RESEARCH ARTICLE

Retinopathy diagnosis technology based on pathological data cluster analysis

Xinyu Xiao[1,*], Yongfeng Qin[2], Meng Jiang[1]

[1]School of Medical Technology, [2]School of General Education, Wuzhou Medical College, Wuzhou, Guangxi, China.

Received: April 3, 2025; **accepted:** August 22, 2025.

With the increasing awareness of people's health, retinopathy has become the focus of medical attention because of its high incidence rate and serious harm. However, the traditional diagnostic techniques for retinal cellopathy have the problem of poor accuracy. To improve the accuracy of retinopathy diagnosis, this study proposed a diagnostic technique for retinal cellopathy based on pathological data cluster analysis. This technology constructed a retinopathy diagnosis network using feature maps of constant size, avoiding the situation of ignoring small blood vessels during down sampling operations. The importance of retinal vessels in the diagnosis of retinopathy was analyzed, and a diagnostic network integrating the characteristics of retinal vessels was designed to achieve the retinopathy diagnosis. The results showed that the average diagnostic accuracy rate of the proposed method in ten experiments reached 93.82%, which was significantly better than that of traditional methods. The proposed retinopathy diagnosis technique based on pathological data clustering analysis could improve the diagnostic accuracy of retinopathy. This study provided a new technological means for the accurate diagnosis of retinopathy, which was of great significance for promoting the development of related medical diagnostic techniques.

Keywords: pathology; clustering; data analysis; retinopathy; diagnostic technique.

[*]**Corresponding author:** Xinyu Xiao, School of Medical Technology, Wuzhou Medical College, Wuzhou, Guangxi 543000, China. Phone: +86 15200862692. Email: 15200862692@163.com.

Introduction

The improvement of social quality of life has made people pay more attention to physical health. At present, cataract, glaucoma, turbid vitreous, and other diseases have become one of the main diseases that damage human eye health. In terms of retinopathy, as of 2013, the World Health Organization pointed out that the number of retinopathy patients worldwide reached 347 million, and the number of retinopathy patients in most countries was on the rise with 187 million people still undiagnosed. Eighty percent of patients with retinopathy come from low- and middle-income countries. Complications such as blindness and fundus hemorrhage caused by retinopathy greatly endanger the quality of patients' life. However, the current diagnostic techniques based on image processing have low accuracy and are difficult to handle complex data. The traditional machine learning methods have cumbersome algorithms and insufficient flexibility. Moreover, there is a lack of unified standards for measuring digital indicators of fundus blood vessels. These drawbacks have limited the development of

diagnostic techniques for retinopathy [1]. Although the current diagnosis of retinal diseases benefits patients, there are still many challenges. In the image processing of retinopathy, due to the interference of vascular features such as optic disc, bleeding, and background texture, the system often finds it difficult to extract accurate vascular network structures. Meanwhile, there is no unified medical standard for the measurement of fundus vascular digital indicators [2]. The current diagnosis of retinopathy is mostly based on a small number of pictures of machine learning method diagnosis, which lacks robustness, not suitable for a large number of complex data. In addition, retinopathy image segmentation still faces many challenges. The contrast between the blood vessels and the background is weak, the arterial blood reflex is strong, and the edges are blurred, which lead to the large blood vessels being easily missegmented. Further, the illumination is affected by camera parameters, shooting conditions, and the characteristics of the retinopathy itself, which increases the difficulty of vascular segmentation. The image components are also complex. In normal images, the optic disc and background textures are easily misjudged as blood vessels. There are lesions such as exudation and bleeding in the images, which further increases the difficulty of accurate segmentation. Also, it is time-consuming and labor-intensive for specialists to label high-resolution images, while the number of samples in public databases is limited, which seriously restricts the performance of models trained on labeled data. The digital measurement standards for fundus vascular morphology have not yet been unified. Due to the limited generalization ability of models trained on a small number of samples, diagnostic methods based on segmentation results have poor robustness when facing a large amount of complex and variable clinical data [3].

Cluster analysis of pathological data is a current research hotspot. With the help of massive pathological data and high-speed graphics processing unit (GPU) devices, it is possible to solve complex problems. The diagnosis of retinopathy involves multiple techniques. Traditional machine learning methods require the design of algorithms for each feature, which is labor-intensive and has poor robustness. However, the end-to-end training process of pathological data cluster analysis can automatically mine abstract image features, bringing convenience to classification [4, 5]. Many studies explored the application of pathological data cluster analysis in the diagnosis of retinopathy. Sinha *et al*. proposed a retinal image enhancement method based on fuzzy C-means clustering, which grouped the brightness levels into multiple clusters and assigned a cluster membership to each brightness level. These membership values were mapped to the corresponding initial values, and the green channel of the modified image was equalized using adaptive histogram equalization to obtain an enhanced image. This method could preserve the natural features of anatomical retinopathy images and improve the overall information conveyed by retinal images, thereby achieving effective disease diagnosis [6]. The detection of diabetes retinopathy is a time-consuming and laborious process, which requires ophthalmologists to examine and evaluate the digital color fundus photography images of the retina and identify the diseases related to vascular abnormalities caused by diabetes. Kumari *et al*. proposed a diabetic retinopathy detection method based on an automatic decision-making ResNet feedforward neural network and analyzed and mapped missing connections of retinal arterioles, microaneurysms, venules and central fovea, cotton wool spots, macula, outer lines of the optic disc, and hard exudates and hemorrhages in color and black-and-white images through integrated mapping technology. Accurate calculations when obtaining vector sequences were performed to identify diabetic retinopathy processes [7]. Manual observation of retinopathy makes the detection process more complicated, tedious, and error prone. Dayana *et al*. designed a retinopathy severity classification method based on feature fusion and optimized an

integrated refined deep residual network (RDRN), which first eliminated noise and enhanced image contrast through preprocessing. Then, the optic disc was segmented using a spatial attention U-Net based on dilated convolution. The blood vessels were segmented using an entropy-based hybrid technique. Subsequently, the lesion area was segmented using a fusion U-Net model based on the attention mechanism and a weighted focal loss function. Features were extracted and fused using a two-layer fusion network, and retinopathy was classified using an RDRN that integrated a squeeze excitation module and a tunicate spider monkey optimization algorithm [8]. Retinopathy image segmentation is an important step to assist doctors in diagnosing retinopathy and provide information about blood vessels. Compared with manual segmentation, segmentation based on pathological data cluster analysis can complete the task of retinopathy image segmentation more quickly. With people's concern for vision health and the continuous increase in the number of retinal diseases, more people need retinal examinations. The research on retinal image segmentation algorithms based on pathological data cluster analysis is becoming increasingly important. However, image segmentation algorithms based on image processing have difficulty dealing with complex and variable retinopathy images, and the images captured by different types of cameras have significant color differences. Therefore, algorithms may need to be designed with different parameters for different situations. Algorithms based on traditional machine learning require manual design of features that are also obtained through some simple algorithms, which has no guarantee if these algorithms can extract the required features well or these features are the most suitable features for pixel classification in the image. Different theories have also proposed different feature selection. The design of these features is a complex task [9, 10].

Although there are many achievements in retinopathy diagnosis, the existing research still has certain limitations, which include that some methods still have difficulty accurately extracting the fine vascular features and lesion areas when dealing with complex lesion images, while some others rely on specific datasets or preprocessing methods and have insufficient universality. There are still some methods that lack in-depth analysis of the time complexity and space complexity of the algorithms, making it difficult to be efficiently applied in actual clinical scenarios. This study proposed a retinopathy diagnosis technique based on pathological data cluster analysis using a large-scale dataset of retinopathy images. The proposed method designed a diagnostic network that integrated retinal vascular characteristics by analyzing pathological data and combining cluster analysis methods. Image segmentation was employed to divide the retinal image into two parts of blood vessels and background. Pathological data cluster analysis was used to extract the abstract features in the image before the classification and diagnosis being carried out through the deep learning network. This study would promote medical scenarios and be helpful for the early and accurate diagnosis and effective treatment of retinopathy by improving the accuracy of diagnosis, reducing the risk of complications of retinopathy, and effectively improving the health status of patients, which would have a positive and profound impact on improving the health level of the human eyes.

Materials and methods

Segmentation of retinopathy image

The image segmentation of retinopathy was the segmentation of retinal cell images collected by a fundus camera into two parts, including blood vessel and background. However, the current image segmentation algorithm for retinopathy based on image processing technology is still unable to match the segmentation effect of doctors. Especially in the case of exudate and hemorrhage in the image of retinopathy, the segmentation effect is obviously reduced. The common image segmentation methods of retinopathy can be summarized as follows.

$$R = \left(f_n \, o \, f_{n-1} \, L \;\; f_1 \right)\left(I \right) \quad (1)$$

where R was the result of segmentation. I was the input image of retinopathy. f_i was the operation of the image and the synthesis of these operations. For the algorithm based on image processing technology, each operation f was the image processing step defined in an algorithm such as matching filtering, Gabor filtering, *etc*. For the algorithm based on pathological data clustering analysis (PDCA), it could be divided into two steps as below.

$$V = \left(f_k \, o \, f_{k-1} \, L \;\; f_1 \right)\left(I \right) \quad (2)$$

$$R = \left(f_n \, o \, f_{n-1} \, L \;\; f_{k+1} \right)\left(V \right) \quad (3)$$

where V was the feature matrix composed of the feature vectors of each pixel. From f_1 to f_k was the feature extraction of the algorithm, which was also a pre-defined operation. However, from f_{k+1} to f_n was the classifier part. After setting the parameters, these operations could be learned according to the training samples, and the parameter values would be updated during the learning process. For deep learning, all operations of image were obtained through learning.

Image segmentation of retinopathy

Image segmentation of retinopathy was to divide the image into non-overlapping or regularly overlapping small pieces and used the algorithm to process one small piece at a time. After all the image blocks were processed, they were spliced into a whole to complete the segmentation of the retinopathy image. The pixel-by-pixel segmentation algorithm could also be regarded as an image segmentation algorithm, which was a special case of the block window size of 1. The whole image segmentation algorithm could also be regarded as a special case, that was, the block window was the whole image itself. In this way, the segmentation process of the image of retinopathy could be abstracted into the following models.

$$P_{w,w'} = partition\left(I \right) \quad (4)$$

$$R_{P_{w,w'}} = \left(f_n \, o \, f_{n-1} \, L \;\; f_1 \right)\left(P_{w,w'} \right) \quad (5)$$

$$R = partition^{-1}\left(R_{P_{w,w'}} \right) \quad (6)$$

where $partition$ was the function of dividing image I into image blocks. $P_{w,w'}$ was the ordered image block sequence obtained by dividing. w' was the size of the image block to be divided. w was the size of the target window to be segmented from the image block. R_P was the segmentation result of the image block. $partition^{-1}$ was the inverse operation corresponding to $partition$. The segmentation result was spliced according to the sequence of blocks to form a complete segmentation image. The PDCA algorithm was used to analyze the algorithm time complexity of retinopathy image segmentation. In this algorithm, the computation was mainly concentrated in data layer and full connection layer. For the down sampling layer and activation layer, the computational complexity was not on the same order of magnitude as the data, so it was ignored [11]. When the window was not one-dimensional, the full connection layer could not be used because it would destroy the position relationship of the matrix itself. Therefore, only data analysis and manipulation were needed. To study the influence of w and w' on the algorithm complexity, the parameters of PDCA must be fixed. However, due to the need to get different w and w', the fixed network could only let w and w' show a one-to-one correspondence. Therefore, to solve this problem, only the diagnostic network for retinal lesions was used with the front-end part of the network keeping the same. By adjusting the fault diagnosis parameters, the experimental requirements were obtained. Compared with the front data

layer, the calculation amount of a few layers of reverse diagnosis could also be ignored. Therefore, in a forward operation, the time complexity of the diagnosis network of retinopathy was as follows.

$$T\left(w,w'\right)=\sum_{l=1}^{L}r^{l}c^{l}k_{w}^{l}k_{h}^{l}n^{l}\left[O\left(M\right)+O\left(A\right)\right] \quad (7)$$

where $l\left(1\le l\le L\right)$, l, and L were the deepest data layers. r^{l} and c^{l} were the number of rows and columns of input matrix in the diagnosis layer l, so for the first layer, it was the number of rows and columns in w'. k_{w}^{l} and k_{h}^{l} were the width and height of diagnosis core in the diagnosis layer l. n^{l} was the number of diagnosis cores in the diagnosis layer l. $O\left(M\right)$ was the complexity of one floating-point multiplication. $O\left(A\right)$ was the complexity required for this method by doubling the floating-point number. In a fixed diagnosis network of retinopathy, r^{l} and r^{1} were linear. Similarly, r^{l} and c^{l} were also linear. Then the equation would be the follows.

$$T\left(w,w'\right)=\sum_{l=1}^{L}\left(a^{l}\right)^{2}r^{l}c^{l}k_{w}^{l}k_{h}^{l}n^{l}\left[O\left(M\right)+O\left(A\right)\right] \quad (8)$$

where $r^{l}=a^{l}r^{1}$. Then, the time complexity could be expressed as below.

$$T\left(w,w'\right)=r^{1}c^{1}\left\{\sum_{l=1}^{L}\left(a^{l}\right)^{2}r^{l}c^{l}k_{w}^{l}k_{h}^{l}n^{l}\left[O\left(M\right)+O\left(A\right)\right]\right\} \quad (9)$$

The above formulas showed that, for a fixed size diagnosis network of retinopathy, $T\left(w,w'\right)$ was only related to $r^{l}c^{l}$, that was, the size of w'. It was noted that the number of image blocks needed to be divided for a retinopathy image was related to w because the ultimate goal was to segment all pixels in the image, and only w could be segmented each time. For segmentation of a whole retinopathy image, the time complexity was as follows.

$$T_{I}\left(w,w'\right)=\frac{I_{w}I_{h}}{w_{w}w_{h}}w_{w}'w_{h}'\left\{\sum_{l=1}^{L}\left(a^{l}\right)^{2}r^{l}c^{l}k_{w}^{l}k_{h}^{l}n^{l}\left[O\left(M\right)+O\left(A\right)\right]\right\} \quad (10)$$

where I_{w} and I_{h} were the width and height of the retinopathy image. w_{w} and w_{h} were the window width and height of each output segmentation result. w_{w}' and w_{h}' were the width and height of the retinopathy image block. Equation 10 demonstrated that $w_{w}=w_{h}=1$, and its time complexity was very high, especially for the high-resolution image of retinopathy. The time complexity of the whole image segmentation algorithm was the lowest. The segmentation algorithm based on PDCA was between the two. The time complexity of retinal vascular segmentation algorithm based on PDCA was much lower than that of pixel-by-pixel segmentation algorithm, especially when w_{w} and w_{h} values were relatively large. In terms of time complexity, algorithms based on PDCA were more complex than whole image segmentation. However, in terms of spatial complexity, it was much lower than whole image segmentation, which was because the forward operation of each retinopathy diagnosis network only required running one image block rather than the entire image. Meanwhile, the image block size could be adjusted according to the performance of the program deployment machine to avoid network failure caused by insufficient memory or insufficient display memory. A large number of samples were needed to complete the training, while overfitting the training data should be avoided to help cross the local best point. The Visual Geometry Group (VGG) network (http://www.robots.ox.ac.uk/~vgg/research/very_deep/) and the depth residual network (https://arxiv.org/abs/1512.03385) were trained on the ImageNet database (http://www.image-net.org/), which currently contained more than 1.4×10^{7} labeled samples of retinopathy images. For the image of retinopathy, there were few labeled images in the international open databases. Digital Retinal Images for Vessel Extraction (DRIVE) (https://www.isi.uu.nl/Research/Databases/DRIVE/) only contained 40 labeled retinopathy images. Challenge in

Automated Segmentation of the Retinal Vasculature (CHASE_DB11) (https://blogs.kingston.ac.uk/retinal/chasedb1/) only had 28 labeled retinopathy images. STructured Analysis of the Retina database (STARE) (https://cecas.clemson.edu/~ahoover/stare/) had 20 labeled retinopathy images. Moreover, the samples in these databases were taken under different conditions. In DRIVE database, 33 pictures were of normal people with the remaining seven pictures as mild retinopathy, while most of the pictures in STARE were of retinopathy [12]. The information contained in the samples was quite different, which was not conducive to network convergence. In this case, training the whole image segmentation diagnosis network of retinopathy required more sample expansion and faster convergence speed of the neural network to reduce the overfitting problem caused by too many network iterations. However, for the network of image segmentation, these problems did not exist. During training, the image block of retinopathy input into the diagnosis network of retinopathy was obtained by sampling from the original image, and the number of samples could be controlled at will, which meant that, based on the method of PDCA, each original sample could get at least $\frac{I_w I_h}{w_w w_h}$ samples, and the samples could be overlapped, which would produce more samples. Enough samples allowed the design of a very deep diagnostic network of retinopathy to extract more abstract features. The segmentation method based on image segmentation could change the data distribution of training samples, so that the depth diagnosis network could learn more useful features. An original retinal image could be expressed as a distribution p , then the distribution of the blocks could be expressed as $p_{w'}$. The image block was obtained by sampling from the original image. In other words, the distribution p could be obtained by synthesizing the limited sampling $p_{w'}$ as follows.

$$p = g\left(p_{w_1'}, p_{w_2'}, \mathrm{L}\ , p_{w_n'}\right) \tag{11}$$

$$w_i' - w_j' \neq \phi, i \neq j \tag{12}$$

$$\bigcup_{i=1}^{n} w_i' = I \tag{13}$$

$$w_I' \mathrm{I}\ \ w_j' = \phi, i \neq j \tag{14}$$

where ϕ was an empty set, meaning no pixels. In this way, equation 12 restricted all image blocks from having two identical image blocks, which was a prerequisite for ensuring that only a limited number of samples could sample the entire image. Equation 13 ensured that a limited number of image blocks could be synthesized into the whole image. When the restriction equation 14 was added, the method of image segmentation became non-overlapping between any two different image blocks. When training the diagnosis network of retinopathy, full image segmentation was needed to take the whole image as the input of the network or use some methods to cut to expand the sample. But, in essence, it was still full image segmentation. Based on the segmentation of image blocks, it was not necessary to put all the blocks in the image into the diagnosis network of retinopathy with equal frequency for training. If the designed network was not ideal in some local segmentation effects such as in the video disk, the number of training samples of image blocks including the video disk could be increased to enhance the local learning of the network. The method of image segmentation was different in training stage and testing stage. In the training stage, the training dataset was generated by image segmentation to improve the segmentation effect of the diagnosis network of retinopathy. Therefore, in this stage, it was not simply to block the image in a non-overlapping area, but to reduce the proportion of false segmentation as much as possible. The areas prone to misclassification in the diagnosis network of retinopathy often occurred in the

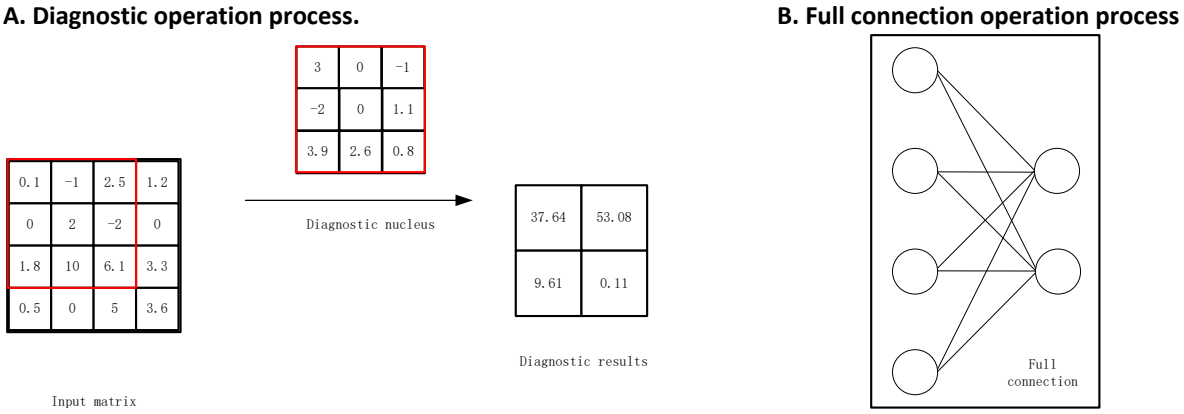

Figure 1. The diagnostic and full connection operation processes.

edge of blood vessels, the edge of the optic disc, and other areas. Therefore, the sampling strategy adopted in this research was to randomly sample a small amount, and then mostly sample the vascular area, the edge of the optic disc, and the area with lesions. In the test phase, all the blocks needed to meet equation 13 to ensure that all the pixels could be segmented. Therefore, during testing, the nonoverlapping block method could be used and followed by inputting the network block by block. Concatenating the results or overlapping method could be used.

Design of diagnosis network for retinopathy

The diagnosis operation in the diagnosis network of retinopathy was shown in Figure 1A. The diagnosis core slide from the starting position to the last position of the input matrix with a fixed window size. Every time the position was changed, a diagnosis result would be produced. Therefore, for the input two-dimensional matrix, the sequential sliding diagnosis operation would make the output result matrix retain its spatial position information, and its relative position would not change, which was very important for image segmentation. However, in the diagnosis network of retinopathy, the full connection layer was usually used at the end (Figure 1B). Each neuron in the latter layer was connected with each neuron in the former layer. Each connection meant that a matrix operation was needed. For the retina image, which had a high resolution, the computation was too large to run on a general computer [13]. This kind of full connection was designed for the task of image classification, which could make the final feature vector more fully nonlinear, thus forming more abstract features. However, the final feature vector dimension of image classification was very low, even if a full join operation was used, its computation was still at an acceptable level. Meanwhile, the full join operation needed to transform the feature into vector form, which would lose the information of the relative position of the elements in the matrix. The number of neurons in the full connection operation was obtained by calculating the length of the current eigenvector, which meant that, when training the network, one network could only accept and process one size of input image and needed to be retrained after changing the size. By replacing the last full connection operation with the diagnosis operation, the original common retinopathy diagnosis network became the whole retinopathy diagnosis network. Full connection operation and diagnosis operation were linear transformation of input data, so it was feasible to replace full connection with diagnosis. Meanwhile, in the process of diagnosis, the output was still a two-dimensional matrix and could maintain the original relative position. In addition, compared with the full connection, the diagnosis operation was a kind of sparse connection. Each element of the output layer was only connected with the local element of the input layer, and the diagnosis core

controlled the number of local elements. This sparse connection could greatly reduce the amount of computation. The mapping relationship between output layers was obtained by sharing a diagnostic kernel with fewer parameters. Compared with fully connected layers, the parameters included in the model would be greatly reduced, which was beneficial for network training and could reduce the possibility of overfitting problems [14, 15].

Image preprocessing of retinopathy

In the retinopathy dataset, the color and contrast of different retinal images were quite different. A diagnostic network for retinopathy was needed to learn the invariant features of color and contrast changes for regression. Therefore, by adding preprocessing steps, it was possible to enhance the regression friendly features in retinal images, remove interference from different colors and contrasts, reduce the number of layers in the retinopathy diagnosis network, and improve the accuracy of regression. The retina area in the image was aligned to the center of the image and cut off the surrounding background area as much as possible. Due to the different background areas around the retina region in different cameras in the dataset, there was a problem of inconsistent normalized size of the retina region. Therefore, it was necessary to make the retina region as large as possible in the image. In addition, maximizing the retinal area could reduce the influence of background as image storage must be rectangular, while the retinal region was disc shaped. Therefore, the background area should be pure black, but a small amount of random noise was inevitably generated during shooting and in the background area. Gaussian filtering was used in this study to smooth the random noises as follows.

$$g(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (15)$$

where x and y were the distance between the pixels and the horizontal and vertical coordinates of the filtering center point. σ was the parameter controlling the smoothing degree during filtering. The larger the σ, the smoother the image. The dataset used in this research had high resolution and needed to filter out noise as much as possible, so the setting was relatively large. Since the surrounding pixels were all 0, the noise pixels in the image filtered by Gauss were very low. Only a reasonable threshold was needed to segment the retina region and background region. After locating the upper, lower, left, and right boundaries of the retina area in the image, the retina area was cut out as the region of interest. Since different cameras were used to take the images in the database, preprocessing was required using PDCA method, which could standardize the histogram of each image to a similar level. After the clustering analysis of pathological data, the image needed to be cut to a uniform size before it could be used as input for training. Since the full connection layer was used in the network, the input size of the layer needed to be fixed for calculation.

Action layer of diagnostic network for retinopathy

The fully connected layer and dropout layer were shown in Figure 2. The role of the full connection layer was to project the feature map space extracted from the front layer of the neural network into the target space. The number of output units of the last full connection layer was generally related to the target of the neural network. If it was a classification, it was the same as the number of categories. The calculation process of full connection layer could be expressed as follows.

$$y = wx \quad (16)$$

where w was the weight matrix connecting input and output in this layer. Offset had been included in w. x was column vector, and x_0 = 1. In the diagnosis network of retinopathy, because the output of the anterior layer was a two-dimensional matrix, all its elements needed to be arranged as column vectors. There were connections between all neurons in two fully connected layers, while there were no

connections between neurons in the same layer. So there were many redundant parameters in this connection because not all neurons represented parameters that were related to all input elements. Many parameters could easily lead to overfitting, so a penalty term was needed to the loss function. However, this redundant parameter was also beneficial for learning more complex features, which was because the diagnostic layer was a local connection within the scope of the diagnostic template and could only be calculated using data within a local range. The fully connected layer calculated using all input data, making feature fusion more thorough.

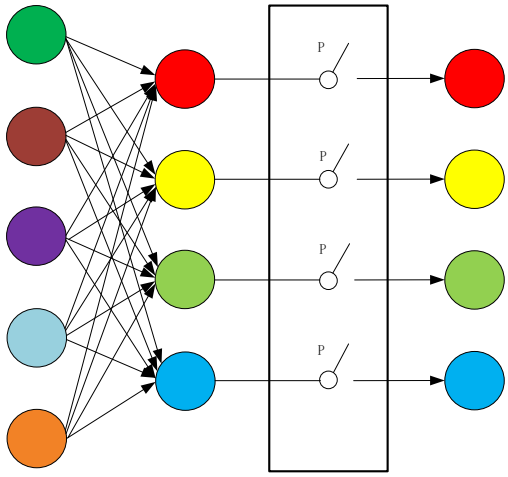

Figure 2. Schematic diagram of full connection layer and dropout layer.

The function of dropout layer was to randomly ignore some outputs with a set probability, thereby preventing overfitting and increasing the generalization ability of the network. The number of neurons in the full connection layer was often large, so in the neural network, the characteristics of the final output might be determined by a few neurons. That was, these few neurons had a large weight in the weight matrix, while other neurons had a small weight. After adding the dropout layer, the output of some neurons was randomly discarded with a certain probability. When the effective output was discarded, the loss function of the neural network would be very high. According to the optimization principle, it would learn the effective features in the current situation, and the discarded neuron output was completely random. Therefore, the neural network was forced to make the output of each neuron effective characteristics. The calculation process of dropout layer was as follows.

$$y = f\left(x', p\right) \tag{17}$$

where p was the set drop probability. The larger the p, the more output was discarded in the dropout layer, and the more sparse the output resulted. The output of discarded neurons was 0, which would force each feature of the output to have stronger expression ability. However, it also destroyed the combined features of the input layer, so that the model could not get more features, and the overall expression ability would decline. Moreover, the setting was also related to the number of neurons contained in the layer. When there were many neurons, a large discard probability could be appropriately selected. Partial output was discarded, causing only some parameters in the entire connection layer to participate in the operation, which reduced the parameters in the entire neural network and was beneficial for preventing overfitting and improving the generalization ability of the neural network. The dropout layer generally only worked in the training phase but would be removed in the test phase. All effective features in the network could be combined to obtain better features, thereby improving the classification performance. Euclidean distance loss function layer was the deepest layer of the whole diagnosis network of retinopathy, which mainly calculated the difference between the output value of the diagnosis network of retinopathy and the value of artificial annotation. The main reason for using the Euclidean distance loss function instead of the cross entropy function was to predict the severity of retinopathy rather than simply classify it into corresponding classes [16, 17]. Moreover, when the predicted value was different from the

marked value, the loss function value of different predicted values should be different. The Euclidean distance loss function was expressed as follows.

$$Loss = \frac{1}{2N}\sum_{n=1}^{N} \| \hat{y}_n - y_n \|_2^2 + \lambda \| w \| \quad (18)$$

where *N* was the number of samples. y was the predicted value of samples. $\hat{y}_n$ was the labeled value of samples. λ was the penalty factor. *w* was the parameter in the model.

Comparative analysis

The comparative experiments were conducted using two E5 2620 V3 CPUs, GTX Titan X graphics processor, 128 GB RAM under Ubuntu 16.04LTS 64-bit operating system. Caffe was employed as deep learning library with cuDNN5.1 acceleration tool. The model training mainly relied on the ImageNet database with some data from DRIVE, CHASE_DB11, and STARE for verification and testing. A total of 30,000 images were randomly selected from the ImageNet database for model training and 5,126 images for model testing, while 40, 28, and 20 images were screened out from the DRIVE, STARE, and STARE databases, respectively, for model verification. The dataset used in this study included 35,126 images of retinopathy with 25,810 (73.48%) in category 0 that meant no obvious retinopathy and no abnormal blood vessels or exudation in the fundus; 2,443 (6.96%) in category 1, indicating mild non-proliferative retinopathy, presenting as retinal microhemangiomas or a small number of bleeding points; 5,292 (15.07%) in category 2 as moderate non-proliferative retinopathy with retinal hemorrhage, hard exudation, and cotton-like spots visible, but not reaching the severe standard; 873 (2.48%) in category 3, indicating severe non-proliferative retinopathy with ≥ 20 retinal hemorrhage in each quadrant or at least 2 quadrants showing venous beaded changes or 1 quadrant demonstrating retinal microvascular abnormalities; 708 (2.01%) in category 4, indicating proliferative retinopathy with the appearance of retinal neovascularization, vitreous hemorrhage, or tractional retinal detachment. For databases of DRIVE, CHASE_DB11, and STARE, their annotation systems differed from the classification criteria of this study, therefore, category alignment was achieved through secondary annotation by experts during data fusion. Since the goal of this research was to distinguish normal and proliferative retinopathy, only category 0 and category 4 needed to be included. So category 1 accounted for 97.33%, while category 4 only accounted for 2.67%. The number of these two types of samples in the dataset was very unbalanced. Therefore, image transformation methods such as mirror image, rotation, translation, and brightness change were applied to make the two types of samples balanced. In this study, proliferative retinopathy samples were defined as negative samples, while normal samples were defined as positive samples. The test samples included 39,533 normal samples and 1,206 proliferative retinopathy samples. All sample sizes were normalized to 512 × 512. The convolution kernels were all set to 3 × 3 (Figure 3). The initial learning rate was 1×10^{-6}, and Nesterov was used as the network optimization method. The number of samples input each time during training was 48, and the probability of dropout layer setting was 0.5. After preprocessing the image data, the parameters for the retinopathy diagnostic network were initiated. The trained retinal image segmentation model was used for initialization of blocks 1, 2, and 3, while the other layers were initialized randomly. Block 1 was mainly used to extract the basic features such as edges and textures of the image, while block 2 was used to extract the mid-layer features such as the shape and direction of blood vessels in the image, and block 3 was used to extract high-level features such as the characteristics of retinopathy in images. Image transformation methods such as mirroring, rotation, translation, and brightness change were used for data augmentation to balance normal and proliferative retinopathy samples. After shuffling the samples, they were input into the diagnostic network one by one for training. The accuracy rate was taken as the measurement

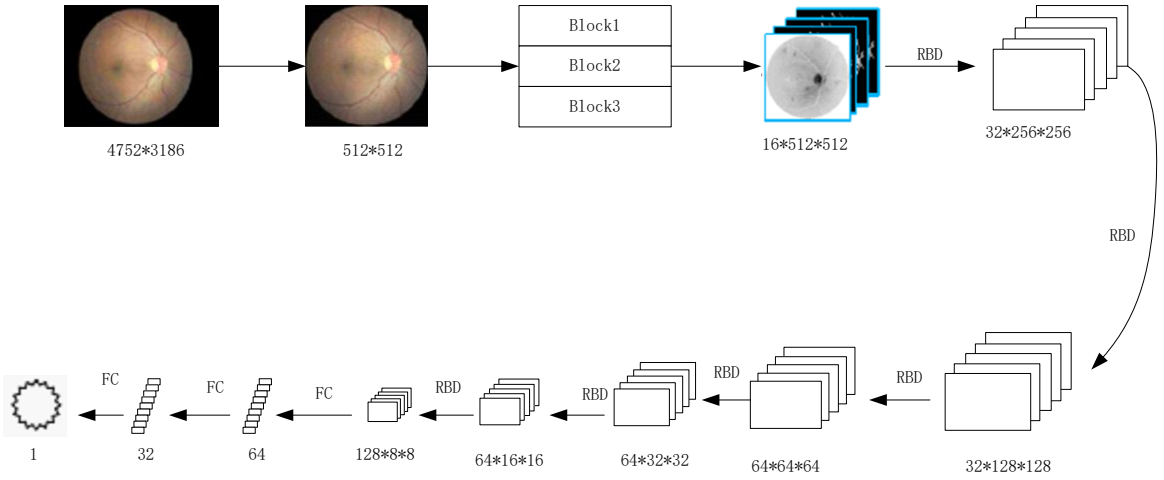

Figure 3. Diagnostic network framework.

standard. The model was verified using the images retrieved from the DRIVE, CHASE_DB11, and STARE databases to evaluate the performance and generalization ability of the model. A total of 5,126 images retrieved from the ImageNet database were used for model testing. Ten independent experiments were performed to verify the stability and generalization ability of the model using 10-fold cross-validation. For each experiment, the training set and the test set were randomly selected from the dataset at a ratio of 8:2. To avoid data bias, stratified sampling was adopted to ensure that the proportion of each type of sample was consistent with that of the original dataset. In each experiment, four types of samples were dynamically enhanced, and three times the number of samples was generated through mirroring, rotation, and brightness adjustment. The network weights were reinitialized in each experiment to avoid the interference of parameters from the previous training. Different random seeds (0 - 9) were used to control the data partitioning and enhancement process to ensure the reproducibility of the experiment. Each experiment followed the standardized process, and the training rounds were set at 200 rounds.

Statistical analysis

SPSS 26.0 (IBM, Armonk, New York, USA) was employed for statistical analysis. An independent sample t-test was used to verify whether there was a significant difference in diagnostic accuracy between PDCA and image processing-based retinopathy diagnosis techniques. The *P* value less than 0.05 was defined as a significant difference between the two groups.

Results and discussion

The experimental results showed that the accuracy of the retinopathy diagnosis model based on image processing was less than 70% with the lowest accuracy of only 35.45% and the average accuracy of ten experiments as 58.29%. Such results might be because traditional methods relied on manual design of features, making it difficult to effectively capture the complex features of retinopathy, while they had poor robustness against interference factors such as uneven illumination and noise. In contrast, the proposed pathological data cluster analysis diagnostic technique performed significantly better. The diagnostic accuracies of the ten experiments were all stable at above 90% with an average accuracy rate of 93.82%. The results demonstrated that there was a significant difference in the diagnostic accuracy between the retinopathy diagnosis technologies based on

Table 1. Comparison of proposed method with traditional image processing method.

Experiment	The diagnostic accuracy of retinopathy based on		*P* value
	Pathological data clustering analysis	Image processing	
1	93.09%	68.76%	< 0.05
2	92.63%	54.69%	< 0.05
3	95.02%	65.49%	< 0.05
4	94.26%	63.28%	< 0.05
5	92.98%	54.26%	< 0.05
6	93.45%	35.45%	< 0.05
7	96.34%	64.37%	< 0.05
8	94.16%	48.49%	< 0.05
9	93.68%	63.28%	< 0.05
10	92.64%	64.86%	< 0.05

pathological data cluster analysis and based on image processing ($P < 0.05$) (Table 1). The results indicated that the PDCA technology could significantly improve the accuracy of retinopathy diagnosis, which might be because the proposed model could automatically mine complex lesion features and had an end-to-end training mechanism. Traditional diagnostic techniques for retinopathy suffer from difficulties in feature extraction and low diagnostic accuracy. This study aimed to break through the limitations of traditional diagnostic methods through PDCA technology and proposed a retinopathy diagnosis technique based on PDCA. The results of ten independent experiments showed that the average accuracy rate of the proposed technology reached 93.82%, significantly superior to traditional image processing methods, proving its effectiveness. In addition, another advantage of segmentation methods based on PDCA was that it could control the sampling position of image blocks, thereby reducing the problem of retinopathy diagnosis networks biased towards a certain category due to class imbalance. In retinopathy images, it was necessary to segment blood vessels and background, but the proportion of these two types of pixels in retinopathy images was severely imbalanced. In the three current available databases, the background class contained more than 85% of pixels. Therefore, for the whole image segmentation algorithm, additional processing might be required to overcome the problem of class imbalance. For algorithms based on PDCA, sampling could be done in a certain way to reduce areas in the image where there were less or no blood vessels. For example, in the boundary part of the image, there were basically no blood vessels. So, sampling of this part could be omitted. This study adopted an adaptive sampling strategy, which excluded areas without blood vessels at the image boundaries and prioritized sampling areas with dense blood vessels, optic disc edges, and lesion areas, resulting in a significant increase in the proportion of vascular pixels in the training samples. Compared with traditional segmentation methods, this strategy significantly improved the effectiveness of vascular segmentation and effectively reduced the impact of category bias on the final diagnosis. The retinopathy diagnosis technology based on PDCA had more network parameters for related analysis and could capture more complex lesion patterns. The design of the size-invariant feature map avoided the loss of small blood vessels in down sampling. However, this study still had certain limitations. The sample size of international public databases was relatively small. Although the sample size had been expanded through data augmentation, the generalization ability in extremely rare lesion types still required more clinical data for verification. In the future, multi-center medical

imaging data can be combined to further optimize the network's ability to identify rare cases.

References

1. Elta G, Enestvedt B, Sauer B, Lennon A. 2018. ACG clinical guideline: Diagnosis and management of pancreatic cysts. Am J Gastroenterol. 113(4):464-479.
2. Buxey K, Sia C, Bell S, Wale R, Wein D, Warrier S. 2017. Clostridium colitis: Challenges in diagnosis and treatment. Anz J Surg. 87(4):227-231.
3. Vuurberg G, Hoorntje A, Wink L, van der Doelen B, van den Bekerom M, Dekker R, *et al*. 2018. Diagnosis, treatment and prevention of ankle sprains: Update of an evidence-based clinical guideline. Brit J Sport Med. 52(15):956.
4. Zhou Q, Huang S, Zhang F, Li S, Li C, Xi Y, *et al*. 2017. Micrornas: A novel potential biomarker for diagnosis and therapy with non-small cell lung cancer. Cell Proliferat. 50(7):e12394.
5. Casal C, Infantes J, Risalde M, Díez-Guerrier A, Domínguez M, Moreno I, *et al*. 2017. Antibody detection tests improve the sensitivity of tuberculosis diagnosis in cattle. Res Vet Sci. 112:214-221.
6. Sinha S, Bhandari AK, Kumar R. 2023. Low quality retinal blood vessel image boosting using fuzzified clustering. IEEE Transactions on Artificial Intelligence. 5(6):3022-3033.
7. Kumari AA, Bhagat A, Henge SK, Mandal SK. 2023. Automated decision making ResNet Feed-Forward Neural Network based methodology for diabetic retinopathy detection. Int J Adv Comput Sci Appl. 14(5):1-12.
8. Dayana AM, Emmanuel WS, Linda CH. 2023. Feature fusion and optimization integrated refined deep residual network for diabetic retinopathy severity classification using fundus image. Multimedia Systems. 29(3):1629-1650.
9. Zhu Y, Shang Y, Shao Z, Guo G. 2017. Automated depression diagnosis based on deep networks to encode facial appearance and dynamics. IEEE T Affect Comput. 9(4):578-584.
10. Sepantafar M, Maheronnaghsh R, Mohammadi H, Radmanesh F, Hasani-Sadrabadi M, Marzieh E, *et al*. 2017. Engineered hydrogels in cancer therapy and diagnosis. Trends Biotechnol. 35(11):1074-1087.
11. Tsochatzidis L, Zagoris K, Arikidis N, Karahaliou A, Costaridou L, Pratikakis I. 2017. Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach. Pattern Recogn. 71:106-117.
12. Liu M, Zhang J, Adeli E, Shen D. 2017. Landmark-based deep multi-instance learning for brain disease diagnosis. Med Image Anal. 43:157-168.
13. Peng K, Ma L, Zhang K. 2017. Review of quality-related fault detection and diagnosis techniques for complex industrial processes. Automatica Sinica. 43(3):349-365.
14. Fan Z, Zhou S, Garcia C, Fan L, Zhou J. 2017. pH-responsive fluorescent graphene quantum dots for fluorescence-guided cancer surgery and diagnosis. Nanoscale. 9(15):4928.
15. Feng X, Moy A, Nguyen H, Zhang Y, Zhang J, Fox M, *et al*. 2018. Raman biophysical markers in skin cancer diagnosis. J Biomed Opt. 23(5):1-10.
16. Wang M, Gendreau J, Gemelas J, Capulong D, Lau C, Mata-Diaz S, *et al*. 2017. Diagnosis and management of malignant melanoma in store-and-forward teledermatology. Telemed E-Health. 23(11):877-880.
17. Boni A, Basini A, Capolupo L, Innocenti C, Corti M, Cobianchi M, *et al*. 2017. Optimized pamam coated magnetic nanoparticles for simultaneous hyperthermic treatment and contrast enhanced MRI diagnosis. Rsc Adv. 7(70):44104-44111.