax classifier for final decision to determine whether the input data block contains microaneurysms.

2.2.3. Faster R-CNN. In 2022, Gao Weiwei et al. proposed an improved Faster R-CNN micro aneurysm automatic detection algorithm based on multi feature scale fusion [3].

First of all, improve the fusion technology of multiple feature scales, integrate deep and shallow feature maps, make them match and add in spatial dimensions, generate new feature maps rich in more semantic information, and set anchor boxes of different scales on different feature layers to enhance the detection ability of small targets, which solves the problem of insufficient utilization of small target features caused by the original Faster RCNN using only the last layer of feature maps for regional candidate network (RPN) calculation. Secondly, aiming at the double quantization error problem in the original RoI Pooling, the region of interest (RoI) leveling pooling technology is proposed. This technology accurately calculates the pixel value of each point in the RoI region through the bilinear interpolation method, effectively eliminates the quantization error, improves the accuracy of feature mapping, and thus enhances the detection accuracy of small targets. Finally, in order to balance the gradient contribution of difficult and easy samples in training, the original Smooth L1 loss function is optimized and the balanced L1 loss function is designed. Through the gradient scaling mechanism, the new loss function reasonably adjusts the gradient proportion of samples with different difficulty, reduces noise interference, and makes the model training process more stable and efficient.

In terms of network design and implementation, VGG16 is used as the basic feature extraction network, and combined with multi feature scale fusion technology to improve. The optimized RPN and RoI Pooling components are used to generate more accurate candidate areas.

In 2023, Yang Li et al. proposed an improved method for detecting fundus microaneurysms embedded in CBAM (Convolutional Block Attention Module) based on Fast R-CNN [5].

This model replaces the original feature extraction network VGG16 of Faster R-CNN with ResNet50, and combines FPN for feature extraction to retain more low-level feature information. At the same time, CBAM attention module is introduced into the residual unit of ResNet50 to enhance the network's attention to important features and key areas through channel attention and spatial attention. In addition, the FPN module is improved by adding fusion factors to optimize the extraction ability of small target features. SGD optimization algorithm is used for training, and appropriate initial learning rate, batch size and iteration times are set. The loss function of the network consists of classification loss and regression loss, which jointly guide the training process of the model to improve the detection performance of fundus microaneurysms.

2.3. Based on local structure awareness

In 2021, Jiakun Deng et al. proposed a retinal microaneurysm (MA) detection method based on local structure awareness, which combined local gradient descriptor (RGD) and traditional features (saliency and texture features), and used a gradient lifting decision tree (GBDT) for classification [6].

In the preprocessing stage, the data set is preprocessed to enhance the image quality and extract the green channel information. Then, the candidate regions are extracted by double-threshold segmentation and morphological processing. In the feature extraction phase, salient features including mean, standard deviation, third moment, energy, entropy and contrast are extracted from candidate regions. Local texture features are extracted based on gray level co-occurrence matrix (GLCM), including correlation, negative moment, variance, entropy, angular second moment and contrast. A new local structural feature, ring gradient descriptor (RGD), is proposed to distinguish MAs and blood vessels by calculating the gradient difference between the paper candidate regions and their surrounding regions, and effectively distinguishing microaneurysms from non-diseased regions.

2.4. YOLO series models

You Only Look Once(YOLO) series models are popular in the field of target detection because of their efficiency and accuracy.

2.4.1. YOLOv4. In March 2022, Gao Weiwei proposed an improved YOLOv4 automatic detection algorithm embedded in SENet (Squeeze and Extraction Networks) [7].

This paper uses the Kaggle Sugar Website dataset, which contains a large number of eyeground images with labels. In the aspect of model construction, the improved fast fuzzy C-means (IFFCM) clustering algorithm is used to optimize the prior box parameters of the target samples to improve the matching degree between the prior box and the feature map. The SENet module is embedded in the backbone network of YOLOv4 to enhance the confidence of MAs by strengthening key information and suppressing background information. The spatial pyramid pooling (SPP) structure is added to the network neck to enrich the feature expression ability and help to separate important context information. In the process of model training and optimization, online data enhancement is used, including random rotation, flipping, filtering and other operations, to improve the model generalization ability. The small batch random gradient descent method (MSGD) with driving factor is used for training, and appropriate learning rate attenuation strategy is set. Ablation experiments the paperre conducted to verify the impact of different improvement strategies on the model performance and optimize the network structure.

2.4.2. YOLOv8. In 2023, Bowei Zhang et al. proposed an improved micro aneurysm (MA) detection model (MA-YOLO) based on SwinIR image super-resolution reconstruction and YOLOv8 target detection framework [8].

This paper uses SwinIR to reconstruct the FFA image from the original 768 × 768 pixels to 1536 × 1536 and 2304 × 2304 pixels. This step enhances the image details and makes the characteristics of microaneurysms clearer. A MA detection layer is added between the neck and head of YOLOv8 to process the shallow feature map from the P2 layer of the backbone network and fuse it with the deep feature map to enhance the model's ability to locate small targets. Transfer and fine tune the model between data sets with different resolutions (original image, 1536 × 1536 pixel and 2304 × 2304 pixel super-resolution image) by using transfer learning, increase the number of data samples and improve the generalization ability of the model. The Wise IoU loss function is used to replace the original CIoU loss function of YOLOv8. Through the dynamic focusing mechanism and focusing coefficient, the excessive punishment of geometric factors on the loss value is alleviated, and the generalization performance of the model in the case of unbalanced sample distribution is improved.

2.5. U-Net series model

U-Net model is outstanding in the field of medical image segmentation because of its unique U-shaped structure.

2.5.1. AOSLO net model based on UNet and Efficient Net. In 2022, Qian Zhang et al. proposed an AOSLO net model based on UNet and EfficientNet, which is used to automatically segment retinal microaneurysms from adaptive optical scanning laser ophthalmoscope (AOSLO) images [9].

This paper prepares the dataset by itself, generates enhanced images, and enhances the visibility of MA boundaries by averaging frame images, inverting colors, and performing local mean filtering. The model AOSLO-net adopts the UNet structure, the encoder part uses the pre trained EfficientNet-b3, and the decoder part uses the standard UNet decoder. The encoder is responsible for extracting MA features at different levels, and the decoder integrates these features into segmentation results. The data set is divided into training set, verification set and test set by cross validation. Increase the diversity of training samples through data enhancement (such as flipping, rotation and scaling) to avoid over fitting. The combined loss function, including binary cross entropy and Dice loss, is used to optimize the model performance. The output image is binarized, and small fragments smaller than the threshold of a specific area are cleared. Select the three models that perform best in the verification set, and improve the performance of the model by taking the union of their outputs.

2.5.2. Task learning combined with U-Net. In 2023, Cui Yongjun and others proposed a multi task learning combined with U-Net micro aneurysm image segmentation method [10].

The network model takes U-Net as the backbone and combines the multi task learning architecture. The encoder part adopts the vgg-16 network without the full connection layer, and adds the full connection layer at the end of the encoder as the output layer of the side task. The convolutional attention module (CBAM) is introduced into the encoder, and the channel attention module and the spatial attention module are used to improve the network's feature extraction ability for microaneurysms. In the decoder part, CARAFE module is used instead of the traditional up sampling method to make better use of the surrounding information and reorganize features. The image segmentation of microaneurysms is the main task, and the presence detection of microaneurysms is the secondary task. Through multi task learning, the model can simultaneously optimize the segmentation and classification performance in the training process, and improve the generalization ability of the model. In order to balance the training of the main task and the sub task, the uncertainty the weight loss function is introduced to enable the model to automatically learn the weight of the loss function suitable for the two tasks. In the loss function design, Focal Loss function is used for the main task (segmentation task) to optimize the imbalance between the focus and the background, so that the model pays more attention to the small aneurysms. The secondary task (classified task) uses a two category cross entropy loss function, and at the same time, a constraint term is added to prevent over fitting.

3. Analysis of experimental results

3.1. Datasets

In this study, most of the articles used public datasets such as ROC, e-ophtha-MA and other standard fundus image datasets to verify the effectiveness of the proposed method in fundus image microaneurysm detection and segmentation tasks. These datasets are widely used in the field of fundus image analysis, which helps to compare the performance of different methods.

3.2. Evaluation index

In order to comprehensively evaluate the performance of the proposed method, a variety of evaluation indicators are used, including but not limited to:

Precision: it measures the proportion of the samples predicted by the model as positive samples that are really positive samples.

Sensitivity or recall: measure the ability of the model to correctly identify all positive samples (i.e. micro aneurysms).

F1 score (F-Score): the harmonic average of precision rate and recall rate, used to comprehensively evaluate the performance of the model.

3.3. Analysis of experimental results

In the two articles in Table 1 based on the same data set, the F-score shows that Xinpeng Zhang et al.'s method of SODNet based on deconvolution neural network (DNN) is significantly better than Jiakun Deng et al.'s method based on local structure awareness.

Table 1. Comparison of algorithm sensitivity and F-score scores based on the same dataset

Database	Methods	Sensitivty against FPIs							F-score
		1/8	1/4	1/2	1	2	4	8	
ROC	SODNet[1]	0.713	0.730	0.751	0.757	0.767	0.800	0.800	0.760
	Based on local structure awareness [6]	0.083	0.104	0.200	0.257	0.344	0.394	0.468	0.264
e-ophtha-MA	SODNet[1]	0.890	0.892	0.895	0.901	0.920	0.930	0.930	0.908
	Based on local structure awareness [6]	0.335	0.424	0.496	0.578	0.634	0.668	0.696	0.547

This is because the SODNet based on random undersampling can detect small targets with unbalanced data in less training samples, and deconvolution can restore the erasure details of the convolution layer feature map, and adjust the size of the final feature map to the original training image. The improved zero fill operation can reduce the noise of training samples and suppress the background. Therefore, the network can achieve good feature representation capability.

Table 2 lists the accuracy, sensitivity and F-score scores of various methods in detail to comprehensively evaluate the model performance. In view of the fact that the comprehensive advantages of the model cannot be accurately judged only by the single index of accuracy or sensitivity, the F-score score is still used as the main basis for judging the advantages and disadvantages of the model. Under this standard, the CNN based method proposed by Muhammad Mateen et al. stood out, with its F-score up to 96%, showing excellent performance. In contrast, the other CNN based method adopted by Yang Li et al. performed mediocrely, with a F-score of only 75.1%, showing a significant difference. It is worth noting that both the highest score and the lowest score are derived from the CNN based model, which indicates that even if the same model framework is adopted, different adjustment and optimization strategies will lead to huge differences in results. In contrast, YOLO series models have shown relatively stable performance and become a trustworthy choice.

Table 2. Comparison of method accuracy, sensitivity and F-score scores

method	R(%)	P(%)	F-score(%)
Feature embedding[3]	87	95	96
Improved Faster R-CNN based on multi feature scale fusion[4]	90.89	94.32	92.57
Embedded CBAM based on Fast R-CNN[5]	78.2	72.2	75.1
YOLOv4[7]	89.77	87.13	88.51
YOLOv8[8]	88.23±0.11	97.98± 0.06	92.85± 0.09

DOI: 10.54254/2755-2721/97/20241465

4. Conclusion

At present, the artificial intelligence technology for diabetic retinopathy (DR) microaneurysm screening based on fundus images, although showing great potential and value, still faces multiple challenges in practical application:

When using the deep learning model to screen microaneurysms, different subtle changes in the same model led to huge differences in results. Even the same model framework may lead to significant differences in the final screening results due to subtle differences in parameter settings, training strategies, data pre-processing, etc. For example, F-score values can differ by 20% in different improvements based on CNN models. This uncertainty brings great difficulties to the optimization and debugging of the model, and also affects the reliability and stability of the model in practical applications.

When different authors use DR microangioma screening technology based on fundus images, they usually use different public data sets or self collected data sets for training and testing, which makes it difficult to directly compare and evaluate the results of different studies. This not only affects the reproducibility of technology, but also hinders the further development and promotion of technology.

At present, most DR micro aneurysm screening models based on fundus images are based on convolutional neural network (CNN). Although CNN has achieved remarkable success in the field of image processing, with the continuous development of technology, the demand for model innovation is increasingly urgent. At the same time, the effect of other non CNN models in the screening of microaneurysms is often poor, which is difficult to compare with CNN models. Therefore, how to explore new model architecture and algorithms on the basis of maintaining the advantages of CNN model to improve the accuracy and efficiency of screening is one of the problems to be solved urgently.

In conclusion, the artificial intelligence technology for DR microaneurysm screening based on fundus images has made significant progress, especially the introduction of deep learning methods, which has greatly improved the detection, segmentation and positioning performance of microaneurysms. In the future, with the continuous development and innovation of technology, the paper has reason to believe that artificial intelligence technology will play a more important role in DR microaneurysm screening, and provide more powerful support for early diagnosis and treatment of diabetic patients.

References

[1] Zhang, X., Wu, J., Peng, Z., & Meng, M. (2020). Sodnet: small object detection using deconvolutional neural network. IET Image Processing.

[2] Almasi, R., Vafaei, A., Kazeminasab, E., & Rabbani, H. (2022). Automatic detection of microaneurysms in optical coherence tomography images of retina using convolutional neural networks and transfer learning. Scientific Reports.

[3] Gao, W. W., Yang, Y. L., Fang, Y., Fan, B., & Song, N., (2023). An Improved Faster R-CNN Algorithm for Automatic Detection of Retinal Microaneurysms with Multi-Feature Scale Fusion. Acta Photonica Sinica, 52(4), 220-231.

[4] Mateen, M., Malik, T. S., Hayat, S., Hameed, M., Sun, S. (2022). Deep Learning Approach for Automatic Microaneurysms Detection. Sensors (Basel, Switzerland), 22(2), 542. https://doi.org/10.3390/s22020542.

[5] Yang, L., Shao, H., Cui, W, C. (2024). An Improved Faster R-CNN Method for Detecting Microaneurysms in Fundus Images with Embedded CBAM. Yangtze River Information and Communications, 37(5): 40-43.

[6] Deng, J., Tang, P., Zhao, X., Pu, T., Qu, C., & Peng, Z. (2022). Local Structure Awareness-Based Retinal Microaneurysm Detection with Multi-Feature Combination. Biomedicines, 10(1), 124. https://doi.org/10.3390/biomedicines10010124

[7] Gao, W, W., Shan, M, T., Song, N., Fan, B., & Fang, Y. (2022). An Improved YOLOv4 Algorithm for Automatic Detection of Microaneurysms in Fundus Images with Embedded SENet. Journal of Biomedical Engineering, 39(4), 8.

[8] Zhang, B., Li, J., Bai, Y., Jiang, Q., Yan, B., & Wang, Z. (2023). An Improved Microaneurysm Detection Model Based on SwinIR and YOLOv8. Bioengineering (Basel, Switzerland), 10(12), 1405. https://doi.org/10.3390/bioengineering10121405
[9] Zhang, Q., Sampani, K., Xu, M., Cai, S., Deng, Y., & Li, H., et al. (2021). Aoslo-net: a deep learning-based method for automatic segmentation of retinal microaneurysms from adaptive optics scanning laser ophthalmoscope images. arXiv e-prints.
[10] Cui, Y, J., Lei, K, J., & Ma, Q, M. (2024). A Microaneurysm Image Segmentation Method Combining Multi-task Learning with U-Net. Electronic Design Engineering. (15), 190-195. doi: 10.14022/j.issn1674-6236.2024.15.040.