2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

An Automatic Petechia Dots Detection Method on Tongue*

Chunqi Qian, Hongyu Gu, Zhecheng Yang, Chuanchi Wang, Jingqing Hu, Hong Chen

***Abstract*—Tongue diagnosis with features like tongue coating, petechia, color, size and so on is of great effectiveness and convenience in traditional Chinese medicine. With the development of image processing techniques, automatic image processing can reduce hospital inspection for patients. However, there are ubiquitous problems of inadequate accuracy in petechia dots detection with previous methods. In this paper, we propose a method of petechia dots detection on tongue based on *SimpleBlobDetector* function in OpenCV library and support vector machines model, which improves the detective accuracy. We test 128 clinic tongue images and select 9 of the images with plentiful petechia dots for further experiments. Our method achieves mean value of false alarm rate 4.6% and missing alarm rate 11.8%, which have 19.4% and 8.2% reduction respectively compared to previous work.**

***Clinical Relevance*—The method can provide detailed information of tongue, which assists doctors to investigate curative effect.**

I. INTRODUCTION

In traditional Chinese medicine, tongue diagnosis is practically applied in evaluating patients' physical condition. The abnormal changes of color, tongue coating and appearance of petechia dots on the surface of tongue have a strong correlation with illness, even some incubated diseases in organs inside body [1]. However, diagnosis is conducted by doctors according to what they see with their naked eyes, and diagnosis results are usually affected by environments, such as light on tongue. What is more, doctors are unable to monitor patient's recovery process to adjust the treatment plans delicately without hospitalization. With the development of image processing technology and portable devices, it is possible to obtain detailed information of petechia dots on the surface of tongue with the tongue pictures taken when each clinic visit, such as their area and number, which enables doctors find detailed change of patients' tongue and adjust their treatment plans accordingly.

Petechia dots are tiny bluish, purplish or darkish dots on the tongue surface which have diverse shapes and distributions [2]. Many efforts have been made to distinguish petechia dots accurately in recent years. Some works focus on sole or a few petechia dots with Top-hat high-pass filtering and Gaussian histogram [1]. Researchers in [3] explore an analytical method based on watershed segmentation to identify small amounts of dots. As the algorithms rely on predetermined threshold, the method could only find out if the petechia dots exist. Besides,

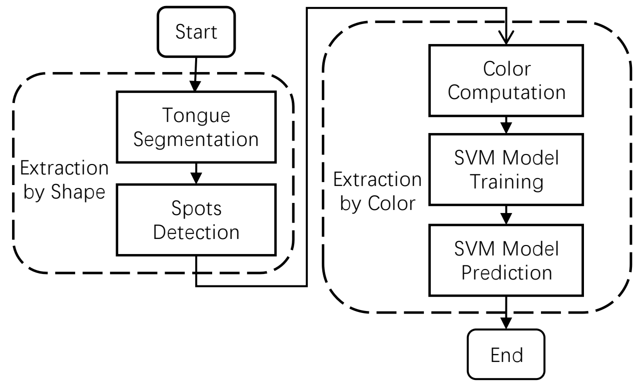


Figure 1. The propose method of petechia detection.

the total area and number of petechia dots are not calculated, which are important for doctors to judge the disease status and curative effect. For a more promising application, a method focusing on dense dots on the tongue surface is proposed in [4], in which a convolutional mask is used to highlight the objective dots and a color gamut is set to exclude irrelevant dots [4]. Region growing algorithm, which can tell homogenous dots with semblable gray value, is also adopted in petechia dots extraction [5]. However, the algorithm finds a dense mass of petechia dots instead of sole petechia dots. Authors in [6] use *SimpleBlobDetector* function and K-means to achieve the defined detection accuracy of 89.9%, false alarm rate of 6% and missing alarm rate of 10.1% [6]. However, the approach has false alarm rate of 24% and missing alarm rate of 20% when applied in our clinic tongue images. Authors in [7] and [8] adopt support vector machines (SVM) in tongue feature extraction effectively.

In this paper, we put forward a method for petechia dots detection that coordinates *SimpleBlobDetector* function and SVM to realize a higher recognition accuracy of petechia dots. Because the color of petechia dots is different from that of other part of tongue, and the shape of petechia dots is similar with circle, they can be extracted from tongue image by shape and color [9]. Considering these characteristics, we adopt both *SimpleBlobDetector* function and SVM model in our works.

The complete process of our method is shown in Fig.1, which includes two main steps. The first step is to extract valid area and locate all circular dots from the tongue image preliminary. Then, region growing algorithm is adopted to extract color information of petechia dots to establish datasets for SVM model training. With a generated binary classifier,

* This work is supported by the National Science and Technology Major Project from Minister of Science and Technology, China (Grant No. 2018AAA0103100), Guangzhou Foshan Science and Technology Innovation Project (No. 2020001005585), and National Natural Science Foundation of China (No. U19B2041), partly supported by Beijing Engineering Research Center (No. BG0149) and Tsinghua National Laboratory for Information Science and Technology (No. 042003266).

C. Qian, H. Gu, Z. Yang, H, Chen are with School of Integrated Circuits, Tsinghua University, Boulder, Beijing 100084 China (corresponding author to provide phone: 010-62795097; e-mail: hongchen@tsinghua.edu.cn).
C. Wang, J. Hu are with Institute of Basic Theory for Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700 China

the real petechia dots can be extracted from all the dots on tongue.

II. THE PROPOSED METHODS

A. Automatic Tongue Segmentation

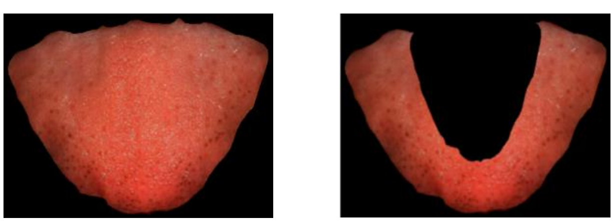
Because of the existence of fissure and texture on the tongue, the detection of petechia dots is prone to be influenced and the detection accuracy is low as a result. To reduce the influence of the non-ideal factors, we first segment the tongue image.

Considering that petechia dots aggregate on the periphery of tongue [2] while fissure and texture exist on the middle and the rear of tongue, we design a mask with background color to cover the invalid area. The mask is a miniature of tongue image, given that the margin of tongue can be retained integrally with the mask. The experimental results show that when the mask has the size of 0.8 times the length and 0.5 times of the width of the tongue, the fissure and texture can be covered without any loss of petechia dots as shown in Fig. 2, in which the original tongue image is on the left, and the tongue with mask is on the right.

B. Petechia Dots Extraction with SimpleBlobDetector Function

Generally, petechia dots are circular dots that separate on tongue. Thus, *SimpleBlobDetector* function is used to recognize petechia dots as many as possible in our method.

As said above, the color of petechia dots is different from that of other parts of the tongue. We find that the difference of saturation (S) between petechia dots and non-petechia dots is more significant than that of hue (H) or value (V) in HSV color space and that of grey value of R-component, G-component and B-component in RGB color space. *SimpleBlobDetector* is a function of OpenCV library, which can locate all circular dots according to the parameters of *threshold*, *repeatability*, *area*, *circularity*, *interiaRatio* and *convexity,* as shown in Table I. The values of these parameters are determined experimentally except for *maxThreshold*, *minArea* and *maxArea* in our work.



(a) Original tongue image (b) Tongue image with mask

Figure 2. Example of a mask on tongue. The length of mask is 0.8 times and the width of mask is 0.5 times of that of the tongue.

TABLE I. PARAMETERS OF SIMPLBLOBDETECTOR FUNCTION

Parameter Type	Chosen Value
thresholdStep	1
minThreshold	0
maxThreshold	-0.901**mean_S*+270.44
minArea	*width**0.1452-52.975
maxArea	*width**0.1452+12.025
minCircularity	0.3
minConvexity	0.5
minInertiaRatio	0.1
minDistBetweenBlobs	5

The parameter of *ThresholdStep* determines step of thresholds of images' binarization, which is used to obtain a series of binary images from gray image of tongue. The smaller *thresholdStep* is, the more selected dots *SimpleBlobDetector* function obtains. When *ThresholdStep* is 1, nearly all dots on the tongue image can be located. The parameter of *minThreshold* stands for the minimum value of thresholds for the series of binary images. As petechia dots are darker than other part of tongue, we can set *minThreshold* 0. The parameter of *minCircularity* represents the minimum circularity of chosen dots and the range is [*minCircularity*,1]. The parameter of *minConvexity* relates with ratio of area and convex hull area, whose range is also [*minConvexity*,1]. The parameter of *minIntertiaRatio* is the ratio of minimum and maximum rotation inertia. The more of the three parameters above close to 1, the more likely the chosen dots are circles. The experiments show that when *minCircularity*, *minConvexity*, *minIntertiaRatio* and *minRepeatability* are 0.3, 0.5, 0.1 and 2 respectively, nearly all the dots can be selected. The parameter of *minDistBetweenBlob* represents the distance between selected dots, and can be set to a proper value (5 in our work) to avoid overlapping.

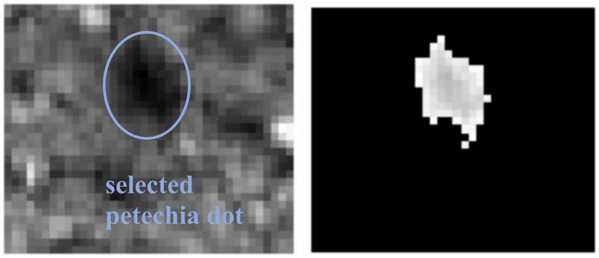
Since the value of *thresholdStep* is chosen to be small enough (1 in our work) to circle dots as many as possible including non-petechia dots. To exclude the non-petechia dots, we set *maxThreshold* to exclude the dots whose gray value are smaller than that of petechia dots. It is observed that *MaxThreshold* is a function of the mean value of S-component (*mean_S*). Hence, we first calculate the value of *mean_S*. Note that most of the tongues are covered by tongue coating, it is necessary to remove the tongue coating when calculating the value of *mean_S*. With the value of *mean_S*, we find that the maximum threshold (*maxThreshold*) of each tongue image is linear to *mean_S*, in our work *maxThreshold= -0.901*mean_S+270.44* for 9 tongue images with petechia dots. Similarly, *minArea* and *maxArea*, which determine the size of petechia dots area, are the function of the width of tongue as shown in Table I.

C. Acquisition of Color Information

As many non-petechia dots are not reddish or violet like petechia dots, petechia dots can be selected by their colors. To eliminate the negative effects brought by heterogeneous color of tongue, the *S*-component of the tongue image could not be used directly. The gray value difference between petechia dot and the surrounding area is different from that between the non-petechia dots and their surrounding area.

To obtain the color information of selected petechia dot, we need calculate the area of petechia dot and its average gray value. We adopt region growing algorithm to obtain the area information, which is an image segmentation algorithm. With a seed point, each pixel around the seed point is compared by gray value and determined whether the pixel should be accounted into seed points according to the comparison result. If they are homogeneous, the pixel would be treated as a seed point. The process stops until all the pixels are compared. When adopting the region growing algorithm, we should consider three key factors as follows.

- The location of seed points. Working together with *SimpleBlobDetector*, the region growing algorithm set



(a) A petechia dot with irregular margin (b)Extraction result by region growing
Figure 3. Example of region growing.

the pixel with the highest gray value inside a petechia dot as the seed point, because the growing direction is from higher to smaller gray value.

- The condition to include the adjacent pixels. Region growing algorithm has an iteration process comparing the pixels surrounding the seed point in four directions. If the color difference between the compared pixel and the mean value of connected region is smaller than a threshold, with which most of the petechia dots in the tested images are selected, the pixel can be counted into the connected region.
- Terminating condition of iteration process of the algorithm. The process of adding adjacent pixels to connected region repeats till no new pixel can be account into it.

Figure 3(a) shows a dot, which is not circular, from a tongue image, and Fig. 3(b) illustrates the extracted petechia dots by region growing algorithm. From the results we find our algorithm can extract petechia dots along the edge effectively.

When we use SVM to classify petechia dots and non-petechia dots by color, the results from the region growing algorithm could not be directly fed to SVM classifier because part of tongue edge may be in shadow. And extra color can be added into the color range in a single tongue image. Moreover, petechia dots are of difference in color in different tongue images. To solve these problems, we use gray difference between petechia dots and their surrounding area to obtain a quantitative color gamut for each petechia dot.

D. Classification of Color with SVM

As petechia dots have similar color such as bluish, purplish and darkish in different tongue images [2], the training procedure of SVM can be used to find color boundary of petechia dots and the trained SVM model is applied to choose the real dots.

SVM map the input training data into a higher dimensional feature space through some previous configured mapping function, which can be Linear, Poly, Sigmoid and so on, is also called kernel functions. In this high-dimensional feature space, the algorithm constructs an optimal separating hyperplane (that maximizes the margin) in accordance with trained data, which we can use as SVM model. The model can be used to divide the test data and get accuracy [10].

When we apply SVM in identifying petechia dots to get the SVM model, we should build a dataset of color information for training. First, we use all the dots detected by *SimpleBlobDetector* from 5 tongue images and obtain gray difference by region growing. Note that the gray difference classification is done in HSV color space. Because the value of H in HSV color space, which represents color hue, is close between petechia dots and non-petechia dots, we use components S and V in HSV color space to setup SVM model, which stands for saturation and value respectively.

The classification results by SVM model are shown in Fig. 4, in which the orange dots represent petechia dots while the blue ones are non-petechia dots. Because the value of S and V is significantly different between every single image, we standardize the results by equation (1):

$$value' = \frac{value - \mu}{\sigma} \tag{1}$$

where the $value$ is original data, $value'$ is processed data, μ is the mean value of all the data, and σ is standard deviation.

III. Results and Discussion

To train the SVM model for petechia dots detection, we divided S and V data into two groups (*V1, V2, S1, S2*) evenly, *V1* and *S1* are used as training set, and *V2* and *S2* are the test set. We feed the training set to models with linear kernel function, polynomial kernel function and radial basic function (RBF) respectively. The results with different kernel functions are shown in Table. II, from which we find that RBF kernel has the best performance with 88.4% accuracy.

In this work, we detect petechia dots from 128 tongue images with *SimpleBlobDetector* function and our trained SVM model. The results of each step are shown in Fig. 5, in which three original images are chosen and illustrated in Fig. 5(a), the extraction results by shape are presented in Fig. 5(b), and the extraction results by both shape and color are shown in Fig. 5(c). The *SimpleBlobDetector* firstly extracted circular dots regardless of color difference to avoid missing petechia dots, and followed by a region growing algorithm and SVM

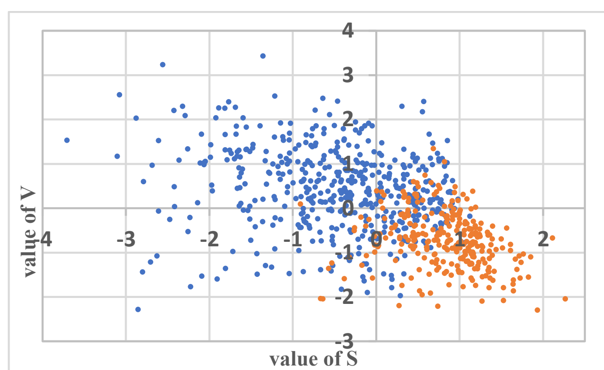
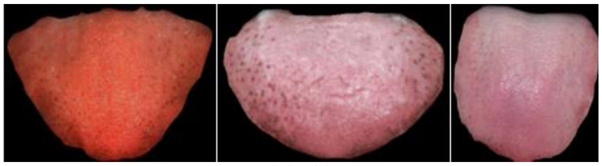


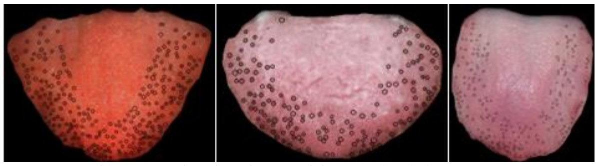
Figure 4. The Result of Classification

TABLE II. Application of Different Kernel

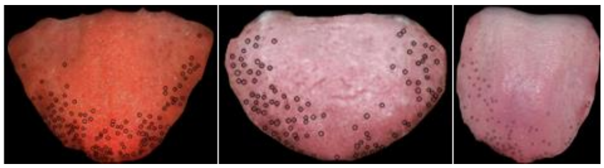
Kernel Type	Related Parameter	Value	Accuracy
Linear	C	0.0057	88.1%
Poly	Coef0	4	88.1%
	Gamma	0.011	
	Degree	0.99	
RBF	C	157	88.4%
	Gamma	0.082	



(a) Original images



(b) Extraction by shape with *SimpleBlobDetector*



(c) Extraction by both shape and color with *SimpleBlobDetector* function and trained SVM model

Figure 5. The results of petechia dots detection

model to exclude non-petechia dots by color difference. From the experimental results we can find our method is efficient in identifying petechia dots.

From the detection results of the 128 tongue images, 50 images have less than 15 petechia dots, 30 images have petechia dots ranging from 16 to 25, 25 images have 26 to 35 petechia dots, and 13 images have more than 36 petechia dots in which 9 pictures are selected to be petechia dots by doctors. The data of 9 selected tongue images are shown in Table III. Affected by texture, fissure, tongue coating and capillary on the tongue surface, some area of tongues images with no petechia dot are misclassified. Thus, we can determine petechiae tongue by the number of selected dots for every image.

In the experiments, the number of misclassified negative samples are recorded as false negative (FN). The number of correctly classified positive samples are recorded as true positive (TP) and FP is false positive samples. Therefore, the total number of detected dots is FN+TP. We define false alarm rate as FP/(TP+FP) and missing alarm rate as FN/(TP+FN). The test results are shown in Table. III, from which we can see that with our method the mean values of false alarm rate and missing alarm rate are 4.6% and 11.8% respectively.

TABLE III. DETECTION RESULTS

Image Number	Detected Petechia Dots	Actual Petechia Dots	False Detected Dots	Miss Detected Dots	False Alarm Rate	Missing Alarm Rate
1	96	90	2	8	2.1%	8.9%
2	53	51	3	5	5.7%	9.8%
3	86	76	0	10	0%	13.2%
4	87	87	7	7	8%	8%
5	87	76	4	15	4.6%	19.7%
6	53	47	1	7	1.9%	14.9%
7	46	44	4	6	8.7%	13.6%
8	38	37	2	3	5.3%	8.1%
9	42	40	2	4	4.8%	10%
Average					4.6%	11.8%

TABLE IV. COMPARISON WITH METHOD IN REFERENCE [6][a]

	Reference [6]	This work
Algorithm	SimpleBlobDetector and K-means	SimpleBlobDetector and SVM
Color Extraction	——	Region Growing
False Alarm Rate	24.0%	4.6%
Missing Alarm Rate	20.0%	11.8%

a. All the tested tongue images are taken from the clinic.

When applying the method of reference [6] to our tested images, only images 1,5,6,8 can give normal results. The comparison between method in reference [6] and our method is shown in Table. IV, which shows that with our method the false alarm rate and missing alarm rate is decrease by 19.4% and 8.2% respectively compared with that with method in [6].

IV. CONCLUSION

The paper proposes a petechia detection method that extracts petechia dots by both shape and color. We use *SimpleBlobDetector* function to elect petechia dots preliminary by shape and SVM model to extract petechia dots by color. The experimental results show that our method has false alarm rate 4.6% and the missing alarm rate 11.8% among our selected images, which have 19.4% and 8.2% decrease compared with previous work [6]. In future work, we will further improve the algorithm with more clinic tongue images.

REFERENCES

[1] J. Li, K. Wang, H. Zhang and X. Wu, "A Feature Extraction Method for Recognition of Petechia Dot in Tongue Image," *2009 First International Conference on Information Science and Engineering*, Nanjing, 2009, pp. 3560-3563.

[2] J. Xu. Color Atlas of Tongue Diagnosis in Traditional Chinese Medicine. Shanghai: Publishing house of Shanghai university of TCM, 2009.

[3] J. Li, H. Zhang, K. Wang and W. Zuo, "An automated feature extraction method for recognition of Petechia Spot in Tongue Diagnosis," *2009 International Conference on Future BioMedical Information Engineering (FBIE)*, Sanya, 2009, pp. 69-7.

[4] X. Li, C. Sha, H. Zhang and Z. Fu, "Spotted tongue recognition via cascade filtering," *2013 6th International Conference on Biomedical Engineering and Informatics*, Hangzhou, 2013, pp. 166-171.

[5] Y. Chu, "The Research of Extraction of the Tongue Image's Feature & Analysis Technology of Healthy Information Based on Traditional Chinese Medicine Inspection", M.S. thesis, Dept. Elect. Eng., Xiamen Univ., Fujian, China, 2008.

[6] S. Wang, K. Liu, L W. Tongue, "Tongue spots and petechia recognition and extraction in tongue diagnosis images," *Computer Engineering &Science*, vol. 39, no. 6, Jun. 2017

[7] C. Wan, Y. Zhang, C. Xia, P. Qian and Y. Wang, "Fissured Tongue Image Recognition Based on Support Vector Machine," 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Suzhou, China, 2019, pp. 1-5.

[8] S. E. and M. A., "Feature Extraction of Tongue Diseases Diagnosis Using SVM Classifier," 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 2019, pp. 260-263.

[9] L. Lo, M. C. Hou, Y. Chen, J. Y. Chiang and C. Hsu, "Automatic Tongue Diagnosis System," *2009 2nd International Conference on Biomedical Engineering and Informatics*, Tianjin, China, pp. 1-5, 2009

[10] V. N. Vapnik, "An overview of statistical learning theory," in IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 988-999, Sept. 1999.