

Microaneurysm (MA) Detection via Sparse Representation Classifier with MA and Non-MA Dictionary Learning

Bob Zhang and Kakhri Karray

Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, ON, N2L 3G1 Canada
{yibo,karray}@pami.uwaterloo.ca

Lei Zhang and Jane You

Department of Computing
The Hong Kong Polytechnic University
Kowloon, Hong Kong
{cslzhang,csyjia}@comp.polyu.edu.hk

Abstract—Diabetic retinopathy (DR) is a common complication of diabetes that damages the retina and leads to sight loss if treated late. In its earliest stage, DR can be diagnosed by microaneurysm (MA). Although some algorithms have been developed, the accurate detection of MA in color retinal images is still a challenging problem. In this paper we propose a new method to detect MA based on Sparse Representation Classifier (SRC). We first roughly locate MA candidates by using multi-scale Gaussian correlation filtering, and then classify these candidates with SRC. Particularly, two dictionaries, one for MA and one for non-MA, are learned from example MA and non-MA structures, and are used in the SRC process. Experimental results on the ROC database show that the proposed method can well distinguish MA from non-MA objects.

Keywords—diabetic retinopathy; microaneurysm; Sparse Representation Classifier

I. INTRODUCTION

World Health Organization (WHO) has estimated that in 2000 there were 171 million people worldwide with diabetes and this number will increase to 366 million by 2030 [1]. This makes diabetes among the leading causes of disabilities, death and economic hardship in the world. DR is a microvascular complication of diabetes and is also responsible for 4.8% of the 37 million cases of blindness in the world, estimated by WHO [1]. The earliest sign of mild non-proliferative DR (first stage of DR) is microaneurysm (MA), which is a round intra-retinal lesion ranging from 10 to 100 micrometers in size and red in color. The detection of MA is crucial in preventing further progression of DR.

Existing MA detection algorithms can be divided into three categories, mathematical morphology [2-5], filter based [6-7] and supervised learning based [4, 8-10]. The Morphological Top Hat Transform [2-5] is the most common morphological operation used to detect MA [2, 4-5, 10]. Filter based algorithms aim to produce an appropriate kernel to match the MA, where high responses can be obtained. These commonly used filters include multi-scale Gaussian filters [6] and wavelet filters [7]. In supervised learning, Artificial Neural Networks [8-9], Support Vector Machine [10], k-NN [4] and pixel classification [4, 8] have been used. Nonetheless, accurate MA detection is still a challenging problem in practice.

In this paper we propose a novel two-step method to detect and classify MA. In the first step, multi-scale Gaussian correlation filters are used to roughly locate the MA candidates. Two dictionaries are learned from example MA and non-MA objects to characterize the MA and non-MA structures. With the learned MA and non-MA dictionaries, sparse representation based classifier (SRC) is then applied to the candidates to distinguish MA objects from non-MA objects. The principle of SRC is to represent a test sample as a linear combination of the training samples, or a dictionary of atoms learned from the training samples. By requiring that the representation coefficients are as sparse as possible, only the coefficients associated with the samples from the same class as the testing sample will be significant, and hence the class label of the testing sample can be identified. The SRC has been successfully used in face recognition [11]. In this paper, we propose to use SRC for MA detection.

The rest of the paper is organized as follows. Section 2 presents the proposed method. Experimental results are presented in Section 3, and conclusion is made in Section 4.

II. PROPOSED METHOD

The proposed method contains two steps, candidate detection and classification. Candidate detection is accomplished with multi-scale Gaussian correlation filtering and region growing, while classification is accomplished by dictionary learning and SRC.

A. Candidate MA Detection

In this step we want to identify all possible MA candidates in a retinal image. Since MA exhibits a Gaussian shape, we can use the Gaussian function G to detect it. The correlation coefficient is a good measure of the resemblance between the Gaussian function and an MA X . If the two entities match well, the correlation coefficient will be high, and vice versa. The correlation function is defined as:

$$\text{coef} = \frac{\sum_m \sum_n (X_{mn} - \bar{X})(G_{mn} - \bar{G})}{\sqrt{\left(\sum_m \sum_n (X_{mn} - \bar{X})^2\right) \left(\sum_m \sum_n (G_{mn} - \bar{G})^2\right)}}. \quad (1)$$

where \bar{X} and \bar{G} are the mean values of X and G .

Since MA vary in size, different sigma values for the Gaussian kernel are required. The first step in candidate detection involves applying a sliding neighborhood filter with multi-scale Gaussian kernels to the fundus image to calculate the correlation coefficient for each pixel. Based on our experiment experience, we chose five scales to represent MA of different sizes with the sigma of the Gaussian function being 1.1, 1.2, 1.3, 1.4 and 1.5. The maximum coefficients from the five responses are combined to form a final response. Fig. 1 shows the final response, where (a) is an input retinal image and (b) is the response. The brighter spots in Fig. 1(b) have a higher coefficient and are more likely to be true MAs.



Figure 1. (a) is a fundus image in the ROC dataset and (b) is its final response to multi-scale correlation filters.

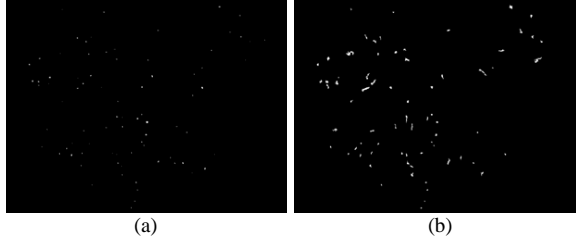


Figure 2. (a) The result after thresholding and vessel removal; (b) region growing applied to (a).

A threshold is then applied to the response map to eliminate candidate pixels with low coefficients. As MA cannot occur on blood vessels, the candidates on them are also removed. The candidates that remain (refer to Fig. 2(a)), however, do not represent the true MA size. Therefore, we apply region growing based on [4]. We first calculate the grayscale pixel intensity value(s) for each candidate. Denote by I_{green} the green channel, and by I_{bg} the background of it computed by applying a median filter of size 25×25 to I_{green} . Denote by $I_{darkest}$ the lowest intensity in I_{green} . A threshold t is computed as:

$$t = I_{darkest} - c \cdot (I_{darkest} - I_{bg}) \quad (2)$$

In this paper c is set to 0.5. The value t is then used to segment I_{green} into a binary image where region growing

begins from the pixel with intensity $I_{darkest}$ and terminates when no more pixels are connected to it. If the resultant region is greater than 120 pixels, it is discarded as no true MA can be that large. Fig. 2(b) shows the result of region growing on the detected candidates in Fig. 2(a).

B. Sparse Representation Based Classification

SRC has been successful used in robust face recognition [11]. Here we propose to use it to classify MA candidates. Given a test sample and a set of training samples, the idea of SRC is to represent the test sample as a linear combination of the training samples, while requiring the representation coefficients are as sparse as possible. If the test sample is from class k , then among its representation coefficients over all the training samples, only those from the samples in class k will be significant while others will be very small, and hence the class label of the test sample can be determined. In practice, the l_1 -norm minimization is used to solve the sparsest linear representation of the test sample over the training samples. More information about SRC can be found in [11].

MA detection is actually a two-class classification problem: MA and non-MA. Therefore, we propose to find two sub-dictionaries for SRC: the MA dictionary and the non-MA dictionary. In the ROC training dataset [12], there are $n=336$ manually marked MA. We can take the detected candidates not marked as MA as the non-MA structures, whose amount is $m=67457$. Let $A = [a_1, a_2, \dots, a_n]$ and $B = [b_1, b_2, \dots, b_m]$, where $a_i, i=1, 2, \dots, n$, and $b_j, j=1, 2, \dots, m$, are 121×1 column vectors containing the pixel values from the 11×11 windows that cover an MA or non-MA object at its center. The columns of A and B are normalized to have unit l_2 -norm.

Since the sizes of A and B are quite large, it is not efficient to use them as the MA and non-MA dictionaries. Here we propose to learn two much more compact dictionaries from A and B . By taking A as the training dataset, we learn a compact MA dictionary Φ from it by solving the following minimization problem:

$$J_{\{\Phi, \Lambda\}} = \arg \min_{\{\Phi, \Lambda\}} \left(\|A - \Phi \Lambda\|_F^2 + \gamma \|A\|_1 \right). \quad (3)$$

where $\|\cdot\|_F$ and $\|\cdot\|_1$ are the Frobenius norm and l_1 -norm, Λ is the representation coefficient matrix of A over Φ , and γ is a scalar to balance the l_2 -norm and l_1 -norm terms. We require that each column of Φ is a unit vector.

(3) is a joint optimization problem of Φ and Λ , and it can be solved by alternatively optimizing Φ and Λ . From some initialization of Φ , we can calculate Λ by solving $\arg \min_{\Lambda} \left(\|A - \Phi \Lambda\|_F^2 + \gamma \|A\|_1 \right)$ with some standard convex optimization algorithm; then by fixing Λ , Φ can be easily updated by solving an l_2 -norm minimization problem

$\arg \min_{\Phi} \|A - \Phi A\|_F^2$. With the updated dictionary Φ , A can then be renewed; the above procedures can be iteratively implemented until J in (3) converges. The convergence of such an alternative optimization is straightforward because in each iteration J can only decrease.

Similarly, for the non-MA training dataset B , we can train a compact dictionary Ψ . In our experiment, the size of both Φ and Ψ is 121×20 , and we let $D = [\Phi, \Psi]$. For each MA candidate vector, denoted by y , we use D to represent it by solving the l_1 -norm minimization problem:

$$J_{\alpha} = \arg \min_{\alpha} \left(\|y - D\alpha\|_2^2 + \lambda \|\alpha\|_1 \right), \quad (4)$$

where α is the sparse representation coefficient vector of y over D , and λ is a constant. α can be solved by convex optimization like in [11]. Since $D = [\Phi, \Psi]$, we can let $\alpha = [\alpha_{\Phi}, \alpha_{\Psi}]$, where α_{Φ} is the representation of y over Φ , and α_{Ψ} is the representation of y over Ψ .

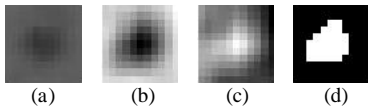


Figure 3. Example of SRC based MA detection. (a) is a cropped MA object; (b) is its reconstruction error r_{MA} and (c) is r_{NMA} ; (d) is the detection result by applying the rule in (5). We see that the MA object is well detected.

If we use the MA dictionary Φ and the associated coefficients α_{Φ} to reconstruct y , the reconstruct error will be $r_{MA}(y) = \|y - \Phi \alpha_{\Phi}\|_2$, and if we use Ψ and α_{Ψ} to reconstruct y , the reconstruct error will be $r_{NMA}(y) = \|y - \Psi \alpha_{\Psi}\|_2$. Intuitively, if the input object y is an MA object, it could be well represented by Φ so that the error r_{MA} will be small. If y is not an MA object, r_{MA} will be big, while the reconstruction error r_{NMA} can be small. Therefore, the magnitude of r_{MA} and r_{NMA} can reflect the membership of y . We use the following rule to classify y :

$$\begin{cases} y \text{ is a MA} & \text{if } r_{MA} \leq r_{NMA} \\ y \text{ is not a MA} & \text{if } r_{MA} > r_{NMA} \end{cases}. \quad (5)$$

Fig. 3 illustrates an example of MA detection by using SRC. Fig. 3(a) is a cropped image from ROC highlighting an MA object. Fig. 3(b) is its r_{MA} image and 3(c) is its r_{NMA} image. Fig. 3(d) is the result by applying the rule in (5). We see that MA object is well detected.

III. EXPERIMENTAL RESULTS

A series of experiments were conducted on 50 images in the public retinal image database provided on the ROC competition website [12]. 20 of them were used in learning the MA and non-MA dictionaries and the other 30 were used for testing. The images were all taken with Topcon NW 100, NW 200 or Canon CR5-45NM ‘non-mydratic’ cameras at the default resolution. All images are in JPEG compressed format with sizes from 768×576 to 1394×1392 pixels. Each image comes with a reference standard that marks every MA agreed upon by the consensus of 4 experts.

The performance of our proposed method for MA detection is evaluated by plotting sensitivity curve against the average number of false positives per image (FROC), as shown in Fig. 4. Sensitivity is the number of true MA correctly detected while false positive is the number of non-MA detected as MA. The proposed method is compared with the Math Morph [4-5] and the candidate detection method [2, 4-5, 10], which consists of shade correction, morphological top hat transform and region growing operations. Shade correction normalizes the image with the help of a median filter. Retinal blood vessels are eliminated with top hat transform and region growing is applied to the remaining candidates after the previous operations.

TABLE I. AVERAGE SENSITIVITY OF DIFFERENT METHODS AT VARIOUS FALSE POSITIVE POINTS FOR TEST DATA

FPs/scan	1/8	1/4	1/2	1	2	4	8	Avg.
Proposed Method ^a	4.6	6.4	10.0	12.7	15.0	19.7	28.9	13.9
Candidate Detection ^a	0.2	0.5	0.9	1.8	3.7	7.3	10.9	3.6
Math Morph ^a	1.3	2.5	5.0	7.2	8.7	10.1	12.1	6.7

a. Sensitivity values are scaled in 10^{-2} .

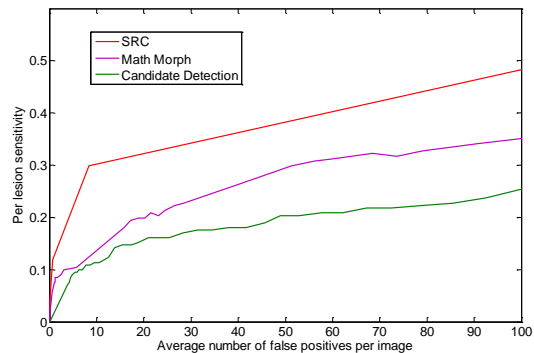


Figure 4. FROC curves of the proposed method in comparison with the candidate detection and Math Morph using ROC images.

As can be seen in Fig. 4, the proposed method has a higher sensitivity compared with other methods under the same false positives per image. Table 1 lists the sensitivity of 1/8, 1/4, 1/2, 1, 2, 4 and 8 false positives per image for all methods derived from Fig. 4. In this table the proposed method has a greater sensitivity from 1/8 to 8 with the

average of all eight points being 0.139 (proposed method), 0.067 (Math Morph) and 0.036 (candidate detection).

IV. CONCLUSION

In this paper we proposed a new method to detect microaneurysm (MA) from color retinal images. The proposed method consists of two steps, candidate detection using multi-scale Gaussian correlation filtering and candidate classification via sparse representation based classifier (SRC). In candidate detection, various Gaussian scales were used to match MA of different sizes. In the second step, an MA dictionary and a non-MA dictionary were first learned, and then used to represent and classify the candidates via SRC. From the experimental results on the ROC dataset, it can be concluded that the proposed method is effective and efficient for MA detection and a vital tool in DR diagnosis. In future work, we will further improve the performance of the proposed method and test it on larger datasets.

ACKNOWLEDGEMENT

This research is supported by the Hong Kong Polytechnic University internal research fund (G-YG79).

REFERENCES

- [1] World Health Organization, "Prevention of blindness from diabetes mellitus," 2006.
- [2] S. Pradhan, S. Balasubramanian, and V. Chandrasekaran, "An Integrated Approach Using Automatic Seed Generation and Hybrid Classification for the Detection of Red Lesions in Digital Fundus Images," *Proceedings: International Conference on Computer and Information Technology Workshops*, pp. 462–467, 2008.
- [3] T. P. Karnowski, V. P. Govindasamy, K. W. Tobin, E. Chaum, and M. D. Abramoff, "Retina Lesion and Microaneurysm Segmentation using Morphological Reconstruction Methods with Ground-Truth Data," *International Conference of IEEE Engineering in Medicine and Biology Society*, August 2008.
- [4] M. Niemeijer, B. van Ginneken, J. Staal, M. S. A. Suttorp-Schulten, and M. D. Abramoff, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 584–592, 2005.
- [5] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation of microaneurysms in fluorescein angiograms of the ocular fundus," *Computers and Biomedical Research*, vol. 29, pp. 284–302, 1996.
- [6] M. Niemeijer, B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A. Mayo, Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, and M. D. Abramoff, "Retinopathy Online Challenge: Automatic Detection of Microaneurysms in Digital Color Fundus Photographs," *IEEE Transactions on Medical Imaging*, vol. 29, pp. 185–195, 2010.
- [7] G. Quellec, M. Lamard, P. M. Josselin, G. Cazuguel, B. Cochener, and C. Roux, "Optimal wavelet transform for the detection of microaneurysms in retina photographs," *IEEE Transactions on Medical Imaging*, vol. 9, pp. 1230–1241, 2008.
- [8] G. G. Gardner, D. Keating, T. H. Williamson, and A. T. Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool," *Brit. J. Ophthalmol.*, vol. 80, pp. 940–944, 1996.
- [9] M. Garcia, C. Sanchez, M. I. Lopez, A. Diez, and R. Hornero, "Automatic Detection of Red Lesions in Retinal Images Using a Multilayer Perceptron Neural Network," *Proceedings: International Conference of IEEE Engineering in Medicine and Biology Society*, pp. 5425–5428, 2008.
- [10] G. B. Kande, T. S. Savithri, P. V. Subbaiah, and M. R. N. Tagore, "Detection of red lesions in digital fundus images," *Proceedings: International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 558–561, 2009.
- [11] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, pp. 210–227, 2009.
- [12] Retinopathy Online Challenge, <http://roc.healthcare.uiowa.edu/>, 2008.