

# Pattern Classification for Dermoscopic Images Based on Structure Textons and Bag-of-Features Model

Yang Li<sup>1</sup>, Fengying Xie<sup>1</sup>(✉), Zhiguo Jiang<sup>1</sup>, and Rusong Meng<sup>2</sup>

<sup>1</sup> Image Processing Center, School of Astronautics,  
BeiHang University, Beijing 100191, China

oskorliyang@163.com, {xfy\_73,jiangzg}@buaa.edu.cn

<sup>2</sup> General Hospital of the Air Force, PLA, 30 Fucheng Road, Beijing 100036, China  
mrs5822@163.com

**Abstract.** An effective method of pattern classification for dermoscopic images based on structure textons and Bag-of-Features (BoFs) model is proposed in this paper. Firstly, the pattern structures of images were enhanced. Secondly, images with obvious directivity were rotated to align their principal directions with horizontal axis, and Otsu method was used to obtain interesting regions. The intensity values of each pixel in the interesting region and its neighborhood composed patch vector. For each pattern, patch vectors of training images were clustered to generate K structure textons and a dictionary with 5 K elements was obtained. Then BoFs model was applied to obtain texton histograms for training and testing images respectively. Finally, a nearest neighbor classifier with chi-square distance was adopted to classify. The experimental results shows that our enhancement method is beneficial to pattern classification and correct classification rate achieves 91.87%.

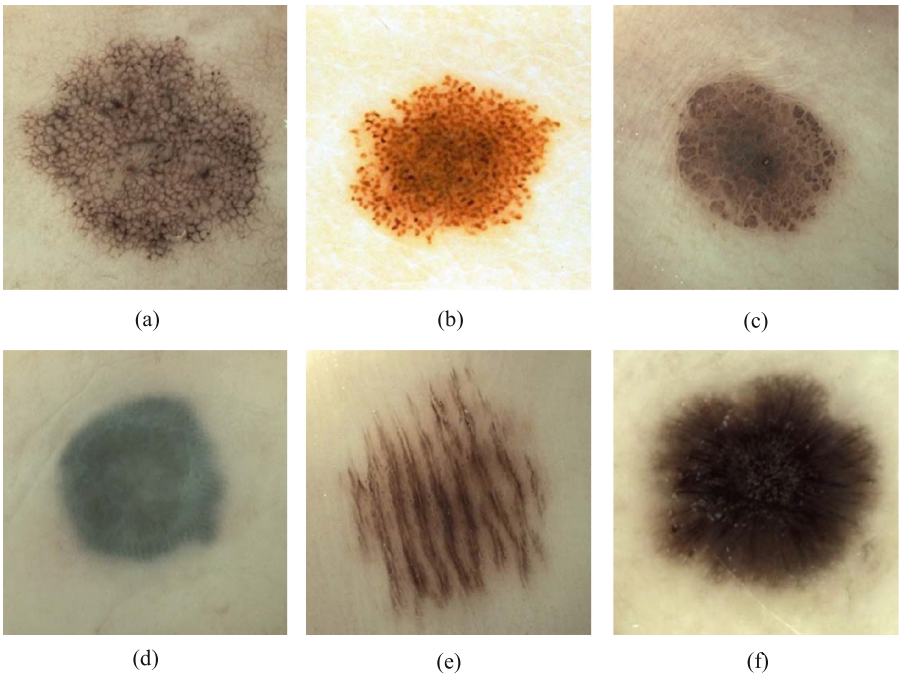
**Keywords:** Dermoscopic image · Computer-aided diagnosis · Pattern classification · Texton

## 1 Introduction

As one of the most rapidly increasing cancers, especially in the United States, Australia and Japan with dry and sunny conditions, skin cancer has been getting more attentions [1]. In the last two decades, digital dermoscopy, a non-invasive skin imaging technique, has been widely used to collect dermoscopic images and many dermatologists turn to computerized analysis for dermoscopic images to improve the accuracy in the diagnosis of pigmented skin lesions (PSLs) [2,3]. Recently, computer-aided diagnosis (CAD) system based on image processing has been developed and become an active area of research, especially about melanoma which is a highly malignant skin cancer.

Now, although there are many algorithms based on computerized analysis to classify the benign and malignant melanoma objectively, the last diagnosis

results still need dermatologists to decide subjectively. Therefore, some objective indices such as color, shape and pattern types are more significant to assist dermatologists. Among these objective indices, pattern type is very important for dermatologists to diagnose melanoma using classic approaches like pattern analysis, Menzies method, 7-point checklist and CASH (color, architecture, symmetry, homogeneity). During the 2000 Consensus Net Meeting on Dermoscopy (CNMD), 7 principal patterns (global pattern) were related to the diagnosis of melanoma [4]. The 7 global patterns of the lesions are: Reticular Pattern, Globular Pattern, Cobblestone Pattern, Homogeneous Pattern, Parallel Pattern, Starburst Pattern and Multicomponent Pattern. Multicomponent pattern is with a combination of three or more above 6 patterns and highly suggestive of malignant melanoma [5]. An illustration of the first 6 patterns is presented in Fig. 1.



**Fig. 1.** An illustration of the six patterns in skin lesions. (a) Reticular pattern: it is the most common global pattern present in melanoma with net-like texture. (b) Globular pattern: it presents itself as small aggregated globules and may have different colors. (c) Cobblestone pattern: It is similar to Globular pattern but they are large and closely aggregated. (d) Homogeneous pattern: it presents uniform texture in the lesion. (e) Parallel pattern: it is found on the palms and soles due to the particular anatomy of these areas. (f) Starburst pattern: it is characterized by the presence of streaks in a radial arrangement (Color figure online) [5].

A few methods of pattern classification for dermoscopic images have been reported. In [6], Tanaka et.al. used 110 texture features to classify skin tumor into three patterns: homogeneous pattern, globular pattern and reticular pattern. Gola et. al. [7] also classified melanoma into the three patterns using a method based on edge detection, mathematical morphology and color analysis. Sadeghi et. al. [8] proposed a texton-based approach with the joint probability distribution of filter responses to detect five patterns. In [9], a pattern analysis by color and texture features is modeled by Markov random field. Sáez et.al. [10] improved this method and used Gaussian model, Gaussian mixture model and BoFs model to detect globular, homogeneous and reticular pattern for the whole lesion.

BoF is a feature extraction scheme used originally for text classifiers which was introduced to the imaging domain by treating textons as words [11], especially about texture classification [12,13]. In this paper, we proposed a novel and effective pattern classification method for dermoscopic images based on structure textons and BoFs model. Our methodology, while similar to [8], is different in that: (i) the structure of pattern is enhanced, (ii) texton is based on patch exemplars directly rather than filter banks.

The rest of the paper is organized as follows. In Sect. 2, the dataset for experiments is described in detail. In Sect. 3, we introduce the proposed pattern classification for dermoscopic images. Evaluation and results are presented in Sect. 4. Finally, Sect. 5 gives the conclusions.

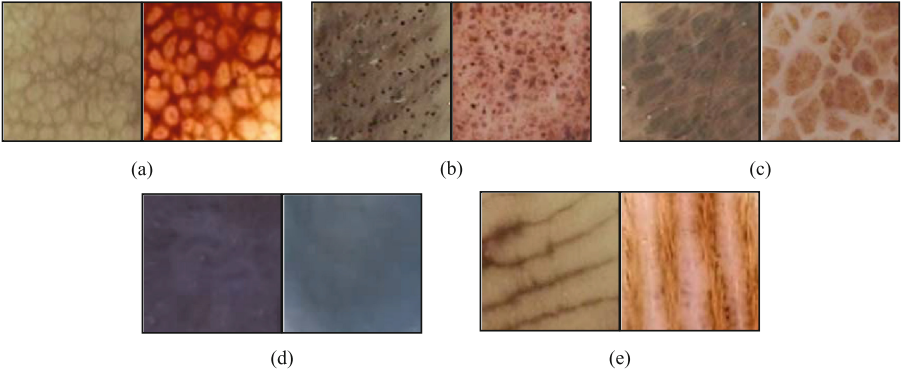
## 2 Materials

We obtained 75 Caucasian dermoscopic images from <http://www.dermoscopy.org/default.asp> and <https://dermoscopy.k.hosei.ac.jp/DermoPerl/>, and 115 Xanthoderm dermoscopic images from General Hospital of the Air Force of the Chinese People’s Liberation Army. A set of 128 pixels  $\times$  128 pixels lesion images representing the 5 patterns are extracted from the 190 images. Each pattern type has 50 lesion images. The details of database are shown in Table 1 and an illustration is displayed in Fig. 2. Because we cannot collect adequate quantity of the dermoscopic images with starburst pattern, the starburst pattern is excluded in our experiments.

**Table 1.** The details of database.

	Reticular	Globular	Cobblestone	Homogeneous	Parallel
Caucasian	13(12) <sup>3</sup>	20(18)	18(12)	24(21)	21(12)
Xanthoderm	37(25)	30(20)	32(24)	26(22)	29(24)
Total	50(37)	50(38)	50(36)	50(43)	50(36)

<sup>3</sup>Number of 128  $\times$  128 lesion images (number of dermoscopic images)



**Fig. 2.**  $128 \times 128$  lesion image samples of each type of pattern (the left is Xanthoderm and the right is Caucasian). (a)~(e) are reticular, globular, cobblestone, homogeneous, and parallel pattern respectively.

### 3 Pattern Classification Method

#### 3.1 Enhancement of Lesion's Pattern Structure

From Fig. 2, it is not hard to find that, for one thing, the color belonging to the same pattern from different lesions has large differences, for another, similar color may exist in different pattern lesions. Therefore, color features are inappropriate for pattern classification. On the contrary, different patterns have huge disparities on the structure. Therefore, we tried to find features to represent the structure of each pattern. In order to avoid the influence of different contrasts in same pattern lesion, it is necessary to enhance the lesion's pattern structure.

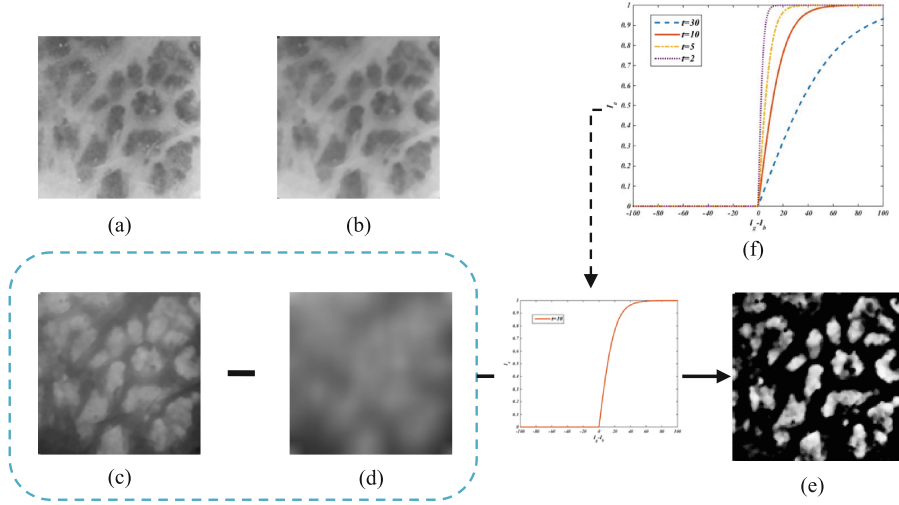
For each lesion image, firstly grayscale image with 256 brightness levels (0 to 255) was obtained from the color image and then a median filter with size  $3 \times 3$  was employed to remove the Gaussian white noise and impulse noise from the grayscale image. Secondly image  $I_g$  was yielded through inverting the grayscale image and a blurred image  $I_b$  was obtained from the image  $I_g$  through a  $21 \times 21$  moving average filter. Then, a response function is defined to enhance the lesion's structure as follows:

$$I_e = \frac{2}{1 + \exp\left(-\frac{\max(0, I_g - I_b)}{t}\right)} - 1 \quad (1)$$

where  $t$  is a factor to determine overall steepness of the function.

There are two parameters: the size of average filter and  $t$ . As shown in Fig. 3, (a) is original grayscale image, (b) is the denoising result using median filter, (c) is the inverting result of (b). In order to reduce the brightness values of all pixels in pattern structure regions as far as possible, we changed the size of average filter and found that  $21 \times 21$  can obtain satisfactory results. Fig. 3 (d) shows the blurred image converted from (c). We obtained the difference values between image (c) and image (d). It is easy to understand that the values of almost all

the pixels in the pattern structure regions are positive and oppositely those in background region are negative. We set the negative values to 0 and used Eq. 1 to enhance the pattern structure. For the parameter  $t$ , the smaller the value of  $t$  is, the more threshold-like the function is and the image is over-enhanced, as shown in Fig. 3(f). Through experiments, we found that most of the difference values are less than 50. So we set  $t$  with a value of 10. Fig 3(e) is the enhancing result. Comparing (e) with (a), it can be seen that the pattern structure regions are highlighted greatly.



**Fig. 3.** Result of enhancing the lesion’s pattern structure. (a) Grayscale image, (b) denoising result, (c) inverting result from (b), (d) blurred image, (e) enhancing result and (f) the curves of Eq. 1 with different  $t$ .

### 3.2 Creation of Texton Dictionary

Different patterns have different structures. We assume that different patterns are composed of different textons (local structure). Thus, for each pattern, we can densely extract local structure features pixel by pixel over the images, and then obtain the textons by simple clustering method. In a patch with  $n \times n$  size, the brightness values of a central pixel and its neighborhood can directly represent the local structure. Before this, we do two steps as follows:

#### (1) Finding Principal Direction of the Texture Pattern

A good texture feature must have rotational invariance. However, the patch-based presentation method is not rotationally invariant [13]. In order to address this problem, we rotate the lesion image according to its the principal orientation.

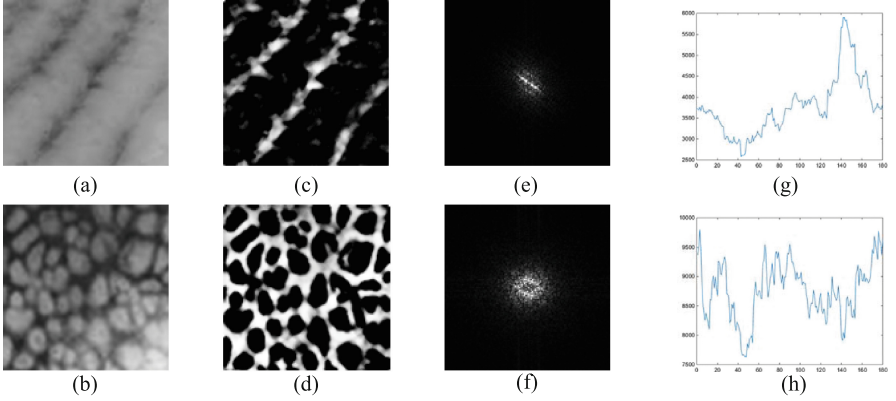
The Fourier spectrum is ideally suited for describing the dominant direction for texture image [14]. For a grayscale image  $I(x, y)$ , its Fourier transform is defined as:

$$F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) \exp\left(-\frac{j2\pi(ux + vy)}{N}\right), u, v = 0, 1, \dots, N - 1 \quad (2)$$

and its Fourier spectrum is defined as:

$$P(u, v) = F(u, v) \overline{F(u, v)} \quad (3)$$

where  $\overline{F}$  is the conjugate complex of  $F$ . As shown in Fig. 4, (a) and (b) are two original grayscale lesion images with different patterns, (c) and (d) are the enhanced image of (a) and (b), (e) and (f) show the Fourier spectra of (c) and (d) respectively. It can be seen that the texture of image (c) has obvious directivity and in its spectra, the main energy is along the principal direction of image (c). But the energy distribution of the Fourier spectra for image (d) is uniform at each direction. In other words, the image (d) does not have principal direction.



**Fig. 4.** Examples of Fourier spectrum for lesion images.(a) (b) lesion images. (c) (d) enhanced images. (e) (f) Fourier spectrum of (c) and (d). (g) (h) the  $P(\theta)$  function curves of (e) and (f).

In order to describe this characteristic, we express the spectrum  $P(u, v)$  in the polar coordinates  $P(r, \theta)$ , the original point of which locates at the center of spectra image. For each direction  $\theta(\theta \in [0^\circ, 1^\circ, 2^\circ, 3^\circ, \dots, 180^\circ])$ , we evaluate the integral  $P(r, \theta)$  over  $r$  to obtain a 1-D function  $P(\theta)$ , which is defined as follows:

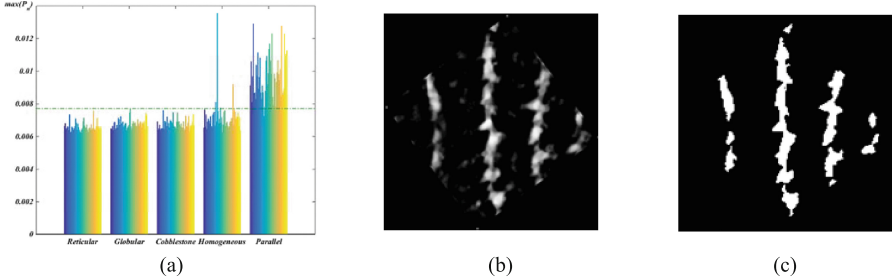
$$P(\theta) = \sum_{r=0}^W P(r, \theta) \quad (4)$$

where  $W$  is the radius of a circle centered at the origin. Fig. 4(g) and (f) are the  $P(\theta)$  function curves for (e) and (f) respectively.

As mentioned above, not all texture images have obvious directivity. We only need to rotate the images with obvious directivity. We normalize  $P(\theta)$  using following formula:

$$P_n(\theta) = \frac{P(\theta)}{\sum_{\theta} P(\theta)} \quad (5)$$

In fact, from Fig. 2, among the 5 patterns, only parallel pattern has obvious texture directivity. We calculate the maximum value of  $P_n(\theta)$  for each enhanced lesion image in our database and find that the value for most of enhanced lesion images is less than 0.0077, except for the parallel pattern, as shown in Fig. 5(a). If the maximum value of  $P_n(\theta)$  for an enhanced lesion image is more than 0.0077, the image is rotated  $\theta_{max}$  (the value of  $\theta$  corresponding to the maximum value of  $P_n$ ) degree clockwise to align the principal direction of this image with horizontal direction. Fig. 5.(b) is the rotating result of Fig. 4(c).

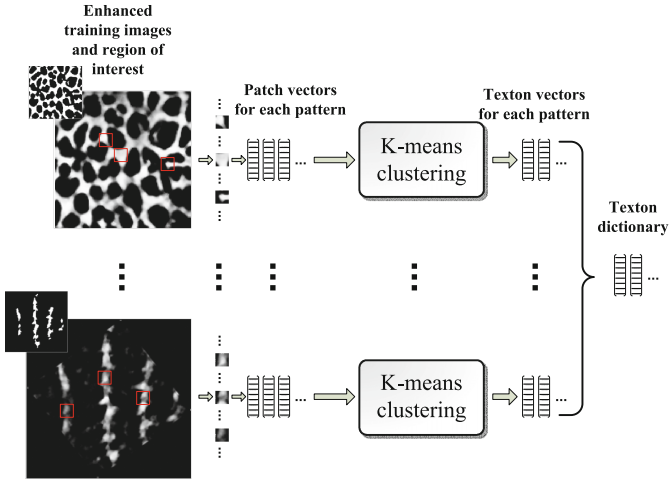


**Fig. 5.** An example of image rotation and interesting region extraction. (a) The maximum values of  $P_n(\theta)$  for each enhanced lesion image. (b) The rotating result of Fig. 4(c). (c) The binary result of (b).

## (2) Extracting Region of Interest

From Fig. 2, it can be seen that the texture for each pattern lesion image is non-homogeneous. The background regions of images after enhanced have very low intensities and the rotated images are also filled with 0 value. Therefore, if we use all of pixels in the image to generate textons, it would have a negative effect on the result. Because we are more concerned about the pattern structure, the enhanced images are binarized using Otsu method to extract the interesting region, as shown in Fig. 5(c). In addition, this operator can reduce the calculation time.

We create texton dictionary using the method proposed by Varma [13] in the interesting regions shown in Fig. 5(c). The process is illustrated in Fig. 6. For the training samples of each pattern, we extract the brightness values of a pixel located in interesting region and its neighborhood within the  $n \times n$  window as

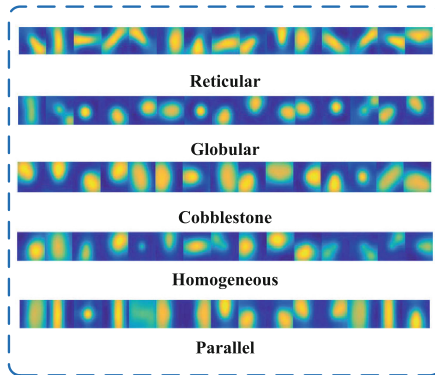


**Fig. 6.** Creation of texton dictionary.

a patch vector  $\mathbf{x}$ . We normalize  $\mathbf{x}$  via the Weber’s law [13]:

$$\mathbf{x} \leftarrow \frac{\mathbf{x}}{\|\mathbf{x}\|_2} \log\left(1 + \frac{\|\mathbf{x}\|_2}{0.03}\right) \tag{6}$$

And then all patch vectors are clustered using K-means method to obtain K textons for each training pattern lesion images. At last, we can obtain 5K textons in total. Fig. 7 is the visualization of texton dictionary, where the value of K for each pattern is 15 and the patch size is  $13 \times 13$ . It can be seen that there are disparities among the structure textons of the 5 patterns.



**Fig. 7.** The visualization of texton dictionary.



### 3.3 Construction of BoFs and Classifier

Now, the texton dictionary has been created and then the BoFs can be constructed by the frequency of occurrences of the textons. For a lesion image, each patch vector is labeled based on its closest element in the texton dictionary via Euclidean distance. The BoFs (texton histogram) of a lesion image is formed by counting the frequencies of texton labels of the lesion image.

In the learning stage, we obtain the category model database by calculating the BoFs of training lesion pattern images. In the testing stage, firstly the BoF of a testing lesion image is obtained based on the texton dictionary and then the testing lesion image is classified using a nearest neighbor classifier, where we use the chi-square statistic, which is a good metric in texture classification, to computer the distance between the two BoFs. The equation is defined as follows:

$$\chi^2(\mathbf{h}_t, \mathbf{h}_d) = \frac{1}{2} \sum_{k=1}^{5K} \frac{(h_t(k) - h_d(k))^2}{h_t(k) + h_d(k)} \quad (7)$$

## 4 Evaluation and Results

To evaluate the performance of our method, correct classification rate is computed. In addition, a 3-times 5-fold cross-validation is used. In a cross-validation, 40 images are randomly selected as training samples and the rest 10 images are as testing samples for each pattern. All steps of the proposed method were implemented using Matlab R2013a on the PC with 3.40 GHz Intel® Core<sup>T</sup> i7 processor and 8GB DDR3 SDRAM.

We compared our method with the state-of-the-art texton learning methods: LBP [15] and MR8 [12]. The experimental parameters of these two methods are given as follows:

**LBP:** We obtained the rotationally invariant, uniform LBP texton dictionary with 1 to 5 scales, which are  $LBP_{8,1}^{riu2}$ ,  $LBP_{8,1+12,1.5}^{riu2}$ ,  $LBP_{8,1+12,1.5+16,2}^{riu2}$ ,  $LBP_{8,1+12,1.5+16,2+20,3}^{riu2}$ , and  $LBP_{8,1+12,1.5+16,2+20,3+20,4}^{riu2}$  respectively. The Matlab code of LBP is downloaded from <http://www.cse.oulu.fi/CMV/Downloads/LBPmatlab>. The experiments are carried out on both original grayscale images and enhanced images (not rotated). The results are shown in Table 2;

**MR8:** The maximum response 8(MR8) filter bank also has rotationally invariant and multiscale characters. It consists of 38 filters but only maximum 8 filter responses were saved. In [12], this method yielded better results than any other filter bank. The Matlab code of MR8 comes from <http://www.robots.ox.ac.uk/~vgg/research/texclass/with.html>. The experiments are also carried out on both original grayscale images and enhanced images (not rotated). The filter size is  $49 \times 49$ (default) and the texton number for each pattern is 5, 10, 15 20, 25, 30, 35 respectively. The results are shown in Table 3.

From Tables 2 and 3, it can be seen that using the LBP method, the maximum average correct for grayscale images and enhanced images is respectively 86.27%

**Table 2.** The average correct classification results(%) using LBP texton dictionary.

scale	1 scale	2 scale	3 scale	4 scale	5 scale
Gra.(Grayscale)	72.93 ± 0.92	80.27 ± 2.31	84.67 ± 1.97	<b>86.27±1.01</b>	85.87 ± 0.23
Enh.(Enhanced)	80.80 ± 0.00	85.73 ± 0.61	88.80 ± 1.06	<b>89.60±1.20</b>	89.33 ± 1.22

**Table 3.** The average correct classification results(%) using MR8 texton dictionary.

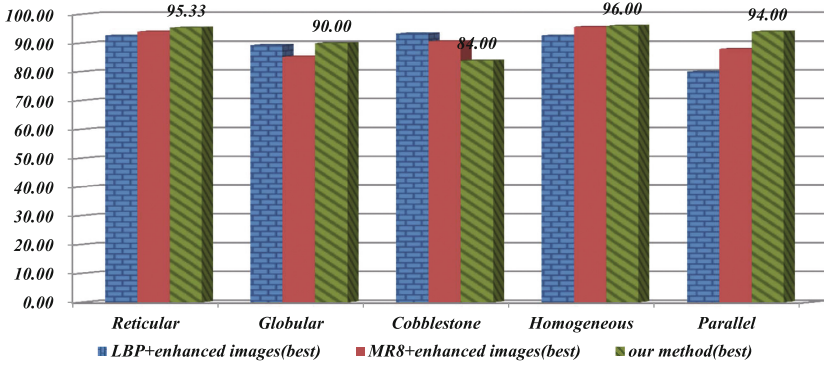
K	5	10	15	20	25	30	35
Gra.	67.87 ± 0.46	71.60 ± 2.08	73.60 ± 0.00	73.33±0.23	<b>76.53±0.46</b>	74.53 ± 1.85	76.00 ± 0.69
Enh.	87.07 ± 1.01	89.20 ± 0.69	89.60 ± 0.69	89.33±1.85	89.60 ± 1.83	90.40 ± 0.69	<b>90.93±0.92</b>

and 89.60 % at 4 scale and the method based on MR8 achieves 76.53 % maximum average correct for grayscale images and 90.93 % for enhanced images. In addition, we can find that the average classification results are obviously improved when using enhanced images to classify, especially for MR8. This shows that our enhanced method is beneficial to classify lesion images. For the proposed method, the texton number K for each pattern and the size of patch are two important parameters. In the experiments, the values of K are as the same as the values for MR8 and the patch size is set to 3 × 3, 5 × 5, 7 × 7, 9 × 9, 11 × 11 and 13 × 13 respectively. The results are shown in Table 4.

**Table 4.** The average correct classification results(%) using our method.

K	5	10	15	20	25	30	35
3×3	77.73 ± 1.51	81.07 ± 1.01	83.87 ± 1.97	83.73 ± 2.20	83.87 ± 0.23	85.07 ± 0.61	83.87 ± 0.61
5×5	84.13 ± 1.40	86.67 ± 1.40	87.73 ± 1.22	87.87 ± 1.15	88.67 ± 1.22	88.40 ± 0.69	90.13 ± 0.61
7×7	88.67 ± 0.83	89.33 ± 0.46	89.47 ± 1.01	90.13 ± 0.83	89.73 ± 0.92	90.53 ± 0.83	89.73 ± 0.23
9×9	89.33 ± 1.85	91.07 ± 1.01	91.33 ± 0.61	91.47 ± 1.40	90.67 ± 0.61	90.80 ± 0.00	90.80 ± 1.74
11×11	88.80 ± 1.06	90.93 ± 0.23	90.80 ± 1.06	91.33 ± 0.23	90.67 ± 0.23	91.60 ± 1.06	91.60 ± 0.40
13×13	89.87 ± 0.83	90.93 ± 1.01	<b>91.87±0.83</b>	91.33 ± 0.61	90.80 ± 0.40	91.47 ± 0.23	91.60 ± 1.06

From Table 4, it can be seen that in our experiments, the average correct is more than 90 % when the size of patch is more than 9 × 9 and the value of K is more than 10. Among those, there are 12 groups, the results of which are better than the best result of compared method (90.93 %). When K is 15 and patch size is 13 × 13, our method obtains best result (91.87 %). Besides, the best average accuracies of these three methods for each category of dermoscopic images are shown in Fig. 8. Clearly, in most cases our method is better than other methods. Especially for parallel pattern, the correct is more than other two methods 6 % and 14 % respectively, which shows that directly rotating lesion images to make them have same principal direction can get better result than the features including rotationally invariant.



**Fig. 8.** The best average accuracies of these three methods for each category of dermoscopic images.

## 5 Conclusions and Future Works

In this paper, we presented an effective classification method based on textons and BoFs model to classify 5 patterns (reticular, globular, cobblestone, homogeneous, and parallel) for dermoscopic images. Firstly, a response function is defined to enhance the pattern structure. Secondly, the enhanced lesion images with obvious directivity are rotated to align principal directions with horizontal direction. Then we use Otsu method to extract interesting region and obtain patch vectors for each lesion image. For each pattern, the patch vectors of training lesion images are clustered to generate  $K$  structure textons and a texton dictionary with  $5K$  elements is obtained. We adopt the theory of BoFs to obtain texton histograms for training images and testing images respectively. At last, a nearest neighbor classifier with chi-square distance is adopted to classify. The experimental results show that our enhancement method is beneficial to lesion pattern classification and our classification method, which correct classification rate achieves 91.87%, outperforms the LBP and MR8 methods.

In this paper, the 5 pattern lesion images with the size of  $128 \times 128$  are manually extracted from dermoscopic images. So our future works mainly include two aspects: (i) designing an effective method to automatically segment dermoscopic images to obtain lesion region; (ii) adding starburst and multi-component pattern into the experiments to improve the abilities of assistant diagnoses. In addition, the recognition of malignant and benign lesions is also an important future work.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (Grant nos. 61471016, 61371134 and 61271436).

## References

1. Siegel, R., Ma, J., Zou, Z., et al.: Cancer statistics. *CA Cancer J. Clin.* **64**(1), 9–29 (2014)
2. Binder, M., Schwarz, M., Winkler, A., et al.: Epiluminescence microscopy: a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Arch. Dermatol.* **131**(3), 286–291 (1995)
3. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. *Artif. Intell. Med.* **56**(2), 69–90 (2012)
4. Argenziano, G., Soyer, H.P., Chimenti, S., et al.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. *J. Am. Acad. Dermatol.* **48**(5), 679–693 (2003)
5. Argenziano, G., Soyer, H.P., Giorgi, V., Piccolo, D., Carli, P., et al.: Interactive Atlas of Dermoscopy. EDRA Medical Publishing, Milan (2000). [www.dermoscopy.org](http://www.dermoscopy.org)
6. Tanaka, T., Torii, S., Kabuta, I., et al.: Pattern classification of nevus with texture analysis. *IEEJ Trans. Electr. Electron. Eng.* **3**(1), 143–150 (2008)
7. Isasi, A.G., Zapirain, B.G., Zorrilla, A.M.: Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. *Comput. Biol. Med.* **41**(9), 742–755 (2011)
8. Sadeghi, M., Lee, T.K., McLean, D., et al.: Global pattern analysis and classification of dermoscopic images using textons. In: *SPIE Medical Imaging*. International Society for Optics and Photonics, pp. 83144X–83144X-6 (2012)
9. Serrano, C., Acha, B.: Pattern analysis of dermoscopic images based on Markov random fields. *Pattern Recognit.* **42**(6), 1052–1057 (2009)
10. Sáez, A., Serrano, C., Acha, B.: Model-based classification methods of global patterns in dermoscopic images. *IEEE Trans. Med. Imaging* **33**(5), 1137–1147 (2014)
11. Tsai, C.F.: Bag-of-words representation in image annotation: a review. *Int. Sch. Res. Not. Artif. Intell.* **2012**, 1–19 (2012)
12. Varma, M., Zisserman, A.: A statistical approach to texture classification from single images. *Int. J. Comput. Vis.* **62**(1–2), 61–81 (2005)
13. Varma, M., Zisserman, A.: A statistical approach to material classification using image patch exemplars. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(11), 2032–2047 (2009)
14. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: *Digital Image Processing Using MATLAB*. Publishing house of electronics industry, Beijing (2005)
15. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)