A Fast Localization and Extraction of Microaneurysm for Early Detection of Diabetic Retionopathy

Yun Cheng * **Weirong Liu** * **Chenglong Wang** * **Xin Gu** **
Yijun Cheng ** **Shengnan Wang** * **Jun Peng** *

* *School of Computer Science and Engineering, Central South
University, Changsha, China 410083 (e-mail: pengj@csu.edu.cn)*
** *School of Automation, Central South University, Changsha, China
410083*

Abstract: Diabetic Retinopathy (DR) is regarded as one of the leading causes of blindness globally. Microaneurysms (MAs) detection is essential to the computer aided diagnosis of DR at an early stage. However, the automatic detection of MAs is still a challenging problem as they are too tiny to be recognized and hard to be distinguished from other similar lesions. Therefore, we propose an efficient localization and extraction method for MAs, where the edge detection and Random Forest are utilized to enhance the accuracy of detection results. Finally, the proposed method is evaluated based on the public retinal image database MESSIDOR. Numerical results show that high accuracy and timely detection can be obtained with the proposed solution.

Keywords: Microaneurysm; Diabetic Retionopathy; CLAHE; Edge detection; Random Forest

1. INTRODUCTION

Diabetic Retinopathy (DR) is a chronic progressive disease caused by Diabetes, and it is regarded as one of the leading causes of blindness globally, see e.g., Gargeya and Leng (2017). Thus, the timely detection and early treatment of DR is of great significance. Among various features detected during the DR diagnosis, microaneurysms (MAs) are said to be the first important indication of DR, see e.g., Tavakoli et al. (2013). Regular screening of MAs is an essential method for ophthalmologists to detect diabetic patients at an early stage, see e.g., Rahim et al. (2016). However, this process is so subjective and expensive when considering the large number of DR patients globally that plenty of patients can not get timely detection.

In order to realize cost-effective and accurate detection of DR, the automated screening systems have been proposed and attracted much research attention in recent years. Typically, the automated screening systems are supported by many image analysis technologies, among with image segmentation and image recognition are two key steps.

The image segmentation is usually the front step of color fundus image analysis, which tends to satisfy the requirements of sensitivity. The methods based on mathematical morphology have been extensively used as they provide a professional analysis and processing to MAs' shape and construction. For example, Warfield et al. (1996) applied a bilinear Top-Hat transformation and matched filtering into segmentation, and got the final segmentation result by the adaptive entropy thresholding operation. Because it can alleviate the problem of uneven illumination in the color fundus image, until now, the Top-Hat is still a useful step of image enhancement process in image preprocessing, such as in Eftekhari et al. (2019). Then, in Habib et al. (2016), region growing was presented to tidy up the scattered candidates caused by matched filtering. What's more, Canny edge detection is an another choice applied in Purwita et al. (2011), because the edges of majority MAs are macroscopic in the preprocessed fundus images.The methods above only consider about the pixel's grayscale distribution in gray fundus images. However, as circled in yellow rectangles in Fig. 1, MAs are tiny in size and similar to other retinal structures or lesions which leads to a high false positive. To cope with this problem, a simplest solution is mentioned in Purwita et al. (2011), other retinal structures such as blood vessel, optic disk, macula, fovea and other retinal lesions such as exudates and hemorrhages were subtracted.

With the development of machine learning, a large number of classifiers are designed to realize the subtraction more precisely. It is called image recognition (or called image classification), and it is regarded as another key step whose target is to decrease the false positive by separating MAs from non-MAs. One popular type of classifiers is based on the binary decision tree. For instance, Manjaramkar and Kokare (2017) adopted the classification and regression tree (CART) algorithm for true MAs classification. Habib et al. (2016) used an ensemble classifier called Random Forest with 79 features ranked according to their importance. Another efficient classification technology used in image recognition is Support Vector Machine (SVM). It has been widely applied in the problem of dichotomy. Kande et al. (2009) used multiple SVMs to classify the candidate red lesions from other dark segments. Akram

et al. (2013) presented a hybrid classifier which combines the Gaussian mixture model, SVM and an extension of multimodel mediod based modeling approach in an ensemble to improve the accuracy of classification. More influential statistical geometrical features of MAs need to be taken into account, because these classification methods depend too much on whether the selected features are so accurate that non-MAs can be removed and MAs can be retained.
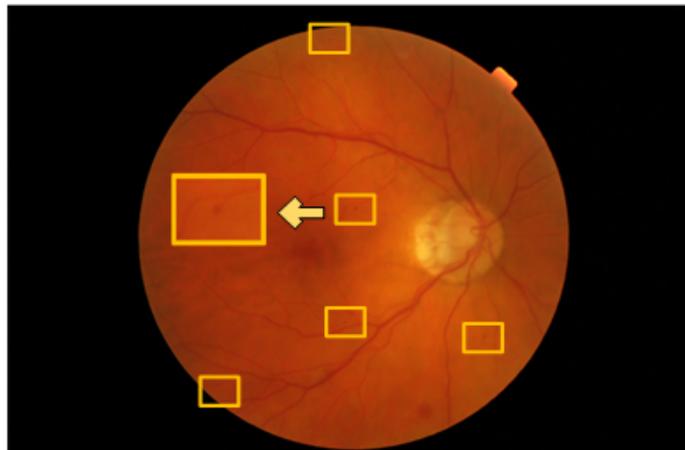
Fig. 1. A color fundus image with MAs. (The regions circled in yellow rectangles are the locations of all MAs in this sample image, and the arrow points to an enlarged image of this area.)

In order to address the above challenges, we combine the image segmentation and image classification. We adopt Canny edge detection and holes filling to get candidate MAs and use Random Forest classifier to subtract the non-MAs in which 24 influential statistical geometrical features are mined so as to enhance the specificity. The contributions of this paper are three-fold:

- Firstly, the Contrast Limited Adaptive Histogram Equalization (CLAHE) and Top-Hat are combined for image enhancement during the preprocessing period, in order to enhance the differences between MAs and non-MAs in the grayscale images.
- Secondly, the Random Forest is utilized to further extract the true positive among the candidates. To improve the accuracy of classifiers, we investigate 24 influential features to enhance the specificity.
- Extensive experiments are conducted to verify the performance of our proposed method with the designs above. The experimental results show that the detection accuracy can be improved.

The rest of the paper is organized as follows. Section 2 describes the design of our proposed approach and presents the details of two key steps. Section 3 introduces the database and evaluation metrics, and shows the analysis of the experimental results. Finally, in Section 4, we conclude this paper and discuss the future work.

2. PROPOSED APPROACH

We take apart the approach into three stages, and the detailed flowchart of the approach in this paper is showed

in Fig. 2. The first stage is image preprocessing in which CLAHE and Top-Hat play the roles of enhancing the contrast of lesions and background. In the second stage, candidate MAs are located by Canny edge detection and holes filling. Compared with matched filtering, it makes the approach faster because region growing is needless to the candidates. Finally, the candidates are divided into MAs and non-MAs by an ensemble classifier called Random Forest for lower false positive. The resultant images are binary which can realize the pixel-based comparison with the 'gold standard' images divided by ophthalmologists manually.
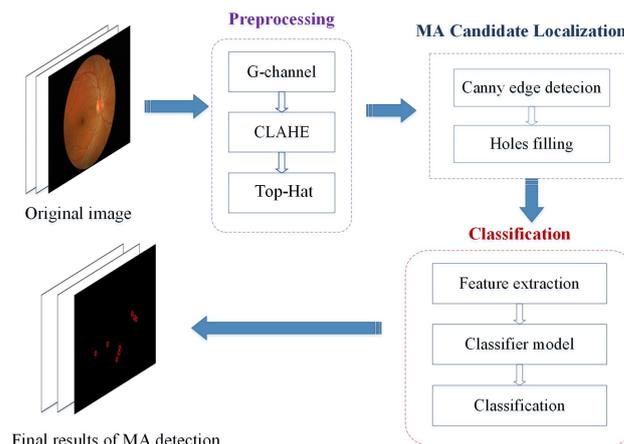
Fig. 2. The detailed flowchart of the proposed approach.

2.1 Image Preprocessing

The color fundus images in the database are various due to the difference of fundus cameras, the angle of photography, the evenness of illumination and the different retinas. So it is necessary to preprocess the images. Through channel selection, image enhancement and Top-Hat transformation, the data of the regions of interest (ROI) with the strongest contrast are obtained. And these three methods are described in the following paragraghs in order.

The first step is channel selection because color fundus images are RGB images. Fig. 3 shows the comparison of the clarity of retinal structures in the three channels and there is a flagrant contrast in the green channel (G-channel). Therefore, the image in the G-channel is selected and converted to gray. The grayscale image is shown in Fig. 4(a) and called I_{gray}.

Before further preprocessing, there is a noticeable operation called mask extraction. Because the image is rectangular while the ROI is circular, which means that the ROI is encircled by black pixels. Therefore, we want to make sure that they are black all the time because of the processing of all pixels in the image in the subsequent two steps. We apply threshold segmentation to extract mask and set the *threshold* to 0.015. The mask image is shown in Fig. 4(b) and called I_{mask}.

The second step is image enhancement. The histogram is used to make statistics of the distribution of pixels' gray value in the whole image. Histogram Equalization (HE) means to equalize the color of the whole image through the
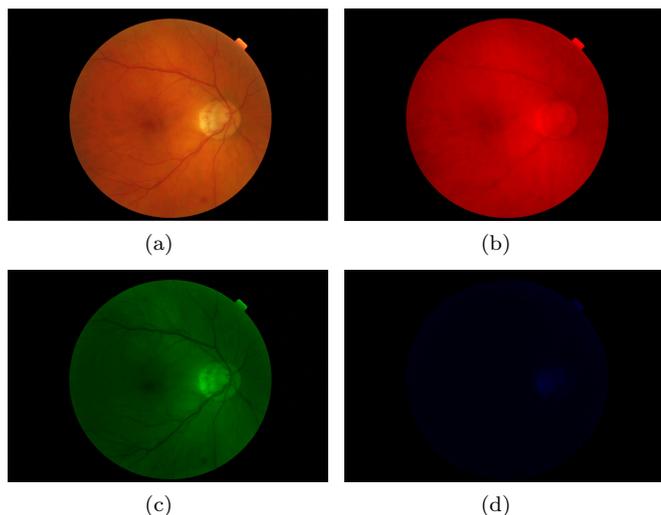
Fig. 3. The comparison of the clarity of retinal structures in the three channels. (a)The original image. (b)The image in red channel. (c)The image in green channel. (d)The image in blue channel.

gray value's distribution reflected in the histogram, so that the color of the whole image will not be concentrated in a limited range and it is often used for image enhancement. The difference between Adaptive Histogram Equlization (AHE) and HE is that AHE achieves the function of enhancing the image's edges to facilitate segmentation by changing the process of the whole image into the process of pixel's neighbor rectangular areas' cumulative distribution function. Most importantly, the difference between CLAHE and AHE is the limit to the contrast so that it effectively restricts the amplification of noise. In detail, the contrast limit is applied to each pixel's neighborhood in CLAHE, so as to obtain the corresponding transformation function which is used to reduce the noise enhancement from AHE. And CLAHE trims the histogram at the specified threshold before calculating the cumulative distribution function in the neighborhood. The CLAHE image is shown in Fig. 4(c) and called I_{clahe}.

Color fundus retinal image contains MAs and other retinal structures such as blood vessel, optic disk, macula, fovea and other retinal lesions, but they coexist with a large area of background, which influences the extraction of MAs. Based on the MAs' particular shape and size, the Top-Hat is utilized as the third ster to transform the augmented images, so as to further make the outlines of MAs clearer. Through the balance of the whole image illumination, this step provide basis for the following image segmentation. The principle of the Top-Hat is to subtract the open operation image from the original image, which can be represented by (1). The image after Top-Hat is shown in Fig. 4(d) and called I_{tophat}.

$$I_{tophat} = I_{clahe} - I_{clahe} \bullet se \tag{1}$$

In this equation, se is a structural element that used in the open operation on images and open operation is defined as corrosion followed by expansion.
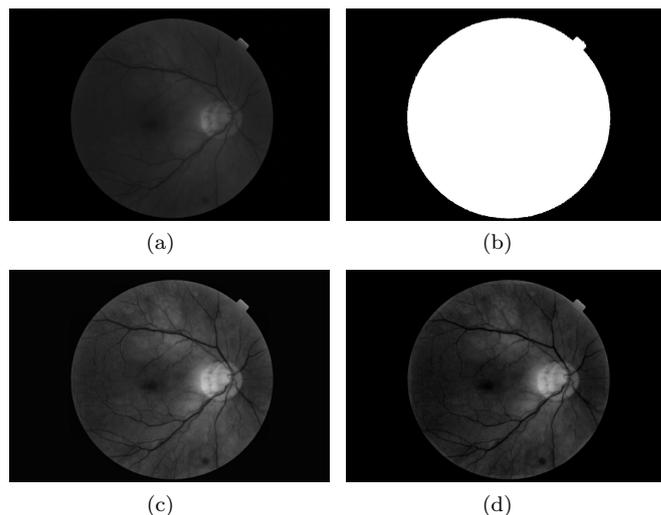
Fig. 4. The results obtained in each step of preprocessing. (a)The grayscale image in G-channel. (b)The template of the grayscale image. (c)The image which is dealt with CLAHE after covered by the template. (d)The image which is dealt with Top-Hat.

2.2 Detection and Localization of Candidate MAs

The candidate MAs are extracted using the methods based on image segmentation. The edges of the object regions in the image are detected firstly, and then the holes are filled to obtain the candidate MAs.

Edge detection is to locate the position where the gray intensity has a great changes in the image. The MAs appear as small and black circle regions in the I_{gray}, so most MAs have obvious boundaries. The Canny operator is one of the most classical and effective algorithms in image edge detection. Because it is sensitive to the edge of the objects in the images while it also can suppress the noise. The Canny edge detection based method has four steps.

The first step is Gaussian Blur, and it is used to reduce the noise. The edges of the object regions and the noise in the images are both the positions where the gray intensity changes obviously, so it is easily to identified the noise as the false edges when the edge detection is carried out by gradient operator. Thus, the first step is to remove the noise. The Gaussian filtering is used for denoising, which means the weighted average of gray intensity of all pixels and their neighboring pixels is calculated according to the template coefficients which decrease with the increase of the distance from the central pixel.

The second step of Canny edge detection is the calculation of the gradient and angle. The gradient value is used for comparison and the role of the angle is to determine the direction. The Canny operator is used to convolve the images and gain the gradient of each pixel in the x direction called G_x and the gradient in the y direction called G_y. And the gradient is calculated according to (2) and the angle is calculated according to (3).

$$G = \sqrt{Gx^2 + Gy^2} \tag{2}$$

$$\theta = \arctan \frac{G_y}{G_x} \tag{3}$$

The third step is non-maximum suppression which turns the edges' width with multiple pixels into a single pixel. It further eliminates noise, and more importantly, it refines the edge. Specifically, the gradient of the pixel is compared with the neighboring two pixels along the positive and negative gradient directions. If the gradient value of the current pixel is the largest, the pixel is retained as the edge point, otherwise, the pixel will be suppressed and its gradient value is set to 0.

Finally, there are another two thresholds: the low threshold and high threshold. Then dual threshold processing is utilized to enhance the edge structure and continuity. If the gradient values are higher than the high threshold, they are set as strong edges. They are removed if they are less than the low threshold, the rest pixels are set to weak edges. And if there is a strong edge pixel in the weak edges' 8 neighbors, it is retained, otherwise, it is eliminated.

Canny edge detection is carried out on the sample image and the edges are shown in Fig 5(a). Holes filling is usually combined with edge detection and the object regions obtained by holes filling are candidate MAs, as shown in Fig 5(b).

(a)

(b)
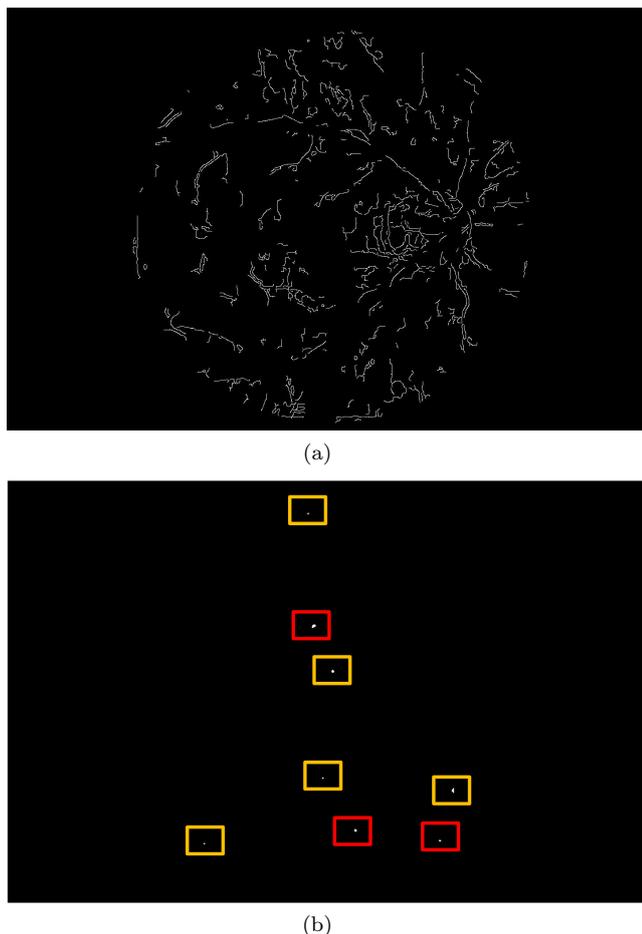
Fig. 5. The results of candidate MAs of the sample image. (a)The segmentation result of Canny edge detection. (b)The segmentation result of holes filling. (The regions circled in yellow rectangles are the locations of true MAs and the regions circled in red rectangles are non-MAs in the sample image.)

2.3 Feature Extraction and Classification

In this paper, we choose the Random Forest to classify the candidate MAs to increase specificity. The Random Forest algorithm consists of three stages: feature extraction, model training and classification of the candidate results. The specific stages of the Random Forest are shown in Fig 6.
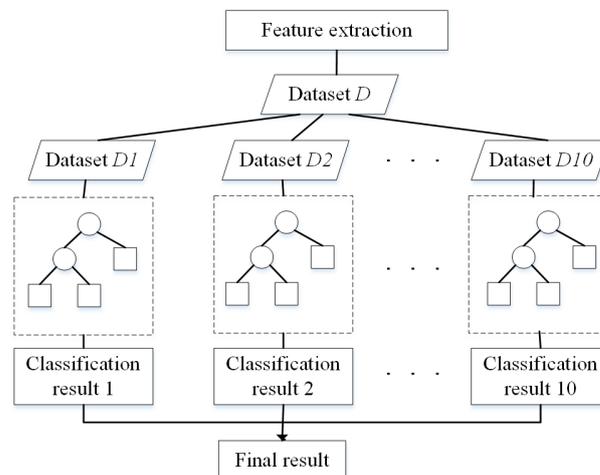
Fig. 6. The structure of Random Forest.

A. feature extraction

At first, we need to gain training samples and training labels. We take the object regions in the 'gold standard' images in the training set as the positive samples whose labels are 1. And the remaining object regions in the candidate images after removing the positive samples are taken as the negative samples whose labels are 0. By doing this, we define the training samples and get the training label called *train_label*.

Then, a 24-dimension feature vector is extracted of each training sample, and the 24 features are comprised of 12 features based on region structure and 12 features based on pixel intensity. In detail, the 12 structural features are obtained by using the 'regionprops' command in MATLAB, which are 'Area', 'EulerNumber', 'MajorAxisLength', 'MinorAxisLength', 'Eccentricity', 'Orientation', 'EquivDiameter' and the length, width, area, circumference and length-width ratio of 'BoundingBox'. And the 12 features based on pixel intensity are the maximum, minimum, mean and variance of the pixel intensity of the object regions from I_{gray}, I_{clahe} and I_{tophat}. According to these data, the training data called *train_data* is obtained which is the 'Dataset D' in Fig 6.

B. training of the Random Forest classifier model

The decision trees in the Random Forest are different, because the bootstrap method is used to obtain the different training samples. And the different training samples come from the data set D containing m samples given in the former step. We select a sample from D one time randomly and copy it into D' so that it can still be selected in the next sampling. After repeating this process m times, we get the data set D' containing m samples which may contains duplicate samples. In this way, we gain the training samples of a decision tree.
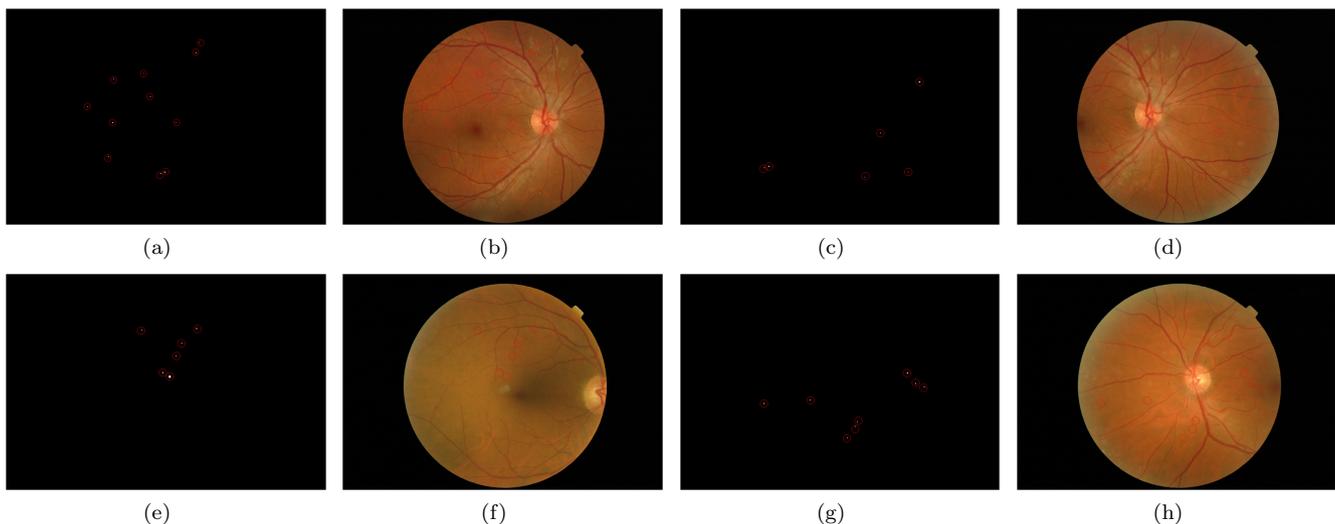
Fig. 7. The extraction results of four images for instance. (a)The extraction result of the first image. (b)The reflection of the extraction result of the first image in original image. (c)The extraction result of the second image. (d)The reflection of the extraction result of the second image in original image. (e)The extraction result of the third image. (f)The reflection of the extraction result of the third image in original image. (g)The extraction result of the forth image. (h)The reflection of the extraction result of the forth image in original image.

In this paper, we need to build 10 decision trees. Thus, 10 D' are obtained by sampling according to the above method and they are marked as D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9 and D_{10}. Finally, a Random Forest model with 10 decision trees is established.

C. classification rule

We apply bagging method to classify the candidate MAs. Firstly, we extract the 24 features for the test set images as same as the training set and gain the test data called $test\_data$. Then, input the $test\_data$ to the Random Forest model and integrate 10 decision trees' classification results obeying the principle of 'the minority is subordinate to the majority'. In other words, if there are no less than 5 results show that the candidate MAs belongs to true MAs, their labels are positive (1), otherwise, they are marked as negative (0). Finally, the negative ones are removed in the final extraction results. There are four images' final extraction results shown in Fig 7.

3. EXPERIMENTAL RESULTS

3.1 Database

We test the proposed approach on a public database called e_ophtha_MA. E_ophtha_MA comprises 148 color fundus images with MAs, and each of them has an original image and a manually annotated image called 'gold standard', see e.g., Wu et al. (2017). In this paper, 98 images are used as the training set and the remaining 50 images as the test set.

E_ophtha_MA is one of a sub database of the MESSI-DOR (methods for evaluating segmentation and indexing techniques dedicated to retinal ophthalmology) database. MESSIDOR was established by the French Ministry of Defense in the TECHNO-VISION project funded in 2004 and is the largest open database of fundus images up to now.

3.2 Evaluation Metrics

For the problem of binary classification, we can divide the prediction results into four categories according to the combination of their true label and the classifier prediction label. The four categories are defined as follows: True Positive (TP): both the true label and the predicted label are positive; False Positive (FP): the true label is negative, while the predicted label is positive; True Negative (TN): both the true label and the predicted label are negative; False Negative (FN): the true label is positive, while the predicted label is negative.

In this paper, TP is identified as the correctly classified MAs; FP is identified as the misclassified background; TN is identified as the correctly classified background; FN is identified as misclassified MAs. Then, calculation formulas of three commonly used evaluation metrics are calculated according to these data, namely sensitivity (Sn), specificity (Sp) and accuracy (Acc). And the specific calculation formulas are shown in (4), (5) and (6).

$$Sn = \frac{TP}{TP + FN} \tag{4}$$

$$Sp = \frac{TN}{TN + FP} \tag{5}$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \tag{6}$$

In this paper, Sn is interpreted as the proportion of the correctly extracted MAs to the total MAs in the 'gold standard' image. Sp is interpreted as the proportion of the correctly extracted background pixels to the total background pixels in the 'gold standard' image. Acc is interpreted as the proportion of the correctly classified pixels to the total pixels in the 'gold standard' image.

3.3 Results Analysis

We divide 50 images from the test set into five groups according to their similarity, so each group has 10 images. And the average of Sn, Sp and Acc of each group are shown in Table 1.

Table 1. The evaluation metrics Sn, Sp and Acc of the final extracted results in five groups.

Number	Sn	Sp	Acc
Group1	85.32%	99.54%	99.54%
Group2	90.25%	97.79%	99.79%
Group3	88.61%	99.90%	99.90%
Group4	92.34%	98.71%	99.71%
Group5	90.11%	98.33%	99.33%

The Sn varies between groups from 85.32% to 92.34%. Because the color fundus images in the database are various due to the angle of photography, the evenness of illumination and so on. Though image preprocessing can alleviate these phenomena, it can not solve these problems completely. Except these, different images have different retinal structures, and the surrounding structures make a great influence on the extraction of MAs, especially in the stage of Canny edge detection.

In order to make the experimental results comparable, we selected two other papers using the same database for training and testing for comparison. And the comparison to the others' method is showed in Table 2.

Table 2. The comparison between the proposed method and two other methods.

	Sn	Sp	Acc
Habib et al. (2016)	about 80%	-	-
Ren et al. (2017)	90.8 %	99.5%	-
proposed method	89.33%	99.65%	99.65%

As we can see in Table 2, first of all, the proposed method has higher sensitivity, suggesting that Canny edge detection algorithm can play a better role of image segmentation when compared with matched filtering. Secondly, the proposed method has higher specificity, suggesting that the Random Forest is effective on removing non-MAs. In addition, there are only 24 features chosen for training. Thus, it is faster because there are 79 features in Habib et al. (2016). In a word, the proposed method in this paper combines the advantages of Canny edge detection and Random Forest. It is a fast and effective method that can be applied to the MAs' localization and extraction, and is helpful to the early diagnosis of DR.

4. CONCLUSION

In this paper, we have investigated the MAs detection problem, which is of significance to the computer aided diagnosis of DR at an early stage. An efficient localization and extraction method has been proposed with the Canny edge detection and Random Forest. By doing this, the detection performance is improved, in terms of the detection accuracy and time efficiency. In future work, the proposed approach for MAs detection will be tested based on some new databases.

REFERENCES

Akram, M.U., Khalid, S., and Khan, S.A. (2013). Identification and classification of microaneurysms for early detection of diabetic retinopathy. *Pattern Recognition*, 46, 107–116.

Eftekhari, N., Pourreza, H.R., Masoudi, M., Ghiasi-Shirazi, K., and Saeedi, E. (2019). Microaneurysm detection in fundus images using a two-step convolutional neural network. *Biomedical Engineering Online*, 18, 1–16.

Gargeya, R. and Leng, T. (2017). Automated identification of Diabetic Retinopathy using deep learning. *Ophthalmology*, 124, 962–969.

Habib, M.M., Welikala, R.A., Hoppe, A., Owen, C.G., Rudnicka, A.R., and Barman, S.A. (2016). Microaneurysm detection in retinal images using an ensemble classifier. In *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–6.

Kande, G.B., Savithri, T.S., Subbaiah, P.V., and Tagore, M.R.M. (2009). Detection of red lesions in digital fundus images. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 558–561.

Manjaramkar, A. and Kokare, M. (2017). Decision trees for microaneurysms detection in color fundus images. In *2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT)*, 1–4.

Purwita, A.A., Adityowibowo, K., Dameitry, A., and Atman, M.W.S. (2011). Automated microaneurysm detection using mathematical morphology. In *2011 2nd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*, 117–120.

Rahim, S.S., Jayne, C., Palade, V., and Shuttleworth, J. (2016). Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening. *Neural Computing and Applications*, 27, 1149–1164.

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

Tavakoli, M., Shahri, R.P., Pourreza, H., Mehdizadeh, A., Banaee, T., and Toosi, M.H.B. (2013). A complementary method for automated detection of microaneurysms in fluorescein angiography fundus images to assess diabetic retinopathy. *Pattern Recognition*, 46, 2740–2753.

Warfield, S.K., Mulkern, R.V., Winalski, C.S., Jolesz, F.A., and Kikinis, R. (1996). An image processing strategy for the quantification and visualization of exercise-induced muscle MRI signal enhancement. *Journal of Magnetic Resonance Imaging : JMRI*, 11, 525–531.

Wu, B., Zhu, W., Shi, F., Zhu, S., and Chen, X. (2017). Automatic detection of microaneurysms in retinal fundus images. *Computerized Medical Imaging and Graphics*, 55, 106–112.