J Electr Eng Technol.2017; 12(1): 420-427
http://dx.doi.org/10.5370/JEET.2017.12.1.420

ISSN(Print) 1975-0102
ISSN(Online) 2093-7423

WLDF: Effective Statistical Shape Feature for Cracked Tongue Recognition

Xiao-qiang Li†, Dan Wang* and Qing Cui*

Abstract – This paper proposes a new method using Wide Line Detector based statistical shape Feature (WLDF) to identify whether or not a tongue is cracked; a cracked tongue is one of the most frequently used visible features for diagnosis in traditional Chinese Medicine (TCM). We first detected a wide line in the tongue image, and then extracted WLDF, such as the maximum length of each detected region, and the ratio between maximum length and the area of the detected region. We trained a binary support vector machine (SVM) based on the WLDF to build a classifier for cracked tongues. We conducted an experiment based on our proposed scheme, using 196 samples of cracked tongues and 245 samples of non-cracked tongues. The results of the experiment indicate that the recognition accuracy of the proposed method is greater than 95%. In addition, we provide an analysis of the results of this experiment with different parameters, demonstrating the feasibility and effectiveness of the proposed scheme.

Keywords: Tongue image, Cracked tongue recognition, Support vector machine

1. Introduction

Diagnosis based on condition of the tongue [1-3] is one of the most important and valuable diagnostic methods in traditional Chinese medicine (TCM) and has been widely used in clinical analysis and applications for thousands of years. According to TCM theory, the tongue reflects the condition of the organs, providing novel evidence for diagnosis. Whenever there is a complex disorder in vivo, many diagnostic possibilities exist and examining the tongue may instantly clarify the main pathological processes.

In recent years, with ubiquitous application of advanced computer technologies in medical science, we cannot ignore new possibilities to improve TCM with computer technologies [4-22]; these possibilities for improvement have drawn attention from myriads of researchers. Chiu [4, 5] built a computerized tongue examination system (CTES) based on computerized image analysis. The colors of the tongue and the thickness of its coating were identified using chromatic and textural algorithm. Pang et al. [6] used Bayesian networks to build tongue diagnosis model which provided a systematic and objective diagnostic standard for the tongue diagnosis. Zhang et al. [7] proposed an in-depth systematic tongue color analysis system for medical applications. Using the tongue color gamut, tongue foreground pixels are first extracted and assigned to one of 12 colors representing this gamut. The ratio of each color for the entire image is calculated and forms a tongue color feature vector. Many other papers [8-11] also discussed the impact of color on the tongue analysis. Before the process of tongue classification, an exact region of the tongue must be extracted from an image, [12-13] try to proposed automated tongue segmentation method. Some papers [14-18] designed image analysis system for feature extraction and diagnostics. Pang et al. [18] even present a tongue-computing model for the diagnosis of appendicitis based on quantitative measurements that include chromatic and textural feature. In addition, the teeth mark is caused by long-period compression of teeth because of the enlargement of tongue body. The teeth-marked tongue usually appears along with spleen deficiency [19]. Shao et al. [20] presented a better teeth mark recognition method than the work of [21, 22], which are concentrated on features of convex, the change of brightness of tongue. During the past several years, certain achievements have been made in tongue color analysis and diagnostic classification technologies.

TCM explains that some cracks on the tongue can be categorized into a particular kind of crack, possessing a peculiar characteristic in that these cracks are more or less obvious radial cracks on the tongue surfaces. Cracked tongues refer to the surface of the tongue covered with various kinds of cracks or lines in deep or shallow shape caused by the fusion or separation of the ligula papillae [23]. Under normal circumstances, the surface of the tongue should be smooth and show no apparent cracks [1]. Obvious radial cracks that appear on the surface suggest a deficiency of qi-blood and the consumption of yin by excessive heat and, sometimes, blood stasis. For instance, a deep crack in the center reaching the tip reflects hyperactivity of heart fire.

† Corresponding Author: School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China. (xqli@shu.edu.cn)
* School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China
Received: October 31, 2015; Accepted: September 22, 2016

Fig. 1. The process to obtain target image. Tongue image with (a) and without (b) undesired black background; (c) the target image in blue box.

Notwithstanding the fact that cracked tongues play a pivotal role in clinical practice of TCM, unfortunately, little work has been done to digitally analyze any part of it. Liu and Zhang extract tongue cracks to some degree using a wide line detector (WLD) [24, 25], which has been considered quite an effective method until now. However, Liu and Zhang only concentrate on identifying tagged pixels of cracks with certain thresholds. Using this method, no effective classifier, such as an SVM, classifies tongues, which makes their classification results unconvincing and too specific for general use. On the other hand, cracks tagged by drawing white lines is not pragmatic under all circumstances, because cracks often appear so complicated and intertwined that it is difficult to tag them accurately using this method. Moreover, to the best of our knowledge, no prior method can identify whether or not a tongue is cracked.

In this study, we present a recognition scheme that can be regarded as further extension and application of the method proposed in [24, 25], including the stages of preprocessing, feature extraction, and classification. In the preprocessing stage, we adjust the sample images into a proper size based on the bilinear interpolation method, and extract our target images. Then we extract the WLDF from a set of preprocessed images, which is considered the most pivotal part in this scheme. Finally, we test the features on an SVM classifier on the basis of the training set of samples previously classified by a TCM veteran practitioner.

The rest of this paper is organized as follows. Section 2 describes the cracked tongue recognition scheme. In Section 3, we present and analyze the experimental results of the proposed method. Section 4 offers our conclusion.

2. Cracked Tongue Recognition Scheme

2.1 Preprocessing

Preprocessing is one of the prime stages for our scheme, the purpose of which is to make the captured samples meet the ends of recognition. The captured tongue sample images after segmentation are still quite large, and usually have a large black background. First, we ruled out the undesired black background (Fig. 1). We then resized the images to a specific size of 300*L (L is the ratio of height to width of the original image). The final target image was then selected right in the center of the image (Fig. 1), on which further feature extraction could be exerted. The position and size of the target image is given by Eq. (1).

$$\begin{aligned} W_t &= 0.6 \times W_r \\ H_t &= 0.6 \times H_r \\ S &= (0.2 \times W_r, 0.2 \times H_r) \end{aligned} \tag{1}$$

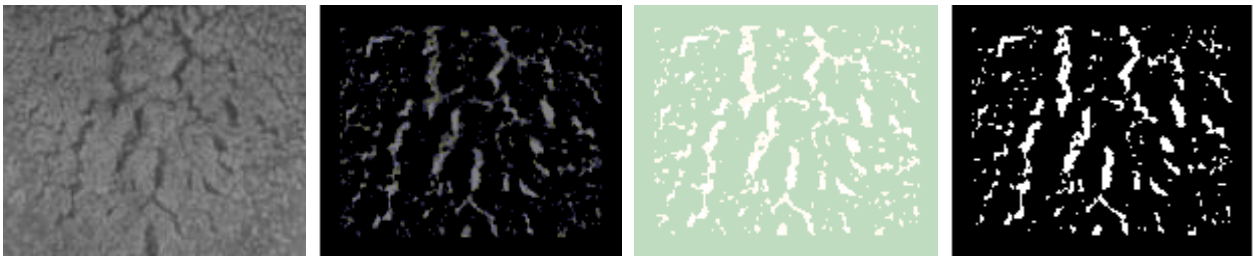
where W_t and H_t are the width and the height of the target image, W_t and H_r are the width and the height of the resized image, respectively, and S is the origin of coordinates of the target image relative to the resized one. We chose such a region as our target image because it could include majority of cracks on the tongue surface while avoiding the effective factors on the extreme top and extreme bottom of the tongue body.

2.2 Wide Line Detector Based Statistical Shape Feature (WLDF) Extraction

In this section, we will discuss WLD in detail. Then we will deliver the strategy to extract the WLDF for recognition. Finally, the classification strategy will be introduced.

2.2.1Wide Line Detector (WLD)

We modified and improved the method proposed in [25] to extract the tongue crack features on the basis of their properties. This method, called the wide line detector (WLD), extracts a line completely by using a nonlinear filter based on the initial line length response via an inverse-Gaussian (IG) weighting mask. They improved the method proposed in [24] using a Gaussian weighting mask.



(a) (b) (c) (d)

Fig. 2. Wide line detected scheme: (a) Target image; (b) line-strength image of (a); (d) histogram equalization of (c); (d) binary image of (c).

$$IG(x,y,x_0,y_0,r)=\frac{1-e^{-((x-x_0)^2+(y-y_0)^2)/2r^2}}{(2r+1)^2-\sum_{\substack{y_0-r\le y\le y_0+r\\ x_0-r\le x\le x_0+r}} e^{-((x-x_0)^2+(y-y_0)^2)/2r^2}} \tag{2}$$

$$c(x,y,x_0,y_0)=\begin{cases} IG(x,y,x_0,y_0,r)\times sech^5(((I(x,y)-I(x_0,y_0))/t) & I(x,y)>I(x_0,y_0) \\ IG(x,y,x_0,y_0,r) & I(x,y)\le I(x_0,y_0) \end{cases} \tag{3}$$

Here, we discuss the WLD and the improvements made so that it to be applicable to our WLDF extraction. The WLD uses an IG mask given by Eq. (2), where (x_0, y_0) is the coordinate of the center, (x, y) is the coordinate of any other pixel within the mask, and r is half of the side-length of the quadrate mask. Instead of the circular mask in [24], we chose the quadrate mask to make the WLD more feasible for coding while maintaining the accuracy of the results.

With the IG mask, the WLD first calculates the intensity of the center of the mask and groups pixels that have similar intensities to the center into a weighted mask having similar brightness (WMSB). In this method, we measure the similarity by Eq. (3). Where

$$\sec h(x)=2/\left(e^{-x}+e^{x}\right) \tag{4}$$

$I(x, y)$ is the brightness of the pixel (x, y) of the tongue body image, t is the brightness contrast threshold, and c is the output of the weighting comparison. This function measures the smoothness, and does not have too large of an effect on c when a brightness of the pixel changes slightly, especially when it is near the threshold.

After obtaining the value of c for every pixel within the mask, the mass of the WMSB in the center of the mask is given by

$$m(x_0,y_0)=\sum_{\substack{x_0-r\le x\le x_0+r\\ y_0-r\le y\le y_0+r}} c(x,y,x_0,y_0) \tag{5}$$

The initial line strength L is the inverse WMSB mass obtained by using the following rule:

$$L(x_0,y_0)=\begin{cases} g-m(x_0,y_0), & m(x_0,y_0)<g-\delta \\ 0, & others \end{cases} \tag{6}$$

where g is the geometric threshold equal to 0.5, and δ is the output threshold. We limit g to a constant value instead of the values used in the method proposed in [25], which significantly releases the burden of calculation. In addition, we added the output threshold δ here to make the WLD more flexible for different cases. As a normalized quadrate mask is used, the maximum value of m is not larger than but very close to the unity and, thereby, the line strength ranges between zero and one. Fig. 2(b) shows the line-strength image.

Here we see that the WLD needs three parameters: brightness contrast threshold t, half side-length of the quadrate mask r, and output threshold δ. According to [25], it is the standard deviation of all the non-zero pixels in the target image given by Eq. (7).

$$t=\left(\frac{\sum_{i=1}^{N}\left(I(x_i,y_i)-\overline{I(x_m+y_m)}\right)^2}{N-1}\right)^{1/2} \tag{7}$$

$$\overline{I(x_m+y_m)}=\sum_{i=1}^{N} I(x_i,y_i)/N$$

N is the number of non-zero pixels of the target image. Although there have been discussions about the relationship between the mask radius r and the width of detected lines, it is still difficult to determine a good value using only the theory for every case, because of huge variety among the tongue cracks. We can attempt to determine best values of r and δ by performing experiments and analyzing the results. The analysis about selecting r and δ will be offered in the section of experimental results.

2.2.2 WLDF extraction strategy

First, we exerted the WLD on the pixels within the

preprocessed target image shown in Fig. 2(a) of a given sample, and obtained a line-strength image shown in Fig. 2(b). The value of the pixels could be extracted from one of the R, G, B, and I channels of the target image. According to our experimental results, we found that the B and G channels have better performance in recognition. Therefore, we could choose one of the B and G channels as our object of research.

Second, we performed histogram equalization to the line-strength image and Fig. 2(c) describes the result because the line-strength image is rather a dark image with low contrast. The purpose of histogram equalization is to stretch the gray levels of the line-strength image, and the foreground and the background are more distinct from each other. Hence, some normal textures on the surface detected as fake cracks could be better ruled out in the next step of image binarization.

Third, we binarized the histogram equalized line-strength image using the threshold of Otsu [26], and the set of pixels involved in each bright connected regions in this binary image was regarded as our detected region (DR) shown in Fig. 2(d), namely, a potential crack.

We focused on two features of the detected regions, one is the maximum length of each detected region denoted by d, and the other is the ratio denoted by f between the maximum length and the area of the detected region. The two features are described in the following equations:

$$d_n = \max_{(x_i,y_i),(x_j,y_j)\in DR_n} \{\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}\} \tag{8}$$

$$f_n = d_n / \mathrm{card}(DR_n) \tag{9}$$

where card (DR_n) counts the number of pixel elements in a certain DR. Because tongue cracks are long and thin, and usually appear radial, f is generally much larger than $1/\pi$ (the area of big dot).

Here, we define the WLDF as a vector given by (r, $d1$, $d2$, $d3$, $f1$, $f2$, $f3$, S(d), S(f)) where r is the half side-length of the quadrate mask; $d1$, $d2$ and $d3$ represent the three largest maximum length of all the detected regions; $f1$, $f2$ and $f3$ represent the ratio between maximum length and area of the correspond detected region with maximum length; S(d) is the standard deviation of d of all the detected regions; and S(f) is the standard deviation of f in all the detected regions. The feature of the half side-length of the mask is limited to a number proportional to the width or height of the target image, which is given by

$$r = round(p \times \max\{W_t, H_t\} \tag{10}$$

where p is a certain coefficient of proportionality and W_t and H_t are the width and height of the target image, respectively. Because r is proportional to the side-length of the target image, r presents the effect of the size of the target image in WLDF. Moreover, r is one of the key parameters in WLD since the performance of WLD is affected depending on the value of r. The value of r is also proportional to p, hence, p is one of the direct effective factors to r. When the target image is considered, p becomes the only optional variable controlling the value of r because W_t and H_t are fixed. We will discuss the coefficient of proportionality p in the section of experimental results.

We extracted (d_1, d_2, d_3) into WLDFs to evaluate the length of the detected regions, in others word, three regions with largest length is selected. We describe these regions as the following:

$$\begin{aligned} d_1 &= \max\{U(d)\} \\ d_2 &= \max\{U(d) - \{d_1\}\} \\ d_3 &= \max\{U(d) - \{d_1\} - \{d_2\}\} \end{aligned} \tag{11}$$

Where $U(d)$ is the universal set of lengths of all the detected regions in one target image.

Similarly, (f_1, f_2, f_3) determines whether or not the detected regions(DR1, DR2, DR3) with lengths (d_1, d_2, d_3) are thin, and is given by

$$\begin{aligned} f_1 &= d_1 / \mathrm{card}(DR_1) \\ f_2 &= d_2 / \mathrm{card}(DR_2) \\ f_3 &= d_3 / \mathrm{card}(DR_3) \end{aligned} \tag{12}$$

We denote the universal sets of ratios between maximum length and the areas of all the detected regions in one target image by $U(f)$. $S(d)$ and $S(f)$ are given by

$$S(d) = \sqrt{\frac{1}{N}\sum_{d_i\in U(d)} (d_i - d_m)^2} \tag{13}$$

$$S(f) = \sqrt{\frac{1}{N}\sum_{f_i\in U(f)} (f_i - f_m)^2} \tag{14}$$

where N is the number of all the detected regions in $U(d)$ or $U(f)$ U(f)., d_m and f_m are the mean values of the elements of U(d) and $U(f)$, respectively.

We selected $S(d)$ and $S(f)$ as features of the WLDF so that tongues having small cracks but being fairly smooth in non-cracked region should be recognized as cracked tongues. The standard deviation describes a divergence of data, and if the value of d or f is much larger than the mean value, the standard deviation should fluctuate a lot. Therefore, even though it is a small crack on a smooth surface, the two standard deviations could make it rather outstanding.

2.3 Classification strategy

In our recognition scheme, we have already presented the method of preprocessing and feature extraction earlier in this article. We chose the Support Vector Machine (SVM) with a linear kernel as our classifier. We do not focus on feature selection here because the number of the

features is not large.

3. Experimental Results and Discussion

This section refers to the performance of classification based on the WLDF. Here we considered a set of samples consisting of 245 non-cracked tongues and 196 cracked tongues. We selected all the samples from clinical practice and Chinese medicine experts categorized them so that we could make sure that they are representative. We applied the SVM with a linear kernel as the classifier. We organized this section as follows: First, we implemented the experiment with given parameters of $p(r)$ and δ, and then we presented a general performance of this recognition scheme on a different channel (R, G, B, and I) of the sample images. Second, we discussed the performance of different values of the parameters of $p(r)$ and δ. In addition, we also presented experiment results of other classifiers such as LDA and KNN, and conducted experiment of feature selection based on SVM-REF to investigate which feature is better for cracked tongue recognition.

3.1 Performance With Given Parameters on the R, G, B, and I Channel

As parameter r is proportional to the coefficient of p, we set the value of parameter p to 1/15. The parameter δ is limited to the value of 0.1. We extract the WLDF, respectively, of the R(Red), G(Green), B(Blue), and I(Intensity) channel of every given sample as four sets of vectors with their class symbols. For each set, we adopt the leave-one-out cross-validation (LOOCV) method to test the performance of classification. We evaluate our classification performance using five categories: Accuracy (AC), True Positive Rate (TPR), True Negative Rate (TNR), Sensitivity (SE), and Specificity (SP). The cracked samples are considered as positives, and the non-cracked ones are considered as negatives. AC is the correct rate for the entire set of test samples. TPR is the correct rate for positives, and TNR is the correct rate for negatives. SE is given by

$$SE = TPR/(TPR + 1 - TNR) \tag{15}$$

When SE is larger, we can ensure that less positive samples are misjudged as negative ones. SP is given by

$$SP = TNR/(TNR + 1 - TPR) \tag{16}$$

When SP is larger, we also can ensure that less number of negative samples are misjudged as positive samples.

Table 1 shows the results of the experiment, respectively, on the R, G, B, and I channels. As Table.1 shows, it is easy to see that the G and B channels outperform the other channels. The R channel has poor performance on the TPR, which indicates that the R channel is not very efficient for correct classification of positive samples. The G, B, and I channels have rather balanced performance on both the TPR and TNR. As shown in Table 1, although the B channel has a slight advance over the G channel, we cannot identify which is better because the set of samples is not very large. Based on the previously discussed analysis, this recognition scheme can reach an accuracy of at least 95% on the given set of samples.

Table 1. Experiment result with parameter $p = 1/15$ and $\delta = 0.1$

Channel	AC	TPR	TNR	SE	SP
R Channel	0.8589	0.8631	0.9959	0.9952	0.8791
G Channel	0.9461	0.9710	0.9751	0.9750	0.9711
B Channel	0.9502	0.9751	0.9751	0.9751	0.9751
I Channel	0.9295	0.9668	0.9627	0.9628	0.9667

3.2 Performance Analysis with Varying Values of δ and $p(r)$

As mentioned above, δ and $p(r)$ are the key parameters for WLDF extraction. This section mainly concentrates on the performance of classification with different δ and $p(r)$, which is demonstrated as the form of curves of classification accuracy. Using the method mentioned in 3.1, we calculate the average accuracy of the LOOCV test for each δ and $p(r)$. We performed the test on the B channel of the samples.

3.2.1 Analysis for a constant p value

We limit p to the value of 1/15, and δ changes in [0, 0.38] by the step of 0.02. Fig. 3 shows the results of the experiment. The accuracy is high around the range of δ between 0 and 0.24, as illustrated in Fig. 3. When δ is beyond 0.24, there is a sharp decline in the accuracy.

3.2.2 Analysis for a constant δ value

We limit δ to the value of 0.1, and p changes in [0.02 0.40] by the step of 0.02. Fig. 4 shows these results. As shown in Fig. 4, the accuracy is high when p is in the range of 0.06 and 0.20. When δ is greater than 0.20, the accuracy has a novel decline. When δ is less than 0.06, the accuracy decreases.

Based on Fig. 3 and Fig. 4, it is clear that one parameter has a range with good classification accuracy, while the other is fixed. Hence, for each set of samples, there is a certain value of $p(r)$ and δ at which maximum accuracy can be achieved; this concept needs further research.

3.3 Experiment results with different classifiers

As a performance comparison, we also consider testing

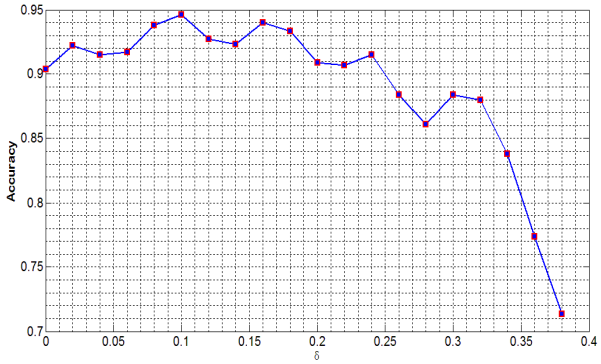


Fig. 3. The accuracy of the proposed method with the constant parameter p and varying δ in [0, 0.38]

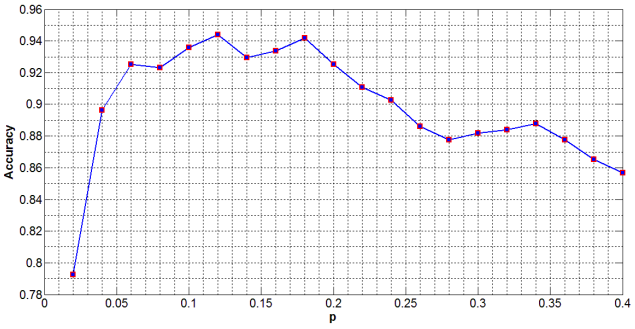


Fig. 4. The accuracy of the proposed method with the constant parameter δ and varying p in [0.02 0.40]

the KNN classifier [27] and the LDA classifier [28] on the same feature and datasets as SVM classifier. Experimental results on the R, G, B, and I channels are showed in Fig. 5. It is easy to see that SVM classifier performs better than KNN classifier and LDA classifier. Take B channel for example, the accuracy of SVM, KNN and LDA is 95.02%, 88.7% and 87.83%. So, SVM classifier is appropriate to be applied in the cracked tongue recognition task.

3.4 Feature selection based on SVM-REF

Feature selection is an important and necessary step in machine learning and data mining. By removing most irrelevant and redundant features from the data, feature selection helps improve the performance of learning models by alleviating the effect of the curse of dimensionality, enhancing generalization capability, speeding up learning process and improving model interpretability. There are many feature selection methods. Compared with other feature selection methods, SVM-RFE is a scalable, efficient wrappers method [29].

Although there are only nine features used in our method, we conduct experiment of feature selection based on SVM-REF. The results are presented in Fig.6, demonstrates that different feature subsets lead to different forecast accuracy. Comparing to the total accuracy on the subset of less features, the results using SVM on the 9 features obtain higher accuracy than others.

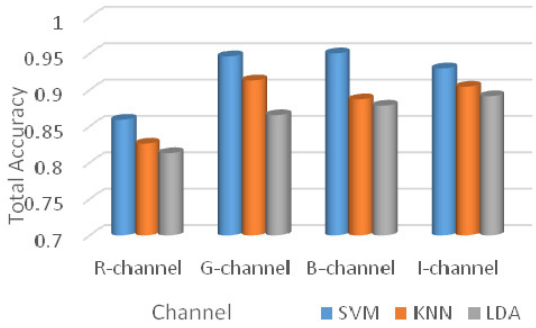


Fig. 5. Performance comparison between different classifiers

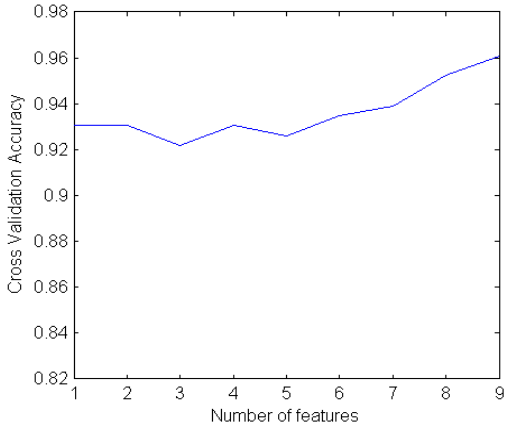


Fig. 6. The results of the cross validation accuracy versus retained feature number by SVM-RFE.

The results also demonstrates that these 9 statistical shape features elaborately designed in the proposed method is appreciate and effective.

4. Conclusion

In this study, we proposed an effective feature to identify cracked tongues, the WLDF. We modified and improved some of the details of the WLD proposed in [25] so that it can be more feasible for extracting the WLDF features applied in the proposed method. To be specific, the geometric threshold is limited to a constant number to reduce the amount of computations; the output threshold is introduced here to make the WLD more flexible in complex situations. The WLDF is defined as a vector (r, d_1, d_1, d_2, d_3, f_1, f_2, f_3, $S(d)$, $S(f)$).

A lot of experiments were exerted on a set of 245 non-cracked tongues and 196 cracked tongues. We chose SVM as our classifier, which was proved to produces better results than KNN and LDA classifier. As a result, we obtained a high accuracy of above 95%, and found that the WLDF performs better on the B, G, and I channels than it does on the R channel. Moreover, we discussed the effects of different $p(r)$ and δ, which demonstrates that there is a certain range of values for $p(r)$ and δ that perform well in

the classification. In addition, experimental results of feature selection demonstrates that these 9 statistical shape features is appreciate and effective.

In the future, we will try out more feature and classifycation methods to make our research more feasible in clinical practice. Specifically, we aim to identify the quantity and type of cracks on the tongue's surface rather than only identifying whether or not a tongue has cracks.

References

[1] B. Kirschbaum, Atlas of Chinese Tongue Diagnosis. Eastland, Seattle, WA, 2000.

[2] H. Yiu. Fundamental of Traditional Chinese Medicine. Foreign language Press, Beijing, 1992.

[3] N. Li, D. Zhang and K. Wang. Tongue Diagnostics, Beijing, China: Shed-Yuan Publishing, 2006.

[4] C.C. Chiu, H.S. Lin, S.L. Lin, A Structural Texture Recognition Approach for Medical Diagnosis through Tongue. Biomedical Engineering, Application, Basis and Communication 7, 143-148, 1995.

[5] C. C. Chiu, A Novel Approach based on Computerized Image Analysis for Traditional Chinese Medical Diagnosis of the Tongue. Computer Methods and Programs in Biomedicine 61, 77-89, 2000.

[6] B. Pang, D. Zhang, N. Li and et al. Computerized Tongue Diagnosis based on Bayesian Networks. IEEE Trans. Biomedical Engineering 51, 1803-1810, 2004.

[7] B. Zhang, X. Wang, J. You and et al. Tongue Color Analysis for Medical Application, Evidence-Based Complementary and Alternative Medicine Volume 2013, Article ID 264742, 11 pages.

[8] C. H. Li and P. C. Yuen, "Tongue image matching using color content," Pattern Recognition, vol. 35, no. 2, pp. 407-419, 2002.

[9] X. Wang and D. Zhang, An Optimized Tongue Image Color Correction Scheme, IEEE Tran. On Information Technology in Biomedicine, 14(6): 1355-1364, 2010.

[10] L. Zhuo, P Zhang, B. Chen and et al. Automatic Tongue Color Analysis of Traditional Chinese Medicine Based on Image, In: 13th International Conference on Control, Automation, Robotics & Vision, Singapore, pp. 637-641, 2014.

[11] X. Wang and D. Zhang, "A high quality color imaging system for computerized tongue image analysis," Expert Syst. Appl., vol. 40, no. 15, pp. 5854-5866, 2013.

[12] B. Pang, D. Zhang, K.Q. Wang.The Bi-Elliptical Deformable Contour and Its Application to Automated Tongue Segmentation in Chinese Medicine. IEEE Trans. Medical Imaging 24, 946-956, 2005.

[13] Z. Fu, W. Li, X. Li and et al. Automatic tongue location and segmentation. In: IEEE IntConf Audio Lang Image Process, Shanghai, China, pp 1050-1055, 2008.

[14] X. Wang, B. Zhang, Z. Yang, et al. Statistical Analysis of Tongue Images for Feature Extraction and Diagnostics, IEEE Tran. On Image Processing, vol. 22, no. 5, pp. 5336-5347, 2013.

[15] Y. Wang, Y. Zhou, J. Yang and Q. Xu, "An image analysis system for tongue diagnosis in traditional Chinese medicine," in Computational and Information Science,, vol. 3314 of Lecture Notes in Computer Science, pp. 1181-1186, 2004.

[16] R. Kanawong, T. O. Ajayi, T. Ma, et al. "Automated tongue feature extraction for ZHENG classification in traditional Chinese medicine" Evidence-Based Complementary and Alternative Medicine, vol. 2012, Article ID 912852, 14 pages, 2012.

[17] S. Yamamoto, N. Tsumura, T. Nagaguchi, et al., "Regional image analysis of the tongue color spectrum," International Journal of Computer Assisted Radiology and Surgery, vol. 6, no. 1, pp. 143-152, 2011.

[18] B. Pang, D. Zhang, K. Wang. Tongue image analysis for appendicitis diagnosis. Inform Science, 175(3): 160-176, 2005.

[19] J. F. Li, N. M. Li, K. Q. Wang, et al, Extracting feature of teeth-marked tongue image, Proc. of diagnosis section of China Society of Integrated Traditional Chinese and Western Medicine, pp.100-105, Beijing, 2009.

[20] Q. Shao, X. Q. Li, Z. C. Fu, Recognition of teeth-marked tongue based on gradient of concave region. In: IEEE IntConf Audio Lang Image Process, Shanghai, China, pp 968-972, 2014.

[21] S.D. Zhong, Y.K. Wei and Z.G. Xie, Research on a tooth-marked rapid locating method based on convex hulls, Microcomputer Information, 25(3): pp. 312-314, 2009.

[22] Y.T. Zhang, Research on analysis method of tongue and teeth-marked tongue, Doctor Dissertation, Beijing University of Chinese Medicine, 2005.

[23] Y. Xin, X.Z. Guo, L.S. Zhang, Tongue Diagnosis (Chinese-English). Tianjin Science and Technology Translation Publisher, 2001.

[24] L. Liu, D. Zhang, J. You, Detecting Wide Lines using Isotropic Nonlinear Filter. IEEE Trans. Image Processing 16, 1584-1595, 2007.

[25] L. Liu and D. Zhang, Extracting Tongue Cracks Using the Wide Line Detector, ICMB 2008, LNCS 4901, pp. 49-56, 2008.

[26] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. System, Man, and Cybernetics, 9(1): pp. 62-66, 1979.

[27] C. Domeniconi, J. Peng and D. Gunopulos, Locally adaptive metric nearest-neighbor classification, IEEE Trans. PAMI, 24(9): pp. 1281-1285.

[28] C.-J. Liu., H. Wechsler. Gabor Feature Based Classification Using the Enhanced Fisher Linear

Discriminant Model for Face Recognition. IEEE Trans. Image Processing. 11(4): pp. 467-476, 2002.

[29] Kai-Bo Duan, Jagath C. Rajapakse and et al, "Multiple SVM-RFE for Gene Selection in Cancer Classification With Expression Data" IEEE Trans. on Nanobioscience, 4(3): pp. 228-234, 2005.

Xiao-qiang Li He received the Ph.D. degree in computer science from Fudan University, Shanghai, China, in 2004. He is now associate professor of computer science at Shanghai University, Shanghai, China. His research interests include image processing, pattern recognition, computer vision, and machine learning.

Dan Wang She received the B.S. degree in computer science from Shanghai University. Now she is a M.S. candidate of computer science at Shanghai University, Shanghai, China. Her research interests include computer vision and machine learning.

Qing Cui He received the B.S. degree in computer science from Yantai University, Yantai, China, in 2014. Now, he is a M.S. candidate of computer science at Shanghai University, Shanghai, China. His research interests include computer vision and machine learning.