Sensing and Imaging (2021) 22:8
https://doi.org/10.1007/s11220-021-00332-8

ORIGINAL PAPER

Texture Analysis of Tongue Coating in Traditional Chinese Medicine Based on Transfer Learning and Multi-Model Decision

Qingxin Xiao[1] · Hui Zhang[1] · Jing Zhang[1] · Li Zhuo[1]

Received: 25 August 2020 / Revised: 27 November 2020 / Accepted: 9 January 2021 / Published online: 10 February 2021

Abstract
The tongue coating texture is one of the basic characteristics of tongue manifestation in traditional Chinese medicine (TCM), which is generally divided into three types: curdy coating, greasy coating, and non-corrosive coating. In this paper, a texture analysis method of tongue coating in TCM is proposed by using transfer learning and multi-model decision. Firstly, tongue coating texture is pre-classified by using the transfer learning strategy to pre-train and fine-tune initial network model trained on large-scale dataset with the tongue coating dataset. Then, the multi-model decision is made by comparing the classification accuracy of different deep network models including InceptionNet V3, ResNet50, and MobileNet V1 to further optimize the texture analysis results of tongue coating. The experimental results show that the proposed texture analysis method of tongue coating can achieve better classification accuracy, which has practical meanings for assisting the clinical diagnosis and research for TCM.

Keywords Traditional chinese medicine · Tongue coating · Texture analysis · Transfer learning · Multi-model decision

1 Introduction

Tongue diagnosis is an important part of the diagnosis of traditional Chinese medicine (TCM). It refers to understanding the physiological functions and pathological changes of the human body by observing changes in tongue manifestation [1, 2], in which the tongue coating texture reflects the growth and decline of yang and wet turbidity. The characteristic of tongue coating texture is a kind of tongue manifestation with more abstract meaning, which is generally divided into three categories: curdy,

✉ Hui Zhang
huizhang@bjut.edu.cn

[1] Signal and Information Processing Laboratory, Beijing University of Technology, Beijing 100124, China

greasy and non-corrosive in the tongue diagnosis of TCM. In the clinical diagnosis, the tongue manifestation usually depends on the physicians' subjective judgment, in which the environmental factors such as different light sources and brightness can affect the physicians in making an accurate diagnosis. Therefore, it has important practical significance for promoting the objective research of tongue diagnosis in TCM to automatically analyze the tongue coating texture by computer [3–5], and the processing accuracy may determine the validity of automatic analysis of subsequent tongue manifestation.

However, texture analysis of tongue coating for automatic tongue diagnosis in TCM has strong semantic and abstract characteristics in describing the texture types. And clinically, it is difficult to obtain abundant and diverse tongue data, as well as the tongue coating texture for different diseases tends to have smaller inter-class distances and larger intra-class differences. Therefore, it is a very challenging task to obtain an analysis model with precision up to clinical application standards by using only traditional handcraft features or universal deep features.

In this paper, we proposed a practical and scalable scheme for texture analysis of tongue coating method by using transfer learning and multi-model decision. Firstly, tongue coating texture is pre-classified by using the transfer learning strategy to pre-train and fine-tune initial network model trained on large-scale datasets with the tongue coating dataset. Then, the multi-model decision is made by comparing the classification accuracy of the different deep network model to further optimize the texture analysis results. The contributions of this paper are as follows: (1) We build a transfer learning-based deep model trained on tongue manifestation dataset, to exploit the domain knowledge off-line to predict the classes of tongue coating texture. (2) By incorporating the powerful and complementary abilities from multiple networks, the diversity of the final model is improved so that the classifier can tackle various unpredictable difficulties.

In the remainder of this paper, the related works is briefly reviewed in Section II. We detail the proposed method in Section III and report the experimental results in Section IV. The final section draws conclusions.

2 Related Works

Images analysis technology is the main approach to automatically analyze the characteristics of tongue manifestation, which is an objective method of tongue diagnosis developed in the 1980s. Since then, many researchers [3–10] have made great efforts on tongue characterization including tongue color correction, tongue image segmentation, and tongue manifestation analysis, etc.

For the texture analysis of tongue coating in TCM, the basic idea is to regard it as a classification problem. The traditional classification method of tongue coating texture is to combine the handcrafted-features with classifier or similarity measure. For example, Wei et al. proposed a classification method based on improved subspace for tongue coating by dividing the tongue area into different image sub-blocks to analyze the density of the texture structure in order to obtain the curdy and greasy type and index of the entire tongue image, the classification accuracy is 83% [11].

Zhai et al. proposed a tongue coating method based on weak edge and Gabor filtering energy [12], the recognition accuracy can achieve about 91% for self-built tongue manifestation datasets.

In recent years, various deep learning networks have been proposed, such as VGG [13], GoogLeNet [14] and ResNet [15], etc., which have been widely applied in the fields of image classification, image recognition, target detection, etc., especially for medical image processing and have achieved superior performance over traditional methods [16, 17]. Because deep learning allows computational models that are composed of multiple processing layers to learn data representation with multiple levels of abstraction. Therefore, applying the theoretical knowledge of deep learning to the texture analysis of TCM tongue images can obtain more accurate classification accuracy. For example, Fu et al. applied GoogLeNet for tongue coating classification and achieved 92% classification accuracy [18]. In our previous work [19], we designed a light-CNN network to classify tongue color and coating color, and the accuracy rates in tongue color and coating color classification were 91.58% and 90.53% respectively. Existing related works have proved that deep learning networks can significantly improve the feature representation ability to produce higher classification accuracy. However, there are no more works to apply the deep learning network into the texture analysis task of tongue coating.

3 Proposed Method

Because the tongue coating texture in the TCM sense has more abstract meaning, it may suboptimal to directly utilize the existing deep network including traditional learning and training strategies to mining powerful descriptive texture features from tongue image samples. Considering that transfer learning is an optimization strategy, it is possible to make the progress of the target task modeling faster or improve its performance through the knowledge transfer from the basic network of the well-informed large-scale training set. Besides that, in view of the multi-model decision can overcome the deficiencies in single-model discrimination to obtain higher classification accuracy, a texture analysis method of tongue coating in TCM is proposed by using transfer learning and multi-model decision. Firstly, tongue coating texture is pre-classified by using the transfer learning strategy to pre-train and fine-tune initial network model trained on large-scale datasets with the tongue coating dataset. Then, the multi-model decision is made by comparing the classification accuracy of different deep network models including InceptionNet V3, ResNet50, and MobileNet V1 to further optimize the texture analysis results of tongue coating.

The overall architecture of the proposed method includes two stages: pre-classification of the tongue coating texture based on transfer learning and texture analysis of tongue coating based on the multi-model decision as shown in Fig. 1. Since transfer learning is an optimization strategy in deep learning, the classification model trained in large-scale data sets can be used as the initial network, and classification accuracy of the model can be improved by using the knowledge transfer method to execute the new learning tasks. In the stage of pre-classification, the pre-training networks trained in the large-scale ImageNet dataset is fine-tuned to n initial

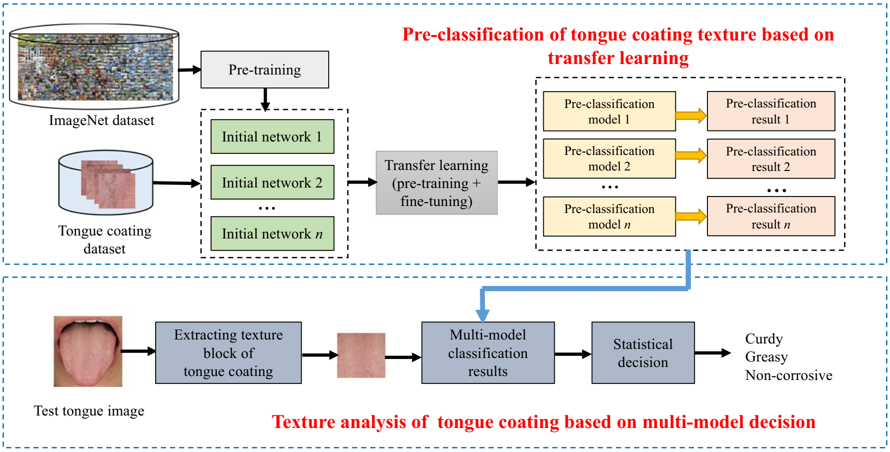


Fig. 1 The proposed overall architecture of texture analysis of tongue coating

networks by using the transfer learning with pre-training + fine-tuning to obtain n pre-classification models. Then, tongues coating texture are pre-classified by using pre-classification models to obtain n pre-classification results. In the stage of texture analysis of tongue coating, because the classification results trained by different network maybe have different results for the same test sample, the multi-model decision method is utilized to improve the classification accuracy. Specially, the tongue texture is classified with trained multiple models to produce multiple classification results after extracting the texture block of tongue coating from test tongue image. Then, the tongue coating texture, such as curdy, creasy, and non-corrosive, is analyzed from multiple classification results by using statistical decision method.

3.1 Pre-classification of Tongue Coating Texture Based on Transfer Learning

In this section, we will pre-classify the tongue coating texture by applying the transfer learning strategy to carry out network model's fine-tuning and classifier training. Next, we will introduce the two aspects of the comparison and selection of deep network models and the pre-classification of the tongue coating texture.

3.2 Comparison and Selection of Deep Network Models

Until now, there are various deep learning networks. It is necessary to compare and select the appropriate deep network models for the classification of the tongue coating texture. Due to the limited number of tongue images, the training strategy using pre-training combined with fine-tuning can effectively improve the classification accuracy. The mainstream convolutional networks mainly include AlexNet, VGGNet, InceptionNet, and ResNet, etc., whose classification performance in the ImageNet competition keep rising. Here, AlexNet is the earliest network model, which uses a 5-layer convolution and 3-layer full connection and the network structure is shallow [20]; VGGNet improves the AlexNet network, which uses 3×3 convolution to replace the 11×11 and 7×7 convolution kernels and deepens them in depth [13]; InceptionNet introduces the Inception module into the network infrastructure and innovates in the network width [14]; ResNet adopts the method of jump connection in the network architecture, and the accuracy of which is also greatly improved [15]. Since InceptionNet and ResNet have the optimal classification performance among the four networks, we will select the two network architectures as experimental networks. In addition, considering that InceptionNet and ResNet have more model parameters as well as MobileNet with lightweight network architecture is an efficient model for mobile and embedded devices [21], the MobileNet will be introduced into the experiment in order to comprehensively compare the performance of network models with different scales. As a result, we will select these three networks as the basic network to verify the effectiveness of the fine-tuning method to improve network accuracy.

3.3 Pre-classification of the Tongue Coating Texture Guided by Transfer Learning Strategy

"Pre-training + fine-tuning" is a commonly used transfer learning method in deep learning [22]. Firstly, the existing network is initialized by using the network parameters trained in the large-scale dataset; Then, the parameters of network model are fine-tuned with tongue images dataset to classify the tongue coating texture.

(1) **Pre-training** In this part, the initial models are trained with the images in the ILSVRC-2012 dataset (normalized to a size of 224×224) as input into the convolutional network. In the network pre-training process, we adopt the unsupervised pre-training method based on layer-by-layer greedy training, i.e., training from the first hidden layer to the last one of the training network, and taking these trained network parameters as the initial value of the overall network parameters.
(2) **Fine-tuning** In order to obtain a classification model of the tongue coating texture, a certain number of images are randomly selected from the tongue images dataset as training samples, which are preprocessed and normalized into the size 224×224. Next, the network is initialized by using the weights of the pre-trained model. Then, the classification model of tongue coating texture is generated by fine-tuning the convolutional network model after a large number of iterative processes. The specific process is further described as follows:

Step 1 Prepare *Ntrain* training samples and *Ntest* test samples of tongue images;
Step 2 Modify the output category of the last layer of the network;

Step 3 Adjust the configuration parameters of the network as well as appropriately reduce the learning rate and number of iterations of the training;

Step 4 Start network training after loading the parameters of the pre-training networks model to pre-classify the tongue coating texture.

(3) **Pre-classification** Next, we utilize the softmax classifier to select the class with the highest probability as the category of the current sample by calculating the probability that the current sample belong to each category. The softmax regression model is an extension of the logistic regression model on the multi-classification problem. Compared with training multiple logistic regression to do multi-classification problems, softmax regression is more suitable for situations where the categories are mutually exclusive. Let m training set samples to be $\{(\boldsymbol{x}^1, y^1), \ldots, (\boldsymbol{x}^m, y^m)\}$, and the dimension of vector $\boldsymbol{x}^i$ is $n+1$, the class label y can take c different values. The cost function $\mathrm{J}(\theta)$ of the regression model is denoted as [23]:

$$J(\theta) = \frac{1}{m}\sum_{i-1}^{m}\sum_{j=1}^{c} 1\left\{y^i = j\right\} \times \lg \frac{\exp^{(\boldsymbol{\theta}_j^T \mathbf{x}^i)}}{\sum_{l=1}^{c} \exp^{(\boldsymbol{\theta}_l^T \mathbf{x}^i)}} \tag{1}$$

where $\mathbf{1}\{\cdot\}$ indicates an explicit function, $\mathbf{1}\{\text{a true statement}\}=1$, $\mathbf{1}\{\text{a false statement}\}=0$, θ represents the overall parameter of the network model, θ_j^{T} is the transpose of the parameter θj, and c is the number of class labels.

3.4 Texture Analysis of Tongue Coating Based on Multi-model Decision

Since different models have different performance for the same dataset, a multi-model decision is utilized to analyze the texture of tongue coating in this section. Firstly, one tongue coating sample is evaluated by several classification models such as InceptionNet V3, ResNet50, and MobileNet V1 individually. Then, the maximum number of classes is taken as the current discriminant result by calculating the number of this image classified into each class. Furthermore, if there are two or more maximum number classes, the discriminant result of the single classification model with the highest test accuracy on all test sets is selected as the final result to complete the multi-model decision. The detailed decision-making process is shown in Algorithm 1.

Algorithm 1. Multi-model decision for texture analysis of tongue coating

Input: Tongue coating image *I*, parameters *Mi* the number of times image *I* classified as a certain class by multi-model, where i=1, 2,…, c and c is the number of categories.
Output: Texture analysis result.
Process:

```
Initialize the basic parameters Mi (i=1,2,…,c) is 0;
Multiple model discrimination calculation Mi;
Search for the maximum value in Mmax;
if Mmax is unique then
    Find the corresponding category j which get the maximum value Mmax;
    Output the analysis result;
else
    Find all models corresponding to the maximum value Mmax;
    the model with the highest accuracy on all test dataset is set to the final decision model;
    Output the analysis result;
end
```

4 Experimental Results and Analysis

In order to verify the effectiveness of the proposed texture analysis of the tongue coating, we conduct two experiments, (1) accuracy comparison of the network direct training and network fine-tuning (transfer learning); (2) the accuracy comparison of single model decision and multi-model decision. The tongue image dataset includes 488 cases collected from the Department of Traditional Chinese Medicine of Xuanwu Hospital of Capital Medical University. In the 488 tongue image data, the male to female ratio of the subjects was 3:5, and the age range was 6–98 years old, with an average age of 54 years and a median age of 52 years. According to the distribution of the color and texture of the tongue coating on each tongue image, the standard image block samples are selected by TCM experts and labeled with the corresponding texture class label. The network architecture of this experiment was implemented using the Keras and Tensorflow deep learning architecture. The system is configured as Intel(R) Core(TM) i7-6700 CPU 3.40 GHz, 16G RAM, TITAN X (Pascal) graphics card. We adopted the Stochastic Gradient Descent (SGD) algorithm to train the network and the iteration epoch is set to 50. The specific training parameters are shown in Table 1.

Table 1 Network training parameter settings

Weight attenuation	Batch	Learning rate	Momentum
1e−6	64	0.001	0.9

4.1 Tongue Coating Texture Dataset

A part of tongue image samples is shown in Fig. 2. Since the raw tongue image contains the face, lips, and other parts, excessive tongue color information and the difference in the shape of the tongue will affect the analysis of the tongue coating. Therefore, the standard image blocks for tongue coating texture classification are extracted from the tongue body region after the tongue segmentation process. The standard block samples are shown in Fig. 3, which include curdy, greasy, and non-corrosive coating.

Since the tongue diagnosis of TCM relies heavily on the experience of physicians, it is difficult for non-professionals to give the correct judgment of tongue coating texture with the naked eye alone. Therefore, the data collection method and data annotation in the experimental dataset are all verified by professional TCM practitioners. Under the guidance of the physicians, the square area of the tongue containing the part of the tongue is selected as the sample block. More concretely, our tongue coating texture blocks include 5792 cases of curdy coating, 7363 cases of greasy coating and 4926 cases of non-corrosive coating. Each class is divided into a training dataset and a test dataset with a ratio of 4:1, which shown in Table 2.

4.2 Accuracy Comparison of Network Direct Training and Network Fine-tuning

The selected convolutional neural networks in this paper have three network models: InceptionNet V3, ResNet50 and MobileNet V1. In this experiment, the three networks were trained by direct training and the network fine-tuning of the tongue coating respectively. The experimental results are shown in Fig. 4.

It can be seen that the three network models of InceptionNet V3, ResNet50 and MobileNet V1 have significantly improved the accuracy of tongue coating texture before and after fine-tuning. Here, after the fine-tuning of InceptionNet V3, the accuracy rate is improved from 93.7344 to 95.2592%, and the average accuracy can be increased by about 1.5248%; after the fine-tuning of ResNet50, the accuracy rate is increased from 93.1522 to 96.3127%, the average accuracy can be increased by about 3.16%, which is the highest among the three models; after the fine adjustment of MobileNet V1, the accuracy rate is improved from 90.6016% to 95.8969%, the average accuracy can be increased by about 5.295%. After three fine-tunings, the average accuracy can be increased by about 3.3266%. It can be proved that the transfer learning

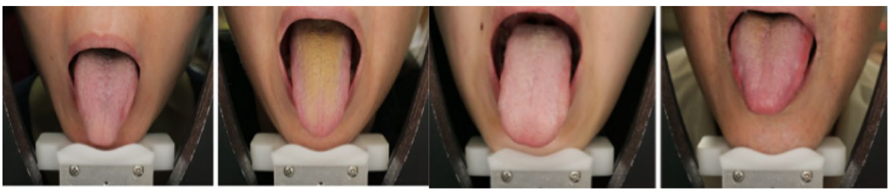
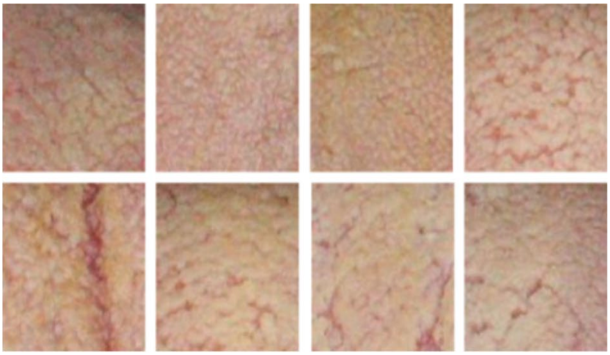
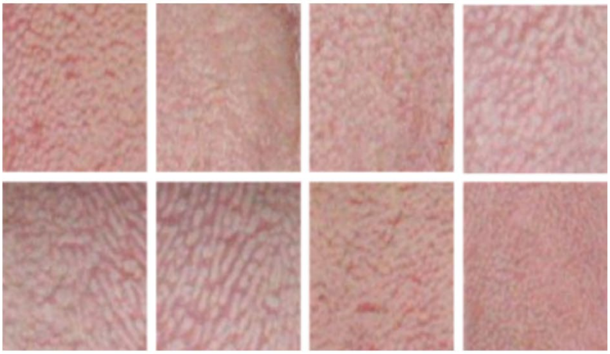


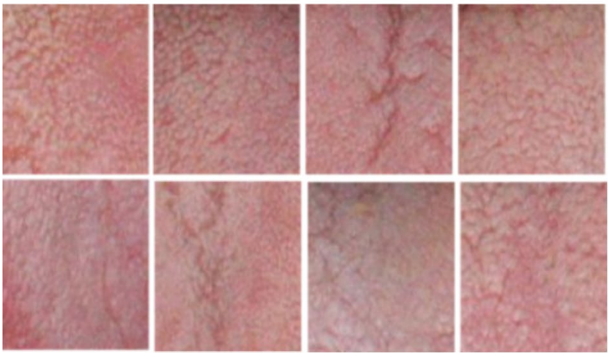
Fig. 2 A part of the tongue images



(a) Curdy coating



(b) Greasy coating



(c) Non-corrosive coating

Fig. 3 Experimental samples of the tongue coating texture

Table 2 Data sample distribution of the tongue coating texture

Texture category of tongue coating	Non-corrosive	Curdy	Greasy	Total
Training dataset	3941	4598	5890	14,429
Test dataset	985	1194	1473	3607

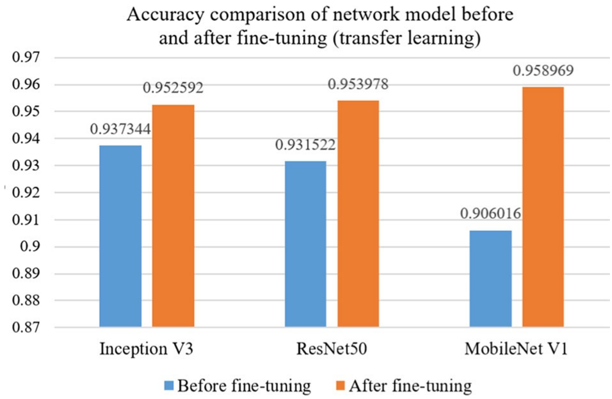


Fig. 4 Comparison of the network before and after fine-tuning

strategy of pre-training + fine-tuning can learn more powerful descriptive features so as to improve the accuracy of the texture analysis of tongue coating.

Besides that, we compared the operation time of different methods by calculating the running time of the model obtained by pre-tuning and post-tuning of each model on all test set images. The experimental results are shown in Table 3, which can be seen that the network models trained in both the network direct training and the network fine-tuning have similar test times on all test data sets because the overall number of parameters of the network model has not changed. Furthermore, due to different network model parameters, the test time of different network models presents different results, in which InceptionNet V3 has the maximum test time and ResNet50 is in the middle. Since MobileNet V1 is a lightweight CNN, its test time is significantly lower than ResNet50 and InceptionNet V3.

4.3 Accuracy Comparison of Single Model Decision and Multi-model ecision

In the multi-model decision making, InceptionNet V3, ResNet50 and MobileNet V1 have a total of six classification models in the network direct training and network fine-tuning. Since the effectiveness of fine-tuning has been proved in Sect. 3.2, the network fine-tuning has a higher classification accuracy than the network directly trained. Therefore, three classification models after network fine-tuning are selected for multi-model decision. The experimental results are shown in Table 4, which can be seen that the proposed multi-model decision approach achieves better classification performance than single model on the four indicators of accuracy, precision, recall and F1-score.

Table 3 Run time (in second) comparison of the tongue coating texture with different models

Network model	InceptionNet V3	ResNet50	MobileNet V1
Before fine-tuning (s)	584.9512	375.2604	242.3035
After fine-tuning (s)	538.5446	379.6757	234.7661

Table 4 Comparison of single model and multi-model decision

	InceptionNet V3	ResNet50	MobileNet V1	Multi-model decision
Accuracy	0.952592	0.953978	0.958969	0.969781
Precision	0.954496	0.951593	0.957494	0.968664
Recall	0.947665	0.956397	0.958600	0.969589
F1-score	0.950323	0.953603	0.957774	0.969027

This is because multi-model decision uses multiple network models to identify tongue coating texture to avoid the deficiencies of single model recognition, and has higher test accuracy.

In terms of test time, the test time for an image of single model and multi-model decision is shown in Table 5. It can be seen that the test time becomes longer and the sum of the time of the three models is similar because the images are classified simultaneously by multiple models although the classification average accuracy is improved. That illustrates the multi-model decision method improve the accuracy at the cost of time.

5 Conclusions

Aiming at the application of tongue coating analysis in tongue diagnosis of TCM, we avoid building the more sophisticated models and propose a practical and scalable scheme based on transfer learning and multi-model decision. Firstly, transfer learning strategy is introduced into the pre-classification of tongue coating texture to ensure that the classification model obtained by training has higher accuracy and reliability, in which the network model trained in large-scale ImageNet dataset is used as the initial network model fine-tuned with the tongue image dataset. Then, texture analysis of tongue coating is made by multi-model decision by selecting the optimal results from InceptionNet V3, ResNet50 and MobileNet V1. The experimental results have shown that the transfer strategy is also adaptable to the texture features of tongue coating. On the three convolutional networks of InceptionNet V3, ResNet50 and MobileNet V1, network model after fine-tuning is higher than the direct training network, which presents the effectiveness of network fine-tuning. At the same time, the classification accuracy of the proposed multi-model decision is higher than that of single models. In the future, we will further improve the analysis accuracy of the tongue coating texture by improving the dataset and optimizing the network structure. Through this research, it

Table 5 Run time (in second) comparison of single model and multi-model decision

Network model	InceptionNet V3	ResNet50	MobileNet V1	Multi-model decision
Running time (s)	0.1493	0.1052	0.0651	0.3256

can promote the development of intelligent analysis of TCM tongue diagnosis, and it has practical meanings for assisting the clinical diagnosis and research for TCM.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Grant Number 61871006) and the Science and Technology Development Program of Beijing Education Committee, China (Grant Number KM201810005026).

References

1. Tania, M. H., Lwin, K., & Hossain, M. A. (2018). Advances in automated tongue diagnosis techniques. *Integrative Medicine Research, 8*(1), 42–56.
2. Zhu, W. F. (2002). *Diagnostics of traditional Chinese medicine*. Beijing: China Traditional Chinese Medicine Press.
3. Shen, L. S., Cai, Y. H., & Zhang, X. F. (2007). *Collection and analysis of tongue images of traditional Chinese medicine*. Beijing: Beijing University of Technology.
4. Zhang, D. P., Zhang, H., & Zhang, B. (2017). *Tongue image analysis* (pp. 1–335). Heidelberg: Springer.
5. Shen, L. S., Wei, B. G., Cai, Y. H., Zhang, X. F., & Wang, Y. Q. (2003). Image analysis for tongue characterization. *Chinese Journal of Electronics, 12*(2), 317–323.
6. Chiu, C. (2000). A novel approach based on computerized image analysis for Traditional Chinese Medical diagnosis of the tongue. *Computer Methods and Programs in Biomedicine, 61*(2), 77–89.
7. Zhuo, L., Zhang, P., & Cheng, B. (2014). *Automatic tongue color analysis of traditional Chinese medicine based on image retrieval* (pp. 637–641). Singapore: Marina Bay Sands.
8. Jiao, Y., Zhang, X. F., Zhuo, L., Chen, M. R., & Wang, K. (2010). Tongue image classification based on Universum SVM. *The 3rd International Conference on BioMedical Engineering and Informatics* (Vol. 2, pp. 657–660). Yantai: China.
9. Qu, P. L., Zhang, H., Zhuo, L., Zhang, P., & Zhang, J. (2017). Automatic analysis of tongue substance color and coating color using sparse representation-based classifier. *IEEE International Conference on Progress in Informatics and Computing* (pp. 289–294). Nanjing: China.
10. Huang X. D., Zhang H., Zhuo L. Li X. G., Zhang J. (2020). TISNet-Enhanced Fully Convolutional Network with Encoder-Decoder Structure for Tongue Image Segmentation in Traditional Chinese Medicine. *Computational and Mathematical Methods in Medicine*. 2020 (pp. 1–13).
11. Wei, B. G., Shen, L. S., Cai, Y. H., & Zhang, X. F. (2004). on analysis algorithm of taste moss in Chinese Medicine. *Chinese Journal of Electronics, 31*(12A), 2083–2086.
12. Zhai T. T., Xia C. M., Wang Y. Y., Zhu Mulangma. (2016). Identification of tongue coating based on Gabor wavelet transform. *Journal of Computer Applications and Software*, 33(10):162–166.
13. Simonyan, K., & Zisserman, A. (2015). *Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations*. USA: San Diego.
14. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. (2015). Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* . (pp. 1–9). USA: Boston.
15. He, K. M., Zhang, X. Y., Ren, S. Q., & Sun, J. (2016). Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). USA: Las Vegas.
16. Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, G. B., & Kim, N. (2017). Deep learning in medical imaging: general overview. *Korean Journal of Radiology, 18*(4), 570–584.
17. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*(9), 60–88.
18. Fu, S., Zheng, H., Yang, Z., Yan, B., Su, H. Y., & Liu, Y. P. (2017). Computerized tongue coating nature diagnosis using convolutional neural network. *IEEE 2nd International Conference on Big Data Analysis (ICBDA)* (pp. 730–734). Beijng: China.
19. Qu, P. L. (2016). Automatic segmentation and analysis system of traditional Chinese medicine tongue image based on deep learning. *Doctor dissertation*. Beijing University of Technology.

20. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *International Conference on Neural Information Processing Systems*. (pp. 1097–1105). USA: Lake Tahoe, Nevada.
21. Howard, A. G., Zhu, M. L., Chen, B., kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H. (2017). MobileNets: efficient convolutional neural networks for mobile vision applications. *CoRR abs/1704.04861*.
22. Yosinski J., Clune J., Bengio Y., Lipson H. (2014). How transferable are characters in deep neural networks. *Advances in Neural Information Processing Systems*, 3320–3328.
23. Liu, W., Wen, Y. D., Yu, Z. D., Yang, M. (2016). Large-margin softmax loss for convolutional neural networks. *International Conference on International Conference on Machine Learning*. (pp. 507–516). USA: New York