IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 28, 2021, accepted August 11, 2021, date of publication August 17, 2021, date of current version August 31, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3105588

Research on a Vehicle-Mounted Intelligent TCM Syndrome Differentiation System Based on Deep Belief Network

ZHAOHUI NING[1,2], YIHAO WANG[3], LIN WANG[3,4], LIJUAN WU[1], AND JIANQIANG SONG[3]

[1]The First Affiliated Hospital, Henan University of Science and Technology, Luoyang 471003, China
[2]The First Affiliated Hospital, Jiamusi University, Jiamusi 154007, China
[3]School of Information Engineering, Henan University of Science and Technology, Luoyang 471023, China
[4]Henan Qunzhi Information Technology Company Ltd., Luoyang 471003, China

Corresponding author: Lin Wang (linwang@haust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62002102 and Grant 62072121, in part by the Key Technologies Research and Development Program of Henan Province under Grant 202102210169 and Grant 212102210088, and in part by Luoyang Major Scientific and Technological Innovation Projects under Grant 2101017A.

ABSTRACT This article is oriented to the research of the vehicle-mounted intelligent traditional Chinese medicine (TCM) syndrome differentiation system to provide a quick and accurate diagnosis suggestion for smart and mobile medical vehicles. Nowadays, TCM modernization has attached more and more attention due to its remarkable clinical effect. The typical TCM diagnosis process is to induces the syndrome types of a patient from the four-diagnosis symptoms based on syndrome differentiation theory, that is, given a set of symptoms to treat, an overall syndrome representation is learned by fusing all the symptoms effectively to mimic how a doctor induce the syndromes. Therefore, we believe that an overall description of symptoms of a patient is very important for the follow-up treatment and should be handled carefully and properly. However, due to the complexity and diversity of syndromes, most recommended prescription of a patient lacks the explicit ground-truth of syndrome. Therefore, in this article, a new TCM syndrome differentiation method based on multi-label classification method and deep learning method was proposed to learn the implicit symptoms induction process in the real diagnosis. A Deep Belief Network (DBN) was used to reconstruct the TCM diagnosis model based on the Rrestricted Boltzman Machine (RBM) mechanism. Towards symptoms-syndrome groups learning, a symptom vector representing the symptom collection of a patient is constructed as the input of DBN, which was used to train the symptom data and labeled samples to build an unsupervised model at first. And then the parameters of DBN was adjusted gradually based on back propagation method. Secondly, binary classification algorithm was used to convert the multi-label classification problem into several corresponding labels. Each binary classifier model corresponds to a binary classifier to complete the classification from symptoms to corresponding syndrome types. Finally, Experiments were conducted on a self-built TCM clinical records database, demonstrating obvious improvements on the classification accuracy and convergence rate. Further studies should focus on the promotion of effectiveness of the proposed model to be more in line with the actual needs of onboard medical diagnosis and treatment.

INDEX TERMS DBN, TCM differentiation, deep learning, multi-label classification, onboard medical system.

I. INTRODUCTION

In the new century, with the rapid development of internet technology and the introduction of national policies and regulations related to Internet+medical and health care, mobile diagnosis and treatment has been more and more widely used in Chinese residents' daily health services, and intelligent medical vehicles have emerged at the right time. Intelligent mobile vehicles can assist doctors to collect blood pressure, body temperature and other vital signs information

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

to assist medical institutions to expand their service radius to the grass-roots organization such as community, village and nursing home and so on to strengthen the bottom construction of the medical big data of China. Therefore, intelligent medical vehicles is not only of great significance to the promotion of sense of contentment and happiness of Chinese residents but also one of the most important development directions. Traditional Chinese medicine, as the crystallization of the wisdom of the Chinese nation over thousands of years, has its unique charm [1]. It has been proved that the treatment concept of TCM is gradually accepted by the world and the demand of TCM is increasing. Therefore, TCM will be an indispensable part of the vehicle-mounted intelligent diagnosis and treatment system.

Nowadays, people are no longer satisfied with simple data applications such as storage, query, but how to use a large number of data to discover the hidden correlations in the data has been gradually becoming a hot spot for scholars today with the rapid development of computer technology today [2]. Solid theoretical knowledge of medicine and computer technology are the foundations of the modernization of TCM. In the clinical research of TCM, clinical data processing and diagnostic model have become the hot and difficult points in the study of TCM modernization. There are several reasons as follows. Firstly, the diagnosis data comes from complex sources and is stored in various forms, which brings great challenges to the management, application and visualization of TCM data. Secondly, the relationships between symptoms and syndrome types are very complicated and various. Therefore, it is very difficult to dig out more useful information from the massive clinical data of TCM without the help of peripheral technologies such as machine learning, data mining, multi-mode fusion and so on [3]. Making full use of modern information technology to build smart and mobile TCM system is of great significance to develop intelligent vehicle-mounted medicine to safeguard people's health and promote the scientific management level and national innovation ability [4].

In recent years, some emerging technologies have caught people's attention. The initial attempts of these technologies in the field of TCM have achieved great success, which effectively promotes the development of modernization of TCM and allows TCM to make more contributions to mankind [5]. The classic and clinical diagnosis and treatment experience stored in database and books can be converted into formatted standard data by big data analysis technology to provide a scientific expression method for the personalized and fragmented clinical records [6], [7]. Deep learning as a new type of artificial intelligence technology, is able to extract the essential features from massive data and find unknown and hidden relationships between various data and predict the development trend of samples by constructing deep learning models [8]. Deep learning has been used in many applications successfully in the field of western medicine and has been proved that it is able to provide strong technical support for these applications. Similarity, deep learning models are expected to have the same effect in the field of TCM. Deep learning technology imitates the way of human thinking, layers and categorize huge and complex data, extracts general laws from the messy original data, and then continues to refine by summarizing and advancing step by step. In this way, intelligent models can learn by themselves and become smarter and smarter by only learning the original data without pre-given theories and laws.

The vehicle-mounted TCM diagnosis and treatment system has coruscated a different vitality due to combining the academic system of TCM and the modern artificial intelligence technologies, whose service object has gradually changed from "patients" to "patients and sub-health people" and the service scope has extended from the hospital to communities and remote areas, enabling more people to access better health care at more affordable costs and thereby promoting the sustainable development of the Chinese health care industry. Compared with normal medical scenarios, there are fewer medical staff and more patients usually in mobile medical vehicles, therefore, the ability of the mobile medical equipment to handle a large number of task in a short period of time has become the most critical to vehicle-mounted diagnosis and treatment technology. In addition to being able to transmit medical data back to hospital in real time, the algorithms of data processing built into the devices must be efficient and fast to provide a quick and accurate diagnosis suggestion [9]. Deep learning brings new solutions to TCM diagnosis and treatment model, which will promote the smart and mobile TCM systems to be more intelligent than individual human, even a smart team [10]. Due to the powerful feature extraction capability of deep neural network, it is able to represent the complex and various diagnosis and treatment experience of TCM [11]. What's more, online learning methods make the useful knowledge increase as the algorithms run [12], that is, on the basis of learning traditional Chinese physicians' experience, the constructed deep TCM syndrome differentiation model can continue to explore conduct various prescriptions that have tried in the past and predict the effectiveness of herbs in advance. For this reason, it can be foreseeable that the TCM diagnosis and treatment machines will be smarter than human individuals and teams in the near future [10].

Since entering the 21st century, with the advancement of the Internet and information technology, the vehicle-mounted medical system has ushered in new development opportunities. The mobile medical diagnosis and treatment system based on artificial intelligence has attracted the participation of many technology companies [13]. Internet giants such as Google, Amazon, Microsoft, IBM, etc. have all made great achievements in smart and mobile healthcare, which has been greatly benefited human society. For example, the breast cancer diagnosis expert treatment system designed by Google, the diagnostic accuracy in analyzing slices is even higher than that of experienced pathologists [14]. "Doctor Watson" is an auxiliary diagnosis and treatment system designed and developed by IBM, which uses artificial intelligence technology to simulate the treatment ideas of the world's top oncologists and is capable of providing professional treatment suggestions almost the same as those of experts.

The diagnosis and treatment of smart and mobile TCM has also achieved great advantage by simulating the process of TCM syndrome differentiation and treatment though fusing the existing artificial intelligent technologies such as deep learning, data mining, text analysis and so on. For example, Ji *et al.* used semantic analysis technology to discover the internal connection between symptoms in TCM medical records, and proposed underlying pathogenesis hypotheses in 2017 [15]. Li *et al.* analyzed the mode of thinking and therapeutic process of TCM, constructed a deep model to learn the relationships between symptoms and syndrome types, and then syndrome types and prescriptions, forming a whole therapeutic process in 2018 [16]. It means that deep learning algorithms can probably be combined with TCM theories to develop TCM clinical dialectics systems suitable for in-vehicle medicine.

In general, the intelligent TCM syndrome differentiation method in this article had added the process of data processing, and combines with the onboard vehicle-mounted medical system to further improve the level of mobile medical diagnosis and treatment [17]. Compared with western medicine, there is less research on vehicle-mounted intelligent TCM assisted diagnosis and treatment system [18]. This article has analyzed and processed the TCM medical records from classical books and real clinical data, carried out standardized storage and established a TCM syndrome differentiation model though suitable deep learning algorithms. The experiments have proved the proposed method is able to give a higher accuracy rate, which is of great significance to the development of the vehicle-mounted intelligent medical system [19]. The rest of this article is organized as follows. Section II surveys the related works about the proposed method. Section III describes the overall framework and the proposed method. Section IV presents the experiments and evaluates the results. Finally, section V provides some concluding remarks.

II. RELATED WORK

A. DEEP BELIEF NETWORK

Deep learning is a more complex machine learning method and is a new research field. Its appearance has accelerated the development of machine learning and has attracted more and more attention from experts, scholars, and technology companies. Deep learning models also play an important role in many fields such as speech recognition, image recognition, biomedical information processing, abnormal data detection, and smart medical care. After long-term research on traditional neural networks, people have developed the concept of deep learning. A good example of deep learning model is a multi-layer perceptrons with multiple hidden layers. For neural networks, depth often refers to the total number of non-linear combinations in the function obtained during the learning process. A neural network usually consists of a

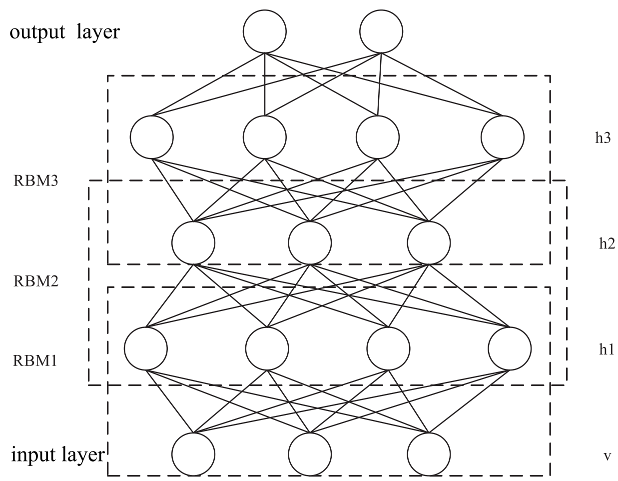


FIGURE 1. Sketch of DBN structure diagram.

hidden layer, an input layer, and an output layer, while a deep neural network has 3 hidden layers, 1 input layer, and 1 output layer, such as a DBN [20].

The concept of DBN was first proposed by Geoffery Hinton in 2006. As one of neural network models, it can be used for unsupervised learning and supervised learning. The purpose of the former is to retain the original features of data as possible while reducing the dimensions; the later aims to minimize the accuracy rate of classification [21]. However, regardless of unsupervised and supervised learnings, the essence of DBN is the process of feature learning to obtain better feature expression. DBN is developed from the RBM and consists of several layers of neurons (as shown in Figure 1). Compared with other traditional network models, DBN is a probability generation model which is more suitable to represent the joint distribution of sample category labels and data [22].

RBM is a typical deep network model that learns the probability distribution of the input data set, its basic network structure consists of visible layer v and a hidden layer h. The explicit nodes are used to accept input data and implicit nodes are used to extract features [23]. Each layer of RBM nodes can be divided into an active state represented by 1 and an inactive state represented by 0 [24] respectively. In addition, RBM also expresses the data as a probabilistic model through learning. Once the model is trained or converges to a stable state through unsupervised learning, it can also be used to generate new data. The above two features make the layer-by-layer training of DBN effective and the training data at the later layer representative by extracting features from the preceding hidden layer, which can be used to solve the problem of insufficient training samples by generating new data. A DBN model consists of several RBMs stacked, and the training process is carried out layer-by-layer from low to high. That is, DBN is composed of multiple layers of dominant and invisible neurons and the invisible neurons are called feature detectors. The bottom layer of the DBN represents the loudness of the data, and each neuron can represent a dimension of the data vector. Pre-training and fine-tuning are the two stages of DBN training [25].

- *Unsupervised Pre-Training Stage:* At this stage, DBN uses the greedy layer-by-layer learning algorithm based on RBM mechanism to generate its model. At first, the bottom RBM is trained with raw input data and takes the features extracted from the bottom RBM as the input of the top RBN. Secondly, repeat the first process to train as many RBM layers as possible. The DBN uses the result of layer-by-layer accumulation of RBM to reconstruct the original features to performs the initial feature vector fitting to achieve gradual adjustment of DBN parameters. At this stage, it is assumed that two vectors are utilized to describe the active states in the visible layer and hidden layer respectively, then visible layer vector can be mapped to the hidden layer through RBM reconstruction and then the mapped vector will be regarded as a new input of the next RBM. Eventually, a completely DBN model that represents the probability distribution of raw data will be obtained. In this article, Gibbs sampling method is used to verify the DBN pre-training based on the above specifications.
- *Supervised Fine-Tuning Stage:* DBN can obtain the optimal solution of each RBM layer in the pre-training stage, but these optimal solutions are only local optimal solutions and cannot represent the globally optimal solution for the entire DBN model structure. Fine-tuning based on Back Propagation (BP) algorithm. At the last top layer of DBN. BP network is set to receive the output feature vector of RBM as its input feature vector and then the entity relationship classifier is supervised to be trained. Therefore, the back propagation network also propagates the error information to each layer of RBM from top to bottom to optimize the whole DBN network.

The whole training process of DBN network can be regarded as the initialization of weight parameters of a deep BP network, so that DBN can overcome the two disadvantages of BP network that is easy to fall into local optimal and has long training time due to random initialization of weight parameters.

B. MULTI-LABEL CLASSIFICATION

The problem of multi-label classification is common in daily life, such as a movie can be classified into action and crime at the same time, a piece of news can be classified into politics and law at the same time. there are also problems in the prediction of gene function in biology, scene recognition, disease diagnosis and so on. In the early days, single-label classification algorithms in supervised learning was used to solve this kind of problems due to the simple structure and small amount of data. However, with the increase of the data volume and structure, new methods for multi-label classification are needed to be sought [26]. At the current moment, there are still several difficulties. First of all, the number of class labels is uncertain. Some samples may have only one class label, while others may have dozens or even hundreds of class labels. what's more, there are dependencies between categories. For example, the sample containing the blue sky class standard mostly contains white clouds. How to solve the dependency problem among the categories is also a big difficulty.

Multi-label classification studies a sample and a set of corresponding labels. Assuming that $\chi = \Re^d$ represents d-dimensional sample space, $\mathcal{Y} = \{y_1, y_2, \cdots, y_q\}$ represents all the possible labels corresponding the task samples. The task of the multi-label classification is to learn a function $h : \chi \rightarrow 2^Y$ from the training set $\mathcal{D} = \{(x_i, Y_i)| \leqslant i \leqslant m\}$. Each $x_i \in \chi$ is a d-dimensional vector and $Y_i \subseteq (Y)$ is a label set for training set. For a test sample $x \in \chi$, the multi-label classifier will predict a label $h(x) \subseteq (Y)$ as label set of the test sample x. At present, there are many multi-label learning algorithms which can be divided into two categories: one is the method based on problem transformation, the other is the method based on algorithm applicability. The former is to transform the data to use the existing algorithm. The later is to extend a particular algorithm, so as to be able to deal with multiple marked data to adapt to the application data.

- *Problem Transformation Methods*

When the label y_i is directly dependent on each other, the most commonly used problem transformation method is the Binary Relevance (BR) algorithm [27], [28]. The effect of this method is better than other classification methods and it is easier to understand. The BR algorithm can convert the multi-label classification problem into several relatively independent binary classification problems. For an unclassified sample instance, it can learn each label and output all the values of the classification results for each label. When Considering the dependencies between labels, the generalization ability of the final model constructed by the Classifier Chains (CC) method is better than that of the model built by BR method [29]. The core idea of CC method is to decompose the multi-label classification problem and transform it into a binary Classifier chain, in which the construction formula of the binary Classifier behind the chain is based on the prediction results of the previous binary Classifier. During model construction, shuffle the order of labels at first and then build the model corresponding to each label from the beginning to end. Suppose a certain classifier A_j has its corresponding label represented by λ_j. When sample x_i is classified on the first sample, it means that the value l cannot be 0, and the output result is $Pr(\lambda_l \mid x\lambda_{j+1})$. The classifier will judge the labels in turn, and use the predicted value of the former x_i as the input prediction $Pr(\lambda_l \mid x\lambda_{j+1})$. When discriminating the last label, the classifier will take all the previous discriminant prediction results into consideration, and finally output the prediction considering all the results [30].

- *Algorithm Adaptation Methods*

These kinds of methods directly apply the existing single-label algorithm to multi-label problem. For example, the core idea of k-Nearest Neighbor (KNN) algorithm is to obtain the k closest instances of the learning instance and then the label set of these instances is used to predict the value of the forecast label set of the learning instance by the Maximum Posteriori Probability (MAP) method. The emergence of the random k-label algorithm solves the Linear Programming (LP) algorithm requiring a large number of labels, and avoids the problem of over-complexity when training the classifier model. In order to obtain and improve the computational efficiency, the PAKEL algorithm uses the LP algorithm to train samples on a random k-label subset of the multi-label classifier and uses an algorithm based on ensemble learning to help it have a sufficient prediction result.

Assuming that a certain label set is represented by Y, the possible second label subset of Y is represented by k^l, and the fourth q label subset is represented by $Y^l(q) \in Y$, and the training data is as follows:

$$D_{Y^l(q)} = \left\{(x_i, \sigma_{Y^l(q)}(y_i \bigcap Y^k(q)))1 \leq i \leq n\right\}. \tag{1}$$

The multi-label classifier of the training data set is:

$$h_D : x \to \tau(D_{Y^l(q)}). \tag{2}$$

The random k-label algorithm uses the LP algorithm on $Y^l(q)$ to obtain the frame of the classification label, and the total number of the full set of labels is represented by C. The samples that need to be predicted in each type of label can be calculated by the following formula:

$$\tau(x, k_j) = \sum_{r=1}^{m}[k_j \in Y^l(q)]. \tag{3}$$

The above formula calculates the actual value of the label under the integration framework. The predicted label subset requires the following formula calculation:

$$y = \left\{k_j \mid \frac{\mu(x, k_j)}{\tau(x, k_j)} > 0.5, 1 \leq j \leq c\right\}. \tag{4}$$

Even the PAKEL algorithm fails to fully recognize the correlation between multi-label and samples. It establishes a classifier corresponding to each different label, which improves the overall training ability of the complex multi-label classification model effectively.

III. TCM SYNDROME DIFFERENTIATION MULTI-LABEL CLASSIFICATION ALGORITHM BASED ON DBN

The clinical data needs to be cleaned to establish a data warehouse before the construction of multi-label TCM syndrome differentiation model though deep learning algorithm based on DBN. The whole process of TCM syndrome differentiation model construction is shown in Figure.2. The TCM clinical syndrome differentiation algorithm refers to establishing a simulated the process of TCM diagnosis, that is summarizing the syndrome types according to the symptoms of the patient. This article proposed a TCM clinical syndrome differentiation based on DBN to meet the rapid demand of vehicle-mounted diagnosis and treatment. A DBN based on RBM mechanism is used to train several classifiers, and each classifier is trained for each type label. When DBN-trained classifier is used to predict the class label for a new symptom instance, the value of all classification models will be output to avoid the lack of syndrome types.

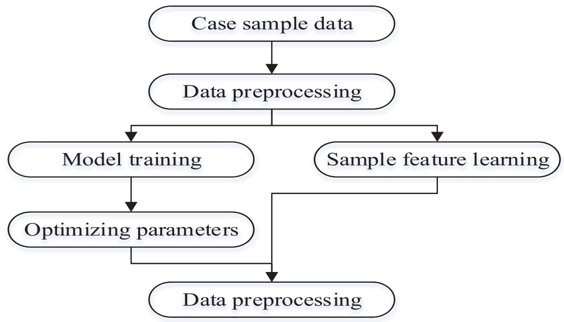


FIGURE 2. Flowchart of TCM syndrome differentiation model construction.

A. CONSTRUCTION OF DBN MODEL BASED ON RBM

Existing TCM clinical syndrome differentiation models need to rely on artificially set parameters to train case samples. Inaccurate parameters set usually leads to excessively long training time and failure to converge. Therefore, the existing eight-principal syndrome differentiation model of TCM requires researchers to have very rich experience to construct a parameter set with high rationality for the problem being solved. Therefore, the existing neural network models for TCM eight-principal syndrome differentiation model are poorly able to learn the characteristics of clinical cases, which is far from practical clinical application.

In this article, DBN based on RBM mechanism is firstly used in TCM clinical syndrome differentiation model to reconstruct the feature vector of the case sample continuously. Gradient descent algorithm is used in this article to avoid the inaccuracy of setting parameters manually. By this way, it is ensured that all network parameters are optimal in the local interval. Finally, the BP neural network in the supervised learning mechanism is used to reversely train the original feature data to ensure that the optimal network parameters are obtained, which are also global parameters.

Suppose the symptoms set of a group of cases is represented by $\mathcal{D} = \{x_1, x_2, \ldots, x_n\}$, and each symptom has a corresponding set of labels $\mathcal{Y} = \{Y_1, Y_2, \ldots, Y_n\}$, which represents the hidden layer feature vector under RBM. It is known that a RBM consists of a hidden layer and a visible layer. The input symptom set $\mathcal{D}$ is reconstructed and the output $\mathcal{D}'$ is calculated. $x' \in \mathcal{D}'$ at $h_j = 1$ is calculated by the following formula:

$$p(h_j = 1) = \sigma(\sum_{i=1}^{m} w_{ij} \cdot v_i + a_i). \tag{5}$$

$$\sigma = \frac{1}{1 + e^{-x}}. \tag{6}$$

$$p(h_i = 1 \mid v) = \sigma(\sum_{i=1}^{m} w_{ij} \cdot v_i + b_i). \tag{7}$$

where w_{ij} represents the connection strength between the hidden layer node i and the visible layer node j, a_i and b_j are the parameters of RBM model, v_i and h_j are the $i-th$ and $j-th$ node of the visible layer and hidden layer respectively. m is the number of input nodes from visible layer into the hidden node h_j. RBM needs to reverse mechanism to construct the original feature vector and the process of solving the feature vector is as follows:

- Solve for $p(h_j = 1 \mid v)$ according to the symptom sample x;
- Get the random number $s \in (0, 1)$. Every time when $s > p(h_j = 1 \mid h)$ appears, $x' = 0$, otherwise it is 1;
- Repeat the above steps to get the output value continuously until the end.

RBM is essentially an energy model and its superposition process is calculating link weights between the visible and hidden layers. It continuously reconstructs the feature vector to adjust the parameters by itself to make RBM to superimpose continuously. Finally, the global optimal weights of the model can be obtained. The hidden layer vector and the sample input vector can be calculated to get the energy function result. The energy function describes the state measurement of the RBM model. The smaller the function value, the less energy the system state will be more stable. The calculation formula is as follows:

$$E(v, h) = -\sum_{i \in v} a_i v_i - \sum_{j \in h} b_j h_j - \sum_{i \in v, j \in h} w_{ij} v_i h_j. \quad (8)$$

Indicates that a_i and b_j denote the offset parameters of the RBM in the visible layer and the hidden layer, which describe the corresponding neurons weights respectively. A scalar and various combinations of variables are linked together through an energy model, and continuous changes of the scalar are training in the model. The joint probability distribution calculation formula is as follows:

$$p(v, h) = \frac{1}{Z} exp(-E(v, h)). \quad (9)$$

$$Z = \sum_{v,h} exp(-E(v, h)). \quad (10)$$

Z is the normalization factor. In this article, Z is set to 10 for calculating the formula by experimental verification. The node state in the visible layer is determined by the joint probability distribution $P(v, h)$. The calculation formula for the state of the hidden layer node is as follows:

$$P(v_i \mid h) = \sigma(\sum_{j \in h} w_{ij} + a_i). \quad (11)$$

$$P(h_j \mid h) = \sigma(\sum_{j \in h} w_{ij} + b_j). \quad (12)$$

The connection weight between the visible layer and the hidden layer in the model is represented by w_{ij}. The network scale of the trained model can restore the test data. The larger the distribution probability of $P(v)$, the better the restoration effect. The calculation formula of distribution probability is as follows:

$$P(v) = \sum_{h} P(v, h) = \frac{1}{Z} exp(-E(v, h)). \quad (13)$$

The RBM training belongs to an unsupervised learning model, as a result, the training data itself is not labeled at the end of learning process. Therefore, in order to ensure the accuracy of the training results, it is necessary to continuously adjust the probability distribution of v, which is the input data and its likelihood function $\theta \{w, a, b\}$ has been continuously adjusted to ensure that the training data can achieve the maximum possible probability under the condition of parameters unanimous. The formula for calculation θ is as follows:

$$\theta' = \theta + \eta \frac{\partial lnP(v)}{\partial \theta}. \quad (14)$$

Derivation of parameter θ of the likelihood function can calculate the maximum probability, and then determines the value of parameter θ with the maximum probability.

The training mechanism of RBM is a model generated by unsupervised learning. The original feature vector in each network node is input from the visible layer of the first layer. The model can obtain the feature vector of the case sample from the visible layer. After the feature vector is reconstructed, another form of the original feature vector can be transmitted to the hidden layer. After the feature vector enters the hidden layer, its mapping will be sent to the input layer again. The input layer will reconstruct the feature vector once again. The above process is repeated continuously, and the weight is updated according to the difference from the input unit to the hidden layer. In the process of RBM accumulation, the input original disease set is continuously reconstructed, and the network parameters are continuously updated until the predetermined number of trainings is reached. DBN can obtain the weights among its various neurons through the above-mentioned RBM training process. Although the above-mentioned RBM generation training only continuously updates the local optimal solution rather than the global optimal solution for the entire network, the local optimal solution is often close to the overall optimal solution. DBN network structure usually set as a BP neural network as the last output layer which to give the final classification results, which will be compared with the real label data to find the difference and to slightly adjust the neurons weights in each layer to find the final global optimal solution for the entire DBN model.

The RBM energy model between the visible layer and hidden layer can be represented by formulas Eq. 8. The input sample data will be stored in the input layer at first and the state of the hidden layer is calculated. Secondly, the neuron state from the input layer to the visible layer is calculated to ensure that the original input case data sample can fit the output feature vector to the greatest extent to obtain the final local optimal parameters. The whole process is described as Algorithm 1. In this algorithm, the input original case data is

Algorithm 1 Feature Learning Algorithm of Case Samples Based on RBM

Input: Training data set $\mathcal{D}$, Label set Y, Parameter space $\theta\{w, a, b\}$, η
Output: Parameter space θ
bengin
The sample data set $\mathcal{D}$ is normalized as [0, 1] to obtain the standard vector $\mathcal{D}'$
Initialize the input layer node v
input $\mathcal{D}'$
for $i = 1$ to Q **do**
  for $j = 1$ to m **do**
    According to the $\mathcal{D}$ get $P(v_i = 1 \mid h)$
    Calculate the hidden state of the i-th layer
    if $z > p(v_i = 1 \mid h)$ **then**
      $x_i' = 0$
      output $\mathcal{D}'$
      To calculate $p(v_i = 1 \mid h)$
    else
      $x_i' = 1$
    end if
    if $z > p(v_i = 1 \mid h)$ **then**
      $x_i = 1$
    else
      $x_i = 0$
    end if
  end for
  Use Equation to update the parameters θ
end for
Return θ

cleaned at first to obtain the feature vector of the symptom set and then the visible layer node in the RBM mechanism is initialized by using the processed symptoms data set to obtain the hidden layer nodes and the number of their feature vectors. The input of each lower layer nodes is a hidden layer nodes with its upper layer nodes, the feature vector of this layer is obtained and the hidden layer nodes of this layer is constructed, and eventually, parameter θ can be updated to achieve the final global optimal value.

B. TCM CLINICAL SYNDROME DIFFERENTIATION ALGORITHM

The TCM clinical syndrome differentiation algorithm simulates the process of TCM consultation by establishing a mathematical model, and the patient's syndrome type is derived from the symptoms through clinical differentiation. Through the research on the theoretical knowledge of Chinese medicine and machine learning knowledge, it is found that the process of Chinese medicine differentiation can be solved by classification methods in information technology. By using the DBN, the TCM clinical syndrome differentiation algorithm establishes the non-linear and complex connection between the types of symptoms and syndrome in TCM. Existing auxiliary consultation systems often fail to consider the relationship between symptoms and syndrome types, resulting in omissions when the system infers syndrome types based on the patient's symptoms. By adopting the label classification method in this paper, the proposed TCM clinical syndrome differentiation algorithm can solve the problem of the complex connection between the types of symptoms and syndrome in TCM.

In this paper, the proposed algorithm firstly utilizes the binary association classification to convert the multiple label classification problem into the multiple corresponding binary classifiers, converts the multi-label into a single-label, and combines the DBN for each binary classifier. The classification model obtained from syndrome training solves the complex relationship between symptoms and syndrome types. Every time for a new symptom instance to be classified, the classifier is used to judge it one by one, and the predicted result is output. The outputs of all the classifiers will form a set, and this set is the final output result. The algorithm can effectively solve the problem of the relationship between symptoms and syndrome review.

The two categories of algorithm adaptation and problem transformation are basically two different categories in the multi-label classification method in which the problem transformation is the most common used method. The binary association classification algorithm is often divided into N classifiers $H_n : x \to \{1, -1\}$ for the multi-label problem to be solved. Each label has a corresponding classifier to process it, and each label has a specially trained classification model for it. The binary association classification algorithm outputs the predicted labels of all the classifiers for the new sample to be classified.

Each instance of the treatment test has its label attribute and characteristic attribute. The multi-label classification algorithm makes predictions based on this, and the final prediction is actually the one that has the highest value in the candidate label set.

$$Y_x = \left\{y_1, y_2, \ldots, y_q\right\} = \max_{y\in\{0,1\}} argP(y_1, y_2, \ldots, y_q \mid x) \tag{15}$$

The DBN-based TCM clinical syndrome differentiation algorithm first uses the mechanism of the RBM to process and train the existing TCM case data. In order to obtain the best parameters of the proposed model, we continuously fit and train the feature vector, then save the model. Use the trained model of the DBN as the classifier for each label, and convert the multi-label problem into a single-label problem. Secondly, use the trained classifier to determine whether a single label is the label of a sample which needs to be predicted, finally determine all the outputs, the whole progress is described as Algorithm 2.

In the algorithm, the number of samples is equivalent to the number of nodes in the input layer of RBM, which is m, and n represents the dimension of the feature vector. The standardization of the input data is completed in the first two steps of the algorithm, and then the connection weight of the

RBM is calculated w, the offset between the hidden layer and the visible layer, and the hidden layer can be obtained by using the conditional distribution probability. The node state of the visual layer is also calculated using the conditional distribution probability, and then the offset parameter and weight parameter of the RBM are continuously updated and saved. The prediction of the sample label is completed at the end. Then return the unclassified samples to the label set.

Algorithm 2 Feature Learning Algorithm of Case Samples Based on RBM

Input: Training data set $\mathcal{D}$, Labels $\mathcal{Y}$, Test sample x, learning rate η, Tag number Q. Iterations T, n
Output: Label y

```
Input training data set D
The standard treatment D get the standard vector x_0
Initialize the connection weights w by using the normal distribution N(0, 0.001)
Initialize the offset b = 0
Initialize the offset a
for i = 1 to T do
    for i = 1 to m do
        Compute the probability in the i-th hidden layer
        Compute P(h_ij {0, 1} by conditional distribution P(h_ij | v_i)
    end for
    for j = 1 to n do
        Compute the probability in the 2i-th hidden layer
        From conditional distribution P(v_2i | h_j) get P(v_2i {0, 1}
    end for
    for i = 1 to n, j = 1 to m do
        Update Learning Rate η, a, b
    end for
end for
save the parameters w, a, b, sample and label x, y
for i = 1 to q do
    Calculate the binary classification results t
    Add the tag to the collection y
end for
Return θ
```

On the basis of DBN, the time complexity of standardizing the input data in the TCM clinical differentiation algorithm is $\mathcal{O}(n)$. The clinical model of TCM can be roughly divided into four steps. The first step is to calculate the state value of the neuron in the hidden layer using with the conditional distribution probability. It can be seen that the state of the layer node is calculated according to the dimension of the input feature vector. This process needs to be performed m times. Calculation, so $\mathcal{O}(n^m)$ is its time complexity; Next is the calculation of the neuron state of the visible layer, which can be obtained from the neuron state in the hidden layer calculated in the first step, thus the time complexity can be represented by $\mathcal{O}(m^n)$; The updating of the parameters is the third step, and its time complexity is $\mathcal{O}(mn)$; The above three steps need to be executed continuously. To complete all the calculations, the algorithm needs to be executed to the maximum number T. Hence the time complexity is $\mathcal{O}(n^m + m^n)\mathrm{T} + \mathcal{O}(mn)\mathrm{T}$. Finally, the test sample needs to be predicted, and each label is performed once, which requires q operations. Therefore, the time complexity of the overall model is $\mathcal{O}(n^m + m^n)\mathrm{T} + \mathcal{O}(q^n)$.

IV. EXPERIMENT

The experiment uses the open-source deep learning framework TensorFlow to build the DBN model. Tensor data structure is used in the experiment for data transfer and storage. The stochastic gradient descent (SGD) function is used in data training process to achieve convergence of algorithm. Dropout tool is also used to avoid over-fitting imagination during the experiment by setting a drop threshold. At the end, the same test sample set and training sample set are applied to three different multi-label classification algorithms to ensure the persuasiveness of the experimental results.

The experimental clinical data mainly come from Henan Sisheng Chinese Medicine Research Institute (河南四圣中医药研究院) and Guangzhou Xiangxue Internet Hospital(广州香雪互联网医院). Each case in the standardized database provides symptoms of TCM four diagnosis and corresponding syndromes, therapeutic principles and methods, prescriptions and herbs. The database is cleaned and formalized and eventually a total of real 12,054 medical cases are obtained as usable symptoms-syndrome groups, which covers most of TCM symptoms and syndromes. What's more, in order to enhance the generalization ability of the model, the distribution of TCM syndrome types contained in the database is basically consistent with the uniform distribution. In the process of formalization of above database, a standard symptom knowledge database based on TCM four diagnosis was constructed, including tongue, face, eyes, body, limbs, pulse, meridians and so on. Each category contains several types of characteristics. For example, the category of tongue can be divided into 5 types such as shape, color, state, moss texture, moss color and each type is classified into several levels according to the actual TCM clinical experience. The symptom sets for each case was converted into a formative input vector according to the symptoms and their priority. To make the experiment results more robust, we randomly split the whole data into parts, the training data is about 70% and test data is about 30%.

When using the RBM mechanism to learn case sample data, it is necessary to initialize the offset and weight of the RBM to determine the number of hidden layer units and the learning rate. The weight matrix and the offsets in the hidden layer and the visible layer can be obtained in the training phase. But their initial values need to be set before the model is trained. The initial values are generally given randomly. The experiment selected random numbers in $N \in (0, 0.01)$ of the normal distribution as the weight vector. In order for the model to have a higher generalization ability, the symmetry between different neurons will be destroyed. It can be seen that the layer offset is initialized according to the formula, and the initial value of the offset of the hidden layer is 0, so that the correct result of the output edge can be easily obtained. The offset parameter of the first visible layer is represented by i. For the a_i feature in the training sample, the proportion of the sample in the active state is represented by p_i.

$$a_i = \log \frac{p_i}{1 - p_i} \tag{16}$$

The weight of case samples will change during the learning process of the model, and the amount of change each time is the learning rate. Usually the value of the learning rate is in a fixed range. When the learning rate is set too large, the phenomenon of rapid and unstable convergence will occur, and finally the model will miss the optimal weight. Setting the learning rate too low will cause the convergence time to be too long and the training speed to be slow. The experiment balances the problem of learning rate setting value by adding state learning rate (moment) to ensure that the optimal weight obtained by the model can be closer to the weight parameter. In the training phase, the adjustment gradient is obtained by integrated with the last adjustment gradient, and the basic idea of the state learning rate is a dynamic learning rate multiplying with the last adjustment gradient. Such that the proposed model can effectively converge faster.

DBN has powerful learning capabilities. It belongs to a relatively deep network structure. It is easy to learn abnormal data and noise generated by sample data during the samples training process, which will make the model appear to be over-fitting. The over-fitting problem will lead to insufficient model generalization performance and inaccurate classification of the sample prediction in the later stage of the model prediction. Deep learning models usually train enough data to avoid this kind of imagination. The effective data collected in actual case data collection is often limited, and the number of samples is small. Sample expansion is carried out according to the distribution law of sample data to ensure that the expanded data is true and effective, and can meet the training requirements of DBNs. Studies have found that the contribution of symptoms to each of its related syndrome types is different. Some symptoms contribute less to the syndrome type, and some symptoms contribute more. If the symptom contribution value is greater than 0.015 but less than 0.02, each symptom sample which corresponding to the syndrome type is distributed on the basis of the power ratio. Verify the existing data according to the power ratio distribution.

Suppose the specific function of the sample data subject to the power rate distribution is:

$$y = ax^{-b} \tag{17}$$

The setting syndrome data is represented by y, the symptom data is represented by x, and the parameter is a,b. Take the logarithm to the base 10 of the formula:

$$\log_{10} y = \log_{10} ax^{-b} \tag{18}$$

Let $y = \log_{10} y$; $x = \log_{10} x$, $\log_{10} a$ be the parameters, and the formula is converted to:

$$y = c - bx \tag{19}$$

Take the logarithm of the data symptom and syndrome type of the sample to verify whether it is a linear equation. On the basis of the secondary energy, we can infer that the sample data follow a power law distribution. We use Matlab to fit the distribution of the sample data, and take the logarithm with base 10 for the sample data. The sample data can better conform to the power rate distribution. The sample data is expanded according to the power probability distribution rule to generate data according to the rule. Select 30% of the sample data as the test sample, and the remaining 70% of the sample data as the training data. Since sample data is prone to noise and other problems, they need to be standardized.

The different attributes of the symptoms in the sample data are quite different, and the units of each element are also different. Therefore, it is necessary to use the min-max normalization algorithm for the unified standardization of the collected original symptom data. The formula is:

$$y = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{20}$$

The original data is represented by x, the maximum value of the sample data is represented by x_{max}, x_{min} is the minimum value of the sample data, and the normalized value is represented by y. Often the distribution of symptoms in patients is scattered and uneven, and some of the symptoms are missing in the collected clinical data sets. The experiment uses NaN to supplement the missing values.

For the training of the weights and offsets between the hidden layer and the visible layer, the computational efficiency of the method of training each sample data one by one will be very low. If the data is large and high-dimensional, the training time will be too long. In practice, samples are often divided into small pieces for calculation, and only one small piece is calculated at a time. This can give play to the convenience of matrix operations and the efficiency of CPU operations. During the experiment, multiple pieces will be trained at the same time. In order to ensure a stable learning speed, the average gradient is selected when the parameters are updated. The formula for calculating the average gradient is as follows:

$$\Delta\theta = (\frac{1}{N_{batch}})^N \sum_{i=1}^{T} \frac{\partial \log P(v_i \mid \theta)}{\partial \theta} \tag{21}$$

Among them $T = batcht^t$, the data length of mini-batchers is N_{batch}. For single-sample learning $N_{batch} = 1$, N_{batch} of them cannot be too large, because too large will lead to low gradient sensitivity, which will cause the model to miss the optimal value.

In this paper, maximum entropy algorithm and Bayesian algorithm have been adopted for comparison to demonstrate the performance of the proposed TCM clinical syndrome differentiation method. Because as the common machine

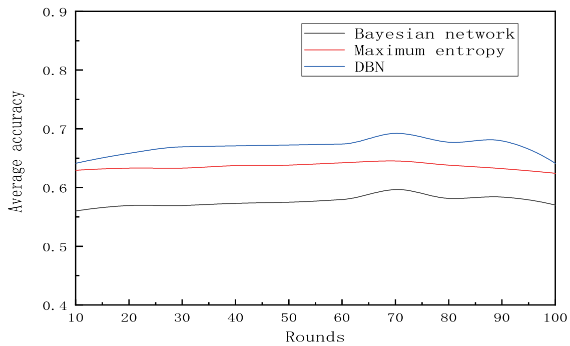


FIGURE 3. Average accuracy.

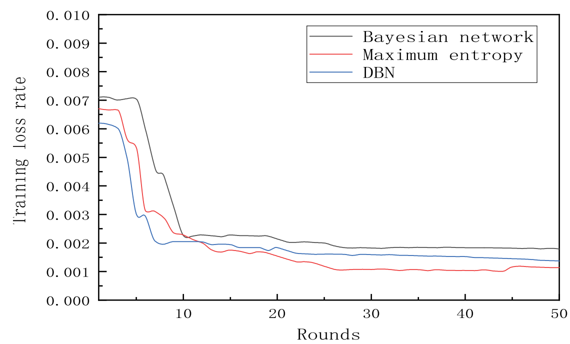


FIGURE 4. Training loss rate.

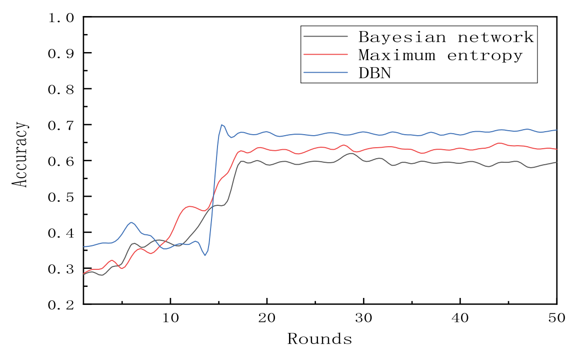


FIGURE 5. Accuracy.

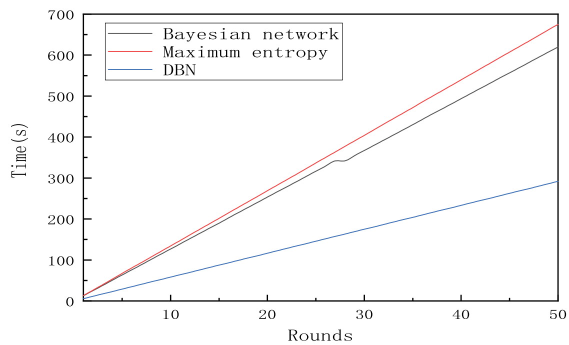


FIGURE 6. Time.

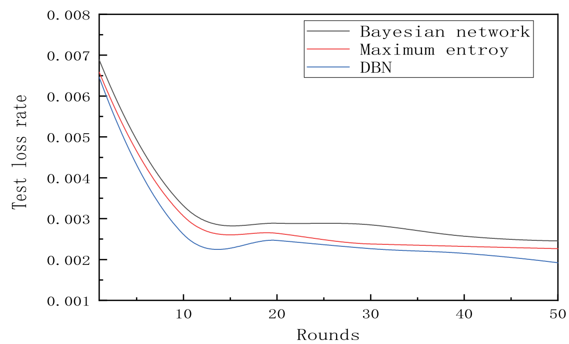


FIGURE 7. Test loss rate.

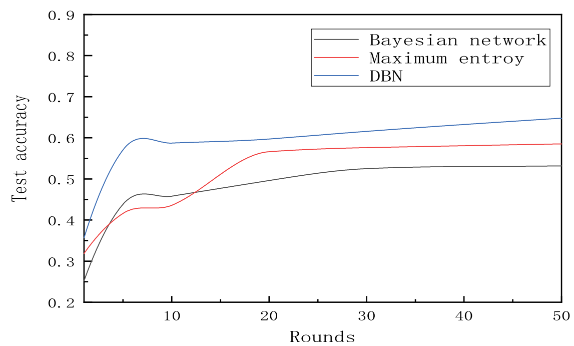


FIGURE 8. Test accuracy.

learning algorithms, the two algorithms are both the criteria to make the statistical properties of selected random variables most in line with the objective situation of TCM data. And they are very suitable to deal with the high-dimensional models with small correlation between different dimensions. When the number of nerves is about 70, the overall performance of the DBN is the best, as shown in Figure 3.

In order to objectively evaluate the performance of the DBN based TCM clinical syndrome differentiation algorithm, we select the commonly used criteria including the accuracy rate, the loss rate, and the training time per round of three indicators, which can evaluate the effect of the algorithm. As shown in Figure 4 to Figure 8. From the Figures 4 to 6, we can clearly see that the DBN based TCM clinical syndrome differentiation algorithm is superior to the compared classification algorithms in terms of the training time and the accuracy. The DBN-based algorithm has just started to look at the loss rate of iterative training as a whole. Compared with the previous two, the loss value of the iterative training is slightly better. The entropy algorithm can set constraints more flexibly to reduce the degree of data fit and the loss; in the accuracy rate, the DBN-based algorithm to start the training phase, the local performance may be inferior to the maximum entropy algorithm and the shell due to constant search for parameters to obtain better accuracy. The Bayesian algorithm is stable and better than the comparison algorithm in the later stage; in the training time of each round, it can be seen that the training time of each round of this algorithm is significantly better than the two comparison algorithms. What's more, from the above Figures 4 and 5, the overall trend is getting better for the loss rate and accuracy with the increasing of training rounds, however, the throughput lines have been slightly fluctuating all the time. The reason for this phenomenon is possibly that the SGD method used in DBN training does not work well when gradient slope is not continuous near the minimum value. leading the algorithm to jump around the minimum points.

From the experimental Figure 7 to the Figure 8, the DBN-based TCM clinical syndrome differentiation algorithm is better than the comparison algorithm in the test loss rate and the test accuracy rate. The maximum entropy algorithm can better meet the constraints, and the accuracy rate is higher than that of Bayeux Si algorithm. The performance of these three indicators can fully explain that the performance of the TCM clinical syndrome differentiation model based on deep confidence network proposed in this paper is better than that of the maximum entropy algorithm and the Bayesian algorithm.

V. CONCLUSION

This article introduces a TCM clinical syndrome differentiation algorithm which is applied to the vehicle-mounted intelligent medical system to provide a fast and accurate diagnosis suggestion. The algorithm firstly uses the RBM mechanism to reconstruct the feature vector of the training clinical samples. In the reconstruction process, the gradient descent method is used to update the network parameters to ensure that the local network structure is optimal, and the parameters of DBN network can also be fitted to the original data to avoid the inaccuracy of manually parameter setting. In the last layer, the network parameters are updated to obtain the global optimal parameters of the whole network. Secondly, the BR algorithm is used to convert the multi-label classification problem into two-categories problem to solve the symptoms and syndrome types. A DBN-based classifier is trained to learn the complex nonlinear relationship between symptoms and syndrome. Experiments have proved that proposed method has good performance on TCM syndrome induction compared with the other two commonly used models. This method can be applied to the vehicle-mounted intelligent TCM clinical medical system probably.

However, there are still some disadvantages in this work. First of all, the clinical samples for DBN training are less than other works which may not cover all exceptions, as a result, the lack of training samples may bring the weak generalization ability of the model. Secondly, the typical symptoms-syndrome that adopted by TCM colleges should be tested to make the experiment results more robust. In the further work, studies should focus on the promotion of effectiveness of the proposed model and extend to the whole therapeutic process of TCM diagnosis and treatment for vehicle-mounted application scenarios.

REFERENCES

[1] F. Yang, J. Yan, S. Li, and C. Yang, "Risk analysis and safety design of mobile smart medical system," in *Proc. IEEE 3rd Inf. Technol. Mechatronics Eng. Conf. (ITOEC)*, Oct. 2017, pp. 163–166.

[2] J. Li, Q. Wu, R. Zheng, J. Zhu, Q. Ge, and M. Zhang, "Randomised block-coordinate Frank–Wolfe algorithm for distributed online learning over networks," *Cognit. Comput. Syst.*, vol. 2, no. 2, pp. 72–79, Jun. 2020.

[3] K. Kansara, V. K. Singh, R. Patel, R. R. Bhavsar, and A. P. Vora, "Numerical investigations of phase change material (PCM) based thermal control module (TCM) under the influence of low gravity environment," *Int. J. Heat Mass Transf.*, vol. 167, Mar. 2021, Art. no. 120811.

[4] S. Ge, Q. Xing, A. Zhang, and Y. Wang, "Effect of traditional Chinese medicine (TCM) on survival, quality of life, and immune function in patients with ovarian carcinoma: A protocol for systematic review and meta analysis," *Medicine*, vol. 100, no. 2, 2021, Art. no. e23904.

[5] D. Xian-Hui, M. Dong-Xue, Z. Tian-Ci, H. Xiao-Ping, X. Li-Jun, L. Ya-Lei, L. Hao, and G. Wei-Juan, "The effect of TCM-induced HAMP on key enzymes in the hydrolysis of AD model cells," *Neurochem. Res.*, vol. 46, no. 5, pp. 1068–1080, May 2021.

[6] T. Panch, P. Szolovits, and R. Atun, "Artificial intelligence, machine learning and health systems," *J. Global Health*, vol. 8, no. 2, Dec. 2018, Art. no. 020303.

[7] L. Badea, "Functional discrimination of gene expression patterns in terms of the gene ontology," in *Biocomputing*. Singapore: World Scientific, 2002, pp. 565–576.

[8] D. C. Stenning, T. C. M. Lee, D. A. van Dyk, V. Kashyap, J. Sandell, and C. A. Young, "Morphological feature extraction for statistical learning with applications to solar image data," *Stat. Anal. Data Mining, ASA Data Sci. J.*, vol. 6, no. 4, pp. 329–345, Aug. 2013.

[9] B. Huang, Z. Chen, S. Su, B. Wang, S. Su, and Q. Wu, "Development of the intelligent vehicle mounted with fumigation system for port container quarantine," *Chin. J. Hygienic Insecticides Equip.*, vol. 25, no. 3, pp. 217–220, 2019.

[10] F. Qu, Q. Zhang, M. Dai, Y. He, J. Wu, X. Zhang, Y. Zhu, Y. Gu, F. Wang, and X. Xu, "An evaluation survey of traditional Chinese medicine learning among international students majoring in conventional medicine: A study from a university in China," *BMC Complementary Med. Therapies*, vol. 21, no. 1, pp. 1–8, Dec. 2021.

[11] G. Martínez-Arellano, G. Terrazas, and S. Ratchev, "Tool wear classification using time series imaging and deep learning," *Int. J. Adv. Manuf. Technol.*, vol. 104, nos. 9–12, pp. 3647–3662, Oct. 2019.

[12] A. Ebigbo, R. Mendel, A. Probst, J. Manzeneder, L. A. D. Souza, Jr., J. P. Papa, C. Palm, and H. Messmann, "Computer-aided diagnosis using deep learning in the evaluation of early oesophageal adenocarcinoma," *Gut*, vol. 68, no. 7, pp. 1143–1145, Jul. 2019.

[13] V. F. S. Fook, J. H. Tee, K. S. Yap, A. A. P. Wai, J. Maniyeri, B. Jit, and H. P. Lee, "Smart mote-based medical system for monitoring and handling medication among persons with dementia," in *Proc. Int. Conf. Smart Homes Health Telematics*. Berlin, Germany: Springer, 2007, pp. 54–62.

[14] S. Ju, Y. Sun, and Y. Su, "Internet of Things smart medical system and nursing intervention of glucocorticoid drug use," *Microprocessors Microsyst.*, vol. 83, Jun. 2021, Art. no. 104008.

[15] Y. Niu and G. Hirst, "Analysis of semantic classes in medical text for question answering," in *Proc. Conf. Question Answering Restricted Domains*, 2004, pp. 54–61.

[16] K. Fujita, M. Akiyama, K. Park, E. N. Yamaguchi, and H. Furukawa, "Linguistic analysis of large-scale medical incident reports for patient safety," in *Proc. MIE*, 2012, pp. 250–254.

[17] A. Banerjee and S. K. S. Gupta, "Analysis of smart mobile applications for healthcare under dynamic context changes," *IEEE Trans. Mobile Comput.*, vol. 14, no. 5, pp. 904–919, May 2015.

[18] K. R. Dayananda, R. Gomes, and J. Straub, "An interconnected architecture for an emergency medical response unmanned aerial system," in *Proc. IEEE/AIAA 36th Digit. Avionics Syst. Conf. (DASC)*, Sep. 2017, pp. 1–6.

[19] Y. Wang, Q. Wu, J. Zhu, L. Wang, and M. Zhang, "A TCM syndrome differentiation algorithm based on deep belief network," in *Proc. 35th Youth Academic Annu. Conf. Chin. Assoc. Automat. (YAC)*, Oct. 2020, pp. 166–171.

[20] Y. Wang, Z. Pan, X. Yuan, C. Yang, and W. Gui, "A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network," *ISA Trans.*, vol. 96, pp. 457–467, Jan. 2020.

[21] I. Rhee, K. E. Bachman, B. H. Park, K.-W. Jair, R.-W.-C. Yen, K. E. Schuebel, H. Cui, A. P. Feinberg, C. Lengauer, K. W. Kinzler, S. B. Baylin, and B. Vogelstein, "DNMT1 and DNMT3b cooperate to silence genes in human cancer cells," *Nature*, vol. 416, no. 6880, pp. 552–556, Apr. 2002.

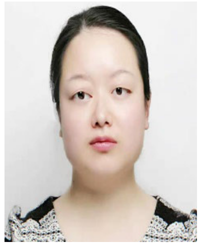
[22] X. Liu, X. Jian, and E. Boerwinkle, "DbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions," *Hum. Mutation*, vol. 32, no. 8, pp. 894–899, Aug. 2011.

[23] M. F. Mukhtar, M. Shiraz, Q. Shaheen, K. Ahsan, R. Akhtar, and W. Changda, "RBM: Region-based mobile routing protocol for wireless sensor networks," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–11, Feb. 2021.

[24] X. Liu, X. Jian, and E. Boerwinkle, "DbNSFP v2.0: A database of human non-synonymous SNVs and their functional predictions and annotations," *Hum. Mutation*, vol. 34, no. 9, pp. E2393–E2402, Sep. 2013.
[25] Y. Yang, "Medical multimedia big data analysis modeling based on DBN algorithm," *IEEE Access*, vol. 8, pp. 16350–16361, 2020.
[26] G. Tsoumakas and I. Katakis, "Multi-label classification: An overview," *Int. J. Data Warehousing Mining*, vol. 3, no. 3, pp. 1–13, 2007.
[27] E. Montañes, R. Senge, J. Barranquero, J. R. Quevedo, J. J. del Coz, and E. Hüllermeier, "Dependent binary relevance models for multi-label classification," *Pattern Recognit.*, vol. 47, no. 3, pp. 1494–1508, Mar. 2014.
[28] M.-L. Zhang, Y.-K. Li, X.-Y. Liu, and X. Geng, "Binary relevance for multi-label learning: An overview," *Frontiers Comput. Sci.*, vol. 12, no. 2, pp. 191–202, 2018.
[29] J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification," *J. Mach. Learn.*, vol. 85, no. 3, pp. 333–359, Dec. 2011.
[30] J. R. Smith, J. Tesic, and R. Yan, "Method and apparatus for model-shared subspace boosting for multi-label classification," U.S. Patent 7 958 068, Jun. 7, 2011.

ZHAOHUI NING was born in Henan, China, in May 1993. She studied in Henan University of Science and Technology, Luoyang, China, from September 2013 to July 2018, majored in clinical medicine and earned a Bachelor of Medicine Degree in five years' time. She is currently pursuing the master's degree in emergency medicine with Jiamusi University. Her current research interests include medical image analysis and emergency care.

YIHAO WANG was born in Henan, China, in March 1995. He received the bachelor's degree in computer science and technology from Henan University of Science and Technology, Luoyang, China, in July 2018, where he is currently pursuing the master's degree in computer science and technology. His current research interests include medical image analysis, deep learning, and reinforcement learning.



LIN WANG received the Ph.D. degree in image and signal processing from CREATIS, INSA de Lyon, Lyon, France, in 2016. In 2017, she joined Henan University of Science and Technology, Luoyang, China, where she is currently a Lecturer with the School of Information Engineering. Her research interests include digital image processing, pattern recognition, scientific visualization, and machine learning.

LIJUAN WU was born in Henan, China, in May 1981. She received the Master of Engineering degree in clinical medicine from Henan University of Science and Technology, Luoyang, China, in July 2008. She currently works as an Associate Chief Physician with The First Affiliated Hospital, Henan University of Science and Technology. Her current research interests include medical image analysis and endoscopy.

JIANQIANG SONG received the M.S. and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2011 and 2019, respectively. Since 2019, he has been a Lecturer with the School of Information Engineering, Henan University of Science and Technology. His research interests include medical image processing, pattern recognition, few-shot learning, and zero-shot learning.