




Article

Local Structure Awareness-Based Retinal Microaneurysm Detection with Multi-Feature Combination

Jiakun Deng ^{1,2,3}, Puyang Tang ^{2,*}, Xuegong Zhao ^{1,3}, Tian Pu ^{1,3}, Chao Qu ⁴ and Zhenming Peng ^{1,3,*} 

¹ School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; 201921050407@std.uestc.edu.cn (J.D.); xgzha@uestc.edu.cn (X.Z.); putian@uestc.edu.cn (T.P.)

² School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

³ Laboratory of Imaging Detection and Intelligent Perception, University of Electronic Science and Technology of China, Chengdu 611731, China

⁴ Department of Ophthalmology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu 610072, China; lucyjeffersonqu@hotmail.com

* Correspondence: pytang@uestc.edu.cn (P.T.); zmpeng@uestc.edu.cn 5 (Z.P.); Tel.: +86-02883288185 (P.T. & Z.P.)

Abstract: Retinal microaneurysm (MA) is the initial symptom of diabetic retinopathy (DR). The automatic detection of MA is helpful to assist doctors in diagnosis and treatment. Previous algorithms focused on the features of the target itself; however, the local structural features of the target and background are also worth exploring. To achieve MA detection, an efficient local structure awareness-based retinal MA detection with the multi-feature combination (LSAMFC) is proposed in this paper. We propose a novel local structure feature called a ring gradient descriptor (RGD) to describe the structural differences between an object and its surrounding area. Then, a combination of RGD with the saliency and texture features is used by a Gradient Boosting Decision Tree (GBDT) for candidate classification. We evaluate our algorithm on two public datasets, i.e., the e-ophtha MA dataset and retinopathy online challenge (ROC) dataset. The experimental results show that the performance of the trained model significantly improved after combining traditional features with RGD, and the area under the receiver operating characteristic curve (AUC) values in the test results of the datasets e-ophtha MA and ROC increased from 0.9615 to 0.9751 and from 0.9066 to 0.9409, respectively.

Keywords: diabetic retinopathy; microaneurysm detection; feature extraction; fundus image analysis



Citation: Deng, J.; Tang, P.; Zhao, X.; Pu, T.; Qu, C.; Peng, Z. Local Structure Awareness-Based Retinal Microaneurysm Detection with Multi-Feature Combination. *Biomedicines* **2022**, *10*, 124. <https://doi.org/10.3390/biomedicines10010124>

Academic Editors: Cecilia Di Ruberto, Alessandro Stefano, Albert Comelli, Lorenzo Putzu and Andrea Loddo

Received: 10 December 2021

Accepted: 3 January 2022

Published: 7 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The number of diabetes patients worldwide is gradually increasing and, with the progression of diabetes, patients may develop DR, which may eventually cause vision loss or even blindness [1]. MA is the initial symptom of DR, and the early identification and timely treatment of retinal MA can prevent further progression of DR. Therefore, it is of great medical significance to realize the automatic detection of MA and assist doctors in the diagnosis of retinal lesions through computer technology.

Color fundus images are the primary way ophthalmologists assess retinal lesions, they judge whether the retina is normal and the grade of DR by visually observing whether there are microaneurysms, hard exudations, soft exudations, hemorrhages, and neovessels in the color fundus images [2]. MA occupies only a few pixels in fundus images and has low local contrast, as shown in Figure 1. At the same time, due to factors, such as the environment and equipment, color fundus images often have different brightness, contrast, and color. Artificial detection of MA is time-consuming, with low accuracy, and easily leads to ophthalmologist fatigue. Therefore, many researchers have studied the automatic detection of MA.

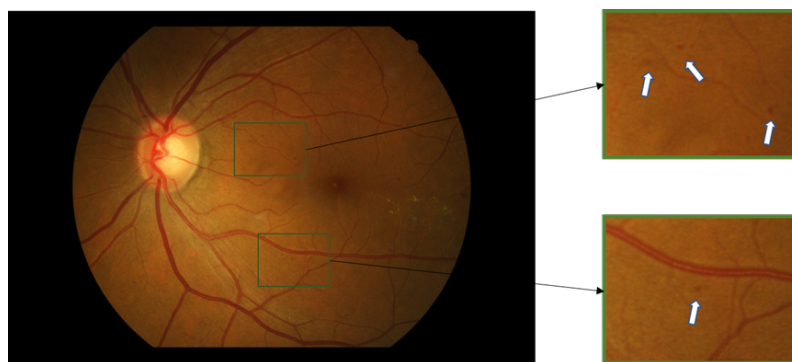


Figure 1. An example of a color fundus image with MAs. The areas with MAs are identified with a green box and zoomed in on the right side of the image, the dark blobs that the white arrow points to are MAs.

The current mainstream MA detection algorithm [3–7] can be briefly summarized as the following three steps: preprocessing, candidate extraction, and candidate classification. The preprocessing methods mainly include color correction, contrast enhancement, reflective elimination, and other image enhancement operations. The main purpose of pretreatment is to better observe the lesions in the retina and prepare for subsequent algorithms. Candidate extraction is primarily to extract image blocks that may contain MA from color fundus images, mainly using morphology, filtering, and other methods. In the step of candidate classification, the accurate detection of MA can be realized by extracting hand-crafted features of each candidate and classifying them with a machine learning classifier.

In this work, we first performed color correction on fundus images, and then used morphological methods to extract MA candidates. We found that MAs and blood vessels were mainly present in the candidate region, and traditional features based on the candidate area are not sufficiently capable and interpretable of distinguishing between them. To make up for the lack of traditional features, a novel local structure feature called ring gradient descriptor (RGD) is proposed, which scans the background around the target in an annular way to find the region most similar to the target and calculates the similarity between the region and the target. Then, a combination of RGD and the saliency and texture features of candidate objects is used by Gradient Boosting Decision Tree (GBDT) for the final candidate classification.

The major contributions of this paper can be summarized as follows.

1. A novel method is proposed for the accurate and reliable detection of microaneurysms with the possibility of applying this method in large screening setups.
2. A simple candidate extraction algorithm based on morphology is proposed to extract the potential MA in fundus images.
3. A new local structure feature RGD is proposed that can describe the local structure of object and its surrounding background and improve the classification performance.

2. Related Works

The detection algorithm of MA can be divided into the physical model-based method, classifier-based method, and deep learning-based detection methods.

Physical model-based MA detection methods are mainly based on the physical characteristics of retinal MA. Joshi et al. [8] employed morphological methods to enhance fundus images and remove blood vessels and then extracted MAs. Zhang et al. [9] proposed a feature-transfer network and local background suppression for MA detection by using the similarity matrix of feature distances to measure the difference between background noise and retinal objects to suppress the local background. Quellec et al. [10] did not perform MA detection in color fundus images but used a lesion template for MA matching in its wavelet-transformed images.

The classifier-based method is the most prevalent method at present, the main process of this method includes candidate extraction, feature extraction, and candidate classification. Orlando et al. [3] employed morphological reconstruction to extract MA candidates. In the process of feature extraction, they constructed a Convolutional Neural Network (CNN) model to extract depth features and combined it with color features, textural features, and geometrical features for candidates classification. Dashtbozorg et al. [4] used a gradient-weighting technique and an iterative thresholding approach to extract MA candidates and used the response of local convergence index filters and the saliency of the candidate area for classification.

Melo et al. [5] used a sliding band filter for MA enhancement and they also used the filter response and the saliency of the candidate area for classification. Antal and Hajdu [6] proposed an ensemble-based framework for MA detection; they selected the optimal results under different preprocessing and candidate extraction methods. Shah et al. [7] removed blood vessels from the green channel and extracted MA candidates using a local thresholding technique. They classified MAs and non-MAs based on statistical features.

The deep learning-based detection method mainly regards MA detection as a segmentation task. Xu et al. [11] improved the U-Net model and achieved pixel-level segmentation of MA. Liao et al. [12] proposed a novel deep convolutional encoder–decoder network for MA detection. Budak et al. [13] used a CNN trained with preprocessed RGB patches to classify MA patches and non-MA patches.

Due to the complex structure of retina and the uneven color and brightness of fundus images, physical model-based methods often have unstable detection effects and low detection accuracy. Classification-based methods tend to have high accuracy because a large number of features are extracted from candidate region images. Deep neural networks have been widely used in the field of computer vision; however, MA detection methods based on deep learning may lead to the existence of over-fitting due to the small amount of data. In addition, deep convolutional neural networks have a large number of parameters and are therefore not easy to use clinically.

In this paper, the classifier-based method was used for MA detection. Previous researchers focused on the target saliency of candidates. In addition, we also paid attention to the local structure of the target and background, and proposed the novel local structure feature RGD to overcome the shortcomings of the saliency features.

3. Materials

We conducted experiments using two publicly available datasets: e-ophtha-MA [14] and ROC [15]. The main specifications of the two datasets are summarized in Table 1.

Table 1. Dataset specifications.

	Image Size (px)	FOV (Degree)	FOV Diameter (px)	NE	NIN	NIP	NMA
ROC	768 × 576, 1394 × 1392	45°	720–1345	4	13	37	336
e-ophtha-MA	1440 × 960, 2544 × 1966	45°	910–1925	1	233	148	1306

FOV: field of view. NE: number of experts. NIN: number of images without MAs. NIP: number of images with MAs. NMA: total number of MAs.

ROC: The Retinopathy Online Challenge (ROC) contains 50 training images and 50 test images, and all MAs were annotated by four experts. The images have different resolutions, ranging from 768 × 576 to 1394 × 1392 pixels with a 45° field of view (FOV). Since the test images do not have MA annotations, only 50 training images were used to verify our proposed algorithm.

E-ophtha-MA: The e-ophtha-MA is a public dataset of color fundus images designed for scientific research in red lesion (MA and small hemorrhage) detection. It contains 233 healthy images and 148 DR images with four resolutions, ranging from 1440 × 960 to 2544 × 1966 pixels with 45° FOV. All images are used to verify our proposed algorithm.

To assess the ability of the new structural features proposed in this paper, the results of the classification between MAs and non-MAs can be evaluated by the receiver operating characteristic (ROC) curve [16] by plotting the true positive rate (TPR) against the false positive rate (FPR) and the area under the ROC curve (AUC). Different TPR and FPR values can be obtained with different thresholds. They are defined as:

$$TPR = \frac{TP}{P} \quad (1)$$

$$FPR = \frac{FP}{N} \quad (2)$$

where P and N correspond to the number of MAs and non-MAs in candidates, respectively. TP is the number of MAs correctly detected, FP is the number of MAs incorrectly detected.

In addition, we evaluated the performance of the detection algorithm at the lesion level. The free-response operating characteristic (FROC) curve [17] was used to evaluate the lesion level of MA detection results of all color images. The abscissa of FROC curve is the average number of false positives per image (FPI), and the ordinate is the sensitivity. The sensitivity represents the proportion of MAs correctly detected by the algorithm. These are calculated as follows:

$$Sensitivity = \frac{TP}{N_{MA}} \quad (3)$$

$$FPI = \frac{FP}{N_i} \quad (4)$$

where N_{MA} is the number of MAs in all fundus images in the test dataset, N_i is the number of images in the test dataset, TP is the number of MAs correctly detected, FP is the number of MAs incorrectly detected. By setting the threshold to classify MA and non-MA, we can obtain pairs $\langle FPI, Sensitivity \rangle$ to draw the FROC curve.

Under the same FPI, higher sensitivity means better detection performance. Similarly, under the same sensitivity, lower FPI means fewer misdetected MA in one image. In order to compare with different methods, we obtained the sensitivity values from FROC curve as the FPI values are 1/8, 1/4, 1/2, 1, 2, 4, and 8. In addition, the average of sensitivity at these seven predefined FPIs (F_{score}) and the partial area under FROC curves between 1/8 and 8 FPI normalized by dividing with the maximum FPI (F_{AUC}) were obtained as the comprehensive evaluation indexes.

Under different detection tasks, F_{score} and F_{AUC} often have different optimal ranges. Even under the same detection task, the number of images also has a great influence on them. In the ROC dataset, the optimal value of these two evaluation indexes should be greater than 0.4, while in the e-ophta-MA dataset, it should be greater than 0.5

4. Methods

A schematic diagram of our method is illustrated in Figure 2. It includes three parts: First, the original image was preprocessed to eliminate the interference of uneven color (see Section 4.1). Second, we extracted the MA candidates (see Section 4.2). Finally, a patch centered on each region was collected to extract target salience and local structural features. The target salience include mean, standard deviation (SD), third moment (TM), energy, entropy, and contrast. The local structural features include the texture feature based on Gray level co-occurrence matrix (GLCM) and RGD. Then, we used the combined features to classify candidates to MAs and non-MAs (see Section 4.3).

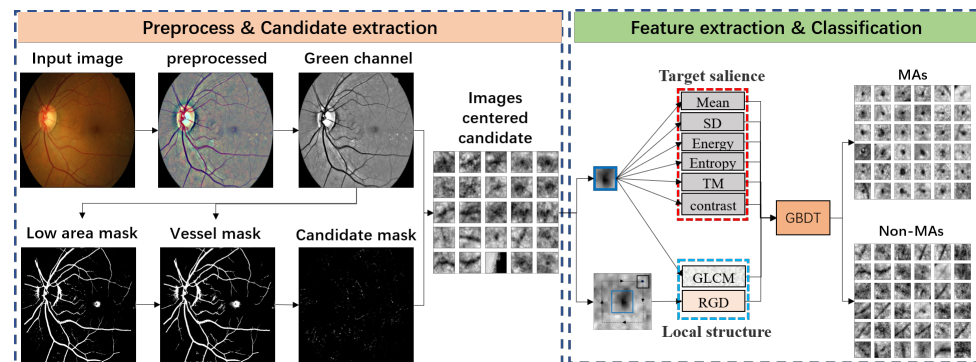


Figure 2. Schematic diagram of the automatic MA detection method, including preprocessing, candidate extraction, feature extraction, and classification.

4.1. Image Preprocessing

Due to the interference of environment and equipment in the process of retinal image acquisition and the differences of people's age and ethnicities [18], the captured fundus images often generally have nonuniform illumination and different colors.

In order to reduce the subsequent computation, we first resized the input image with a scaling factor $\chi/1400$, where χ is related to the width in pixels of the input image. Subsequently, a novel approach proposed by Grinsven et al. [19] was utilized on each channel to enhance the original fundus image with the following equation:

$$I(i, j; \sigma) = \alpha \cdot I(i, j) + \tau \cdot \text{Gaussian}(i, j; \sigma) * I(i, j) + \gamma \quad (5)$$

where $*$ is a convolution operator, σ is the standard deviation of the Gaussian filter, α , τ , and γ are constants. These parameters were set following Grinsven et al. [19], i.e., $\alpha = 4$, $\tau = -4$, $\gamma = 128$ and $\sigma = \chi/30$.

After the image enhancement, there is a great deal of noise in the retina edge. In order to eliminate this noise, we performed region of interest (ROI) detection on it; As shown in Figure 3c, the binary mask of the ROI was obtained by threshold segmentation. Finally, the preprocessed image I_{ce} was measured by means of morphology operations as given in Equation (6):

$$I_{ce}(i, j) = I_{uni}(i, j) \cdot \varepsilon^B(G_{mask}) \quad (6)$$

where I_{uni} and G_{mask} correspond to the enhanced image and ROI mask, respectively. $\varepsilon^B()$ denotes the erosion of an image performed by structural element B with disc type. Figure 3 illustrates the entire preprocessing procedure.

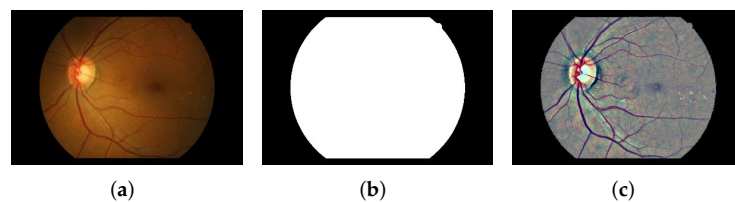


Figure 3. Illustration of the entire procedure for image preprocessing. (a) Original resized image. (b) Binary mask of the ROI. (c) The preprocessed image.

4.2. Candidate Extraction

After color correction, the gray value of fundus image is more uniform, which is more conducive to the extraction of the lesion area by threshold segmentation. An effective candidate extraction method should capture MAs as much as possible and capture fewer non-MAs. To accomplish this, we propose a novel candidate extraction algorithm based on dual-gray threshold segmentation and morphological processing. Figure 4 illustrates the entire candidate extraction procedure.

Due to the main information of microaneurysm being in the green channel [3], we first extracted the green channel G_{ce} from I_{ce} . As the main non-MAs in the retinal image come from blood vessels, we first performed vessel segmentation.

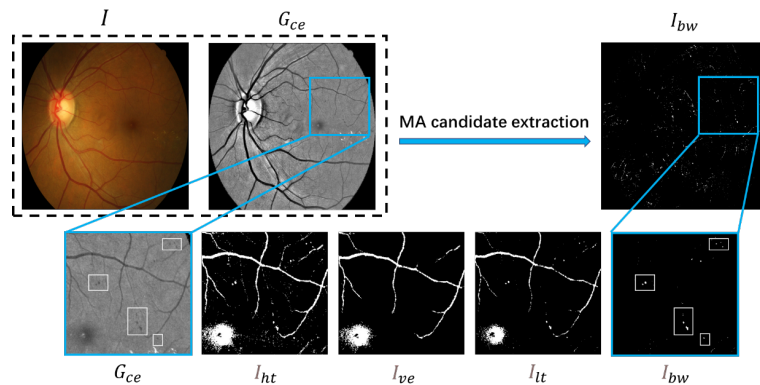


Figure 4. An example to illustrate the candidate extraction method. From top to bottom and left to right: original image I , the green channel of the preprocessed image G_{ce} , the binary mask of candidate area I_{bw} , details in the green channel of the preprocessed image I_{ce} , the first binary image I_{ht} under a higher threshold, the vessels mask I_{ve} , the second binary image I_{lt} under a lower threshold, and the binary mask of candidate area I_{bw} . The areas with MAs are identified with white boxes.

We obtained the first binary image I_{ht} of the low-gray area through threshold segmentation with a higher gray threshold T_h with vessels with more connectivity. The value of T_h ranges from 100 to 115. Then, we reserve the connected domain with an area greater than S through connected domain analysis. The retinal blood vessels mask I_{ve} can be obtained by Equation (7):

$$I_{ve}(i, j) = \begin{cases} I_{ht}(x, y) & \text{if } s(x, y) > S \\ 0 & \text{if } s(x, y) \leq S \end{cases} \quad (7)$$

where I_{ve} corresponds to the vessel mask, and I_{ht} corresponds to the first binary image. $s(x, y)$ correspond to the area of the connected domain in which pixel (x, y) is located. As the retinal vessels occupy a large area, the value of S is set to 400.

After the vessel mask I_{ve} is obtained, we expanded I_{ve} to ensure that the vessel edge can also be eliminated. In addition, we obtained the second binary image I_{lt} with a lower threshold T_l through threshold segmentation, so that fewer non-MAs can be captured. The value of T_l ranged from 90 to 100. Then, the binary mask of candidate area I_{bw} was obtained according to the following equation:

$$I_{bw} = I_{lt} \cdot (1 - \varepsilon^B(I_{ve})) \quad (8)$$

where I_{bw} and I_{lt} correspond to the binary mask of candidate area and the vessel mask, respectively. $\varepsilon^B(\cdot)$ denotes erosion of an image performed by rectangular element B with size $k \times k$. In this article, $k = 5$. Then, we conducted the connected domain analysis for I_{bw} and deleted the pixels whose connected domain area was equal to 1.

We considered each of the connected domains in the binary image I_{BW} as a candidate region for the possible existence of MA. Consequently, we use the center coordinates of each connected domain to extract a certain size image from the preprocessed image I_{ce} for feature extraction and target recognition.

4.3. Feature Extraction and Classification

The salience and local structural features were extracted from each candidate. The local structure features include local texture based on Gray level co-occurrence matrix (GLCM) and our proposed RGD. Then, we combined all the features and used GBDT to classify MA candidates.

4.3.1. Object Saliency

Saliency features of objects have been widely used in classification and detection tasks [20–24]. Since different candidate regions have different sizes and shapes, we extracted image patches of a certain size, which is sufficient to contain one MA from the center of each candidate region as the source of saliency features. The size is set to 11×11 . In this study, six saliency features were extracted from the green channel, including the mean, standard deviation (SD), third moment (TM), energy, entropy, and contrast [25–27].

4.3.2. Local Structures

Local structures include the local texture and RGD. The local texture indicates the homogeneity information of objects [28–31], which is calculated depending on the pixels and their surroundings [32]. We implemented texture feature extraction based on GLCM [33] and six features based on Haralick features [34] were obtained on one offset GLCM matrix. In our work, we selected four different offsets ($0^\circ, 45^\circ, 90^\circ, 135^\circ$) resulting in 24-dimensional Haralick features.

The local texture features used in this article are shown in Table 2. $p(i, j)$ correspond to the (i, j) th entry in normalized GLCM, $p_x(i)$ correspond to the i th entry in the marginal-probability matrix obtained by summing the rows of $p(i, j)$, $p_y(j)$ correspond to the j th entry in the marginal-probability matrix obtained by summing the columns of $p(i, j)$, and $\mu_x, \mu_y, \sigma_x,$ and σ_y are the means and standard deviations of p_x and p_y . $p_{x-y} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j), |i - j| = 0, 1, \dots, N_g - 1,$ and N_g is the number of gray levels.

Table 2. The local texture-feature-based GLCM matrix [34].

Feature Name	Description
correlation	$\sum_i \sum_j \frac{(ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
inverse difference moment	$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i, j)$
difference variance	variance of p_{x-y}
entropy	$-\sum_i \sum_j p(i, j) \log(p(i, j))$
angular moment	$\sum_i \sum_j (p(i, j))^2$
contrast	$\sum_i \sum_j (i - j)^2 p(i, j)$

The object saliency and local texture ignore the relationship between the target and the surrounding background, so their description ability is not enough. By introducing the surrounding background, we can observe a large degree of structural difference between MA and non-MA in this local region.

Vessels are the major component of non-MAs. We defined the $l \times l$ area in the center of the candidate area as the central area A_{ce} , and the adjacent area with width b as the surrounding area A_{sr} . As shown in Figure 5, if the target contained in the candidate region image is MA, we can see that the surrounding region is the retinal background. If the target of the candidate region image is a blood vessel, we can always find a region with the lowest gray mean similar to the central region in the surrounding region. Therefore, we propose a novel local structure feature called a ring gradient descriptor (RGD) to calculate the minimum gradient between the candidate image block and its surroundings to distinguish MAs from vessels.

First, we find the region with the minimum gray mean A_{sm} in the surrounding region in the area around the target through annular scanning, and the size of the scan box is A_{sm} is $b \times b$. Since the blood vessels are multi-directional, we set the moving step of the scan box to 1 pixel each time. Then, the number of scanning K can be calculated as $4(b + l)$ and the minimum gradient between A_{ce} and A_{sm} can be calculated by the following equation:

$$RGD = \frac{1}{b^2} \sum_{i=1}^b \sum_{j=1}^b A_{sm}(i, j) - A_{ce}(i, j) * G(i, j; \sigma) \tag{9}$$

where $G(i, j; \sigma)$ is the Gaussian kernel with standard deviation σ , which has the same size as the central region. $*$ is the convolution operator. Since the gray level of MA is lower in the center and higher in the surrounding area, Gaussian convolution is used instead of calculating the gray level mean, mainly to control the weight of each pixel.

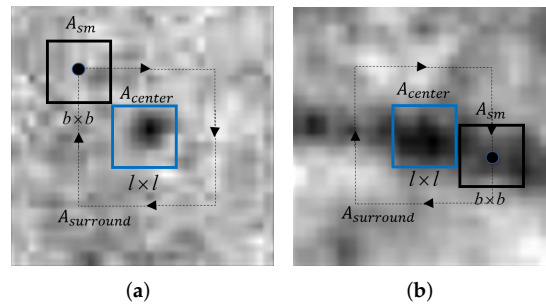


Figure 5. Examples of candidate images. (a) Candidate image containing an MA. (b) Candidate image containing a vessel.

The two most important parameters in the RGD algorithm are l and b , which are determined by the size of MA and the distance from the surrounding blood vessels. The value of σ transforms with l . As shown in Figure 6, different candidates have different optimal parameters. In order to make the model more robust, we calculated 10 RGDs under different parameters as new local structural features as shown in Table 3.

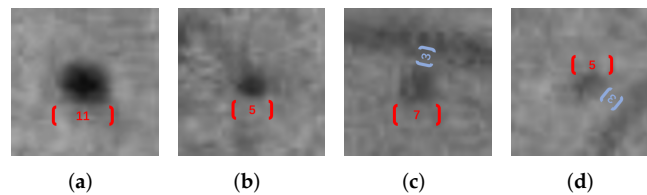


Figure 6. Illustrative example of the optimal parameter settings of different candidates. (a–d) are MAs in different environments, with different pixel widths and distances from the adjacent blood vessel. These two values are marked by red and blue markers respectively.

Table 3. Local structural features.

Parameters	Description										
b	The width of the surrounding area	3	3	3	3	3	5	5	5	5	5
l	The width of the central area	5	7	9	11	13	5	7	9	11	13
σ	The standard deviation of Gaussian kernel	0.9	0.7	0.5	0.5	0.5	0.9	0.7	0.5	0.5	0.5

4.3.3. Classify

To distinguish between MAs and non-MAs, the GBDT classifier [35], which is an ensemble classifier that has been used in general applications, was employed in our work. The increasing popularity of this classifier is mainly attributed to its faster training speed and its robustness. As introduced before, a training set $S_{GB} = \{x^{(i)}, y^{(i)}\}, i = 1, 2, \dots, N$ is constructed by combining features x and corresponding label.

The establishment process of the GBDT is shown in Figure 7, It consists of M base classifiers. The base learner of GBDT is the classification and regression tree (CART). The complete algorithm process of GBDT binary classification algorithm is as follows:

- (1) Initialize the first weak CART:

$$F_0(x) = \log \frac{P(Y = 1|x)}{1 - P(Y = 1|x)} \tag{10}$$

where $P(Y = 1|x)$ is the proportion of MAs in the training sample.

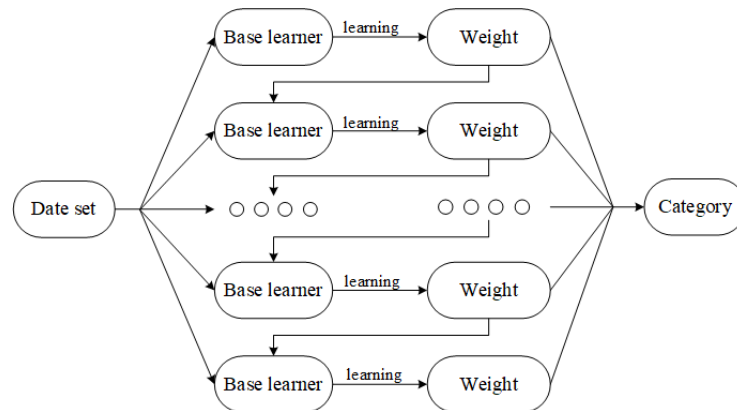


Figure 7. The construction of GBDT [36].

(2) Perform $m(m = 1, 2, 3 \dots M)$ iterations on the base learner:

For $i = 1, 2, \dots, N$, calculate the response value corresponding to the m th tree (negative gradient of the loss function):

$$r_{m,i} = y_i - \frac{1}{1 + e^{-F_{m-1}(x_i)}} \tag{11}$$

For $i = 1, 2, \dots, N$, use CART regression tree fitting data $(x_i, r_{m,i})$ to get the m th regression tree, whose corresponding leaf node area is $R_{m,j}$, where $j = 1, 2, \dots, J_m$ and J_m is the number of leaf nodes of the m th regression tree.

For J_m leaf node region $j = 1, 2 \dots, J_m$, the best fitting value $c_{m,j}$ was calculated:

$$c_{m,j} = \frac{\sum_{x_i \in R_{m,j}} r_{m,i}}{\sum_{x_i \in R_{m,j}} (y_i - r_{m,i})(1 - y_i + r_{m,i})} \tag{12}$$

Update the strong classifier $F_m(x)$:

$$F_m(x) = F_{m-1}(x) + \sum_{j=1}^{J_m} c_{m,j} I(x \in R_{m,j}) \tag{13}$$

(3) Then, the final strong classifier $F_M(x)$ can be expressed as:

$$F_M(x) = F_0(x) + \sum_{m=1}^{J_m} \sum_{j=1}^{J_m} c_{m,j} I(x \in R_{m,j}) \tag{14}$$

(4) Finally, the classification value of sample x can be expressed as:

$$P(Y = 1|x) = \frac{1}{1 + e^{-F_M(x)}} \tag{15}$$

The learning rate of model (ϵ) and the number of levels of trees (C) are the hyperparameters that have the most obvious impact on the accuracy of GBDT model. The prediction accuracy of GBDT is significantly affected by the value of hyperparameter [36]. After plenty of experimentation, the optimal values of ϵ and C were set to 0.03 and 80, respectively.

5. Results

5.1. Candidate Extraction Evaluation

The performance of candidate extraction proposed in this article was evaluated by sensitivity and compared with the previously published candidate extractor algorithms as

demonstrated in Table 4. The proposed method achieved a sensitivity value of 0.51 for the ROC dataset. Although this value did not reach the maximum value, it had a relatively low FPI value, and the performance of the candidate extraction algorithm proposed is better than some algorithms. We obtained a sensitivity value of 0.72 in the ephtha-MA dataset, corresponding to an FPI value of 200.74.

Table 4. Candidate extraction performance using the ROC dataset.

Method	Sensitivity	FPI
Proposed method	0.51	243.38
Shah et al. [7]	0.48	65.00
Dai et al. [37]	0.69	569.39
Adal et al. [38]	0.45	35.2
Walter et al. [39]	0.36	154.42
Zhang et al. [40]	0.33	328.30
Dashtbozorg et al. [4]	0.82	755.50

5.2. Candidate Classification Evaluation

We adopted five-fold cross-validation for model training. We trained three models using target saliency and local texture (TSLT), RGD, and the combined features (CF). The performance of classification based on object salience features and local structure features was evaluated by ROC curve as shown in Figure 8. Their AUC values are shown in Table 5.

Table 5. Validation of different features in MA candidate classification.

Database	Method	AUC
e-optha-MA	TSLT	0.9615
	RGD	0.9566
	CF	0.9752
ROC	TSLT	0.9066
	RGD	0.9205
	CF	0.9409

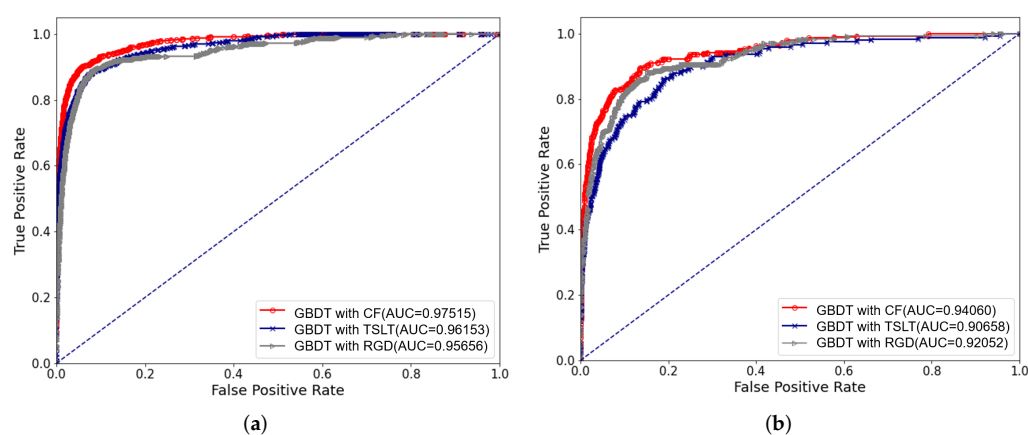


Figure 8. ROC curves of MAs and non-MAs classification on different features. (a) The results of the e-optha-MA dataset. (b) The results of the ROC dataset.

The combined features achieved much higher AUC (AUC = 0.9752 in e-optha-MA, AUC = 0.9409 in ROC) than using traditional TSLT features individually (AUC = 0.9615 in e-optha-MA, AUC = 0.9066 in ROC). The AUC value (AUC = 0.9566) obtained using RGD features were close to those obtained using TSLT features (AUC = 0.9615) in dataset e-optha-MA, and the AUC values (AUC = 0.9205) obtained by using RGD features in dataset

ROC exceeded those ($AUC = 0.9066$) obtained by using TSLT features alone. Therefore, combining traditional features with RGD can improve the classification performance of the model.

Figure 9 presents the FROC curves obtained by not considering the missing MAs from the candidate extraction step and consider. Since many MAs were omitted in the extraction algorithm of candidates, the F_{score} and F_{AUC} values of the final model would be reduced. The values of F_{score} and F_{AUC} in dataset e-optha-MA without considering the omission of MA in the second step were 0.591 and 0.794, which are much higher than the values of 0.434 and 0.583 considering the omission of MA, and the values of 0.349 and 0.519 in dataset ROC are higher than the values of 0.188 and 0.280.

The final results after candidate extraction and candidate classification at the lesion level were compared with other MA detection methods in the ROC and e-optha MA datasets as shown in Table 6. After verification, our algorithm is ahead of some other algorithms in the ROC dataset ($F_{score} = 0.264$, $F_{AUC} = 0.356$), and the F_{AUC} and F_{score} values achieved a leading level in the e-optha MA dataset ($F_{score} = 0.547$, $F_{AUC} = 0.630$). Although we achieved a high detection performance in the e-optha MA datasets, the result in the ROC dataset was poor because of the simplicity of the candidate extraction algorithm.

The sensitivity values of 0.51 and 0.72 in the ROC and e-optha MA datasets were achieved in the candidate extraction step, respectively, which can be said to be the upper limit of the final detection algorithm. The proposed method achieved the sensitivity values of 0.468 and 0.696, respectively, at the two datasets where FPIs is 8, which were close to the upper limit. This shows that our classification algorithm has high performance and the features we extracted were very effective.

In order to more qualitatively display the performance of RGD on MA detection, we used six 224×224 color images containing MAs for pixel-level validation, as shown in Figure 10. The MA candidate region was obtained through dual-threshold segmentation proposed by us. RGD was performed for each pixel in the MA candidate region and the values of b , l , and σ are 5.7, and 0.6, respectively. Then, we normalized the results and segmented them with thresholds of 0.6 and 0.8, respectively. We found that the use of RGD alone was also effective to detect MAs.

Table 6. Performance of different MA detection methods in the ROC and e-optha MA datasets.

Database	Work	Sensitivity against FPIs							F_{score}	F_{AUC}
		1/8	1/4	1/2	1	2	4	8		
ROC	Proposed work	0.083	0.104	0.200	0.257	0.344	0.394	0.468	0.264	0.356
	Chudzik et al. [41]	0.039	0.067	0.141	0.147	0.243	0.306	0.385	0.193	-
	Dashtbozorg et al. [4]	0.435	0.443	0.454	0.479	0.481	0.495	0.506	0.471	0.484
	Eftekhari et al. [42]	0.047	0.173	0.351	0.552	0.613	0.722	0.769	0.461	0.660
	Wu et al. [43]	0.037	0.056	0.103	0.206	0.295	0.339	0.376	0.202	-
	Budak et al. [13]	0.039	0.061	0.121	0.220	0.338	0.372	0.394	0.221	-
	Wang et al. [44]	0.273	0.379	0.398	0.481	0.545	0.576	0.598	0.464	-
	Dai et al. [37]	0.219	0.257	0.338	0.429	0.528	0.598	0.662	0.433	0.553
	Antal and Hjdu [6]	0.173	0.275	0.380	0.444	0.526	0.599	0.643	0.434	0.551
Melo et al. [5]	0.053	0.066	0.077	0.098	0.146	0.208	0.259	0.130	0.185	
e-optha-MA	Proposed work	0.335	0.424	0.496	0.578	0.634	0.668	0.696	0.547	0.630
	Wu et al. [43]	0.063	0.117	0.172	0.245	0.323	0.417	0.573	0.273	-
	Dashtbozorg et al. [4]	0.358	0.417	0.417	0.522	0.558	0.605	0.638	0.510	0.575
	Eftehari et al. [42]	0.091	0.258	0.401	0.534	0.579	0.667	0.771	0.471	0.637
	Chudzik et al. [41]	0.185	0.313	0.465	0.604	0.716	0.801	0.849	0.562	-
	Melo et al. [5]	0.178	0.284	0.383	0.519	0.587	0.587	0.587	0.446	0.551

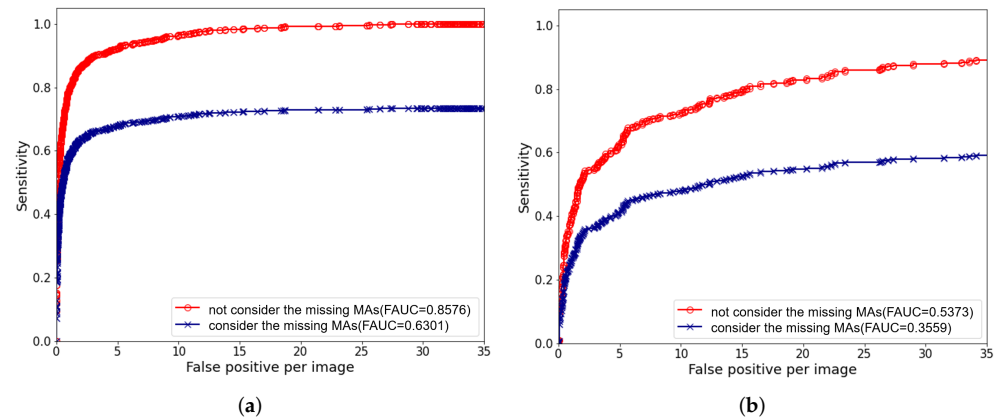


Figure 9. FROC curves of MAs and non-MAs classification. (a) The results of the e-optha-MA dataset. (b) The results of the ROC dataset.

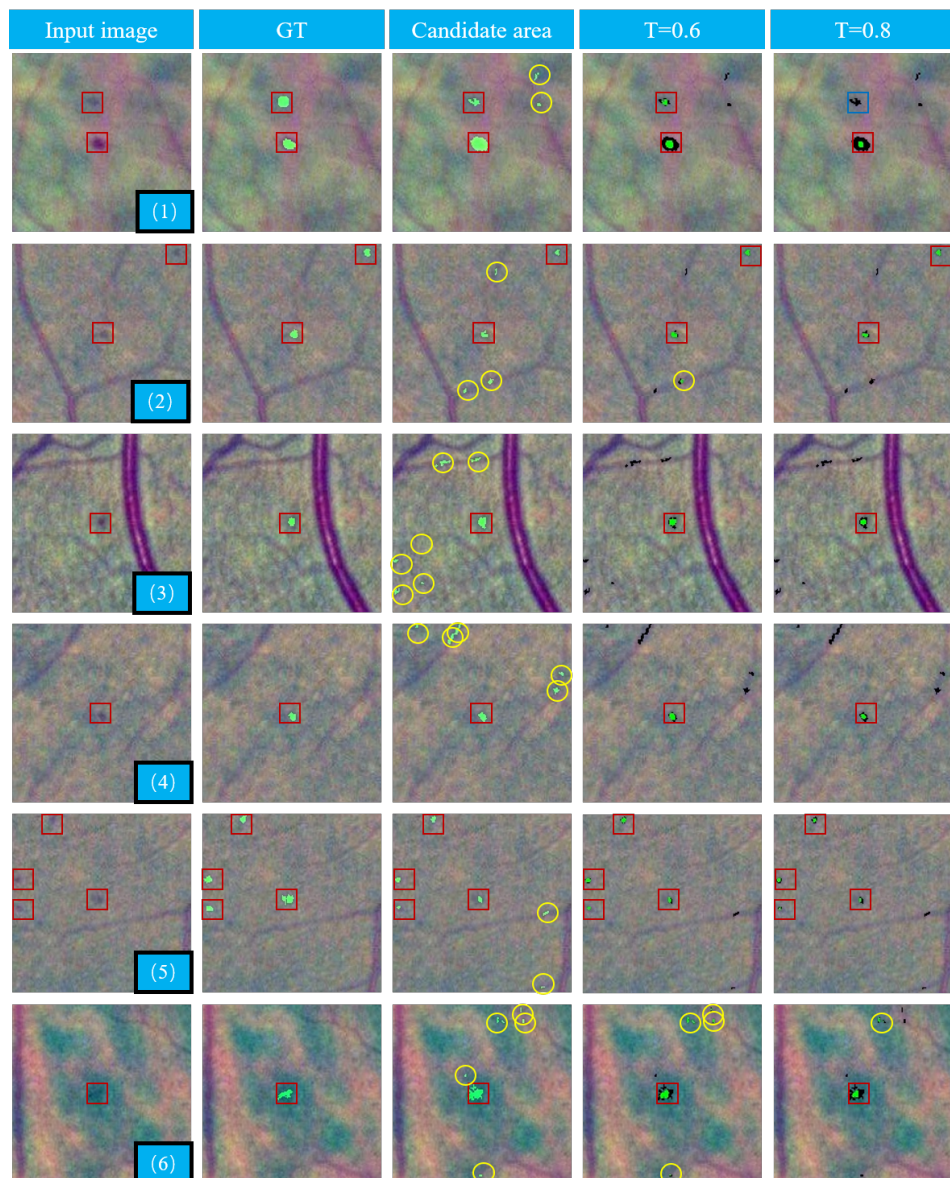


Figure 10. Illustration of pixel level validation. For better visualization, the correctly detected MAs and miss detection MAs are highlighted by red and blue squares, respectively. The FP candidates are highlighted by yellow circles.

6. Discussion

In this paper, we proposed an efficient local structure awareness-based retinal MA detection method with the multi-feature combination (LSAMM). First, the color correction was performed on the images, and simple morphology and threshold segmentation method was used to extract MA candidates. The sensitivity values of this stage in the ROC and e-optha MA dataset, respectively, were 0.51 and 0.72, and their corresponding FPI values were 243.38 and 200.74, respectively.

In the candidate classification step, a novel local structure feature RGD was proposed that can effectively distinguish MA and vessels and improve the performance of classification. After using this, the AUC value in the e-optha MA dataset increased from 0.96153 to 0.97515, and in the ROC dataset, the AUC value increased from 0.90658 to 0.94060. The whole MA detection algorithm proposed achieved a high detection performance in the e-optha MA dataset ($F_{score} = 0.567$); however, its performance in the ROC dataset was mediocre ($F_{score} = 0.264$) as the sensitivity value in candidate extraction step was low.

The candidate extraction algorithm performed worse in the ROC dataset than in the e-optha MA dataset. Nevertheless, the results of candidate classification showed high performance in both datasets. The detection performance of MA can be improved by changing the candidate extraction algorithm and combining RGD with other conventional features in the candidate classification step.

Author Contributions: All authors have contributed substantially to, and are in agreement with the content of, the manuscript. Conception/design, provision of study materials, and the collection and/or assembly of data: J.D. and P.T.; data analysis and interpretation: X.Z. and C.Q.; manuscript preparation: J.D., T.P. and Z.P.; final approval of the manuscript: J.D., P.T., X.Z., T.P., Z.P. and C.Q. The guarantor of the paper takes responsibility for the integrity of the work as a whole, from its inception to publication. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (grant numbers 61775030, 61571096), and the Sichuan Science and Technology Program (grant numbers 22ZDYF2690).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available from the corresponding author upon request.

Acknowledgments: The authors acknowledge the staff at Laboratory of Imaging Detection and Intelligent Perception University of Electronic Science and Technology of China, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital of China.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Long, H.; Chen, B.; Li, W.; Xian, Y.; Peng, Z. Blood glucose detection based on Teager-Kaiser main energy of photoacoustic signal. *Comput. Biol. Med.* **2021**, *134*, 104552. [[CrossRef](#)]
2. Wang, H.; Yuan, G.; Zhao, X.; Peng, L.; Wang, Z.; He, Y.; Qu, C.; Peng, Z. Hard exudate detection based on deep model learned information and multi-feature joint representation for diabetic retinopathy screening. *Comput. Methods Programs Biomed.* **2020**, *191*, 105398. [[CrossRef](#)] [[PubMed](#)]
3. Orlando, J.I.; Prokofyeva, E.; Del Fresno, M.; Blaschko, M.B. An ensemble deep learning based approach for red lesion detection in fundus images. *Comput. Methods Programs Biomed.* **2018**, *153*, 115–127. [[CrossRef](#)]
4. Dashtbozorg, B.; Zhang, J.; Huang, F.; ter Haar Romeny, B.M. Retinal microaneurysms detection using local convergence index features. *IEEE Trans. Image Process.* **2018**, *27*, 3300–3315. [[CrossRef](#)] [[PubMed](#)]
5. Melo, T.; Mendonça, A.M.; Campilho, A. Microaneurysm detection in color eye fundus images for diabetic retinopathy screening. *Comput. Biol. Med.* **2020**, *126*, 103995. [[CrossRef](#)]
6. Antal, B.; Hajdu, A. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 1720–1726. [[CrossRef](#)] [[PubMed](#)]
7. Shah, S.A.A.; Laude, A.; Faye, I.; Tang, T.B. Automated microaneurysm detection in diabetic retinopathy using curvelet transform. *J. Biomed. Opt.* **2016**, *21*, 101404. [[CrossRef](#)]

8. Joshi, S.; Karule, P.T. Mathematical morphology for microaneurysm detection in fundus images. *Eur. J. Ophthalmol.* **2020**, *30*, 1135–1142. [[CrossRef](#)]
9. Zhang, X.; Wu, J.; Meng, M.; Sun, Y.; Sun, W. Feature-transfer network and local background suppression for microaneurysm detection. *Mach. Vis. Appl.* **2021**, *32*, 1–13. [[CrossRef](#)]
10. Quellec, G.; Lamard, M.; Josselin, P.M.; Cazuguel, G.; Cochener, B.; Roux, C. Optimal wavelet transform for the detection of microaneurysms in retina photographs. *IEEE Trans. Med. Imaging* **2008**, *27*, 1230–1241. [[CrossRef](#)]
11. Xu, Y.; Zhou, Z.; Li, X.; Zhang, N.; Zhang, M.; Wei, P. FFU-Net: Feature Fusion U-Net for Lesion Segmentation of Diabetic Retinopathy. *BioMed Res. Int.* **2021**, *2021*, 6644071. [[CrossRef](#)] [[PubMed](#)]
12. Liao, Y.; Xia, H.; Song, S.; Li, H. Microaneurysm detection in fundus images based on a novel end-to-end convolutional neural network. *Biocybern. Biomed. Eng.* **2021**, *41*, 589–604. [[CrossRef](#)]
13. Budak, U.; Şengür, A.; Guo, Y.; Akbulut, Y. A novel microaneurysms detection approach based on convolutional neural networks with reinforcement sample learning algorithm. *Health Inf. Sci. Syst.* **2017**, *5*, 1–10. [[CrossRef](#)] [[PubMed](#)]
14. Decenciere, E.; Cazuguel, G.; Zhang, X.; Thibault, G.; Klein, J.C.; Meyer, F.; Marcotegui, B.; Quellec, G.; Lamard, M.; Danno, R.; et al. TeleOphta: Machine learning and image processing methods for teleophthalmology. *Irbm* **2013**, *34*, 196–203. [[CrossRef](#)]
15. Niemeijer, M.; Van Ginneken, B.; Cree, M.J.; Mizutani, A.; Quellec, G.; Sánchez, C.I.; Zhang, B.; Hornero, R.; Lamard, M.; Muramatsu, C.; et al. Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. *IEEE Trans. Med. Imaging* **2009**, *29*, 185–195. [[CrossRef](#)]
16. Mo, J.; Zhang, L.; Feng, Y. Exudate-based diabetic macular edema recognition in retinal images using cascaded deep residual networks. *Neurocomputing* **2018**, *290*, 161–171. [[CrossRef](#)]
17. Bunch, P.C.; PC, B.; GH, S. A free-response approach to the measurement and characterization of radiographic-observer performance. *Neurocomputing* **1978**, *127*, 124–135.
18. Xie, Y.; Nguyen, Q.D.; Hamzah, H.; Lim, G.; Bellemo, V.; Gunasekeran, D.V.; Yip, M.Y.; Lee, X.Q.; Hsu, W.; Lee, M.L.; et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: An economic analysis modelling study. *Lancet Digital Health* **2020**, *2*, e240–e249. [[CrossRef](#)]
19. Van Grinsven, M.J.; van Ginneken, B.; Hoyng, C.B.; Theelen, T.; Sánchez, C.I. Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1273–1284. [[CrossRef](#)]
20. Peng, Z.; Zhang, Q.; Wang, J.; Zhang, Q.P. Dim target detection based on nonlinear multifeature fusion by Karhunen-Loeve transform. *Opt. Eng.* **2004**, *43*, 2954–2958.
21. Huang, S.; Liu, Y.; He, Y.; Zhang, T.; Peng, Z. Structure-adaptive clutter suppression for infrared small target detection: Chain-growth filtering. *Remote Sens.* **2020**, *12*, 47. [[CrossRef](#)]
22. Wang, G.; Tao, B.; Kong, X.; Peng, Z. Infrared Small Target Detection Using Non-Overlapping Patch Spatial-Temporal Tensor Factorization with Capped Nuclear Norm Regularization. *IEEE Trans. Geosci. Remote Sens.* **2021**. [[CrossRef](#)]
23. Guan, X.; Peng, Z.; Huang, S.; Chen, Y. Gaussian scale-space enhanced local contrast measure for small infrared target detection. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 327–331. [[CrossRef](#)]
24. Liu, Y.; Zhang, P.; He, Y.; Peng, Z. River detection based on feature fusion from synthetic aperture radar images. *Journal of Applied Remote Sensing. J. Appl. Remote Sens.* **2020**, *14*, 016505. [[CrossRef](#)]
25. Sun, L.; Wang, Z.; Pu, H.; Yuan, G.; Guo, L.; Pu, T.; Peng, Z. Attention-embedded complementary-stream CNN for false positive reduction in pulmonary nodule detection. *Comput. Biol. Med.* **2021**, *133*, 104357. [[CrossRef](#)] [[PubMed](#)]
26. Cao, Z.; Kong, X.; Zhu, Q.; Cao, S.; Peng, Z. Infrared dim target detection via mode-k1k2 extension tensor tubal rank under complex ocean environment. *ISPRS J. Photogramm. Remote Sens.* **2021**, *181*, 167–190. [[CrossRef](#)]
27. Zhang, T.; Peng, Z.; Wu, H.; He, Y.; Li, C.; Yang, C. Infrared small target detection via self-regularized weighted sparse model. *Neurocomputing. Neurocomputing* **2021**, *420*, 124–148. [[CrossRef](#)]
28. Wang, Y.; Peng, Z.; Han, Y.; He, Y. Seismic attribute analysis with saliency detection in fractional fourier transform domain. *J. Earth Sci.* **2018**, *29*, 1372–1379. [[CrossRef](#)]
29. Han, Y.; Yang, X.; Pu, T.; Peng, Z. Fine-Grained Recognition for Oriented Ship Against Complex Scenes in Optical Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2021**. [[CrossRef](#)]
30. Zhang, L.; Peng, L.; Zhang, T.; Cao, S.; Peng, Z. Infrared small target detection via non-convex rank approximation minimization joint $l_2, 1$ norm. *Remote Sens.* **2018**, *10*, 1821. [[CrossRef](#)]
31. Zhang, L.; Peng, Z. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sens.* **2019**, *11*, 382. [[CrossRef](#)]
32. Du, J.; Zou, B.; Chen, C.; Xu, Z.; Liu, Q. Automatic microaneurysm detection in fundus image based on local cross-section transformation and multi-feature fusion. *Comput. Methods Programs Biomed.* **2020**, *196*, 105687. [[CrossRef](#)]
33. Wu, P.; Manjunath, B.S.; Newsam, S.; Shin, H.D. A texture descriptor for browsing and similarity retrieval. *Signal Process. Image Commun.* **2000**, *16*, 33–43. [[CrossRef](#)]
34. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H.A. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [[CrossRef](#)]
35. Bikhmetov, T.; Jschke, J. Oil Production Monitoring using Gradient Boosting Machine Learning Algorithm. *IFAC-PapersOnLine* **2019**, *52*, 514–519. [[CrossRef](#)]

36. Zhang, W.; Yu, J.; Zhao, A.; Zhou, X. Predictive model of cooling load for ice storage air-conditioning system by using GBDT. *Energy Rep.* **2021**, *7*, 1588–1597. [[CrossRef](#)]
37. Dai, B.; Wu, X.; Bu, W. Retinal microaneurysms detection using gradient vector analysis and class imbalance classification. *PLoS ONE* **2016**, *11*, e0161556. [[CrossRef](#)]
38. Adal, K.M.; Sidibé, D.; Ali, S.; Chaum, E.; Karnowski, T.P.; Mériaudeau, F. Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning. *Comput. Methods Programs Biomed.* **2014**, *114*, 1–10. [[CrossRef](#)]
39. Walter, T.; Massin, P.; Erginay, A.; Ordonez, R.; Jeulin, C.; Klein, J.C. Automatic detection of microaneurysms in color fundus images. *Med. Image Anal.* **2007**, *11*, 555–566. [[CrossRef](#)] [[PubMed](#)]
40. Zhang, B.; Wu, X.; You, J.; Li, Q.; Karray, F. Detection of microaneurysms using multi-scale correlation coefficients. *Pattern Recognit.* **2010**, *43*, 2237–2248. [[CrossRef](#)]
41. Chudzik, P.; Majumdar, S.; Calivá, F.; Al-Diri, B.; Hunter, A. Microaneurysm detection using fully convolutional neural networks. *Comput. Methods Programs Biomed.* **2018**, *158*, 185–192. [[CrossRef](#)]
42. Eftekhari, N.; Pourreza, H.R.; Masoudi, M.; Ghiasi-Shirazi, K.; Saeedi, E. Microaneurysm detection in fundus images using a two-step convolutional neural network. *Biomed. Eng. Online* **2019**, *18*, 1–16. [[CrossRef](#)] [[PubMed](#)]
43. Wu, B.; Zhu, W.; Shi, F.; Zhu, S.; Chen, X. Automatic detection of microaneurysms in retinal fundus images. *Comput. Med. Imaging Graph.* **2017**, *55*, 106–112. [[CrossRef](#)] [[PubMed](#)]
44. Wang, S.; Tang, H.L.; Hu, Y.; Sanei, S.; Saleh, G.M.; Peto, T. Localizing microaneurysms in fundus images through singular spectrum analysis. *Health information science and systems. IEEE Trans. Biomed. Eng.* **2016**, *64*, 990–1002. [[CrossRef](#)] [[PubMed](#)]