Microaneurysms Detection in Color Fundus Images based on Naive Bayesian Classification

Shengchun Long¹, Jiali Chen¹*, Ante Hu¹†, Haipeng Liu²†, Zhiqing Chen³† and Dingchang Zheng²†

Abstract

Background: As a major complication of diabetes, diabetic retinopathy (DR) is a leading cause of visual impairment and blindness due to delayed diagnosis and intervention. Microaneurysm appears as the earliest symptom of DR. Accurate and reliable detection of microaneurysms in color fundus images has great importance for DR screening.

Methods: A microaneurysms detection method based on Naive Bayesian classifier is proposed for the early diagnosis of DR. First, blood vessels were enhanced and segmented using the analysis of eigenvalues of Hessian matrix. Next, with blood vessels excluded, microaneurysm candidate regions were obtained according to shape characteristics. After image segmented to patches, the features of each microaneurysm candidate patch were extracted, and each candidate patch was classified using Naive Bayes into microaneurysm or non-microaneurysm. The proposed algorithm was trained and tested on e-ophtha MA database, and further tested on another independent DIARETDB1 database. Results of microaneurysms detection on the two databases were evaluated on lesion level and compared with existing algorithms.

Results: The proposed method has achieved better performance compared with existing algorithms. On e-ophtha MA and DIARETDB1 databases, the area under curve (AUC) of receiver operating characteristic (ROC) curve was 0.845 and 0.831, respectively. The free-response ROC (FROC) score on the two databases was 0.362 and 0.207, respectively.

Conclusions: The proposed method based on Naive Bayesian classification of image patches can effectively detect microaneurysms in color fundus images and provide an effective scientific basis for early clinical DR diagnosis.

Keywords: color fundus image; microaneurysms detection; patch; features extraction; Naive Bayesian classifier

Background

Diabetes mellitus has reached epidemic levels worldwide. In 2019, approximately 463 million adults (20-79 years) were living with diabetes; by 2045 this will rise to 700 million[1]. Diabetic retinopathy (DR) is the most common complication of diabetes. Among people with diabetes, the incidence of DR is approximately one third. DR is the leading cause of visual impairment and blindness among working age adults[2]. Clinical evidence shows that early diagnose and clinical intervention of DR can effectively reduce the risk of DR-related vision loss[2].

DR could be diagnosed from resultant abnormalities observed in color fundus images, including microaneurysms (MAs), hemorrhages (HMs), hard exudates (HEs)
and cotton wool spots (CWSs), as shown in Fig. 1. MAs often appear in the fundus posterior pole and as red or dark red isolated small dots with clear borders in color fundus images. The diameter of MA is usually between 15µm and 60µm, or larger, but seldom exceeds 125µm. Compared with other DR-related lesions such as HE and CWS, MA is difficult to detect due to its tiny structure. MA is the earliest symptom of DR. Therefore, accurate and reliable detection of MA in color fundus images is significant for the early screening and diagnosis of DR.

Despite the clinical significance, currently the automatic detection of MA is still difficult, with manual extraction widely used. At present, in clinical DR grading based on color fundus images, mild NPDR (with MA only) and non-DR are classified into the same category[3]. There is a lack of reliable algorithms for MA detection to distinguish early DR.

Computer-aided automatic analysis of color fundus images is highly efficient and has been widely applied in the diagnosis of DR[4]. However, currently studies on computer-aided MA detection are limited. Srivastava et al.[5] proposed a multi-kernel learning to detect MA and HM. Akram et al.[6] extracted dark regions of fundus images by using Gabor filter, and detected MA by combining Gaussian Mixture Model (GMM) with Support Vector Machine (SVM). Ren et al.[7] segmented blood vessels using morphological top-hat transform, extracted candidates of MA by using multi-scale top-hat transform and connected component analysis after removing blood vessels, and finally detected MA by Adaptive Over-sampling Boosting (ADOBoosting). Dai et al.[8] and Dashtbozorg et al.[9] extracted MA candidate regions by analyzing image gradients, and applied the Random Under-Sampling Boosting (RUSBoosting) classifier to achieve the detection of MA. Wu et al.[10] used peak detection and region growing to get MA candidates, then detected MA based on K-Nearest Neighbor (KNN). Javidi et al.[11] and Wei et al.[12] detected MA by dictionary learning. Several other methods of MA detection based on deep learning[13, 14, 15] were proposed.

For clinical application, some limitations of existing algorithms need to be overcome. MA was often detected together with HM recognized as red lesions[16]. The inaccuracy of blood vessels detection could affect MA detection[7]. Deep learning-based algorithms of MA detection were complex with high requirements on computational resources. Moreover, some studies only used one database for testing[12, 15]. Large-scale data on heterogeneous patient cohorts are needed for fully validation.

Considering the limitations of existing methods, in order to achieve accurate MA detection in low computational resources, we aim to propose a new algorithm based on Naive Bayesian classification, and validate it on two different databases. The main framework of this work is depicted in Fig. 2.

Results

ROC curves of the proposed method

The ROC curve and corresponding AUC of MA detection results achieved by 3 classifiers (SVM, Naive Bayes and KNN) are shown in Fig. 3. It can be seen that Naive Bayesian classifier performs better than SVM and KNN, with the AUC of 0.845 and 0.831 on e-ophtha MA and DIARETDB1 databases respectively. Therefore, Naive Bayes was selected for MA detection in this study.
Comparison with different methods

Fig. 4 shows FROC curve of the proposed MA detection method on e-ophtha MA and DIARETDB1 databases compared with existing algorithms.

Table 1 and Table 2 show comparison of Sensitivity between the proposed MA detection method and existing algorithms at different FPI values on e-ophtha MA and DIARETDB1 databases, respectively.

It can be seen from Fig. 4, Table 1 and Table 2 that the proposed method has better performance on MA detection compared to other algorithms. Especially, although the algorithm proposed by Eftekhari et al.[13] performed very well on FROC on e-ophtha MA database, a two-step CNN was used, which was complex and took up a lot of resources.

Typical examples analysis

Typical examples of MA detection results of our experiments on the two databases are shown in Fig. 5 and Fig. 6.

In DIARETDB1 database, MA detection results correspond to different confidences are shown in different colored circles in Fig. 7(a). The evaluation shown in Fig. 7(b) is based on the confidence $\geq 75\%$. As can be seen from Fig. 7, with a standard of ground truth confidence $\geq 75\%$, a part of FPs belong to those labeled with a confidence of less than 75%. For example, the two FPs shown in Fig. 6(b) have a confidence of 50% (shown in Fig. 7(a)).

Discussion

In this study, a novel method is proposed for microaneurysms detection in color fundus images. The effectiveness of the method is proved through experiments on different databases.

Advantages

The accurate segmentation of BV could improve the accuracy of MA detection. It has been proven that using Jerman’s enhanced function[17] can achieve the accuracy of 95.4% in BV segmentation. For the features selection of MA candidate regions, directional local contrast (DLC) and gradient features were creatively used in this study. Especially, DLC was the first time applied in MA detection, which showed promising ability in differentiating MA from other structures (Fig. 8 and Fig. 9). The DLC of center point of each patch, the mean gradient of the entire patch, and the mean gradient on the boundary of MA candidate region were analyzed for MA detection. Through the analysis of characteristics of MA, it is found that the DLC and gradient of MA and non-MA patch in $I_{CLAHE}$ is significantly different. Compared with existing MA detection algorithms based on features extraction[10, 18, 19], using the DLC and gradient features can improve the accuracy of MA detection. In addition, compared with existing MA detection algorithms using deep learning[13, 14, 15] and others[6, 8, 11, 12], operations performed in the proposed MA detection method are easy to calculate and occupy less resources.
Typical misclassification analysis

Despite the high accuracy in MA detection achieved by the proposed algorithm, some misclassified examples affecting accuracy could be seen in Fig. 5 and Fig. 6. There are several causes for misclassification. For FPs, from the white circles in Fig. 5(c) and Fig. 6(c), it can be seen that image noises could lead to erroneous MA detection. Throughout the experiment, most FPs were caused by image noises interference.

For FNs, firstly, as shown by the red circle in Fig. 5(b), the MA has an irregular shape and a relatively large area. It’s a typical example that leads to misjudgment of the algorithm based on the shape features. As the shape features used to distinguish between MA and HM (shown in Table 3), that missed MA might be considered as HM by the classifier. Secondly, as shown by the red circle in Fig. 6(c), this MA was missed in the process of MA candidate regions extraction. Although the applied method for extracting MA candidate regions was efficient, it might miss a small part of MA in candidate regions, which caused the increase of FN in the final MA detection results and affected the overall accuracy of MA detection. The main reasons are as follows: In the first case, MA was attached to BV and removed with it. This case is relatively rare, because MA usually appears at the end of BV which was isolated from the main structure of BV and preserved as an independent slender structure during BV removing, as shown in Fig. 10(d) and Fig. 10(e). In the second case, MA was discarded together with the attached slender segment of BV end in the step of removing slender structures. Thus, the threshold for removing slender structures is important. If it was set too small, some MAs would be removed together. If it was too large, some slender and small BV might still be left. In another case, when MA characteristics are not obvious, as shown by the red circle in Fig. 6(c), this missed MA is too small and has low contrast to background, which looks more like image noise. It is difficult to judge whether it is MA.

MA detection is a difficult task that there is disagreement between different experts in the diagnosis of MA, as shown in Fig. 7. As the DIARETDB1 database shows, from the observation of the image in Fig. 6, it can be seen that the two FPs in Fig. 6(b) are very similar to actual MA, while only two experts label them as MA (confidence of 50%). Because the MA structure is too small, unaided eye observation is also susceptible to image quality. From the viewpoint of clinical application, the algorithmic prediction and clinical diagnosis could be mutually complementary. If there is MA detected in an image, the ophthalmologist can further diagnose whether the image corresponds to DR by MA detection result. Therefore, some FN on lesion level will not affect the MA detection performance on image level, and has little impact on clinical diagnosis of ophthalmologist. Or a lower confidence level can be taken as ground truth for ophthalmologists to screen. Thus, the application of computer-aided MA detection in assisting clinical DR screening deserves further investigation.

Limitations

1) One of the main limitations of this study is the interference of image noises on MA detection. It’s difficult to distinguish MA and image noises in current method.
2) The scale of data for experiment is another important limitation. Two databases with totally 163 (74 in e-ophtha MA and 89 in DIARETDB1) images were used for test. The physiological and pathological factors that might affect MA detection could not be considered in this pilot study.

Future works
For the current limitations above, in the subsequent research, the image denoising operation in preprocessing progress should be improved. In addition, to improve the generalization of the method, more databases that cover more heterogeneous samples are needed for as fully validation as possible, including real samples from hospitals for clinical validation.

Conclusions
Naive Bayesian classification is more accurate than SVM and KNN for MA detection in color fundus images. The proposed computer-aided algorithm showed higher accuracy than existing algorithms compared on the two databases, therefore is promising for clinical application for clinical DR diagnosis.

Methods
This study uses Naive Bayesian classification to detect MA on preprocessed and segmented image patches. Naive Bayesian classification is simple and powerful therefore has been widely used in machine learning algorithms of spam processing, document classification, disease diagnosis and other aspects.

Firstly, color fundus images were preprocessed to enhance the image quality. Then blood vessels (BVs) were accurately segmented by using an improved enhanced method based on Frangi’s[20]. With BVs eliminated, candidate regions of MA were then extracted. Next, processed images were divided into patches according to the location of the MA candidate regions. Features were extracted from each candidate patch. Finally, based on the features, candidate patches were classified using Naive Bayesian classification to differentiate the MA and non-MA. The process is illustrated in Fig. 2. To comprehensively evaluate the algorithm on different patient cohorts, the proposed method was validated on two independent databases (e-ophtha MA and DIARETDB1).

Preprocessing
BV, HMs, and MAs all appear dark red in color fundus images (as shown in Fig. 1). In order to make the dark red regions more obvious, preprocessing progress was performed in the Field of View (FOV).

Compared to the other two channels, the green channel has higher contrast between target area and background, therefore contains the most abundant information. Thus, the first step of preprocessing was to extract the green channel image $I_g$. Then median filter on the green channel image $I_g$ was applied to obtain the background $I_{bg}$. The filter size was larger than the maximal BV width in the fundus image. In this study, the filter size of 50*50 was selected. Finally, the result derived by the median filter $I_{bg}$ was subtracted from the green channel image $I_g$ to eliminate the background and achieve the image with the effect of shade correction,
named $I_{eq}$. The main structure of BVs in the shade-corrected image $I_{eq}$ were clear. Other dark regions that include HM and MA in the original image were also highlighted in $I_{eq}$. But the overall image was dark with low contrast. In order to make the image clearer, the mean gray value of the green channel image $I_g$ in FOV was added on image $I_{eq}$ to maintain in the same gray range as the green channel image. As follows:

$$I_{pre} = I_{eq} + \text{mean}(I_g)$$  \hspace{1cm} (1)

The result of preprocessing ($I_{pre}$) is shown as Fig. 10(b) where BV can be obviously observed.

**Segmentation of blood vessels**

For accurate MA detection, it is important to separate BV and MA which are similar in color and brightness. Thus, BVs were segmented and eliminated to minimize the influence on MA detection. The most widely used BV segmentation methods are machine learning[21] and deep learning[22, 23], which have high accuracy but are complicated and time-consuming. Inspired by Jerman et al.’s work[17], a method of image enhancement based on Hessian matrix eigenvalue analysis was developed to achieve simple, fast, and accurate BV segmentation.

**Blood vessels enhancement**

So far, the most commonly used method for BV enhancement is the Frangi’s[20] enhancement function. This study combined the BV enhancement methods proposed by Frangi and Jerman et al., and constructed a response function to enhance BV by analyzing the eigenvalues of Hessian matrix.

For a preprocessed image $I_{pre}(x, y)$, firstly, Gaussian filter was applied to reduce the noise. Different $\sigma$ values were selected for Gaussian filter to obtain different enhancement functions, and the largest response function was selected to derive BV enhanced image. The value range of $\sigma$ was $3 \sim 6$ and the step was 1 in this study.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$  \hspace{1cm} (2)

The second-order partial derivatives of $x$ and $y$ on the filtered image were computed respectively, then convolved with the preprocessed image $I_{pre}(x, y)$, getting $L_{xx}$, $L_{xy}$, $L_{yy}$ respectively. The Hessian matrix $H$ was composed as follows:

$$H = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix}$$  \hspace{1cm} (3)

Then two eigenvalues $\lambda_1$ and $\lambda_2$ could be obtained from Hessian matrix $H$ as follows, where $tmp$ represented an intermediate variable.

$$tmp = \sqrt{(L_{xx} - L_{yy})^2 + 4 * L_{xy}^2}$$  \hspace{1cm} (4)
\[ \mu_1 = 0.5 \ast (L_{xx} + L_{yy} + \text{tmp}) \quad (5) \]

\[ \mu_2 = 0.5 \ast (L_{xx} + L_{yy} - \text{tmp}) \quad (6) \]

If \(|\mu_1| \leq |\mu_2|\), then \(\lambda_1 = \mu_1, \lambda_2 = \mu_2\). Otherwise, \(\lambda_1 = \mu_2, \lambda_2 = \mu_1\) (that is, to ensure \(|\lambda_2| \geq |\lambda_1|\)).

\(\lambda_2\) was normalized as follows:

\[ \lambda_{\rho} = \begin{cases} \max(\lambda_2) & \lambda_2 > 0 \\ 0 & \text{else} \end{cases} \quad (7) \]

Enhancement function was calculated as follows:

\[ \text{enhancement} = \begin{cases} 0 & \lambda_2 \leq 0 \cup \lambda_\rho \leq 0 \\ 1 & \lambda_2 \geq \frac{\lambda_\rho}{2} > 0 \\ \lambda_2^2(\lambda_\rho - \lambda_2)(\frac{3}{\lambda_2 + \lambda_\rho})^3 & \text{else} \end{cases} \quad (8) \]

The BV enhancement result is shown in Fig. 10(c). Because the BV enhancement process might also enhance some other dark red regions in the original image, MA might be enhanced. The green circles in Fig. 10(c) indicate the MAs contained in the BV enhancement image.

**Blood vessels segmentation**

Firstly, the Ostu’s adaptive threshold method was used to binarize the BV enhanced image in order to obtain the preliminary BV segmentation result, as shown in Fig. 10(d). MAs might be contained in the result, shown as the green circles in Fig. 10(d). To remove MA from BV segmentation result, small regions were excluded to obtain the final BV segmentation results, as shown in Fig. 10(e). Compared with Fig. 10(d), MAs have been removed.

**Extraction of MA candidate regions**

**Further preprocessing**

Due to the tiny structure of MA, the accuracy of MA detection is sensitive to image quality. Therefore, further preprocessing was performed to improve the image quality.

To enhance the contrast between background and highlighted dark regions (including BVs, HMs, and MAs) in the preprocessed image \(I_{\text{pre}}\), the image was first processed with Contrast Limited Adaptive Histogram Equalization (CLAHE) method. A 7*7 Gaussian filter was then applied to reduce the influence of noise. The result \((I_{\text{gauss}})\) is shown in Fig. 11(b).

**Blood vessels removal**

To eliminate the effects of BV on MA detection, on the filtered image \((I_{\text{gauss}})\), the gray value of the corresponding position of the segmented BV was directly set to zero, obtaining \(I_{\text{gsBVO}}\) as shown in Fig. 11(c).
**MA candidate regions extraction**

In the previous step of BV removal, the gray value of BV has been directly set to zero. To further make all the other parts black except for the MA candidates, firstly, the mean gray value of original green channel image in FOV (see Equation 1) was subtracted from $I_{gsBV0}$. Then contrast stretch was used, obtaining $I_{CS}$ as shown in Fig. 11(d). Finally, the Ostu’s adaptive threshold method was used to binarize to obtain the preliminary MA candidate regions $I_{can1}$, as shown in Fig. 11(e).

In BV segmentation, the small regions were excluded to ensure that candidate regions of MA were not included in the segmented BV. As a result, the preliminary MA candidate regions contained some small BV segments which appear as slender structures in Fig. 11(e).

Then, the connected component analysis was used to refine the results of MA candidates. Since MAs appear as dots, the remaining slender structures could be excluded by the ratio (named $R$) of the length of the major axis and the minor axis of each candidate region (or connected component). If the ratio $R$ of a candidate region exceeds the set threshold, the region would be discarded.

The threshold was set based on the mean value of the ratio $R$ of all connected components in the preliminary MA candidate regions. Considering the proportion of the number of the remaining slender structure, through several experiments, 1.2 was finally selected to multiply the mean value, formulated as follows:

$$\text{threshold} = 1.2 \times \frac{1}{n} \sum_{i=1}^{n} R(i), i = 1, 2, \ldots, n$$  \hspace{1cm} (9)

Where $n$ is the total number of connected components in the preliminary MA candidate regions.

Finally, in remaining candidate regions, those with area larger than 500 pixels were discarded, since MAs are often much smaller (maximum MA with 100 ∼ 300 pixels in existing works[9, 12, 10, 24], and 361 pixels in e-ophtha MA database), The final result of MA candidate regions $I_{can}$ is shown in Fig. 11(f), where green circles indicate ground truth of MAs.

**Image patches**

MA is difficult to observe and detect at the whole image level due to its tiny structure. Therefore, the detection of MA was based on image patches.

All images in training set and test set were divided into 25*25 patches according to the location of MA candidate regions, and a suitable classification method was selected to determine whether each test patch contains MA, so as to realize MA detection. MA patches and non-MA patches are shown in Fig. 12.

**Features extraction and classification**

In the process of extraction of MA candidate regions, although the interference of BV on MA detection was eliminated, in addition to image noises in the obtained MA candidate regions, there might also be HM. In order to further separate the true MA from the candidate regions, it is necessary to analyze the difference between MA and HM. As shown in Fig. 1 and Fig. 8, MAs appear as isolated dark red
dots with clear borders, while HMs appear as dark red regions of different sizes and shapes with blurred borders. Thus, MA and HM could be distinguished by mainly using shape features or combined with gradient features.

With the aim of distinguishing between true MA and non-MA, six types of features including color, grayscale, shape, texture, Gaussian filter-based, and gradient were selected for each candidate (or patch). The details are shown in Table 3, where $I_{CLAHE}$ indicates the result of CLAHE on the green channel image $I_g$.

For each MA candidate patch, the directional local contrast ($DLC$) of its center point $p$ was calculated. Assuming that the gray level of the center point $p$ is $I_p$, the DLC of the point $p$ along the direction angle $\theta$ is defined as:

$$DLC_p(\theta) = I_p - \frac{I_p(\theta)}{T_p(\theta)}$$

(10)

$$T_p(\theta) = \frac{1}{r} \sum_{q \in N_p(\theta, r)} I_q$$

(11)

Where $T_p(\theta)$ is the mean gray value of neighboring pixels of point $p$ along the direction angle $\theta$, $r$ is the radius of neighboring region, and $N_p(\theta, r)$ is the set of neighboring pixels of point $p$ along the direction angle $\theta$ within the radius $r$:

$$N_p(\theta, r) = \{(x_q, y_q) | x_q = x_p + k \cos \theta, y_q = y_p + k \sin \theta, k = 1, 2, \ldots, r \}$$

(12)

Finally, the obtained DLC vector of point $p$ is:

$$DLC_p = (DLC_p(\theta_1), DLC_p(\theta_2), \ldots, DLC_p(\theta_n))$$

(13)

Where $n$ is the number of angles, 12 was selected, that is, $360^\circ$ was divided into 12 angles with a step of $30^\circ$. And the select value of $r$ was 12, which was half the patch size. DLC indicates the local contrast characteristics in a neighborhood. A negative DLC value indicates that the gray value of the current pixel is lower than that of the neighborhood. Therefore, DLC could be used to differentiate the structures with different local contrast. Fig. 9 shows DLC distribution of center point of each region of MA, HM, BV and background (BG) shown in Fig. 8.

As shown in Fig. 9, the difference between MA and other regions (HM, BV, BG) in DLC distribution is obvious. The analysis of variance (ANOVA) results showed that the DLC of MA is significantly different from those of HM, BV, and BG ($p < 0.05$ for all).

**Area** is the total number of pixels in candidate region.

**Perimeter** is the total number of boundary pixels that surrounds candidate region.

**Circularity** of each candidate region is defined as follows:

$$Circularity = \frac{\text{Perimeter}^2}{4 \times \pi \times \text{Area}^2}$$

(14)

**Eccentricity** is the ratio of the distance between the two focal points of the circumscribed ellipse of candidate region (focal length) to its major axis length, and it is equal to 0 for a circular region.
**Aspect ratio** is the ratio of the length of major axis to minor axis of candidate region.

**Solidity** is the ratio of area of candidate region to its convex hull area.

**Entropy** indicates image randomness or texture complexity, representing the clustering characteristics of gray distribution, which is defined as follows:

\[
\text{Entropy} = -\sum_{i=1}^{n} (p(i) \log_2 p(i)), \quad i = 1, \ldots, n
\]  

There are \( n \) different gray values in the image, and proportion of each gray value \( i \) is \( p(i) \). When gray level of image is uniform, entropy can reach the maximum.

**Energy** reflects the uniformity of image gray distribution and texture thickness, and it is the sum of square of element values of Gray Level Co-occurrence Matrix (GLCM). It is defined as follows, where \( P \) indicates GLCM.

\[
\text{Energy} = \sum_{i,j} P(i,j)^2
\]

**Homogeneity** reflects the tightness of distribution of elements to the diagonal in GLCM, defined as follows:

\[
\text{Homogeneity} = \sum_{i,j} \frac{P(i,j)}{1 + |i-j|}
\]

**Skewness** reflects the asymmetry of gray value distribution of image, defined as follows:

\[
\text{Skewness} = \frac{E(x-\mu)^3}{\sigma^3}
\]

Where \( x \) is the input gray values of pixels, \( \mu \) and \( \sigma \) represent the mean and standard deviation of the input \( x \), respectively, and \( E \) represents expectation.

As for gradient features, gradient at MA boundary has a abrupt change. As shown in Fig. 8(a), gradient vector on the boundary of MA is divergent, where the absolute value of gradient is large. Whereas the absolute value of gradient at the center of MA is small. As shown in Fig. 8(b), the distribution of gradient vector in the patch of HM is relatively messy. It can be seen from Fig. 8 that the difference in gradient vector distribution is obvious between MA and other structures. Therefore, features based on gradient can be used as reference to distinguish MA from image patches of candidate regions.

Finally, after all the above 44 features extracted from each MA candidate patch, Naive Bayesian classification was used to classify each patch to MA or non-MA.

**Materials and experimental equipment**

The fundus images of two available public retinal image databases (e-ophtha MA[25] and DIARETDB1[26]) were used for experiments.

In e-ophtha MA database, there are four image resolutions ranging from 1440 × 960 to 2544 × 1696 pixels with 45° field of view (FOV). The e-ophtha MA database
contains 233 images without MAs, and 148 images with MAs, which are manually annotated by an ophthalmologist and confirmed by another. The 148 images with totally 1306 MAs were used in this study for experiments.

DIARETDB1 database contains 89 color fundus images with resolution of $1500 \times 1152$ pixels and 50° FOV. Four medical experts marked MA in DIARETDB1 database independently. Considering the disagreement between the four experts annotations, we use the annotations with confidence $\geq 75\%$ (at least three of the four experts agree that there is an MA) as standard for MA labeling, and there are 182 MAs with confidence $\geq 75\%$.

74 images were randomly selected from e-ophtha MA database to derive the training set, which include 574 MAs. The segmented patches on these 574 MA locations were positive samples. To match the number of positive samples, 569 non-MA candidate patches were selected as negative samples. All positive and negative samples made up the training set.

The remaining 74 images in e-ophtha MA database and the whole DIARETDB1 database were used as test set to verify the performance of the proposed MA detection method.

MATLAB 2016a (The MathWorks, Inc., Natick, Massachusetts, USA) was used in the environment of 64bit Windows 10 operating system with 2.9 GHz Intel Core i5 CPU and 16GB memory.

**Experiment procedure**

From each patch of the training set, the aforementioned 44 features were extracted to train the classifier. Then each test patch was classified by the trained classifier to determine whether the candidate was MA based on the 44 features extracted, so as to achieve the purpose of MA detection.

Support Vector Machine (SVM), Naive Bayes (NB) and K-Nearest Neighbor (KNN) were used to perform the experiment. The results were evaluated and compared to find the best classification algorithm for MA detection.

**Evaluation**

Receiver Operating Characteristic (ROC) curve and its corresponding Area Under Curve (AUC) as well as Free-response ROC (FROC) curve were used to evaluate the proposed MA detection method. The closer ROC curve is to the upper left corner, the larger value of AUC is, the better the algorithm is.

The abscissa of ROC curve is False Positive Rate (FPR), which is the proportion of negative samples with positive classification results to all negative samples. The ordinate is Sensitivity, also named True Positive Rate (TPR), which is the proportion of positive samples with positive classification results to all positive samples. The two variables corresponding were calculated as follows:

$$FPR = 1 - \text{Specificity} = 1 - \frac{TN}{TN + FP}$$

$$TPR = \text{Sensitivity} = \frac{TP}{TP + FN}$$
TP indicates that MA is correctly predicted. TN indicates that non-MA is correctly predicted. FP indicates that non-MA is erroneously predicted as MA. And FN indicates that MA is erroneously predicted as non-MA.

Compared with ROC curve, FROC curve has the same Sensitivity on the ordinate and the False Positive per image (FPI) on the abscissa.

As both databases provide the annotation of MA on lesion level, evaluation of the results was performed on lesion level. As aforementioned, the ground truth of DIARETDB1 database with confidence ≥ 75% was used as the standard for evaluation of MA detection results.

First of all, ROC curve and AUC were used to evaluate the performance of 3 classifiers on MA detection, so as to select the best performed classifier, and further analyze its results of MA detection.

Then, FROC curve and FROC score were used to evaluate the performance of the proposed MA detection method and the comparison with existing algorithms.

Lastly, typical examples of MA detection results evaluated on lesion level were analyzed.

Abbreviations
DR: diabetic retinopathy; MA: microaneurysm; HM: hemorrhage; HE: hard exudate; CWS: cotton wool spot; ROC: receiver operating characteristic; AUC: area under curve; FROC: free-response receiver operating characteristic; FPR: false positive rate; TPR: true positive rate; FPI: false positive per image; SVM: support vector machine; NB: naive bayes; KNN: K-nearest neighbor; BV: blood vessel; DLC: directional local contrast; FOV: field of view; CLAHE: contrast limited adaptive histogram equalization; ANOVA: analysis of variance; GLCM: gray level co-occurrence matrix.

Acknowledgements
Authors would like to thank e-optha MA dataset (http://www.eophtha.com) and DIARETDB1 database (http://www.it.lut.fi/project/imageret/diaretdb1/) for providing the fundus images for this work.

Author’s contributions
S.C Long and J.L Chen conceived and designed the experiments. Z.Q Chen provided medical guidance. J.L Chen performed the experiments. J.L Chen and A.T Hu analysed the results. J.L Chen composed the manuscript. H.P Liu and D.C Zheng checked the results and revised the manuscript. All authors read and approved the final version.

Funding
Not applicable.

Availability of data and materials
The data used or analyzed during the study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 College of Computer Science and Technology, Zhejiang University of Technology, 310023, Hangzhou, China.
2 Research Center of Intelligent Healthcare, Faculty of Health and Life Science, Coventry University, CV1 5RW, Coventry, UK.
3 Eye Center, The second Affiliated Hospital of Zhejiang University School of Medicine, 310009, Hangzhou, China.

References
1. Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Umwin, N., Colagiuri, S., Guariguata, L., Motala, A.A., Ogurtsova, K., et al.: Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the international diabetes federation diabetes atlas, 9th edition. Diabetes Research and Clinical Practice 157, 107843 (2019)
2. Stitt, A.W., Curtis, T.M., Chen, M., Medina, R.J., Mckay, G.J., Jenkins, A.J., Gardiner, T.A., Lyons, T.J., Hammers, H.P., Simo, R., et al.: The progress in understanding and treatment of diabetic retinopathy. Progress in Retinal and Eye Research 51, 156–186 (2016)
3. Verbraak, F.D., Abramoff, M.D., Bausch, G.C.F., Klaer, C.C.W., Nijpels, G., Schlingemann, R.O., Der Heijden, A.A.W.A.V.: Diagnostic accuracy of a device for the automated detection of diabetic retinopathy in a primary care setting. Diabetes Care 42(4), 651–656 (2019)
4. Chaturvedi, S.S., Gupta, K., Ninawe, V., Prasad, P.S.: Advances in Computer-Aided Diagnosis of Diabetic Retinopathy. arXiv e-prints, 1909–09853 (2019). 1909.09853

5. Srivastava, R., Duan, L., Wong, D.W.K., Liu, J., Wong, T.Y.: Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels. Computer Methods & Programs in Biomedicine \textbf{138}, 83–91 (2016)

6. Akram, M.U., Khalid, S., Khan, S.A.: Identification and classification of microaneurysms for early detection of diabetic retinopathy: Pattern Recognition \textbf{46}(1), 107–116 (2013)

7. Ren, F., Cao, P., Li, W., Zhao, D., Zaini, O.R.: Ensemble based adaptive over-sampling method for imbalanced data learning in computer aided detection of microaneurysm. Computerized Medical Imaging and Graphics \textbf{55}, 54–67 (2017)

8. Dai, B., Wu, X., Bu, W.: Retinal microaneurysms detection using gradient vector analysis and class imbalance classification. Plos One \textbf{11}(8), 0161556 (2016)

9. Dashtbozorg, B., Zhang, J., Huang, F., ter Haar Romeny, B.M.: Retinal microaneurysms detection using local convergence index features. IEEE Transactions on Image Processing \textbf{27}(7), 3300–3315 (2018)

10. Wu, B., Zhu, W., Shi, F., Zhu, S., Chen, X.: Automatic detection of microaneurysms in retinal fundus images. Computerized Medical Imaging and Graphics \textbf{55}, 106–112 (2017). doi:10.1016/j.compmedimag.2016.08.001. Special Issue on Ophthalmic Medical Image Analysis

11. Javidi, M., Pourreza, H.R., Harati, A.: Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation. Computer Methods and Programs in Biomedicine \textbf{133}, 93–108 (2017)

12. Wei, Z., Chengdong, W., Dali, C., Zhenhu, W., Yugen, Y., Wenyou, D.: Automatic microaneurysms detection based on multi-feature fusion dictionary learning. Computational and Mathematical Methods in Medicine \textbf{2017}, 1–11 (2017)

13. Eftekhar, N., Pourreza, H.R., Masoudi, M., Ghiasishirazi, K., Saeedi, E.: Microaneurysm detection in fundus images using a two-step convolutional neural network. Biomedical Engineering Online \textbf{18}(1), 67 (2019)

14. Chudzik, P., Majumdar, S., Caliva, F., Al-Diri, B., Hunter, A.: Microaneurysm detection using deep learning and interleaved freezing \textbf{10574}, 379–387 (2018). doi:10.1101.12.2293520

15. Budak, U., Şengür, A., Guo, Y., Akbulut, Y.: A novel microaneurysms detection approach based on convolutional neural networks with reinforcement learning algorithm \textbf{5}(1), 14–14 (2017)

16. Seoud, L., Hurtut, T., Chelbi, J., Cheri, F., Langlois, J.M.P.: Red lesion detection using dynamic shape features for diabetic retinopathy screening. IEEE Transactions on Medical Imaging \textbf{35}(4), 1116–1126 (2016)

17. Heran, T., Pernus, F., Likar, B., Spiclin, Z.: Enhancement of vascular structures in 3d and 2d angiographic images. IEEE Transactions on Medical Imaging \textbf{35}(9), 2107–2118 (2016)

18. Veiga, D., Martins, N., Ferreira, M., Monteiro, J.: Automatic microaneurysm detection using laws texture masks and support vector machines. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization \textbf{6}(4), 405–416 (2018). doi:10.1080/21681163.2017.1296379

19. Adal, K.M., Sidibe, D., Ali, S., Chaum, E., Karnowski, T.P., Meriaudeau, F.: Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning. Computer Methods and Programs in Biomedicine \textbf{114}(1), 1–10 (2014)

20. Frangi, A.F., Nissien, W.J., Vincen, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. 130–137 (1996)

21. Orlando, J.I., Prokofyeva, E., Blachko, M.B.: A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. IEEE Transactions on Biomedical Engineering \textbf{64}(1), 16–27 (2017)

22. Oliveira, A., Pereira, S., Silva, C.A.: Retinal vessel segmentation based on fully convolutional neural networks. Expert Systems With Applications \textbf{112}, 229–242 (2018)

23. Guo, Y., Budak, U., Vespa, L., Khorasani, E.S., Şengür, A.: A retinal vessel detection approach using convolution neural network with reinforcement sample learning strategy. Measurement \textbf{125}, 586–591 (2018)

24. Habib, M.M., Welikala, R.A., Hoppe, A., Owen, C.G., Rudnicka, A.R., Barman, S.A.: Detection of microaneurysms in retinal images using an ensemble classifier. Informatics in Medicine Unlocked \textbf{9}, 44–57 (2017). doi:10.1016/j.imu.2017.05.006

25. Decenciere, E., Cazuguel, G., Zhang, X., Thibault, G., Klein, J.C., Meyer, F., Marcotegui, B., Quellec, G., Lamard, M., Danno, R., et al.: Teleophtha: Machine learning and image processing methods for teleophthalmology. Irbm \textbf{34}(2), 190–203 (2013)

26. Kauppi, T., Kalesnykiene, V., Kamarainen, J., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Kalviainen, H., Pietila, J.: The diaretdb1 diabetic retinopathy database and evaluation protocol, 1–10 (2007)

Figures

Tables
Fig. 1 Color fundus image of DR.

Fig. 2 Illustration of the proposed MA detection method.

Fig. 3 ROC curves of 3 classifiers on the two databases. (a) e-ophtha MA database; (b) DIARETDB1 database.
Fig. 4 FROC curves of different methods for MA detection on the two databases. (a) e-ophtha MA database; (b) DIARETDB1 database

Fig. 5 Lesion level evaluation for MA detection results on e-ophtha MA database. (a) Results of MA detection, where green circles indicate TPs, white circles indicate FPs, and red circles indicate FNs; (b) Examples of TP and FN; (c) Examples of FP

Table 1 Comparison of sensitivity at different FPI values for different MA detection methods on e-ophtha MA database.

| Methods           | Sensitivity under different FPI values | Scores |
|-------------------|----------------------------------------|--------|
|                   | 1/8 | 1/4 | 1/2 | 1   | 2   | 4   | 8   |        |
| Eftekhari et al.(2019)[13] | 0.091 | 0.258 | 0.401 | 0.534 | 0.579 | 0.667 | 0.771 | 0.471 |
| Wu et al.(2017)[10]    | 0.063 | 0.117 | 0.172 | 0.245 | 0.323 | 0.417 | 0.573 | 0.273 |
| Veiga et al.(2018)[18] | 0.110 | 0.152 | 0.222 | 0.307 | 0.383 | 0.494 | 0.629 | 0.328 |
| Proposed method      | 0.067 | 0.163 | 0.258 | 0.360 | 0.443 | 0.570 | 0.671 | 0.362 |
Fig. 6 Lesion level evaluation for MA detection results on DIARETDB1 database. (a) Results of MA detection, where green circles indicate TPs, white circles indicate FPs, and the red circle indicates FN; (b) Examples of TP and FP; (c) Examples of FP and FN.

Fig. 7 Analysis of MA detection results compared with ground truth on DIARETDB1 database. (a) MA detection results corresponding to different labeling confidences, where yellow circles indicate labels with confidence ≥ 75%, orange circles indicate labels with confidence ≥ 50%, and brown circles indicate labels with confidence ≥ 25%. (b) Evaluation of MA detection results with ground truth of confidence ≥ 75%, where green circles indicate TPs, white circles indicate FPs, and the red circle indicates FN, corresponding to Fig. 6.

Table 2 Comparison of sensitivity at different FPI values for different MA detection methods on DIARETDB1 database.

| Methods              | Sensitivity under different FPI values | Scores |
|----------------------|----------------------------------------|--------|
|                      | 1/8 | 1/4 | 1/2 | 1    | 2    | 4    | 8    |
| Dai et al.(2016)[8]  | 0.035 | 0.058 | 0.112 | 0.254 | 0.427 | 0.607 | 0.755 | 0.321 |
| Adal et al.(2014)[19]| 0.024 | 0.033 | 0.045 | 0.103 | 0.204 | 0.305 | 0.571 | 0.184 |
| Proposed method      | 0.200 | 0.033 | 0.060 | 0.112 | 0.210 | 0.385 | 0.631 | 0.207 |
Fig. 8  Comparison of different structures in 25*25 patch. (a) MA; (b) HM; (c) BV; (d) Background

Fig. 9  DLC distribution on different structures shown in Fig. 8, where radius indicates the DLC value along the direction angle $\theta$.
Fig. 10 Process of blood vessels segmentation. (a) Green channel image; (b) Result of preprocessing; (c) BV enhanced image (green circles indicate MAs); (d) Preliminary BV segmentation (green circles indicate MAs); (e) Final result of BV segmentation

Fig. 11 Process of MA candidate regions extraction. (a) Result of preprocessing before BV segmentation; (b) Result of further preprocessing; (c) Result of BV removal; (d) Result of contrast stretch; (e) Preliminary MA candidate regions; (f) Final result of MA candidate regions (green circles indicate ground truth of MAs)
Table 3 Descriptions of features for MA detection.

| Feature types          | Symbols | Descriptions                                                                 |
|------------------------|---------|-------------------------------------------------------------------------------|
| Color                  | f1−2    | Mean and standard deviation value of candidate patch in RGB color             |
|                        | f3−4    | Mean and standard deviation value of candidate patch in HSV color             |
|                        | f5−6    | Mean and standard deviation value of candidate patch in CIElab color          |
| Grayscale              | f7−8    | Mean and standard deviation value of candidate patch in $I_g$                 |
|                        | f9−10   | Mean and standard deviation value of candidate patch in $I_{CLAHE}$          |
|                        | f11−22  | Directional local contrast (DLC) of the center point of each candidate region in $I_{CLAHE}$ |
| Shape                  | f23−28  | Area, Perimeter, Circularity, Eccentricity, Aspect ratio, and Solidity of each candidate region |
| Texture                | f29−32  | Entropy, Energy, Homogeneity and Skewness of candidate patch in $I_{CLAHE}$   |
| Gaussian filter-based  | f33−40  | Mean and standard deviation value of candidate patch in corresponding Gaussian filtered result of $I_{CLAHE}$ when $\sigma$ is [1, 2, 4, 8] |
| Gradient               | f41−42  | Mean gradient of candidate patch in $I_{CLAHE}$ (mean value of $dx$ and $dy$) |
|                        | f43−44  | Mean gradient on the boundary of each candidate region in $I_{CLAHE}$ (mean value of $dx$ and $dy$ on the boundary) |

Fig. 12 Examples of patches. (a) MA patches; (b) Non-MA patches