IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 July 2023, accepted 14 August 2023, date of publication 25 August 2023, date of current version 5 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3308602

APPLIED RESEARCH

Tongue Coating Grading Identification Using Deep Learning for Hyperspectral Imaging Data

DONG ZHANG, WENTAI PANG, KEYI WANG, FENGWEN YANG, AND JUNHUA ZHANG

Center for Evidence-Based Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China
Xin-Huangpu Joint Innovation Institute of Chinese Medicine, Guangdong, Guangzhou 510000, China

Corresponding author: Junhua Zhang (zjhtcm@foxmail.com)

This work was supported in part by the Joint Innovation Foundation of Xin-Huangpu Joint Innovation Institute of Chinese Medicine under Grant 2022IR024, in part by the National Natural Science Foundation of China under Grant 82205316, and in part by the Tianjin Key Fields of Traditional Chinese Medicine Science and Technology Project 2021002.

ABSTRACT Tongue diagnosis is one of the four diagnostic methods of traditional Chinese medicine (TCM), which has important value in clinical disease diagnosis and efficacy evaluation. The change in tongue coating is a comprehensive cause of multi-dimensional changes such as color, texture, and substance. However, the color tongue image contains less spectral information, which may lead to the lack of key information in tongue diagnosis. Hyperspectral images can obtain reflection information of tongue images in hundreds of spectral bands. Unlike traditional color images, the rich spectral information can more accurately and sensitively describe and classify tongue coating, and has been widely applied in biomedical images. In this paper, we conducted feature extraction and analysis on hyperspectral images of different tongue coatings, and proposed a spectral-spatial feature deep learning framework to classification and quantitative recognition the tongue coating based on hyperspectral image features. Firstly, 360 hyperspectral images of tongue body were collected, and clinicians were identify all tongue coatings and divided them into 6 different grades. The hyperspectral features of each tongue coating area were extracted respectively. In order to reduce noise interference, singular spectrum analysis was used to preprocess the hyperspectral curve features. Considering the actual situation of tongue coating, a depth learning model was established to analyze the spectral and spatial feature of the hyperspectral tongue image to identified the grading of tongue coating. The experimental results showed that tongue coating with different quantization levels had different hyperspectral features, and the recognition rate of tongue coating quantization level can reach 87.21% using the spatial-spectral features.

INDEX TERMS Traditional Chinese medicine, tongue diagnosis, hyperspectral image, deep learning.

I. INTRODUCTION

Tongue diagnosis is one of the important contents in the traditional Chinese medicine (TCM) with a history of more than 2000 years. The change of tongue can help doctors to diagnose diseases and provide basis for clinical medication. However, the current research on tongue diagnosis mostly relies on the clinician's experience judgment and subjective experience [1], [2]. At the same time, the current tongue image acquisition method was mostly color digital image, which was vulnerable to changes in light sources and interference from the external environment during the imaging process [3]. Therefore, how to realize the objectification of the tongue body had become the focus and difficulty of current research [4].

With the popularization and application of computer image processing technology, the research on tongue body classification, disease diagnosis and efficacy evaluation based on tongue features had been carried out, and achieved satisfactory results [5], [6], [7], [8]. Deepa and Banerjee [9] proposed the CNN Dense net framework to identify the features of the tongue image such as color, texture, tooth markings and the red spots for identifying diabetes mellitus. Zhai et al. [10] transformed tongue images into the hue, saturation and

The associate editor coordinating the review of this manuscript and approving it for publication was Massimo Cafaro.

intensity colour space, and used the dual snake algorithm to obtain the accurate contour of the tongue. Shi et al. [11] explored the data characteristics of tongue and pulse of non-small-cell lung cancer with Qi deficiency syndrome and Yin deficiency syndrome, established syndrome classification model based on data of tongue and pulse by using machine learning methods, and evaluated the feasibility of syndrome classification based on data of tongue and pulse. Tang et al. [12] proposed a novel paradigm by employing artificial intelligence to feature extraction and classification of tongue coating. Rajakumaran and Sasikala [13] devised a new automated method based tongue color image analysis for disease diagnosis and classification. The experimental values pointed out that the presented VGG19-RF model had reached a higher precision. Although these researches had made some achievements, most of them rely on color images which only contain the red, green and blue primary colors. Other imaging information of spectral band may be ignored, and feature collection was incomplete [14]. At the same time, most of the current studies focus on tongue classification or disease auxiliary diagnosis, lacking quantitative indicators which were very important in clinical efficacy evaluation [15].

Hyperspectral images can obtain imaging information under hundreds of different spectral wavelengths [16]. In hyperspectral images, low-level features such as color and texture can be extracted, and high-level semantic features can also be analyzed [17]. Especially, compared to RGB color images, hyperspectral images also have rich spectral features. It had very successful applications in food safety, satellite remote sensing, agriculture, cultural relics protection and other fields [18], [19]. In hyperspectral image analysis, the rich spectral and spatial features combined with machine learning models can effectively complete target detection and classification tasks. Traditional machine learning methods, such as KNN and SVM, rely on manual extraction of spectral features from hyperspectral images for model construction [20], [21]. Although certain results have been achieved, manually extracting features does not fully utilize hyperspectral images [22]. Deep learning is widely used in hyperspectral image classification due to its high-level feature extraction ability. At the beginning, deep learning models were used for spectral feature extraction and analysis, and certain results were achieved [23]. For example, deep belief network and stacked autoencoder was used in an unsupervised model to learn the intrinsic representation of hyperspectral. Spectral information was considered as sequence data in a recurrent neural network (RNN) [24]. Obviously, these models neglects the spatial consistency of hyperspectral data. An increasing number of studies indicate that both spatial and spectral features in deep learning play a key role in hyperspectral classification. And separately extracts spatial and spectral features, has been consistently shown to be an effective framework for hyperspectral classification. 2D-CNN [25] was commonly used for extracting high-level semantic features in hyperspectral images. Wang et al. [26] proposed an end-to-end fast dense spectral–spatial convolution network for hyperspectral classification. Zhong et al. [27] constructed a spectral–spatial residual network to extract fusion features of hyperspectral. Yang et al. [28] proposed a jointly extract spectral and spatial features based on 1-D-CNN and 2D CNN. Sun et al. [29] extracted local spectral features and multiscale spatial features, And feature fusion was carried out to mine the complementary information. On the basis of spatial spectral features, the transformer model was fused and achieved good results [30]. Combining local and global spatial features with spectral features for model training in [31].

At the same time, hyperspectral image technology had also been applied to the study of tongue diagnosis in TCM, and achieved satisfactory results. Zhang et al. [32] used a visible hyperspectral image system with an approximate spectral range of 400-1000nm to predict the tongue colour values and the coating position in TCM, and a stacked autoencoder predict model based on spectral-spatial feature was performed to digital the tongue colour space and the coating. Yan et al. [33] proposed a near infrared spectral identification model based on SVM to distinguish healthy individuals from hepatitis patients. Li and Liu [16] used spectral angle mapper to analyze tongue color based on spectral which can achieve meaningful areas of substances and coatings of tongue. Liu et al. [34] proposed to use hyperspectral technology for tongue diagnosis for the first time in the literature and obtain promising results. Because of the data characteristics of hyperspectral images and tongue diagnosis in TCM, most of the above studies only extract the hyperspectral features of the tongue body for comparison and classification, without combining the spatial characteristics of the image itself for analysis, which may have some impact on the results. However, most of the current research on hyperspectral tongue diagnosis still focuses on the extraction and classification of spectral features, without in-depth analysis of the spectral reflectance differences of different tongue images. Meanwhile, spatial feature information was rarely mentioned in tongue image classification.

In order to make full use of the spectral and spatial features and realize the fast and accurate distinguish of tongue coating grading based on hyperspectral tongue images, a deep learning based framework was proposed. In summary, the contributions of this article are mainly concentrated in two aspects:

(1) In order to solve the problem of less spectral information in color images, this paper used hyperspectral image technology to collect images of tongue body, and obtained the hyperspectral feature distribution of different tongue coatings.
(2) The spectral and spatial features of tongue coating with different quantitative grades were extracted, and the depth learning model was established to quantify the grade of tongue coating, so as to improve the application value of tongue diagnosis in clinical efficacy evaluation.

II. MATERIALS AND METHODS

A. TONGUE SAMPLE PREPARATION

This study had passed the ethics certification of the Medical Ethics Committee of Tianjin University of Traditional Chinese Medicine. 360 residents living in Nankai District of Tianjin China participated in the collection of hyperspectral tongue images. All participants signed informed consent. And all collectors were required not to eat pigmented foods 3 hours before collection, so as not to affect the color of the tongue. At the same time, all collectors had undergone simple training to keep the same posture when sticking out their tongues.

B. TONGUE HYPERSPECTRAL DATA ACQUISITION

Figure 1 showed the collection environment of hyperspectral tongue images. It mainly included the hyperspectral camera, the head fixed part, the computer control and the external light source. The spectral wavelength range was 400-1000 nanometers (nm) with a spectral resolution of 5 nm. The number of spectral bands was 128, and the collected image resolution was 600*400. The fixed part of the head was mainly to ensure that the collector maintains a fixed posture during the acquisition process, so as to ensure that the acquired hyperspectral image was clear and accurate. The computer control part mainly controlled the acquisition parameters and data storage of hyperspectral images, such as exposure time, scanning range and scanning speed, etc. The external light source part was mainly to make the illumination of the collection environment uniform, to ensure that the illumination environment was unified in all collection processes, and to avoid data noise caused by changes in the external illumination environment to the greatest extent [35].

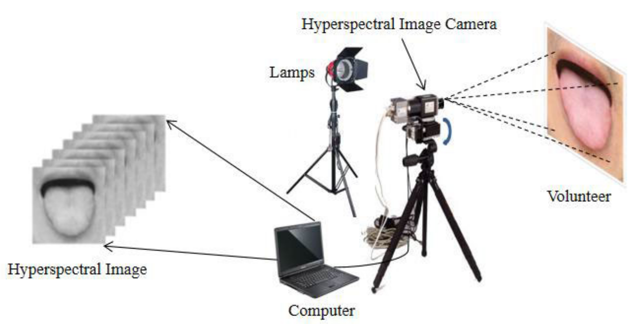

FIGURE 1. Structure of the tongue hyperspectral image acquisition system.

All collectors were collected the hyperspectral tongue images under the acquisition equipment shown in Figure 1. And a three-dimensional hyperspectral tongue cube image data was obtained. The dimension of the cube data was 600*400*128, and the x and y axes represent the spatial resolution of 600*400. The z-axis was the spectral wavelength dimension, which represents the reflectance values at different spectral wavelengths.

In order to further reduce the interference of environmental noise, it was necessary to perform black and white verification on the collected original hyperspectral tongue image data, and normalize the corresponding spectral reflectance [36]. Suppose the original hyperspectral tongue image data was I, the hyperspectral tongue image data after black and white verification was R, the black verification (no light source, the spectral reflectance is 0) was B, the white verification (with the help of the white verification board, make the reflectivity close to 99%) was W, then the calculation process of black and white calibration was:

$$R = \frac{R_0 - D}{W - D} \times 100\%$$

C. DETERMINATION OF TONGUE COATING GRADING

In TCM tongue diagnosis, tongue coating refers to a thin white and moist coating on the tongue. The observation of tongue coating was a very important content and the change of tongue coating can reflect the change of disease to a certain extent [37].

The current research were focus on the classification of tongue coating with color and texture feature, but these results were not suitable for clinical efficacy evaluation [38]. Therefore, a tongue coating grading method was used to describe the dynamic changes of tongue coating which was more conducive to clinical efficacy evaluation [39]. Table 1 introduced 6 different tongue coating grades and specific descriptions in detail. Two clinicians were invited to judge the tongue coating grades of all tongue hyperspectral images. If the results were inconsistent, a third clinical expert was invited to review. Finally, the tongue coating classification standards for tongue hyperspectral images were obtained.

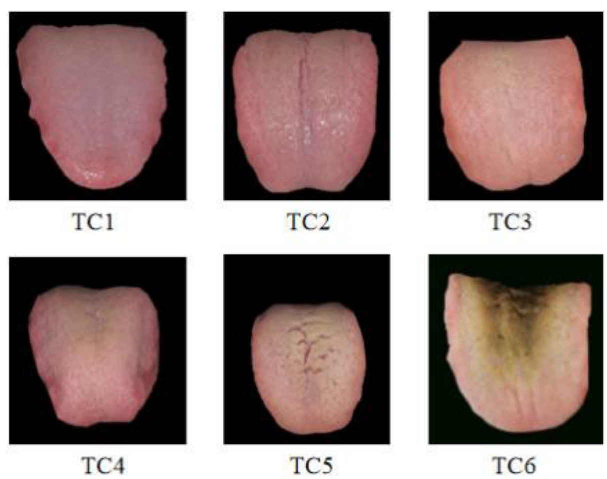

FIGURE 2. Sample images of the six tongue coating categories.

D. METHOD

Figure 3 showed the main technical process of the method, which was mainly divided into hyperspectral data extraction and preprocessing, hyperspectral feature extraction, and classification identification.

TABLE 1. The description of grading of the tongue coating.

No.	Number	Grading of the tongue coating	Tongue coating description
TC1	43	Mirror-approximated tongue coating	There is almost no attachment on the tongue surface, and no obvious position of tongue coating is found
TC2	124	Normal tongue coating	The tongue coating observed is a thin layer of white
TC3	54	White, greasy tongue coating	The thickness of tongue coating is increased compared with that of normal tongue coating, the color is still white, but the texture becomes dense, showing greasy coating
TC4	76	Yellow, tongue coating	The thickness of tongue coating has no obvious change compared with normal, but the color is mainly yellow
TC5	38	Yellow, greasy tongue coating	The thickness of tongue coating is increased compared with that of normal tongue coating, the color is mainly yellow, the texture becomes dense
TC6	25	Gray-black tongue coating	The color of tongue coating is mainly gray and black, showing greasy coating

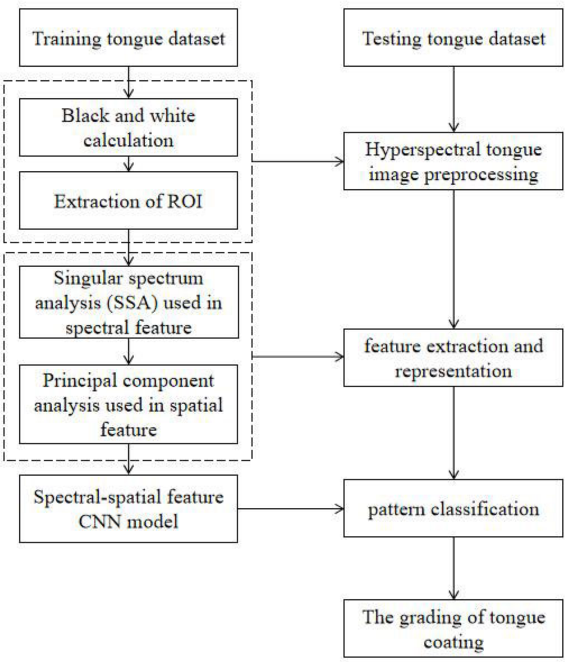

FIGURE 3. The main flow chart of predicting the grading of the tongue coating.

1) SPECTRAL AND SPATIAL FEATURE PREPROCESSING

Tongue coating refers to the thin layer attached to the surface of the tongue. Different populations may have inconsistent tongue coating location and range. Therefore, in the process of tongue coating classification and identification, the tongue coating area was first manually extracted from all hyperspectral tongue images. To ensure that the obtained region of interest (ROI) can represent the characteristics of the tongue coating, the size of the extracted ROI was generally 50*50 pixels, and this operation was implemented by ENVI software.

After obtaining the tongue coating ROI, it also needed to be preprocessed to facilitate subsequent feature extraction and calculation. Since the hyperspectral image contains the spectral and spatial features of the tongue coating, the spectral dimension preprocessing and the spatial dimension preprocessing can be performed for the area of interest of the tongue coating respectively.

Spectral dimension preprocessing mainly refers to calculating the average spectral reflectance value of all pixels in the region of interest. In order to reduce noise interference and data errors, the singular spectrum analysis (SSA) method was selected to process the original spectral reflectance curve [40]. SSA constructed the trajectory matrix of the input signal, and decomposes and reconstructs it to extract the signals of different components in the original signal, such as trend signal, noise signal, etc., which had better performance in data pre-processing. Singular spectrum analysis generally includes four steps of embedding, SVD decomposition, grouping and reconstruction. It was a very effective method for processing nonlinear time series data [41]. In this study, the SSA that the window size was set to 20 and VGD was 5th was used at the spectral curve to smooth the spectral features to further reduce the effects of noise.

The spatial dimension preprocessing was mainly to perform principal component analysis (PCA) processing on the area of interest in the tongue coating, reducing the spectral dimension, enhancing the useful components, and reducing the amount of calculation. The area of interest in the tongue coating after PCA processing was used for subsequent texture feature extraction.

2) THE PROPOSED MODEL

Convolutional neural network (CNN) was a kind of deep learning and had been successfully used in many fields, such as image classification, computer vision, feature extraction, etc [7], [42]. The CNN model was different from traditional neural networks because it learns the multi-layer generation model from the raw data. And the feed forward neural network was been initialized based on the discovered features. What's more, it used back propagation to fine-tune

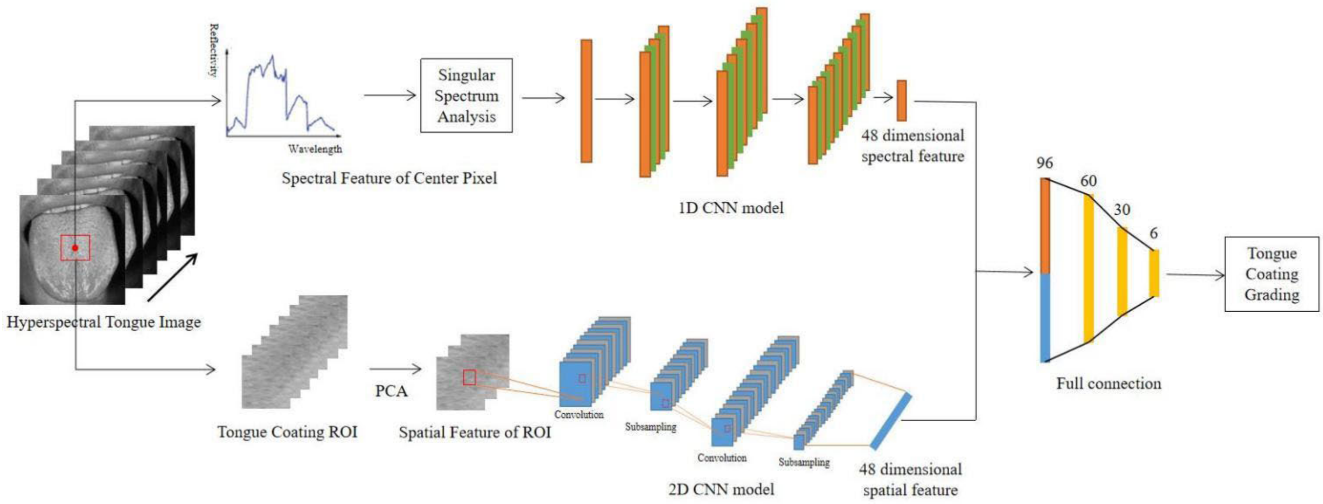

FIGURE 4. Structure of the model.

the network. As seen in Figure 4, a CNN based model was proposed to predict the grade of the tongue coating.

In order to make full use of the spatial and spectral information of hyperspectral tongue images, spectral and spatial features were extracted and fused respectively. In the process of spectral feature extraction, a one-dimensional CNN based model was used to process the average spectral curve of the region of interest. 1D CNN model conducts multilayer of convolution and max pooling operation, a 24 dimensional tongue coating spectral feature vector was obtained. For the region of interest of tongue coating, PCA was used for preprocessing to obtain the input image composed of the first three principal components. A 2D CNN based model includes four layer of convolution and max pooling was proposed. Finally, 48 dimensional spatial feature vectors can be obtained. Finally, the spectral features and spatial features were combined and input into the classification layer. The classification layer consists of three full connection layers, and the outputs of the full connection layers were 60, 30 and 6 respectively. Finally, after extraction and fusion of spectral and spatial features, a prediction model of tongue coating grade based on CNN model was constructed, which can output the prediction results of tongue coating grade.

3) EVALUATION METRICS

The performance of the proposed method was evaluated using statistical metrics including the overall accuracy, precision, recall rate and F1-score [43]. The four metrics can be defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

where true positive (*TP*) indicates the number of tongue coating grading correctly classified, true negative (*TN*) denotes the number of correctly classified as other coating grading. False positive (*FP*) represents the number of misclassified as coating grading, and false negative (*FN*) is the number of misclassified as other coating grading.

III. RESULTS AND DISCUSSION

A. SPECTRAL FEATURE DIFFERENCES OF TONGUE COATING

In the traditional tongue coating classification research, most of them focused on the manual extraction of the color and texture features of the tongue coating area, and the classification of the coating color. The hyperspectral image used in this study can simultaneously obtained the imaging information of tongue in the spectral and spatial dimensions, which had more advantages than the color image information [32], [33].

In recent years, hyperspectral imaging technology had been used widely in clinic. It had been found that the absorption and reflectance of hyperspectral light on biological tissues were mainly affected by hemoglobin and water. The hyperspectral band range used in this paper was 400-1000 nm, and there was no obvious absorption peak of water molecules in this band range, so it was mainly affected by the absorption rate of hemoglobin. The absorbance of hemoglobin between 500-650 nm had a peak, so the reflectivity between 500-650 nm in the tongue coating hyperspectral curve distribution in Figure 5 was lower. After 650nm, the absorbance of hemoglobin decreases rapidly, so the reflectivity on the tongue coating hyperspectral curve increases significantly after 650nm. The distribution of the reflectance of the hyperspectral curve was consistent with the characteristics of the biological tissue [16], [34].

Figure 6 showed the average value of hyperspectral curve characteristics of different grading levels of tongue coating. TC6 had the largest color and thick coating, so its reflectivity value was the lowest. This was mainly because its surface texture was dense and the color was deeper. With the decrease of the level of tongue coating, the color of the tongue coating gradually become lighter and the texture was sparse, so the hyperspectral reflectivity gradually showed an increasing trend.

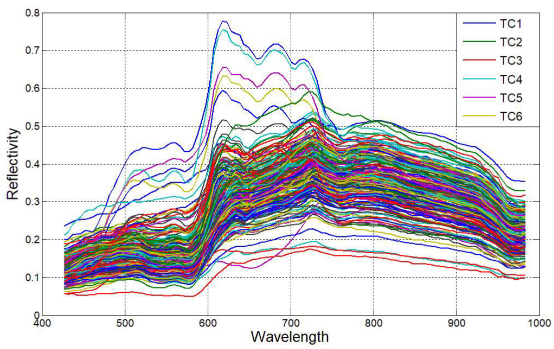

FIGURE 5. Hyperspectral feature distribution of different tongue coating grades.

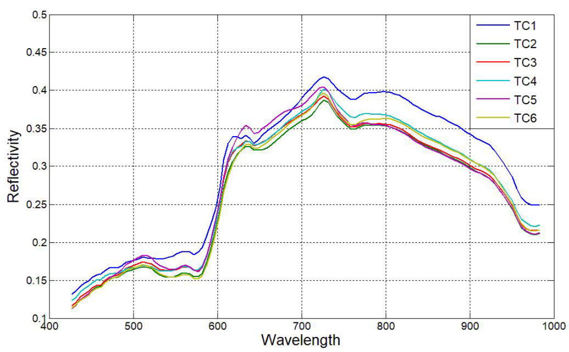

FIGURE 6. Average hyperspectral feature of different tongue coating grades.

B. EXPERIMENT SETUP

The hyperspectral tongue coating model was implemented on the Pytorch deep learning framework. The model trained on the Ubuntu with Intel Core i9 3.7GHz CPU, 128GB RAM, and NVIDIA GTX 3090. And the learning rate was 0.001, batch size was 4.

360 samples of hyperspectral tongue images was divided into the training dataset and the test dataset. In this study, 300 samples was randomly selected as the training dataset. And the rest 60 samples as the test training. Ten fold cross validation was utilized for the model to evaluate the performance.

C. PERFORMANCE OF TONGUE COATING GRADING CLASSIFICATION

In hyperspectral tongue coating images, different degrees of tongue coating can reflect different physical conditions. As the color and texture of the tongue coating deepen, special attention needs to be paid. Table 2 showed the predict result about the proposed methods. As can be seen, the proposed method that fused the spectral and the spatial feature achieved the highest average precision and average recall. The result of only spectral or only spatial feature was decline and the only spatial get the worst in the proposed method.

TABLE 2. Performance of different feature fusions to predict the grade of tongue coating.

Feature	Spectral	√		√
	Spatial		√	√
Proposed	Accuracy	0.7233	0.6849	0.7526
	Precision	0.8342	0.7827	0.8531
	Recall	0.8114	0.7638	0.8368
	F1-score	0.8226	0.7731	0.8449

TABLE 3. Performance with color images to predict the grade of tongue coating.

	Hyperspectral images	Color RGB images
Accuracy	0.7526	0.6852
Precision	0.8531	0.7324
Recall	0.8368	0.7135
F1-score	0.8449	0.7228

The classification of tongue coating was a comprehensive and complex task. Observing and analyzing the classification of different tongue coatings, it can be found that there were differences in color, texture and material composition changes among different levels of tongue coatings. For example, compared to TC3, TC4 had a significantly yellow tongue coating and a layer of tongue coating covering, resulting in a denser texture and texture. When classifying tongue coating grades, if only spectral features were used, spatial feature changes such as texture may be ignored. Similarly, if only spatial features used, the changes in material composition of different tongue coatings cannot be described. Therefore, combining spectral and spatial features can describe tongue coating more comprehensively from multiple dimensions, resulting in more accurate classification results.

D. COMPARE WITH COLOR RGB TONGUE IMAGES

To further illustrate the value and advantages of hyperspectral images in tongue coating classification, we simultaneously collected corresponding color RGB images. The 2D CNN model was applied to classify color RGB tongue coating grading. Compared with the application of spectral-spatial features, the performance of the RGB model did not achieve the proposed method. This was because color RGB images lack rich spectral features, which play an important role in tongue coating classification. During the recognition process of different levels of tongue coating, the changes in spectral features were more obvious and sensitive compared to features such as color and texture. From the classification results,

TABLE 4. Performance of different feature fusions to predict the grade of tongue coating.

Feature				
	Spectral	√		√
	Spatial		√	√
Proposed	Accuracy	0.7233	0.6849	0.7526
	Precision	0.8342	0.7827	0.8531
	Recall	0.8114	0.7638	0.8368
	F1-score	0.8226	0.7731	0.8449
PCA+PLSR	Accuracy	0.7343	0.7217	0.7538
	Precision	0.7864	0.7349	0.7681
	Recall	0.7625	0.7221	0.7469
	F1-score	0.7743	0.7284	0.7574
PCA+SVM	Accuracy	0.7594	0.7669	0.7683
	Precision	0.7891	0.7745	0.7956
	Recall	0.7761	0.7554	0.7674
	F1-score	0.7825	0.7648	0.7812
PCA+ANN	Accuracy	0.6467	0.5631	0.6928
	Precision	0.6325	0.6157	0.6872
	Recall	0.6224	0.6272	0.6791
	F1-score	0.6274	0.6214	0.6831

TABLE 5. Performance of different methods to predict the grade of tongue coating.

	Accuracy	Precision	Recall	F1-score
Proposed	0.7526	0.8531	0.8368	0.8449
CNN	0.7125	0.7326	0.7539	0.7431
SAE[32]	0.7217	0.8028	0.8221	0.8123
3D CNN	0.7325	0.8335	0.8346	0.8340
MIL CNN[7]	0.7389	0.8456	0.8337	0.8396
Deep CNN[14]	0.7443	0.8486	0.8242	0.8362

it can also be seen that using hyperspectral images for tongue coating classification has obvious advantages over color RGB images.

E. COMPARE WITH THE COMPARISON ALGORITHM

1) COMPARISON WITH DIFFERENT FEATURE CLASSIFIERS

In order to verify the practicability of the proposed model, some evaluation indicators were needed to evaluate the prediction results. The average precision and average recall were calculated to evaluated the predict performance. In addition, in order to further verify the performance of the model, it was also necessary to compare the results with the comparison experiment. The principal component analysis with partial least squares regression (PCA+PLSR) [44], principal component analysis with the support vector machine (PCA+SVM) [45], [46], and principal component analysis with the artificial neural network (PCA+ANN) were applied on the spectral and spatial features to predict the grading of tongue coating [47]. In these comparison feature classify, PCA was used to the spectral curve and the flattening spatial region to obtain the spectral-spatial feature [48], [49]. Table 3 listed the classification performance results of three classic machine learning models and our method. From the results, it can be seen that traditional machine learning methods mainly rely on feature extraction and processing, but the extracted features were fewer compared to deep learning models and cannot fully express the different changes in tongue coating.

2) COMPARED WITH OTHER MACHINE LEARNING MODEL

Compared with traditional machine learning methods, deep learning can serve as input and feature extraction for the entire hyperspectral image, and can obtain higher dimensional feature expression, thereby achieving more accurate classification performance. This paper used commonly deep learning models as comparative experiments to verify the effectiveness and accuracy of the proposed method. Table 5 listed the classification results of five deep learning models and the proposed method achieves the best classification performance.

Due to the independent extraction and fusion of spectral and spatial features from hyperspectral tongue coating images in this model, the multi-dimensional features of tongue coating position can be more clearly expressed. In contrast, deep learning models such as CNN pay more attention to the extraction and analysis of spatial features, but lack the integration of spectral features. Further evidence had been provided that both spectral and spatial features play an important role in the analysis of tongue coating.

IV. CONCLUSION

Compared to color RGB images, hyperspectral images not only contain spatial features but also rich spectral information. In the classification task of tongue coating, relying solely on features such as color and texture cannot achieve ideal results. Therefore, this paper proposes a classification framework that integrates spectral and spatial features using hyperspectral image technology. In order to classify the grade of tongue coating to describe the value of TCM tongue diagnosis in more detail, this paper proposed a two-layer CNN model. Compared with the comparative experiments, the proposed deep learning method get the best precision and recall with the hyperspectral images, which was of great significance for the clinical diagnosis and therapeutic evaluation of TCM tongue diagnosis.

There were also some deficiencies in our research. It was costly for the clinician to mark the tongue coating, so the sample size was not much. And then our model was based on one-dimensional spectra or two-dimensional spatial features without considering three-dimensional hyperspectral images. At the same time, some human physiological indicators should also be tested together to make judgments on human health. In the future work, we will study the semi-supervised classification learning model for hyperspectral images, and the 3D hyperspectral cube will be consider for feature extraction and analysis to increase the accuracy of results.

REFERENCES

[1] D. Meng, G. Cao, Y. Duan, M. Zhu, L. Tu, D. Xu, and J. Xu, "Tongue images classification based on constrained high dispersal network," *Evidence-Based Complementary Alternative Med.*, vol. 2017, Mar. 2017, Art. no. 7452427.

[2] K. Xu, D. Feng, and S. Zou, "Transfer learning-based tongue gestures classification using ultrasound images," *Med. Phys.*, vol. 45, no. 6, p. e234, 2018.

[3] E. Srividhya and A. Muthukumaravel, "Diabetic detection using tongue images based on ANN classification," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 1, pp. 2704–2710, Oct. 2019.

[4] J. Li et al., "A tongue features fusion approach to predicting prediabetes and diabetes with machine learning," *J. Biomed. Informat.*, vol. 115, Mar. 2021, Art. no. 103693.

[5] W. Pang, D. Zhang, J. Zhang, N. Li, W. Zheng, H. Wang, C. Liu, F. Yang, and B. Pang, "Tongue features of patients with coronavirus disease 2019: A retrospective cross-sectional study," *Integrative Med. Res.*, vol. 9, no. 3, Sep. 2020, Art. no. 100493.

[6] M. A. Yousif and J. H. Saud, "Tongue print features extraction by Gabor filters family," *Indian J. Forensic Med. Toxicol.*, vol. 13, no. 1, p. 298, Jan. 2019.

[7] X. Li, Y. Zhang, Q. Cui, X. Yi, and Y. Zhang, "Tooth-marked tongue recognition using multiple instance learning and CNN features," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 380–387, Feb. 2019.

[8] Y. Ota, T. Nakaguchi, and V. Bochko, "Tongue coating analysis via machine learning using texture and color features," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 6, pp. S212–S213, Jun. 2018.

[9] S. N. Deepa and A. Banerjee, "Intelligent decision support model using tongue image features for healthcare monitoring of diabetes diagnosis and classification," *Netw. Model. Anal. Health Inform. Bioinf.*, vol. 10, no. 1, p. 41, 2021.

[10] X. M. Zhai, H. D. Lu, and L. Z. Zhang, "Application of image segmentation technique in tongue diagnosis," in *Proc. Int. Forum Inf. Technol. Appl.*, Chengdu, China, 2009, pp. 768–771.

[11] Y.-L. Shi, J.-Y. Liu, X.-J. Hu, L.-P. Tu, J. Cui, J. Li, Z.-J. Bi, J.-C. Li, L. Xu, and J.-T. Xu, "A new method for syndrome classification of non-small-cell lung cancer based on data of tongue and pulse with machine learning," *BioMed. Res. Int.*, vol. 2021, Aug. 2021, Art. no. 1337558.

[12] Y. Tang, Y. Sun, J. Y. Chiang, and X. Li, "Research on multiple-instance learning for tongue coating classification," *IEEE Access*, vol. 9, pp. 66361–66370, 2021.

[13] S. Rajakumaran and J. Sasikala, "An automated tongue color image analysis for disease diagnosis and classification using deep learning techniques," *Eur. J. Mol. Clin. Med.*, vol. 7, no. 9, p. 1743–1759, 2020.

[14] X. Wang, J. Liu, C. Wu, J. Liu, Q. Li, Y. Chen, X. Wang, X. Chen, X. Pang, B. Chang, J. Lin, S. Zhao, Z. Li, Q. Deng, Y. Lu, D. Zhao, and J. Chen, "Artificial intelligence in tongue diagnosis: Using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark," *Comput. Struct. Biotechnol.*, vol. 18, pp. 973–980, Apr. 2020.

[15] T. Jiang, X.-J. Hu, X.-H. Yao, L.-P. Tu, J.-B. Huang, X.-X. Ma, J. Cui, Q.-F. Wu, and J.-T. Xu, "Tongue image quality assessment based on a deep convolutional neural network," *BMC Med. Informat. Decis. Making*, vol. 21, no. 1, p. 147, Dec. 2021.

[16] Q. Li and Z. Liu, "Tongue color analysis and discrimination based on hyperspectral images," *Computerized Med. Imag. Graph.*, vol. 33, no. 3, pp. 217–221, Apr. 2009.

[17] S. Yamamoto, N. Tsumura, K. Ogawa-Ochiai, T. Nakaguchi, Y. Kasahara, T. Namiki, and Y. Miyake, "Early detection of disease-oriented state from hyperspectral tongue images with principal component analysis and vector rotation," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug./Sep. 2010, pp. 3025–3028.

[18] T. Guo, R. Wang, F. Luo, X. Gong, L. Zhang, and X. Gao, "Dual-view spectral and global spatial feature fusion network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5512913.

[19] F. Luo, T. Zhou, J. Liu, T. Guo, X. Gong, and J. Ren, "Multiscale diff-changed feature fusion network for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5502713.

[20] H. Zhou, F. Luo, H. Zhuang, Z. Weng, X. Gong, and Z. Lin, "Attention multihop graph and multiscale convolutional fusion network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5508614.

[21] Y. Duan, F. Luo, M. Fu, Y. Niu, and X. Gong, "Classification via structure-preserved hypergraph convolution network for hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5507113.

[22] A. O. B. Özdemir, B. E. Gedik, and C. Y. Y. Çetin, "Hyperspectral classification using stacked autoencoders with deep learning," in *Proc. 6th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Jun. 2014, pp. 1–4.

[23] T. Li, J. Zhang, and Y. Zhang, "Classification of hyperspectral image based on deep belief networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 5132–5136.

[24] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017.

[25] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[26] W. Wang, S. Dou, Z. Jiang, and L. Sun, "A fast dense spectral–spatial convolution network framework for hyperspectral images classification," *Remote Sens.*, vol. 10, no. 7, p. 1068, Jul. 2018.

[27] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral–spatial residual network for hyperspectral image classification: A 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, Feb. 2018.

[28] J. Yang, Y.-Q. Zhao, and J. C. Chan, "Learning and transferring deep joint spectral–spatial features for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4729–4742, Aug. 2017.

[29] G. Sun, X. Zhang, X. Jia, J. Ren, A. Zhang, Y. Yao, and H. Zhao, "Deep fusion of localized spectral features and multi-scale spatial features for effective classification of hyperspectral images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 91, Sep. 2020, Art. no. 102157.
[30] E. Xie, N. Chen, J. Peng, W. Sun, Q. Du, and X. You, "Semantic and spatial–spectral feature fusion transformer network for the classification of hyperspectral image," *CAAI Trans. Intell. Technol.*, pp. 1–15, Feb. 2023, doi: 10.1049/cit2.12201.
[31] X. Li, M. Ding, and A. Pižurica, "Deep feature fusion via two-stream convolutional neural network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2615–2629, Apr. 2020.
[32] D. Zhang, J. Zhang, Z. Wang, and M. Sun, "Tongue colour and coating prediction in traditional Chinese medicine based on visible hyperspectral imaging," *IET Image Process.*, vol. 13, no. 12, pp. 2265–2270, Oct. 2019.
[33] W. J. Yan, G. Q. He, S. J. Huang, and L. Qin, "Parameter selection based on support vector machines for tongue spectral identification model," *Appl. Mech. Mater.*, vol. 678, pp. 242–251, Oct. 2014.
[34] Z. Liu, Q. Li, J. Yan, and Q. Tang, "A novel hyperspectral medical sensor for tongue diagnosis," *Sensor Rev.*, vol. 27, no. 1, pp. 57–60, Jan. 2007.
[35] T. Xie, S. Li, and B. Sun, "Hyperspectral images denoising via nonconvex regularized low-rank and sparse matrix decomposition," *IEEE Trans. Image Process.*, vol. 29, pp. 44–56, 2020.
[36] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Super-resolution for hyperspectral and multispectral image fusion accounting for seasonal spectral variability," *IEEE Trans. Image Process.*, vol. 29, pp. 116–127, 2020.
[37] H. Chen, Q. Li, M. Li, S. Liu, C. Yao, Z. Wang, Z. Zhao, P. Liu, F. Yang, X. Li, J. Wang, Y. Zeng, and X. Tong, "Microbial characteristics across different tongue coating types in a healthy population," *J. Oral Microbiol.*, vol. 13, no. 1, Jan. 2021, Art. no. 1946316.
[38] Q. Xiao, H. Zhang, J. Zhang, and L. Zhuo, "Texture analysis of tongue coating in traditional Chinese medicine based on transfer learning and multi-model decision," *Sens. Imag.*, vol. 22, no. 1, p. 8, Dec. 2021.
[39] H. Lu, Z. Ren, A. Li, J. Li, S. Xu, H. Zhang, J. Jiang, J. Yang, Q. Luo, K. Zhou, S. Zheng, and L. Li, "Tongue coating microbiome data distinguish patients with pancreatic head cancer from healthy controls," *J. Oral Microbiol.*, vol. 11, no. 1, Jan. 2019, Art. no. 1563409.
[40] L. J. Rodríguez-Aragón and A. Zhigljavsky, "Singular spectrum analysis for image processing," *Statist. Interface*, vol. 3, no. 3, pp. 419–426, 2010.
[41] R. Mahmoudvand, N. Najari, and M. Zokaei, "On the optimal parameters for reconstruction and forecasting in singular spectrum analysis," *Commun. Statist. Simul. Comput.*, vol. 42, pp. 860–870, Apr. 2013.
[42] X. D. Huang, H. Zhang, L. Zhuo, X. Li, and J. Zhang, "TISNet-enhanced fully convolutional network with encoder–decoder structure for tongue image segmentation in Traditional Chinese Medicine," *Comput. Math. Methods Med.*, vol. 2020, Aug. 2020, Art. no. 6029258.
[43] Q. Xu, Y. Zeng, W. Tang, W. Peng, T. Xia, Z. Li, F. Teng, W. Li, and J. Guo, "Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 9, pp. 2481–2489, Sep. 2020.
[44] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatiospectral information fusion and PLS regression," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1127–1137, Mar. 2015.
[45] Y. Garini, I.-T. Young, and G. McNamara, "Spectral imaging: Principles and applications," *Cytometry A*, vol. 69A, pp. 735–747, Aug. 2006.
[46] S. S. Heydari and G. Mountrakis, "Meta-analysis of deep neural networks in remote sensing: A comparative study of mono-temporal classification to support vector machines," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 192–210, Jun. 2019.
[47] M. Pal and G. M. Foody, "Feature selection for classification of hyperspectral data by SVM," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2297–2307, May 2010.
[48] D. Liu and J. Li, "Data field modeling and spectral–spatial feature fusion for hyperspectral data classification," *Sensors*, vol. 16, no. 12, p. 2146, Dec. 2016.
[49] S. Lorenz, P. Seidel, P. Ghamisi, R. Zimmermann, L. Tusa, M. Khodadadzadeh, I. C. Contreras, and R. Gloaguen, "Multi-sensor spectral imaging of geological samples: A data fusion approach using spatio-spectral feature extraction," *Sensors*, vol. 19, no. 12, p. 2787, Jun. 2019.

DONG ZHANG received the master's degree from the School of Computer Science and Technology, Tianjin University. He is currently an Assistant Engineer with the Institute of Chinese Medicine, Tianjin University of Traditional Chinese Medicine. His current research interest includes computer image processing, especially on machine learning.

WENTAI PANG received the Ph.D. degree from the Tianjin University of Traditional Chinese Medicine and the Ph.D. degree from the Institute of Chinese Medicine, Tianjin University of Traditional Chinese Medicine. His current research interest includes traditional Chinese medicine clinical evaluation, especially on COVID-19.

KEYI WANG received the master's degree from the Institute of Chinese Medicine, Tianjin University of Traditional Chinese Medicine, where she is currently pursuing the Ph.D. degree with the Institute of Chinese Medicine. Her current research interest includes evidence-based Chinese medicine.

FENGWEN YANG received the Ph.D. degree from the Tianjin University of Traditional Chinese Medicine. He is currently a Vice Researcher with the Institute of Chinese Medicine, Tianjin University of Traditional Chinese Medicine. His current research interest includes evidence-based Chinese medicine.

JUNHUA ZHANG received the Ph.D. degree from the Tianjin University of Traditional Chinese Medicine. He is currently a Researcher with the Institute of Chinese Medicine, Tianjin University of Traditional Chinese Medicine. His current research interest includes evidence-based Chinese medicine.