



## Original Article

# Traditional Chinese medicine diagnostic prediction model for holistic syndrome differentiation based on deep learning



Zhe Chen <sup>a,b,c,d</sup>, Dong Zhang <sup>a</sup>, Chunxiang Liu <sup>a</sup>, Hui Wang <sup>a</sup>, Xinyao Jin <sup>a</sup>, Fengwen Yang <sup>a,\*</sup>, Junhua Zhang <sup>a,b,\*</sup>

<sup>a</sup> Evidence-based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>b</sup> Haihe Laboratory of Modern Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>c</sup> National Key Laboratory of Chinese Medicine Modernization, Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>d</sup> State Key Laboratory of Component-based Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, China

## ARTICLE INFO

## Keywords:

Traditional Chinese medicine syndromes  
Deep learning  
Holistic syndrome differentiation  
Expert knowledge  
Artificial intelligence

## ABSTRACT

**Background:** With the development of traditional Chinese medicine (TCM) syndrome knowledge accumulation and artificial intelligence (AI), this study proposes a holistic TCM syndrome differentiation model for the classification prediction of multiple TCM syndromes based on deep learning and accelerates the construction of modern foundational TCM equipment.

**Methods:** We searched publicly available TCM guidelines and textbooks for expert knowledge and validated these sources using ten-fold cross-validation. Based on the BERT and CNN models, with the classification constraints from TCM holistic syndrome differentiation, the TCM-BERT-CNN model was constructed, which completes the end-to-end TCM holistic syndrome text classification task through symptom input and syndrome output. We assessed the performance of the model using precision, recall, and F1 scores as evaluation metrics.

**Results:** The TCM-BERT-CNN model had a higher precision (0.926), recall (0.9238), and F1 score (0.9247) than the BERT, TextCNN, LSTM RNN, and LSTM ATTENTION models and achieved superior results in model performance and predictive classification of most TCM syndromes. Symptom feature visualization demonstrated that the TCM-BERT-CNN model can effectively identify the correlation and characteristics of symptoms in different syndromes with a strong correlation, which conforms to the diagnostic characteristics of TCM syndromes.

**Conclusions:** The TCM-BERT-CNN model proposed in this study is in accordance with the TCM diagnostic characteristics of holistic syndrome differentiation and can effectively complete diagnostic prediction tasks for various TCM syndromes. The results of this study provide new insights into the development of deep learning models for holistic syndrome differentiation in TCM.

## 1. Introduction

The distinctive feature of syndrome differentiation in traditional Chinese medicine (TCM) has been extensively utilized in the clinical diagnosis and treatment of various diseases.<sup>1-3</sup> Holistic syndrome differentiation is a unique clinical diagnostic method based on TCM theory, which enables comprehensive analysis of a patient's overall pathological condition to assess the intricacy of TCM syndrome patterns for clinical practice.<sup>4-7</sup> The TCM holistic syndrome differentiation diagnosis method plays a significant role in the treatment and outcome of various diseases, serving as a crucial component within the distinctive clinical TCM diagnosis and treatment system. The standards for TCM syndrome differentiation continue to exhibit certain variations, leading to a significant

amount of syndrome differentiation data in clinical practice that cannot be reasonably applied.<sup>6-9</sup> When confronted with the complex syndrome diagnosis of multiple diseases, TCM practitioners typically rely on their individual subjective experiences, which poses challenges for the standardization and promotion of TCM clinical syndrome differentiation.<sup>8,10,11</sup>

With growing emphasis on the development of artificial intelligence (AI) for TCM, AI-driven TCM characteristic diagnostic technology of syndrome differentiation may alleviate the scarcity of clinical TCM practitioners, and promote AI development of TCM medical decision-making.<sup>12,13</sup> The development of AI in TCM remains in its infancy, and the establishment of an intelligent infrastructure based on deep learning for TCM syndrome is of more paramount significance.<sup>14-16</sup> Several

\* Corresponding authors at: Evidence-based Medicine Center, Tianjin University of Traditional Chinese Medicine, No. 10 Poyanghu Road, Jinghai District, Tianjin 301617, China.

E-mail addresses: [13682027022@163.com](mailto:13682027022@163.com) (F. Yang), [zjhtcm@foxmail.com](mailto:zjhtcm@foxmail.com) (J. Zhang).

<https://doi.org/10.1016/j.imr.2023.101019>

Received 24 September 2023; Received in revised form 2 December 2023; Accepted 17 December 2023

Available online 19 December 2023

2213-4220/© 2024 Korea Institute of Oriental Medicine. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

TCM research institutions have integrated AI techniques with distinctive TCM syndrome differentiation, and have achieved better predictive performance of the model in several tasks, such as TCM syndrome prediction, tongue diagnosis, and constitution identification, among others.<sup>17–20</sup> TCM syndrome differentiation is not limited to a single disease, and a diagnostic approach solely focused on a particular disease cannot adequately capture the clinical diagnostic characteristics of TCM holistic syndrome differentiation.<sup>21</sup>

Multiple TCM syndromes can coexist during the development of various diseases, displaying complex and dynamic changes with overlapping features across the entire process.<sup>22</sup> To our knowledge, previous research has been limited to several TCM syndromes within a single disease, which significantly constrains the performance and extrapolation of models, and impedes the verification and interactive utilization of multiple TCM syndromes in various diseases.<sup>19,23,24</sup> The dynamic complexity of TCM syndrome differentiation also increases diagnostic challenges in clinical practice for holistic syndrome differentiation of multiple diseases and TCM syndromes.<sup>21,22,25</sup>

Given the complexity of holistic syndrome differentiation involving multiple syndromes, this study introduced an overall syndrome model based on deep learning to predict various TCM syndromes. This model aims to provide intelligent decision support for diagnosing intricate TCM syndromes in clinical practice, thereby enhancing clinical decision-making and offering an intelligent approach for holistic differentiation. This study combines the BERT and CNN models, incorporating the characteristics of TCM clinical differentiation as classification constraints, to create a “holistic syndrome differentiation model” capable of handling multiple classification prediction tasks. Furthermore, unlike previous studies that focused solely on specific TCM syndromes within a single disease, the TCM syndrome prediction model developed in this study can be applied for the comprehensive prediction of complex TCM syndromes, thus providing intelligent support for holistic TCM syndrome differentiation with value and significance for TCM clinical diagnosis.

## 2. Methods

### 2.1. Source of data and collection

In this study, we aimed to identify relevant literature, including clinical practice guidelines and expert consensus in TCM, through computerized searches of both Chinese and English databases, including the China National Knowledge Infrastructure (CNKI), Wanfang Data Knowledge Service Platform, VIP Information, China National Biomedical Literature Service System (Sinomed), PubMed, Embase, and others, conducted from inception to 2022. To comprehensively review TCM expert knowledge on holistic syndrome differentiation, we also conducted searches in TCM textbooks (e.g., Internal Medicine in TCM, Pediatrics in TCM, Gynecology in TCM) as supplements.

During the data retrieval and screening phases, two researchers with TCM backgrounds independently conducted literature searches, literature screening processes, and full-text article reviews to assess eligibility for final inclusion. Cross-checking was performed between the two researchers to validate the results, a third researcher was consulted for verification, and a consensus was reached through multiple group discussions. We collected 21 TCM syndromes, including program syndrome identification (exterior, interior, deficiency, excess, cold, and heat syndromes), Zang disease syndrome identification (heart disease, liver disease, spleen disease, lung disease, and nephropathy), and pathological syndrome identification (excess cold, deficiency cold, deficiency heat, excess heat, qi deficiency, blood deficiency, qi stagnation, blood stasis, phlegm fluid retention, and critical syndrome).

### 2.2. Diagnostic criteria

The standards for TCM symptom normalization were primarily developed based on references such as the “WHO International Standard

Terminologies on TCM,” “Standardization of Common Clinical Symptom Terminology in TCM,” and others. As for the information related to TCM syndromes, the main reference used was the “TCM Diagnosis” textbook, while also drawing upon and referring to “Clinical Terminology in TCM” and “Classification and Coding of TCM Diseases and Syndromes” for standardization and classification.

### 2.3. TCM holistic syndrome differentiation model

This study introduces a novel TCM-BERT-CNN model that builds upon the fusion of the BERT and CNN models with added classification constraints featuring TCM characteristics (Fig. 1). The TCM-BERT-CNN models had a learning rate of 1e-5, a batch size of 20, and 20 epochs.

The BERT model is a pre-training language model of a bidirectional encoder based on a transformer, composed of a self-attention mechanism and feed-forward neural network, which can better extract Chinese semantic features with model improvements.<sup>26–28</sup> The BERT model converts text information into token embeddings and builds segment and position embeddings by automatically learning text semantic information and differences.<sup>27</sup> Because of the advantages of the BERT model, it can achieve satisfactory prediction results when dealing with complex and related semantic knowledge data, such as symptoms and TCM syndromes.<sup>29,30</sup>

The TextCNN model is a fine-tuned variant of the CNN model, and the overall framework is composed of an input layer, word embedding layer, convolution layer, pooling layer, and full connection layer to efficiently extract information features with a relatively simple network structure, which is more suitable for text classification tasks of TCM syndrome differentiation.<sup>31–33</sup> The BERT model provides word embedding for different semantic contexts, and the CNN refines the features of each word from the BERT model into N-gram features. The information related to 21 TCM syndromes is categorized into two constraint groups (‘exterior syndrome’ and ‘internal syndrome’; ‘deficiency syndrome’ and ‘excess syndrome’) based on TCM characteristics, and the remaining 17 syndromes form an overall classification group. The BERT embedding output was obtained after extracting the TCM-specific features, which served as the input for the CNN with hierarchical connections, resulting in the structure of the TCM-BERT-CNN model.

After the tokenization of the sentences, special tokens [CLS] and [SEP] are used to accept all words and separate the two sentences. [CLS] serves as the start token [SEP] as the end token, and the vectors are mapped to multiple categories. A two-dimensional vector was used for category determination in the two constraint groups, and a multidimensional vector was employed in the overall classification group. A multilayer Perceptron was used for vector space mapping, with SOFTMAX and SIGMOID serving as activation functions for the constrained and non-constrained groups, respectively. The SOFTMAX function normalizes probabilities based on the natural exponent ‘e,’ ensuring that the sum of probabilities for binary classification data equals 1. The SIGMOID function defines the value range for multiple classifications and distributes the probability values within the 0–1 numerical range.

This study employs PyTorch 1.10 as the framework and Python 3.6 as the experimental environment, to compare the TCM-BERT-CNN model with the BERT model, TextCNN model, LSTM RNN model, and LSTM ATTENTION model.

### 2.4. Model evaluation metrics

This study focuses on a classification prediction task for different TCM syndromes and employs four types of results: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Because of the diverse TCM syndromes in this study and the presence of imbalanced datasets for some syndromes, accuracy was not the optimal evaluation metric, which led to significant discrepancies in the prediction results from minority samples. Therefore, this study used precision,

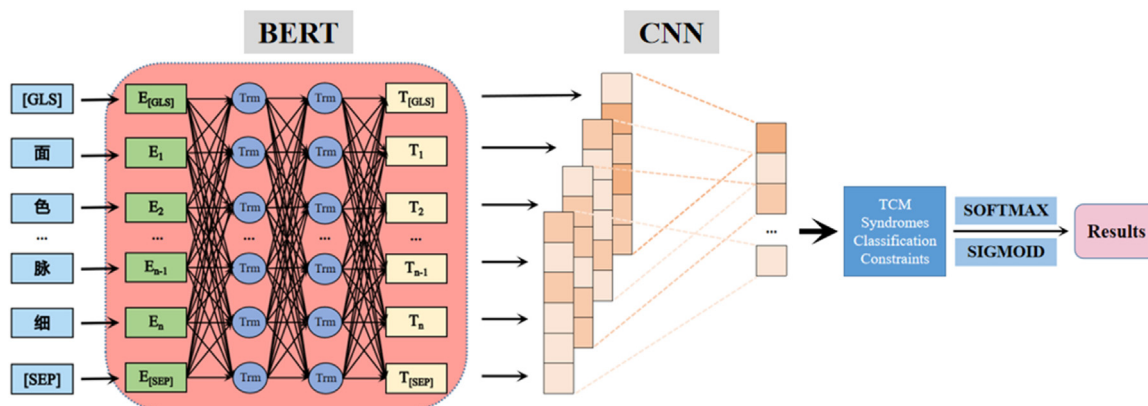


Fig. 1. The Brief structure of the TCM-BERT-CNN model.

recall, and F1 score as measures to assess the model's prediction performance, which are defined as follows:

$$Precision(TP, FP) = \frac{TP}{TP + FP} \quad (1)$$

$$Recall(TP, FN) = \frac{TP}{TP + FN} \quad (2)$$

$$F1 \text{ score} = \frac{Precision * Recall * 2}{Precision + Recall} \quad (3)$$

### 2.5. Data set segmentation and validation

This study utilized a new dataset constructed based on clinical practice guidelines and textbooks in TCM. The performance of the model on the new dataset was assessed using cross-validation. Cross-validation can serve as a method for splitting a dataset and evaluating the deep learning model's understanding of the data. Data overlap was effectively addressed through data partitioning and validation based on cross-validation. In this study, ten-fold cross-validation was used, which involved dividing the dataset into ten subsets, using nine as training sets and one as a test set. After ten cycles of validation, the results from these ten iterations were averaged to assess the accuracy and precision of the model algorithm, effectively mitigating the impact of an imbalanced data distribution on the results.

## 3. Results

### 3.1. Data set information

A total of 6148 samples comprised the final data set with ten-fold cross-validation, and multiple TCM syndromes coexisting and intersecting within a single sample. The classification of dataset distribution for the TCM syndromes were as follows: 362 cases classified as exterior syndrome, 5786 internal syndrome, 2769 deficiency syndrome, 3379 excess syndrome, 1080 cold syndrome, 2923 heat syndrome, 2065 heart disease, 2277 liver disease, 2196 spleen disease, 1011 lung disease, 1123 nephropathy disease, 377 excess cold syndrome, 1741 excess heat syndrome, 701 deficiency cold syndrome, 1187 deficiency heat syndrome, 1954 qi deficiency syndrome, 681 blood deficiency syndrome, 2157 qi stagnation syndrome, 1011 blood stasis syndrome, 2780 phlegm fluid retention syndrome, and 336 critical syndrome.

### 3.2. Model comparison

This study tested five deep learning models for comparison: BERT, TextCNN, LSTM RNN, LSTM ATTENTION, and TCM-BERT-CNN, as shown in Table 1. The results of models are as

Table 1

Results of model evaluation measures.

Model	Precision	Recall	F1
BERT	0.8818±0.0298	0.9083±0.0215	0.8945±0.0195
TextCNN	0.8355±0.0324	0.8428±0.0362	0.8385±0.0258
LSTM RNN	0.8346±0.0313	0.8453±0.0362	0.8394±0.0264
LSTM ATTENTION	0.8557±0.0349	0.8504±0.0273	0.8525±0.0227
TCM-BERT-CNN	0.926±0.0274	0.9238±0.0293	0.9247±0.0239

follows: BERT model with precision (0.8818±0.0298), recall (0.9083±0.0215), and F1 score (0.8945±0.0195); TextCNN model with precision (0.8355±0.0324), recall (0.8428±0.0362), and F1 score (0.8385±0.0258); LSTM RNN model with precision (0.8346±0.0313), recall (0.8453±0.0362), and F1 score (0.8394±0.0264); LSTM ATTENTION model with precision (0.8557±0.0349), recall (0.8504±0.0273), and F1 score (0.8525±0.0227); TCM-BERT-CNN model with precision (0.926±0.0274), recall (0.9238±0.0293), and F1 score (0.9247±0.0239). The ten-fold cross-validation results for each model were relatively stable, indicating that all models exhibited good predictive performance across various TCM syndromes in Fig. 2, and the TCM-BERT-CNN model was higher than the other models in terms of precision, recall, and F1 score.

### 3.3. Model comparison in various TCM syndromes

In the results of these models for 21 TCM syndromes, the BERT model shows that the precision (0.8347–0.9342) with the highest in interior syndrome, recall (0.8586–0.9325) with the highest in heat syndrome, and F1 score (0.8588–0.9311) with the highest in heat syndrome. The TextCNN model shows that the precision (0.7892–0.8746) with the highest in blood stasis syndrome, recall (0.7838–0.8879) with the highest in lung disease, and F1 score (0.8018–0.875) with the highest in nephropathy disease. The LSTM RNN model shows that the precision (0.7989–0.8837) with the highest in qi stagnation syndrome, recall (0.789–0.8868) with the highest in liver disease, and F1 score (0.801–0.8729) with the highest in blood stasis syndrome. The LSTM ATTENTION model shows that the precision (0.7907–0.8945) with the highest in excess cold syndrome, recall (0.8012–0.8901) with the highest in deficiency cold syndrome, and F1 score (0.8148–0.8883) with the highest in deficiency cold syndrome (Table 2–4).

The TCM-BERT-CNN model showed that the precision ranged from 0.8840 to 0.9960, recall ranged from 0.8735 to 0.9926, and the F1 score ranged from 0.8949 to 0.9943 for the TCM syndromes (Table 2–4). Interior, cold, and excess syndromes had higher precision, recall, and F1 scores in principal syndrome identification (Table 2). Lung disease and nephropathy had higher precision, recall, and F1 scores in Zang organ syndrome identification (Table 3). Excess Cold syndrome, deficiency

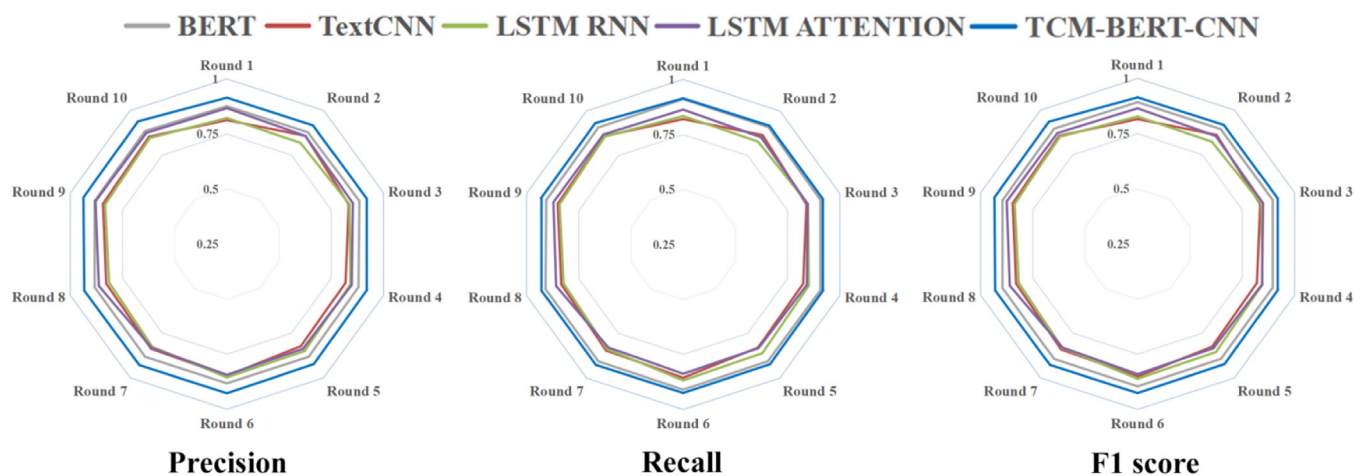


Fig. 2. Ten-fold cross-validation plot of models in precision, recall, and F1 score.

**Table 2**  
Evaluation Measures in principles syndrome identification.

Model	Evaluation Measures	Exterior syndrome	Interior syndrome	Deficiency syndrome	Excess syndrome	Cold syndrome	Heat syndrome
BERT	Precision	0.8995	0.9342	0.9031	0.8464	0.9132	0.9298
	Recall	0.8586	0.9207	0.9101	0.9153	0.8730	0.9325
	F1	0.8785	0.9274	0.9066	0.8794	0.8926	0.9311
TEXTCNN	Precision	0.8518	0.8593	0.8674	0.8640	0.7932	0.8068
	Recall	0.8184	0.8732	0.8608	0.8003	0.8543	0.8689
	F1	0.8347	0.8662	0.8641	0.8309	0.8226	0.8367
LSTM RNN	Precision	0.8077	0.8402	0.8648	0.8715	0.8765	0.8442
	Recall	0.8863	0.8476	0.8478	0.8529	0.8317	0.8829
	F1	0.8452	0.8439	0.8562	0.8621	0.8535	0.8631
LSTM ATTENTION	Precision	0.8500	0.8716	0.8776	0.8314	0.8830	0.7907
	Recall	0.8756	0.8581	0.8508	0.8585	0.8133	0.8673
	F1	0.8627	0.8648	0.8639	0.8447	0.8468	0.8273
TCM-BERT-CNN	Precision	0.9141	0.9960	0.9117	0.9266	0.9562	0.9096
	Recall	0.8913	0.9926	0.9504	0.9566	0.9174	0.9528
	F1	0.9026	0.9943	0.9307	0.9414	0.9363	0.9306

**Table 3**  
Evaluation Measures in Zang organs syndrome identification.

Model	Evaluation Measures	Heart Disease	Liver Disease	Spleen Disease	Lung Disease	Nephropathy Disease
BERT	Precision	0.8530	0.9115	0.8932	0.8955	0.8777
	Recall	0.8832	0.9275	0.9079	0.8895	0.9260
	F1	0.8678	0.9195	0.9005	0.8925	0.9012
TEXTCNN	Precision	0.8203	0.8242	0.8458	0.7892	0.8627
	Recall	0.8038	0.8625	0.8684	0.8879	0.8877
	F1	0.8119	0.8429	0.8569	0.8356	0.8750
LSTM RNN	Precision	0.8033	0.7989	0.8534	0.8127	0.8009
	Recall	0.7986	0.8868	0.8821	0.8367	0.8053
	F1	0.8010	0.8405	0.8675	0.8245	0.8031
LSTM ATTENTION	Precision	0.8843	0.8758	0.7997	0.8631	0.8910
	Recall	0.8449	0.8505	0.8307	0.8624	0.8362
	F1	0.8641	0.8629	0.8148	0.8628	0.8627
TCM-BERT-CNN	Precision	0.8860	0.9121	0.8867	0.9396	0.9300
	Recall	0.9045	0.9295	0.9097	0.9308	0.9412
	F1	0.8951	0.9207	0.8980	0.9351	0.9356

cold syndrome, Qi Stagnation syndrome, and Qi Deficiency syndrome had higher precision, recall, and F1 scores in disease pathological syndrome identification (Table 4).

### 3.4. Symptom feature visualization for the TCM-BERT-CNN model

A semantic feature analysis was conducted for TCM symptom features learned by the TCM-BERT-CNN model to present symptom feature visualization, which was performed to illustrate the model's semantic feature learning and correlation rules between symptoms within each

TCM syndrome. Examples were randomly selected from 21 TCM syndromes for symptom feature analysis and presentation (Fig. 3).

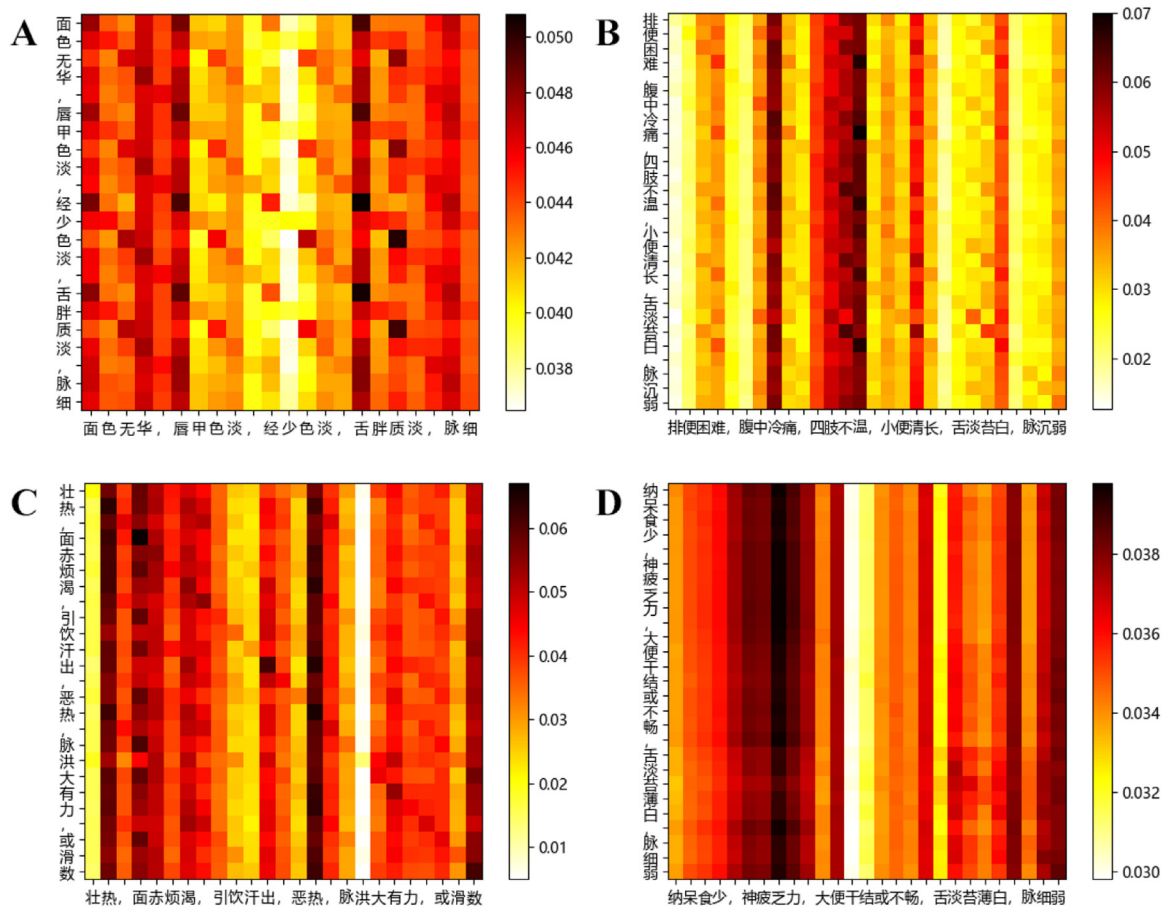
For example, deficiency syndrome can be seen as having “dull complexion, pale lips and nails, light period, pale enlarged tongue, fine pulse”. The symptom feature correlation matrix shows that “dull complexion,” “lips,” “tongue,” and “fine pulse” have higher correlations with other symptom features. “Dull complexion” and “fine pulse” are the main symptom features of deficiency syndrome, with the highest probability of co-occurrence with other symptoms (Fig. 3A).

Cold syndrome included symptoms such as “constipation, cold stomachache, cold limbs, clear abundant urination, pale tongue with white



**Table 4**  
Evaluation measures in disease pathological syndrome identification.

Model	Evaluation Measures	Excess Cold syndrome	Excess Heat syndrome	Deficiency Cold syndrome	Deficiency Heat syndrome	Qi Deficiency syndrome	Blood Deficiency syndrome	Qi Stagnation syndrome	Blood Stasis syndrome	Phlegm Fluid Retention syndrome	Critical syndrome
BERT	Precision	0.8542	0.8402	0.8566	0.8933	0.8870	0.9029	0.8347	0.8562	0.8781	0.8576
	Recall	0.9108	0.9282	0.9102	0.9257	0.9138	0.9164	0.8843	0.8863	0.9283	0.9256
	F1	0.8816	0.8820	0.8826	0.9092	0.9002	0.9095	0.8588	0.8709	0.9025	0.8903
TEXTCNN	Precision	0.8724	0.8584	0.8446	0.8310	0.8014	0.8036	0.7915	0.8746	0.8626	0.8207
	Recall	0.7963	0.8735	0.8126	0.8323	0.8782	0.8088	0.8280	0.8243	0.8746	0.7838
	F1	0.8326	0.8658	0.8283	0.8316	0.8381	0.8062	0.8093	0.8487	0.8685	0.8018
LSTM RNN	Precision	0.8386	0.8681	0.8022	0.8401	0.7996	0.8199	0.8837	0.8621	0.8217	0.8163
	Recall	0.8086	0.8677	0.8132	0.8699	0.8459	0.7890	0.7895	0.8839	0.8398	0.8848
	F1	0.8233	0.8679	0.8076	0.8547	0.8221	0.8042	0.8339	0.8729	0.8306	0.8491
LSTM ATTENTION	Precision	0.8945	0.8537	0.8865	0.8595	0.8720	0.8470	0.8181	0.8881	0.8047	0.8268
	Recall	0.8445	0.8700	0.8901	0.8012	0.8406	0.8356	0.8296	0.8311	0.8796	0.8885
	F1	0.8688	0.8618	0.8883	0.8293	0.8560	0.8413	0.8238	0.8586	0.8405	0.8565
TCM-BERT-CNN	Precision	0.9486	0.9173	0.9513	0.9212	0.9329	0.9097	0.9416	0.9312	0.8840	0.9402
	Recall	0.9594	0.8735	0.8978	0.8874	0.9266	0.9062	0.9336	0.9028	0.9355	0.9008
	F1	0.9540	0.8949	0.9238	0.9039	0.9297	0.9079	0.9376	0.9168	0.9090	0.9201



**Fig. 3.** Symptom Feature Correlation Matrix of TCM syndromes (A: Deficiency syndrome; B: Cold syndrome; C: Excessive Heat syndrome; D: Qi Deficiency syndrome).

fur, sunken weak pulse”, and the symptom feature correlation matrix shows that “cold limbs,” “cold,” and “clear” have higher correlations with other symptoms. Among them, “cold,” “clear,” and “unwarm” are significant features of cold syndrome and have high correlations with each other (Fig. 3B).

In the excessive heat syndrome, included symptoms such as “high fever, reddish complexion with restlessness and thirst, perspiration, aversion to heat, surging or rapid slippery pulse”, and the symptom feature correlation matrix shows that “heat,” “reddish complexion,”

“thirst,” and “rapid” have high correlations with other related symptoms. The model has effectively captured the diagnostic characteristics of excessive heat syndrome, with “heat” being the most prominent symptom, as evidenced by the robust correlations between “high fever” and “aversion to heat” (Fig. 3C).

For qi deficiency syndrome, “torpid intake, lassitude of spirit with lack of strength, dry stool or constipation, pale tongue with thin fur, fine weak pulse” are the manifested symptoms, and the symptom feature correlation matrix indicates that “lassitude of spirit,” “lack of

strength,” and “fine and weak” are highly correlated with other symptoms, in line with the clinical diagnostic criteria for qi deficiency syndrome (Fig. 3D).

## 4. Discussions

### 4.1. Summary of finding

In this study, we constructed five deep learning models, namely, BERT, TextCNN, LSTM RNN, LSTM ATTENTION, and TCM-BERT-CNN, to screen TCM patterns. The results demonstrated the overall experimental performance of the TCM-BERT-CNN model with a precision of 0.926, recall of 0.9238, and F1 score of 0.9247 by ten-fold cross-validation. Additionally, it had a better model performance than the others. The TCM-BERT-CNN model performed better than the other four deep learning models. By evaluating the predictive performance of the TCM-BERT-CNN model for 21 types of TCM syndromes, the model can effectively recognize and predict classification information and demonstrates high precision, recall, and F1 scores, which are better than other models for various TCM syndromes.

The detailed visual results of the TCM-BERT-CNN model confirm the stability of its predictive outcomes. Visualization of symptom correlations across multiple TCM syndromes demonstrated that the TCM-BERT-CNN model effectively learned semantic features and identified the characteristic symptoms of TCM syndromes. While increasing the number of layers in symptom data processing can introduce complexity, this study successfully projected symptom data from lower-dimensional spaces to higher-dimensional spaces for analysis, achieving better model performance in handling various TCM syndromes. The semantic feature presentation of the TCM-BERT-CNN model reveals that information between symptoms in TCM patterns is not independent but exhibits correlations. Different-dimensional data processing indicated that the TCM-BERT-CNN model achieved good predictive results across various TCM syndromes.

Currently, deep learning is undergoing rapid development in TCM.<sup>34–36</sup> The most common application involves predicting various TCM syndromes under a single disease, which significantly limits the model’s extrapolation and fails to encompass the TCM holistic syndromes.<sup>37</sup> The main challenge for TCM text-based deep learning is that current models are unlikely to be able to accommodate TCM holistic syndromes diagnostic methods, and are limited by standards for data processing and artificial intelligence technology of TCM syndrome characteristics.<sup>38–39</sup> Therefore, this study adopted the “holistic syndrome differentiation” to construct the TCM-BERT-CNN deep learning model, without using the common “single disease syndrome differentiation.” Based on the “holistic syndrome differentiation” thinking of TCM, cross-complex predictions were made for 21 syndromes in the principles, Zang organs, and disease pathological syndrome differentiation, thereby achieving a mixed combination prediction of multiple syndrome predictions across various diseases.

This study focuses on intelligent clinical diagnosis decision-making for TCM syndrome differentiation and provides a method for syndrome differentiation diagnosis and treatment for TCM clinical diagnosis. In future research, it will be necessary to integrate expert knowledge rules to construct a new model for intelligent holistic TCM syndrome differentiation based on the integration of rules and deep learning. In conclusion, this study provides intelligent methods for holistic syndrome differentiation to assist in decision-making for TCM clinical diagnosis of TCMs.

### 4.2. Strengths and limitations

The strengths of this study include changing the existing “disease-based” classification prediction model by highlighting the characteristics of “TCM holistic syndrome” for syndrome differentiation and to solve the existing limitations of the TCM syndrome differentiation model. Traditional methods of holistic syndrome differentiation in TCM

depend on doctors, experience accumulation, subjective thinking, and teaching through words and deeds, which also lead to some experiences not being effectively inherited and retained. This study used domain expert knowledge as data, combined with the characteristics of holistic TCM syndrome differentiation based on the integrated model, to innovate the TCM-BERT-CNN model to promote and apply high-level expert knowledge. There are some limitations as follows. The interior syndrome has the highest precision, recall, and F1 score in the model results, which is related to the significant imbalance in the sample. Although this study has predicted 21 TCM syndromes with better performance, the application of special syndrome information such as Six Fu syndrome, “Wei, Qi, Ying and Blood,” and more, is lacking in practice and should be the focus of our research in the future.

### 4.3. Conclusions

This study explored the feasibility of a TCM syndrome differentiation model based on expert knowledge and proposed a TCM-BERT-CNN model with TCM characteristics to complete the end-to-end prediction of 21 TCM syndromes. By comparing the model performance of the five deep learning models, we found that the TCM-BERT-CNN model was better than the other models, which can understand the symptom semantic characteristics of TCM syndromes and is in accordance with the characteristics of TCM holistic syndrome differentiation. The TCM-BERT-CNN model will accelerate the intelligent application of TCM characteristic syndrome differentiation with deep learning, provide modern basic diagnostic equipment for TCM, and guide the clinical diagnosis of TCM.

### Author contributions

Conceptualization: ZC and JHZ; Resources: ZC, DZ, CXL, XYJ, and HW; Methodology, Analysis and Visualization: ZC, DZ, and FWY; Writing: ZC and DZ; Review & Editing: All; Supervision: FWY and JHZ; Project administration: FWY and JHZ.

### Conflict of interest

The authors declare no conflicts of interest in this study.

### Funding

This research was supported by the China Postdoctoral Science Foundation (Certificate Number: 2023M742627), Foundation of State Key Laboratory of Component-based Chinese Medicine (Grant No. CBCM2023201), National Multidisciplinary Innovation Team of Traditional Chinese Medicine (ZYYCXTD-D-202204), National Natural Science Foundation of China under Grant (82205316), National Funded Postdoctoral Researcher Program (No. GZC20231928) and Science and Technology Project of Haihe Laboratory of Modern Chinese Medicine (No. 22HHZYSS00013).

### Ethical statement

Ethical approval was not applicable to this study.

### Data availability

The data that support the findings of this study are available within the article and from the corresponding author upon reasonable request.

### References

- Zhang T, Zhang B, Xu J, Ren S, Huang S, Shi Z, et al. Chinese herbal compound prescriptions combined with Chinese medicine powder based on traditional Chinese medicine syndrome differentiation for treatment of chronic atrophic gastritis with erosion: a multi-center, randomized, positive-controlled clinical trial. *Chin Med*. 2022;17(1):142.

2. Bai R, Yang Q, Xi R, Che Q, Zhao Y, Guo M, et al. The effectiveness and safety of Chinese Patent Medicines based on syndrome differentiation in patients following percutaneous coronary intervention due to acute coronary syndrome (CPM trial): a nationwide Cohort Study. *Phytomed: Int J Phytother Phytopharmacol*. 2023;109:154554.
3. Leung AYL, Zhang J, Chan CY, Chen X, Mao J, Jia Z, et al. Validation of evidence-based questionnaire for TCM syndrome differentiation of heart failure and evaluation of expert consensus. *Chin Med*. 2023;18(1):70.
4. Zhang CT, Cheng HL, Chen KL, Zhang ZP, Lin JQ, Xiao SJ, et al. Progress on prevention and treatment of cerebral small vascular disease using integrative medicine. *Chin J Integr Med*. 2023;29(2):186–191.
5. Wang L, Wu F, Hong Y, Shen L, Zhao L, Lin X. Research progress in the treatment of slow transit constipation by traditional Chinese medicine. *J Ethnopharmacol*. 2022;290:115075.
6. Jiang W, Qi J, Li X, Chen G, Zhou D, Xiao W, et al. Post-infectious cough of different syndromes treated by traditional Chinese medicines: a review. *Chin Herb Med*. 2022;14(4):494–510.
7. Liu ZY, Luo JY, Lu Y, Du SY. Application of traditional Chinese medicine theory in modern traditional Chinese medicine nano-preparation: taking tumor treatment as an example. *Zhongguo Zhong Yao Za Zhi*. 2023;48(6):1455–1462.
8. Liu R, Jiang LJ, Yang Y, Wang CC, Tong X, Xu WM, et al. Study on syndrome differentiation strategy of phlegm and blood stasis syndromes of coronary heart disease based on expert consultation on medical cases. *Ann Palliat Med*. 2021;10(9):9940–9952.
9. Song C, Ye X, Fu H, Lin L, Jin Y, Liu F, et al. Diagnostic and categorization criteria for palpitations below the heart in traditional Chinese medicine: a delphi consensus study. *Altern Ther Health Med*. 2021;27(5):68–72.
10. O'Brien KA, Abbas E, Zhang J, Guo ZX, Luo R, Bensoussan A, et al. Understanding the reliability of diagnostic variables in a Chinese Medicine examination. *J Alternat Complement Med (New York, NY)*. 2009;15(7):727–734.
11. Matos LC, Machado JP, Monteiro FJ, Greten HJ. Can traditional chinese medicine diagnosis be parameterized and standardized? A narrative review. *Healthcare (Basel, Switzerland)*. 2021;9(2).
12. Zhou S, Li K, Ogihara A, Wang X. Perceptions of traditional Chinese medicine doctors about using wearable devices and traditional Chinese medicine diagnostic instruments: a mixed-methodology study. *Digital Health*. 2022;8:20552076221102246.
13. Liu Q, Li Y, Yang P, Liu Q, Wang C, Chen K, et al. A survey of artificial intelligence in tongue image for disease diagnosis and syndrome differentiation. *Digital Health*. 2023;9:20552076231191044.
14. Feng C, Shao Y, Wang B, Qu Y, Wang Q, Li Y, et al. Development and application of artificial intelligence in auxiliary TCM diagnosis. *Evid Complement Alternat Med: eCAM*. 2021;2021:6656053.
15. Duan YY, Liu PR, Huo TT, Liu SX, Ye S, Ye ZW. Application and development of intelligent medicine in traditional Chinese medicine. *Curr Med Sci*. 2021;41(6):1116–1122.
16. Wang Y, Shi X, Li L, Efferth T, Shang D. The impact of artificial intelligence on traditional Chinese medicine. *Am J Chin Med (Gard City N Y)*. 2021;49(6):1297–1314.
17. Li M, Wen G, Zhong J, Yang P. Personalized intelligent syndrome differentiation guided by TCM consultation philosophy. *J Healthc Eng*. 2022;2022:6553017.
18. Li J, Zhang Z, Zhu X, Zhao Y, Ma Y, Zang J, et al. Automatic classification framework of tongue feature based on convolutional neural networks. *Micromachines*. 2022;13(4).
19. Hu C, Zhang S, Gu T, Yan Z, Jiang J. Multi-task joint learning model for Chinese word segmentation and syndrome differentiation in traditional Chinese medicine. *Int J Environ Res Public Health*. 2022;19(9).
20. Gu TY, Yan ZZ, Jiang JH. Classifying Chinese medicine constitution using multimodal deep-learning model. *Chin J Integr Med*. 2022.
21. Jiang M, Lu C, Zhang C, Yang J, Tan Y, Lu A, et al. Syndrome differentiation in modern research of traditional Chinese medicine. *J Ethnopharmacol*. 2012;140(3):634–642.
22. Zhang GD, Chen Q, Tao TM, Yan ZY. Comparison of syndrome differentiation and treatment system between Huangdi Neijing and Treatise on Cold Damage. *Asian J Surg*. 2023.
23. Ding L, Zhang XY, Wu DY, Liu ML. Application of an extreme learning machine network with particle swarm optimization in syndrome classification of primary liver cancer. *J Integr Med*. 2021;19(5):395–407.
24. Huang Z, Miao J, Chen J, Zhong Y, Yang S, Ma Y, et al. A traditional Chinese medicine syndrome classification model based on cross-feature generation by convolution neural network: model development and validation. *JMIR Med Inform*. 2022;10(4):e29290.
25. Zhang H, Ni W, Li J, Zhang J. Artificial intelligence-based traditional chinese medicine assistive diagnostic system: validation study. *JMIR Med Inform*. 2020;8(6):e17608.
26. Vaswani A, Shazeer N, Parmar N, Jakob U, Llion J, Aidan N.G., et al. Attention is All you Need. 2017.
27. Devlin J, Chang M-W, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*. 2018 abs/1810.04805.
28. Cui Y, Che W, Liu T, Qin B, Yang Z. Pre-training with whole word masking for Chinese BERT. *IEEE/ACM Trans Audio Speech Lang Process*. 2021;29:3504–3514.
29. Zhou L, Liu S, Li C, Sun Y, Zhang Y, Li Y, et al. Natural language processing algorithms for normalizing expressions of synonymous symptoms in traditional Chinese medicine. *Evid-Based Complement Alternat Med: eCAM*. 2021;2021:6676607.
30. Ma Y, Sun Z, Zhang D, Feng Y. Traditional chinese medicine word representation model augmented with semantic and grammatical information. *Inf*. 2022;13(6):296.
31. Alshubaily I. TextCNN with Attention for Text Classification. *CoRR*. 2021 abs/2108.01921.
32. Zhang Y, Wallace B.C. A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. 2017.
33. Gong L, Ji R. What does a TextCNN learn? *CoRR*. 2018 abs/1801.06287.
34. Ma S, Liu J, Li W, Liu Y, Hui X, Qu P, et al. Machine learning in TCM with natural products and molecules: current status and future perspectives. *Chin Med*. 2023;18(1):43.
35. Dong X, Zheng Y, Shu Z, Chang K, Xia J, Zhu Q, et al. TCMPR: TCM Prescription Recommendation Based on Subnetwork Term Mapping and Deep Learning. *BioMed Res Int*. 2022;2022:4845726.
36. Kim TH, Kang JW, Lee MS. AI Chat bot - ChatGPT-4: a new opportunity and challenges in complementary and alternative medicine (CAM). *Integr Med Res*. 2023;12(3):100977.
37. Liu J, Dan W, Liu X, Zhong X, Chen C, He Q, et al. Development and validation of predictive model based on deep learning method for classification of dyslipidemia in Chinese medicine. *Health Inf Sci Syst*. 2023;11(1):21.
38. Li D, Hu J, Zhang L, Li L, Yin Q, Shi J, et al. Deep learning and machine intelligence: new computational modeling techniques for discovery of the combination rules and pharmacodynamic characteristics of Traditional Chinese Medicine. *Eur J Pharmacol*. 2022;933:175260.
39. Bao YF, Ding HK, Zhang ZH, Yang KH, Tran Q, Sun Q, et al. Intelligent acupuncture: data-driven revolution of traditional Chinese medicine. *Acupunct Herb Med*. 2023 September 12.