IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 5, 2019, accepted September 17, 2019, date of publication September 26, 2019, date of current version October 28, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2943927

Research on Medical Data Feature Extraction and Intelligent Recognition Technology Based on Convolutional Neural Network

WEIDONG LIU[1], CAIXIA QIN[2], KUN GAO[1], HENG LI[1], ZUEN QIN[3], YAFEI CAO[1], AND WEN SI[4,5]

[1]Shenzhen Traditional Chinese Medicine Hospital, Shenzhen 518033, China
[2]Shenzhen Luohu Hospital Group Luohu People's Hospital, Shenzhen 518001, China
[3]Hechi Traditional Chinese Medicine Hospital, Hechi 547000, China
[4]College of Information and Computer Science, Shanghai Business School, Shanghai 201400, China
[5]Department of Rehabilitation, Huashan Hospital, Fudan University, Shanghai 200040, China

Corresponding Authors: Yafei Cao (caoyafei64@126.com) and Wen Si (siwen@fudan.edu.cn)

This work was supported in part by the Shenzhen Governmental Basic Research under Grant JCYJ20150529150715499, and in part by the Shenzhen medical three projects under Grant SZSM201812066.

ABSTRACT In order to mine information from medical health data and develop intelligent application-related issues, the multi-modal medical health data feature representation learning related content was studied, and several feature learning models were proposed for disease risk assessment. In the aspect of medical text feature learning, a medical text feature learning model based on convolutional neural network is proposed. The convolutional neural network text analysis technology is applied to the disease risk assessment application. The medical data feature representation adopts the deep learning method. The learning and extraction of different disease characteristics use the same method to realize the versatility of the model. A simple preprocessing of the experimental data samples, including its power frequency denoising and lead convolution regularization, constructs a convolutional neural network for medical data feature advancement and intelligent recognition. On the basis of it, several sets of experiments were carried out to discuss the influence of the convolution kernel and the choice of learning rate on the experimental results. In addition, comparative experiments with support vector machine, BP neural network and RBF neural network are carried out. The results show that the convolutional neural network used in this paper shows obvious advantages in recognition rate and training speed compared with other methods. In the aspect of time series data feature learning, a multi-channel convolutional self-encoding neural network is proposed. Analyze the connection between fatigue and emotional abnormalities and define the concept of emotional fatigue. The proposed multi-channel convolutional neural network is used to learn the data features, and the convolutional self-encoding neural network is used to learn the facial image data features. These two characteristics and the collected physiological data are combined to perform emotional fatigue detection. An emotional fatigue detection demonstration platform for multi-modal data feature fusion is established to realize data acquisition, emotional fatigue detection and emotional feedback. The experimental results verify the validity, versatility and stability of the model.

INDEX TERMS Convolutional neural network, medical data feature, intelligent recognition, convolution regularization, data feature fusion.

I. INTRODUCTION

Medical health data is a multi-modal, complex data that continues to grow rapidly. It contains a wealth of information. The challenges related to medical health data include how to quickly and accurately collect and obtain medical health data, and how to efficiently use high-speed networks to conduct medical health data. Reliable and efficient transmission, how to use artificial intelligence related machine learning and deep learning techniques to extract useful information from

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

health medical big data, and develop intelligent applications for the majority of medical staff and ordinary people. Convolutional neural networks are an important structure of deep learning and are mainly used to identify two-dimensional images. The convolutional neural network itself has many performance advantages. For example, what is most praised is its weight sharing. This sharing makes the neural network closer to the biological visual nervous system and reduces the complexity of the network model to some extent. Degree is a sharp reduction in the number of training for weights. In addition, the structure of the secondary extraction feature of the convolutional neural network makes it not required for the input of the network, and even the original image, which is highly tolerant to the translation, scaling or deformation of the image [1]–[3].

Health care big data comes from medical institutions, wireless sensor networks, wearable sensors, environmental data, etc., including structured data, unstructured and other multimodal data. This section introduces research progress on multi-modal data disease risk assessment at home and abroad, including data feature representation learning, analysis based on medical text data, analysis based on medical image data, fatigue detection and sentiment analysis [4]–[6]. The data-driven feature learning method is to use a large number of data training feature learning algorithms to finally obtain the representation method of the original data features, which is simply to let the data speak for itself. There is a supervised learning method, and the training sample data set includes sample data and the label of the sample data, the sample data is passed from the input layer to the output layer by layer, and finally the predicted value of the sample is obtained, which is called forward propagation. Calculate the error between the predicted value of the sample and the true target value of the sample, and perform layer-by-layer propagation from the output layer to the input layer to optimize the connection parameters between the layers in the structure, in order to minimize the cost function value. This process is backward widely used convolutional neural networks [7]–[9], using supervised learning methods to learn the characteristic representation of raw input data. Local connections between adjacent layers in the CNN are a variation of the neural network. CNN is used in image analysis [10], [11], speech recognition [12], text analysis [13]–[15], etc., especially in the field of image analysis, such as face recognition [16], image classification [17], Scene Recognition [18]. The literature [19]–[21] uses structured data for health monitoring and disease risk assessment, and obtains patient disease characteristics based on professionally designed disease feature extraction methods on the dataset, using machine learning. The method stratifies the patient according to the degree of risk of the disease. This method has been widely studied and applied to the clinic. Literature [22]–[24] uses big data analysis methods to learn characteristics from a large number of structured data sets for disease assessment and identification. The convolutional neural network method is a data-driven feature representation learning method. The joint use of word vectors can effectively obtain the context information of text data, obtain the representation of text features, and perform natural language processing tasks such as text classification and text sentiment analysis. There are great advantages in [25]–[27]. Intelligent Recognition Technology Based on Convolutional Neural Network is introduced [28]. The existing methods for intelligently extracting disease-related features from texts require medical experts to design features, but experts designing disease-related features are limited by domain knowledge, and when the disease characteristics change, the model design needs to be modified, which consumes a lot of manpower and time. In the real world, the symptoms of a wide variety of diseases are complex, and the diverse disease risk feature recognition methods limit the versatility and adaptability of feature extraction and intelligent recognition models, which is difficult to achieve in clinical applications.

This paper explores the multi-modal medical data feature representation learning related content by mining information from medical big data and developing intelligent application related issues, and studies the application of multi-modal data features for disease risk assessment, etc. At the same time, how to be intelligent Chemical medical applications provide comprehensive data services for research. The convolutional neural network text analysis technology is applied to the disease risk assessment application. The medical data feature representation adopts the deep learning method. The learning and extraction of different disease characteristics use the same method to realize the versatility of the model. A simple preprocessing of the experimental data samples, including its power frequency lead convolution regularization, constructs a convolutional neural network for medical data feature advancement and intelligent recognition. On the basis of it, several sets of experiments were carried out to discuss the influence of the convolution kernel and the choice of learning rate on the experimental results. In addition, comparative experiments with support vector machine, BP neural network and RBF neural network are carried out. The results show that the convolutional neural network used in this paper shows obvious advantages in recognition rate and training speed compared with other methods. . In the aspect of time series data feature learning, a multi-channel convolutional self-encoding neural network is proposed. Analyze the connection between fatigue and emotional abnormalities and define the concept of emotional fatigue. The proposed multi-channel convolutional neural network is used to learn the data features, and the convolutional self-encoding neural network is used to learn the facial image data features. These two characteristics and the collected physiological data are combined to perform emotional fatigue detection. An emotional fatigue detection demonstration platform for multi-modal data feature fusion is established to realize data acquisition, emotional fatigue detection and emotional feedback. The experimental results verify the validity, versatility and stability of the model.

TABLE 1. Comparison of several medical data feature extraction algorithms.

Number of samples	Pretreatment	Feature extraction	Classification	Recognition rate
29 people	1-40Hz Butterworth fourth-order filter	Linear Discriminant Analysis (LDA) dimension reduction	Euclidean distance method	100
10 people 140	Wavelet denoising	Differential threshold	BP neural network	60%;
Group data	-	Wavelet transform	RBF neural network	100%
-	-	Wavelet transform and independent component analysis	Residual percentage (PRD) correlation value parameter	70%

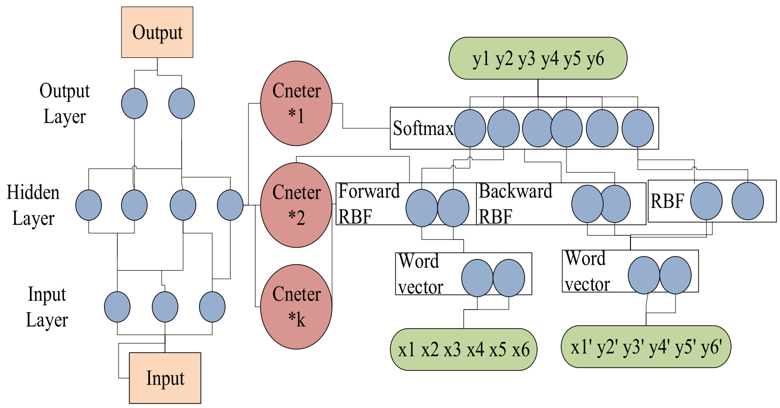

FIGURE 1. RBF neural network structure model.

II. MEDICAL DATA FEATURE EXTRACTION FROM CONVOLUTIONAL NEURAL NETWORKS

A. FEATURE EXTRACTION MODEL CONSTRUCTION OF CONVOLUTIONAL NEURAL NETWORKS

Careful analysis of the development status of medical data feature extraction, we can find that in the identification process, we mainly use various types of filters and wavelet. The methods of feature extraction mainly include wavelet transform and differential threshold. The method, the dynamic feature extraction based on time window, etc., commonly used classifiers include support vector machine, neural network (BP neural network, RBF neural network) and BAB algorithm. The following is a simple comparison of several representative methods as shown in Table 1 below:

The most mature research is the R wave extraction method. Its performance is well tested in the standard database MIT BIH, but it is much lower in the actual clinical data set. As for the P wave and T wave extraction algorithms on MIT-BIH the accuracy rate is still low, and the clinical application is somewhat remote. The essence of the learning rule of BP network is to use the gradient descent method to continuously correct and adjust the weight and closed value of the network through error feedback until the error is the smallest. The network does not need to define the mapping relationship between ''input and output'' in advance, and can learn and store the relationship itself during the training process. The structure of the BP neural network includes an input layer, a hide layer, and an output layer. The topology of RBF neural network is similar to multi-layer feedforward network. It is a three-layer feedforward network, including input layer, hidden layer and output layer. The hidden layer activation function is a radial basis function, which is a kind of local Symmetric nonlinear function. The RBF network maps the input data to a new space through a radial basis function nonlinear transformation. The adjustable parameters are the weight of the linear combiner. As shown in Figure 1.

As shown in Figure 1, the training process of the RBF network mainly includes two stages. The first stage determines the parameters of the activation function, including the central function and the shape parameters. The second stage training usually uses the least squares method to perform linear weight training between the hidden layer and the output layer. Determining the RBF center is crucial. There are two points to note in the selection: First, the center selected in the data needs to pass the sampling test, and it must not be randomly selected. Second, the data cannot be gathered too much when the center is selected, which is easy to cause linear correlation. Form numerical ill conditions.

In this experiment, the convolutional neural network has 11 layers, including 1 input layer, 1 full connection layer,

TABLE 2. Medical data identification network structure based on convolutional neural network.

Name	Types of	Window size	Step size	Output size	Output quantity
Input	Input layer	3600x2	-	3600x2	1
Conv 1	Convolution layer	13x2	1	3588x1	6
Pool1	Activation function	-	-	3588x1	6
Tanh	Pooling layer	2x1	2	1794x1	6
Poo12	Convolution layer	13x1	1	1782xI	12
Conv3	Activation function	-	-	1782x1	12
Poo13	Pooling layer	4x1	4	446x1	12
Conv4	Convolution layer	9x1	1	438x1	24
Tanh	Activation function	-	-	438x1	24
Poo14	Pooling layer	4x1	4	110x1	24
Fc	Convolution layer	9x1	1	102x1	36
Output	Activation function	-	-	102x1	36
Input	Pooling layer	6x1	6	17x1	48
Conv 1	Fully connected layer	-	-	300	300
Pool1	Output layer	-	-	15	15

1 output layer and 8 hidden layers. The hidden layer includes 4 convolution layers and 4 pool layers. Each convolution layer is followed by a pooling layer for secondary feature extraction. After the input medical data is convolved and pooled, the feature map is getting smaller and smaller, but after the hidden layer is propagated, the number of feature maps is gradually increasing. This means that the number of extracted features is increasing, and the ECG signal information can be more fully represented. Finally, all of this information is input to the fully connected layer to classify the medical data features.

The input 3600×2 medical data passes through the first convolutional layer, and the convolution kernel is 13×2, and 6 characteristic maps of 3588×1 $(3600 - 13 + 1, 2 - 1 + 1)$ are obtained. After the first pooling layer, sampling is performed. The kernel is 2×1, and the size of the six feature maps becomes 1794×1 (2588/2, 1/1), see Table 3.3 for details. After passing through the fourth pooling layer, $48 \times 17 \times 1$ feature maps are obtained. At the fully connected layer, 300 neurons are fully connected to the $48 \times 17 \times 1$ feature maps obtained by convolution of the previous layer. The number of connection parameters is $48 \times 17 \times 1 \times 300 = 244800$. Finally, use the softmax function to classify it into 15 categories.

The choice of the activation function can map the features to solve the nonlinear problem. A suitable activation function can increase the training speed of the network and improve the recognition rate of the image.

$$sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

The gain of the sigmoid function in the central region is relatively large, and the gains on both sides are relatively small, so that the features can be well mapped. Therefore, for a long time, sigmoid was widely praised. Similar to sigmoid is a function whose basic expression is:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

B. MEDICAL DATA FEATURE EXTRACTION

In the process of collecting medical data, the waveform is easily damaged due to interference from the participants themselves or the external environment, mainly including power frequency interference, baseline drift and myoelectric interference. Power frequency interference and baseline drift are easy to eliminate due to the difference between the size of the frequency band and the medical data signal, and the myoelectric interference covers the entire medical data signal, making filtering powerless. These effects will cause distortion to the data to a certain extent, and in order to avoid interference from these unrelated factors, we need to take appropriate pre-processing of the medical data signals.

The bilinear transform method is used to convert the analog low pass filter into the Z plane of the digital band stop filter through the S plane. The relationship between analog low pass and band stop is as follows:

$$S = \frac{\varpi_0^2 P}{p^2 + \varpi_0^2} \tag{3}$$

Among them, S and P respectively represent the Laplacian variables of the simulated low-pass prototype and the simulated band-stop, and ϖ_0^2 is the geometric center frequency of the analog band-stop filter, which adopts the bilinear

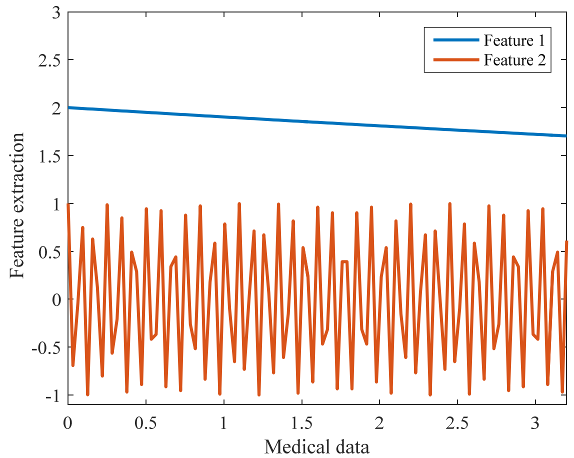

FIGURE 2. Feature extraction verification after denoising medical data signals.

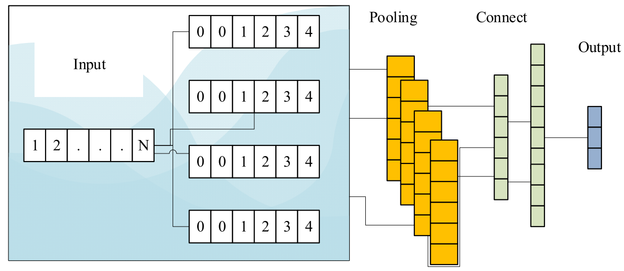

FIGURE 3. Medical data feature extraction network structure of convolutional nerve [12].

transformation design.

$$P = 2fs\frac{z-1}{z+1}, \quad \varpi_0 = 2fs\tan(\frac{\Omega_0}{2}) \tag{4}$$

Therefore, this paper adopts the second-order Butterworth digital notch filter. The filter has a narrow stopband and a flat passband. This structure effectively eliminates the medical data signal without losing the useful signal. The impact of power frequency interference on medical data signals. After the experiment using the above filter, the denoising result is shown in Figure 2 below.

As shown in Figure 2, the data of the same lead has similarities and has its intrinsic properties. Then we should try to use sample training so that each lead can find the convolution kernel that is most suitable for itself. If we use the convolution kernel $1 \times X$, then the different leads will no longer perform convolution operations.

The convolution unit CU has multiple convolutional feature planes, 1D-Cov represents a one-dimensional convolution, and each lead has three convolutional units, and the convolution between the different leads is independent. Since the data selected in the database of this experiment are both II leads and VI leads, there are a total of 6 convolution units, compared to the traditional CNN with only 3 convolution units. Each lead's data is sampled and found to be the most suitable for its own convolution unit, and then aggregated through the fully connected layer for final identification.

$$\begin{cases} f(x) = gE(gD(\bigcup_{i=1}^{2} gC_i(gB_i(gA_i(x_i))))) \\ gD(x) = \phi(Wx+b) \\ gE(x) = \dfrac{1}{1+e^{-(Wx+b)}} \end{cases} \tag{5}$$

where g_D is the fully connected layer calculation function, g_E is the LR layer calculation function, uncle.) is the excitation function, g_A, g_B, g_C are the convolution unit calculation function, and the expression form of the sampling function and the convolution function is similar.

III. FEATURE EXTRACTION AND INTELLIGENT RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORKS

A. FEATURE EXTRACTION MODEL FOR MEDICAL DATA

The convolutional neural network is designed as a five-layer structure: an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer, which are used to learn the characteristic representation of medical text data for risk assessment of diseases. The input layer is to preprocess and digitize the medical record text data and input it into the CNN network structure. As shown in Figure 3, the convolution layer performs a convolution operation on each word, and the beginning and ending words need to be filled with empty words (indicated by 0). After the convolution operation is completed, the data enters the pooling layer and the fully connected layer, and finally reaches the output layer, and the risk assessment result is obtained.

The more pure the corpus of the word vector training, the better the use of professional corpus. In this paper, the clinical record data of all patients in the hospital are extracted from the data center, and the data is cleaned up as a training corpus. The corpus data volume is 300M, the word vector dimension is set to 50, and the word vector is 52100 words after training. In the field of image analysis, the convolutional neural network method is used to express the method of training medical text disease risk assessment, and the learned characteristics are obviously superior to those of the manual design feature method. The training medical text disease risk assessment data has image data. Commonality, convolutional deep learning methods can be used for end-to-end feature representation learning.

All parameters are defined as the parameters of the set $\theta = \{w^1, w^4, b^1, b^4\}$, θ are randomly initialized and updated using the stochastic gradient descent method. The goal of the training is to maximize the log likelihood function.

$$\max \sum_{y \in D} \log p(class_y | D, \theta) \tag{6}$$

where D is the training sample set, class is the real classification of the training samples, and a is used to represent the learning rate. The formula for parameter updating

TABLE 3. Algorithm 1 training medical text disease risk assessment convolutional neural network.

Training convolutional neural network

Input: X: training sample, medical text raw input data
Y: Training sample label, patient input diagnosis output of raw input data: construct CNN, network parameter $\theta=\{w^1,w^4,b^1,b^4\},\theta$, test result.
C=5 //Convolution kernel size
T=50 //sample block size
For i=1, 2...,I //I is the number of iterations
For j=1, 2...,m //m is the number of blocks in the sample
Read in a sample block x and the corresponding label Y
Find the vector representation of the data x
For n=1,2,...,T
Calculate the number of words in the sample
Calculated convolution
Get the maximum Pooling result for N words
The classification output y*x is obtained through the fully connected layer and the softmax classifier.
End for
Gradient descent method to modify the parameter θ
End for
End for

is as follows:

$$\theta=\theta+a\frac{\partial \log p(class_y|D,\theta)}{\partial\theta} \tag{7}$$

The pseudo-code for the disease risk assessment model of the training convolutional neural network is shown in Algorithm 1. The medical text data includes the training set X and the test set X′, and the corresponding annotation data is from the doctor's diagnosis result, and is divided into a training label set Y and a test label set Y′. Algorithm 1 training medical text disease risk assessment convolutional neural network as shown in Table 3.

Based on the medical text data disease risk assessment model proposed in this study, the data center stores all medical related data generated by patients in the hospital, including the patient's subjective condition, doctor's medical record, drug record, treatment cost, and treatment plan. Here, the risk assessment of the disease is performed using only the patient's medical record text information, learning the model based on the patient's historical data, and finally using the learned model to assess whether the patient is in a high risk category. The model mainly consists of three parts:

(1) Learning word vector

Text data analysis first needs to be a standardized representation of the text data, which is to use digital vectors to represent the text. We use all patient history medical text records stored in the data center to perform data cleansing and data preprocessing, and use the processed data as corpus training to obtain word vectors.

(2) Training CNN to learn medical text features

We selected disease data from all patient clinical record text data, and we chose to use several common and serious chronic diseases such as hypertension, cerebral infarction, and coronary heart disease. Extract the subjective condition of the patient with a certain disease, the doctor's consultation record, etc., and preprocess the data as the disease sample data. The disease sample data and the comparative sample data are divided into a training sample set and a test sample set. The sample data is digitally represented by the word vector and then input into the CNN for training and testing, and finally the trained CNN is obtained.

(3) Risk assessment

When conducting a risk assessment, input the medical text data related to the disease such as the patient's subjective condition, preprocess the data, use the word vector for text representation, input the CNN, and output the risk assessment result of the disease.

B. MULTIMODAL MEDICAL DATA FEATURE LEARNING CONVOLUTIONAL NEURAL NETWORK MODEL

The disease risk assessment model of multimodal medical data needs to extract unstructured text data features and structured data features, and fuse the two types of features. A multi-modal disease risk assessment model based on CNN was proposed. Using CNN to learn and extract text features using unstructured text data using data-driven feature learning method, 79 features were extracted according to doctor's suggestion of structured data. In CNN, the feature fusion layer is designed in the structure to determine the connection parameters between structured data features and unstructured data features and classifiers through supervised learning.

Medical data used in multimodal medical disease risk assessment models, including structured data and unstructured data. Among them, the structured data includes basic information of the patient (for example, age, gender, living habits), laboratory data, and the like. The unstructured medical text data is the patient's clinical record text, which mainly includes the patient's self-reported condition, the doctor's medical record and treatment record. Table 4 shows the items included in the multimodal data set, in which the patient's self-reported condition and the doctor's medical record are unstructured data, the patient's statistical information, living habits, test items and results, and disease are structured data.

For structured data, there is a large amount of missing data due to incomplete entry and incomplete patient examination, so structured data needs to be filled. Before data reconstruction, we first use data sets for data preprocessing. Integrating data ensures the correctness and atomicity of the data, for example, using the retrieved height and weight data to obtain a body fat index. Suppose $R^{m\times n}$ is a data matrix, the number of rows m represents the total number of users, and the number of columns n represents the state of the user. Assuming there are k potential factors, the original matrix $R^{m\times n}$ can be approximated as $R = PQ^T$, and each element in the matrix can be represented by $r = p_w^t q$, and we will determine that their values are converted to optimization problems:

$$\min_{p,q}(\sum_{u,v}(r_{uv}-p_u^T q_v)^2+\lambda_1||p_u||^2+\lambda_2||q_v||^2) \tag{8}$$

TABLE 4. Multimodal medical data.

Project	Description	Type of data
Patient statistics	Gender, age, height, weight, etc.	Structured
living habit	Whether smoking, drinking, whether there is a history of genetic disease	Structured
Inspection items and results	Including 682 items, such as blood pressure, heartbeat, etc.	Structured
disease	Patients with diseases such as cerebral infarction, hypertension, diabetes, etc.	Structured
Patient self-reported condition	The patient's own description of the symptoms, onset time and medical history	Unstructured
Doctor's medical record	The doctor asked the patient's medical record.	Unstructured

C. INTELLIGENT IDENTIFICATION OF MEDICAL DATA BASED ON CONVOLUTIONAL NEURAL NETWORKS

Emotional fatigue detection uses multimodal data, and data feature extraction methods need to be designed according to the structure of different data. User basic information and physiological information, including the user's age, gender and other data obtained using statistical methods; body temperature, heart rate, directly use the collected data. We use the data-driven deep learning model to extract features from ECG signals and facial expression images related to user sentiment information, avoiding the problems caused by manual extraction of features. The motion data is collected using the mobile device sensor data. For the vehicle equipment data, a certain closed value is automatically divided into {fast, medium speed, slow speed}, and the temperature and fuel consumption are directly recorded.

(1) Convolution layer

Enter the 3-channel medical data time series data. The convolution kernels of the three convolutional layers C1, C2, and C3 are 50, 40, and 20, respectively. The convolution kernels have the same size and are set to 1×5. Using the formula for the convolution operation, a 3×256 feature map is input to C1. After the convolution operation is completed, 50 3×252 feature maps are obtained; $50 \times 3 \times 84$ feature maps are obtained through C2 operation to obtain $40 \times 3 \times 80$ feature maps; 40 The feature map of 3×40 is operated by C3 to obtain $20 \times 3 \times 36$ feature maps.

(2) Pooling layer

The pooling area sizes of the three pooling layers of P1, P2, and P3 are set to 3, 2, and 2, respectively. The Pl layer input feature map and output feature map number are 50, and the sizes are 3×252 and 3×84 respectively; the P2 layer input feature map and the output feature map number are 40, and the sizes are 3×80 and 3×40 respectively; P3 layer input feature map and output feature The number of pictures is 20 and the sizes are 3×36 and 3×18 respectively.

(3) Fully connected layer

The current layer is set up with 400 neurons, each of which is fully connected to each of the last pooling layer P3, and 400 output values of the fully connected layer are represented as characteristic of the original medical data x.

Multimodal learning can better identify the user's emotions. There are two main methods for multimodal data fusion: feature level fusion and decision level fusion. Feature layer fusion is the fusion of all data features into feature vectors as classification features. The decision layer fusion is categorized separately according to the data of each modality, and then the results of the classification are linearly combined, such as using the average weight.

TABLE 5. Algorithm 2 Unsupervised Training TCA.

Unsupervised training TCA
Input: untagged data set UD;
Output: network structure, parameters of each layer in the network
1:Desired number of convolution layer and pooling layer N;
2: Initialize all weight matrices and bias vectors randomly convolution layer and pooling layer;
3:i--1;
4: if i=1 then
5: The input of C_i is UD;
6: else
7: The input of C_i is the output of P_i
8: end if
9: Greedy layer-wise training C_i;
10: Find parameters for C by cost function;
11:Use the output of C_i as the input of the P_i
12: Max pooling operator;
13: if i<N then
14: return line 4;
15: end if

The sample data set used for unsupervised learning is represented as $Z = (z_1, z_2 \ldots z_n)$ and consists of medical data and facial image data. Supervised training includes a sample data set $X = (x_1, x_2 \ldots x_n)$ and a corresponding set of labels Y, including all kinds of data in emotional fatigue testing. Algorithm 2 Unsupervised Training TCA as shown in Table 5.

IV. EXPERIMENT AND ANALYSIS

When training CNN, you need to define the size of the convolution operation sliding window. The experimental setting sliding window takes 1 to 9 word steps respectively to 2 to verify the effect of the sliding window size on the evaluation

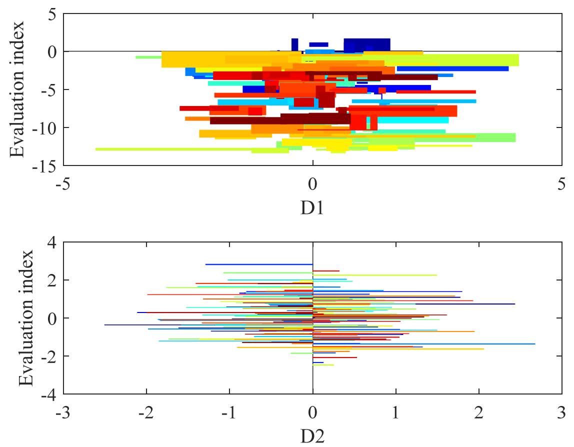

FIGURE 4. CNN convolution sliding window size impact on evaluation indicators.

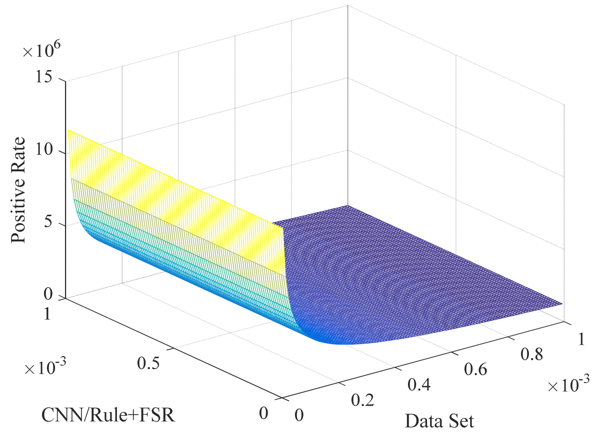

FIGURE 5. CNN and rule + FSR based on the ROC curve on the data set.

TABLE 6. Risk assessment results.

Disease	Data set	Algorithm	Accuracy	Recall rate	Accuracy
Brain infarction	D1	CNN	95.5%	97.7%	96.5%
		RuIe+FSR	84.4%	90.4%	86.9%
	D2	CNN	98.1%	93.0%	95.6%
		RuIe+FSR	95.5%	91.4%	93.6%
Lung infection	D3	CNN	-	-	92.9%
		RuIe+FSR	95.5%	97.7%	93.9%
Coronary artery stiffness	D4	CNN	84.4%	90.4%	-
		RuIe+FSR	-	-	96.5%

indexes of the algorithm. Figure 4 shows the results of the CNN algorithm using the D1, D2 data sets for training tests. The effect is of the sliding window size on the risk assessment indicators. When the sliding window size is set to 5, the accuracy, accuracy, and recall rate are better than other window sizes. The text data driven disease risk model has a CNN algorithm sliding window size of 5, convolution kernel. The number is set to 100.

Our convolution kernel slides over a local area of the image, but in CNN, the filter we typically use slides across the entire matrix. Therefore, the "width" of the filter is usually the same as the width of the input matrix. The height, or region size, may vary, it is a sliding window.

As shown in Figure 4 and Table 6, the comparison results between the proposed CNN method and the Rule+FSR method for textual disease risk assessment are shown. It can be seen from the results that the deep learning model CNN proposed in this paper is applicable to many diseases. Risk assessment, the RuIe+FSR method can only assess risk for a disease. Table 2-3 shows the accuracy, recall, and accuracy evaluation results of the CNN algorithm on the four experimental data sets D1, D2, D3, and D4, and the accuracy of the Rule+FSR method in the D4 data set for risk assessment. Using dataset D1 for risk assessment of patients with cerebral infarction, CNN algorithm has an accuracy of 95.5%, a recall rate of 97.7%, and an accuracy rate of 96.5%. Using dataset D2 for risk assessment of patients with cerebral infarction, using data from patients with hypertension with similar symptoms as comparative data, the accuracy of CNN algorithm was 84.4%, the recall rate was 90.4%, and the accuracy rate was 86.9%. Data set D3 includes data from patients with lung infection and those without lung infection. The evaluation indexes of CNN algorithm are accuracy of 98.1%, recall rate of 93%, and accuracy rate of 95.6%. The CNN algorithm in the D4 dataset for the risk assessment of coronary heart disease heart disease is more than 91%, with an accuracy of 93.6% over the Rule+FSR accuracy of 92.9%.

As shown in Figure 5, it is difficult to distinguish between patients with two diseases. Although the results of the risk assessment using the D2 dataset were lower than the other three datasets, the lowest accuracy among the four evaluation indicators reached g4.4%, which verified the effectiveness of the algorithm in identifying associated diseases, but fine-grained. The feature recognition is yet to be studied. The ROC curve of the risk assessment of the CNN and Rule+FSR algorithms on the D4 dataset is shown in Figure 5. The proposed CNN method is better than the Rule+FSR method. In summary, the experimental results show that the proposed model performance exceeds the rule-based Rule+FSR method. At the same time, the experimental results on multiple datasets also verify the versatility of the model. High-risk assessment of cerebral infarction was performed on structured dataset P1 using three machine learning methods, including naive Bayesian model [5], K nearest neighbor model [7], and decision tree model (Decision Tree) [9].

The experimental results are shown in Figures 6 and 7. The accuracy of the three machine learning models is about 60%, and the decision tree model has the highest accuracy rate of 63%, followed by the naive Bayesian model and finally the K-nearest neighbor model. The recall rate of the naive Bayesian model is up to 0.80, followed by the decision tree model and finally the K-nearest neighbor model. It can be concluded from Fig. 6 that the naive Bayesian model, the decision tree model, and finally the AUC C area under the

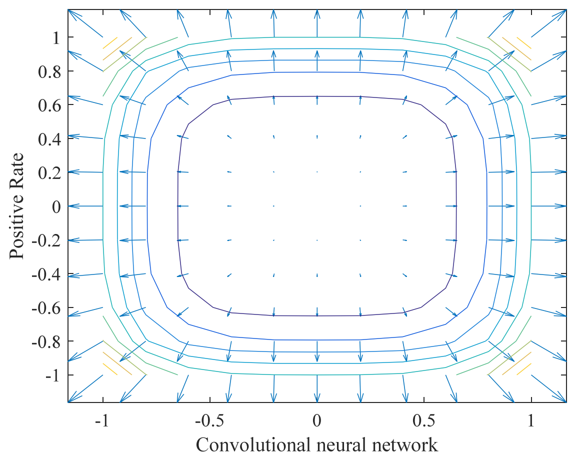

FIGURE 6. ROC curve of structured data convolutional neural network learning.

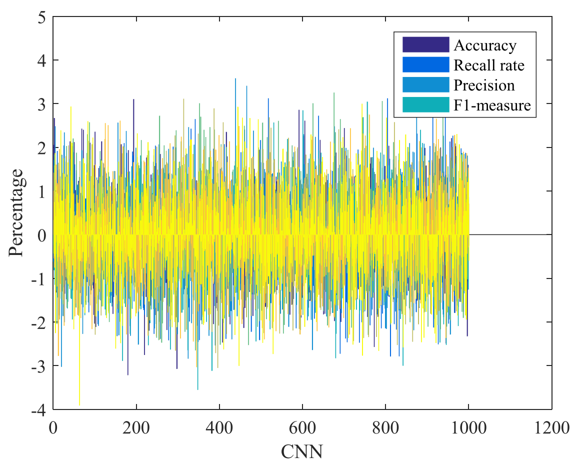

FIGURE 7. Evaluation index of structured data 3 convolutional neural network learning.

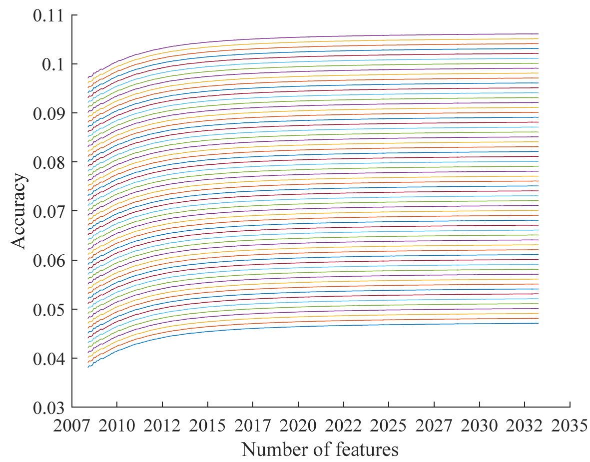

FIGURE 8. Trends in the accuracy of the test set as the number of features increases under the P2 and P3 data sets.

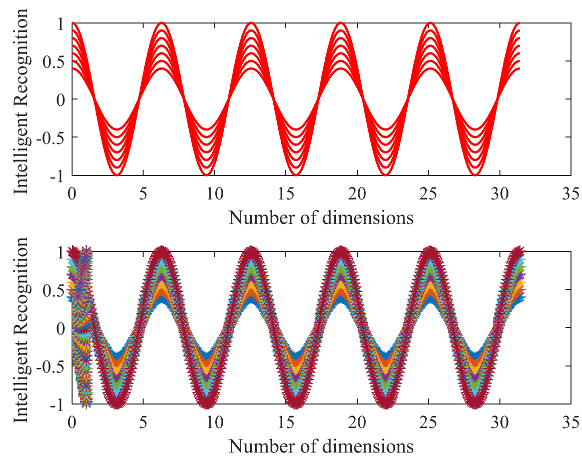

FIGURE 9. Convolutional neural network medical data vector feature intelligent identification verification.

roc curve corresponding to the K-nearest neighbor model are: 0.4950 0.4536 0.6463. The experimental results show that using only the user's age, gender, and structured data such as the laboratory could not accurately determine whether the user is a high-risk group of cerebral infarction. This is because cerebral infarction is a complex symptom and cannot be based on these. Simple features determine whether the disease is at high risk for cerebral infarction.

The number of features of structured data in the multi-modal dataset P3 is fixed. The 79 features related to the disease are determined by professional doctors. The text data features are learned using CNN. The number of features is correct for the P2 and P3 datasets. The impact of the recall rate, the number of features is variable. In the experiment test, the CNN extracted text feature number is changed to 10-120, and the number of iterations is 200. The accuracy of P2 dataset and P3.dataset are shown in Figure 8.

As shown in Figure 8, the figure shows that when the feature number of the text is less than 30, the accuracy of P2 and P3 is significantly smaller than the number of text features greater than 30, because when the number of text features is relatively small, the text cannot be completely described. As shown in Figure 8, the accuracy of the P3 data set is more stable than the accuracy of P2, effectively mitigating the vibration of the accuracy.

As shown in Figure 9, we can see that with the increase of the number of extended word vector dimensions, the accuracy rate first rises and then decreases, and the recall rate shows a trend of decreasing first and then rising; F1-Score is relatively flat, but overall It also shows a trend of rising first and then falling. This is because, compared with the case of using only basic text features, adding an extended word vector and using a real vector instead of the original orthogonal binary text vector can improve the recognition probability of the information extraction algorithm for a single word, and make full use of the words between the words. Similarity relationship, looking for words is similar to the extracted target text in the training corpus as extraction candidate words, which can improve the recognition accuracy and recall rate of the algorithm. At this time, when the word vector dimension is 100, we can see the Score of the extraction algorithm reaches the highest level, and the accuracy and recall rate can reach a high level.

However, with the gradual increase of the dimension of the extended word vector, the dimension of the word vector becomes smaller after it exceeds 100. This is because the increase of the vector dimension does not fully describe the characteristic attribute of the vocabulary itself. When the vector dimension is too large,, the description of the text in the word vector model will cause over-fitting, and Moreover, the weight of the basic text features in the original algorithm is reduced, so that the basic text features can be used for intelligent recognition.

However, due to the small size of the data set used in this paper, especially in the learning process of word vector expansion, the training data can not completely cover all the information of the entire corpus, resulting in lower overall accuracy and total recall rate of the algorithm; The distribution of the targets ''disease'', ''therapeutic means'' and ''medical measurement and examination'' is uneven, and the performance in different aspects is different. In the future improvement, we can consider introducing a larger data set to alleviate the similarity.

V. CONCLUSION

Big data in the health care field is an integral part of the national big data strategic layout. The analysis and excavation of valuable information is also related to the development of universal health care. A medical text feature learning model based on convolutional neural network and intelligent recognition were proposed. The text analysis technique combined with word vector and convolutional neural network was applied to unstructured medical text feature extraction for disease risk assessment. Experiments were carried out on a variety of diseases using data from a patient in a hospital. The results of the experiment demonstrated the feasibility of the proposed model for disease risk assessment and also verified the versatility of the model. Based on the medical text feature learning model, a multimodal medical data feature learning model was proposed. Based on the fusion of text data and structured data, medical data features were extracted for disease risk assessment. Experiments have verified that the multi-modal data fusion disease risk assessment method performs better than the text disease risk assessment method in terms of training time and stability.

Disease risk assessment models and intelligent identification based on text data and multimodal data fusion, while addressing the versatility of disease risk assessment, require more integration expertise in the application of disease risk assessment. The study model improves the performance of risk assessment when dealing with associated diseases, diseases with similar symptom characteristics. Although the tensor convolution self-encoding neural network model takes into account the correlation between image elements of different layers of medical image data, more in-depth research is needed. The application of tensor convolutional self-encoding neural network models in other fields remains to be studied.

REFERENCES

[1] P. Wei, Z. Zhou, L. Li, and J. Jiang, ''Research on face feature extraction based on K-mean algorithm,'' *EURASIP J. Image Video Process.*, vol. 2018, pp. 83–95, Dec. 2018.

[2] H. G. Hong, M. B. Lee, and K. R. Park, ''Convolutional neural network-based finger-vein recognition using NIR image sensors,'' *Sensors*, vol. 17, no. 6, pp. 1297–1306, 2017.

[3] A. Jami and P. S. Heyns, ''Impeller fault detection under variable flow conditions based on three feature extraction methods and artificial neural networks,'' *J. Mech. Sci. Technol.*, vol. 32, no. 9, pp. 4079–4087, 2018.

[4] D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, ''Intelligent constellation diagram analyzer using convolutional neural network-based deep learning,'' *Opt. Express*, vol. 25, no. 15, pp. 17150–17166, 2017.

[5] L. Jing, T. Wang, M. Zhao, and P. Wang, ''An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox,'' *Sensors*, vol. 17, no. 2, pp. 414–421, 2017.

[6] S. Yu, Y. Cheng, S. Su, G. Cai, and S. Li, ''Stratified pooling based deep convolutional neural networks for human action recognition,'' *Multimedia Tools Appl.*, vol. 76, no. 11, pp. 13367–13382, 2017.

[7] D. Peralta, I. Triguero, S. García, Y. Saeys, J. M. Benitez, and F. Herrera, ''On the use of convolutional neural networks for robust classification of multiple fingerprint captures,'' *Int. J. Intell. Syst.*, vol. 33, no. 1, pp. 213–230, Nov. 2017.

[8] H. Yang, Y. Zou, Z. Wang, and B. Wu, ''A hybrid method for short-term freeway travel time prediction based on wavelet neural network and Markov chain,'' *Can. J. Civil Eng.*, vol. 45, no. 5, pp. 35–45, 2017.

[9] H. Yuan, X. Wang, X. Sun, and Z. Ju, ''Compressive sensing-based feature extraction for bearing fault diagnosis using a heuristic neural network,'' *Meas. Sci. Technol.*, vol. 28, no. 6, pp. 65018–65027, 2017.

[10] R.-X. Chen, X. Huang, and L.-X. Yang, ''Rolling bearing fault identification based on convolutional neural network and discrete wavelet tranform,'' *J. Vib. Eng.*, vol. 31, no. 5, pp. 883–891, 2018.

[11] C. Lu, Z. Wang, and B. Zhou, ''Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification,'' *Adv. Eng. Inform.*, vol. 32, pp. 139–151, Apr. 2017.

[12] J. Tang, D. Wang, Z. Zhang, L. He, J. Xin, and Y. Xu, ''Weed identification based on K-means feature learning combined with convolutional neural network,'' *Comput. Electron. Agricult.*, vol. 135, pp. 63–70, Apr. 2017.

[13] D. M. D'Addona, A. M. M. S. Ullah, and D. Matarazzo, ''Tool-wear prediction and pattern-recognition using artificial neural network and DNA-based computing,'' *J. Intell. Manuf.*, vol. 28, no. 6, pp. 1285–1301, 2017.

[14] B. Liu and J. Zheng, ''Partial discharge pattern recognition in power transformers based on convolutional neural networks,'' *High Voltage App.*, vol. 53, no. 5, pp. 70–74, 2017.

[15] W. Ding, J. Gu, Z. Shang, S. Tang, Q. Wu, E. A. Duodu, and Z. Yang, ''Semantic recognition of workpiece using computer vision for shape feature extraction and classification based on learning databases,'' *Optik*, vol. 130, pp. 1426–1437, Feb. 2017.

[16] P. Shi, K. Liang, N. Zhao, and S. An, ''Intelligent fault diagnosis for gears based on deep learning feature extraction and particle swarm optimization SVM state identification,'' *China Mech. Eng.*, vol. 28, no. 9, pp. 1056–1061, 2017.

[17] S. Dong, Z. Yuan, C. Gu, F. Yang, H. Fu, C. Wang, C. Jin, and J. Yu, ''Research on intelligent agricultural machinery control platform based on multi-discipline technology integration,'' *Trans. Chin. Soc. Agricult. Eng.*, vol. 33, no. 8, pp. 1–11, 2017.

[18] R. Chen, S. Chen, M. He, D. He, and B. Tang, ''Rolling bearing fault severity identification using deep sparse auto-encoder network with noise added sample expansion,'' *J. Vib. Eng.*, vol. 30, no. 5, pp. 874–882, 2017.

[19] L. Ruiqi, Z. Xian, Z. Luo, and L. Lin, ''Research on the intelligent judgment of traffic congestion in intelligent traffic based on pattern recognition technology,'' in *Cluster Computing*, 2018, pp. 1–8. doi: 10.1007/s10586-017-1684-8.

[20] Y. Zhou, T. Xu, W. Zheng, and H. Deng, ''Classification and recognition approaches of tomato main organs based on DCNN,'' *Trans. Chin. Soc. Agricult. Eng.*, vol. 33, no. 15, pp. 219–226, 2017.

[21] D. Wang, M. Zhang, Z. Li, J. Li, M. Fu, Y. Cui, and X. Chen, ''Modulation format recognition and OSNR estimation using CNN-based deep learning,'' *IEEE Photon. Technol. Lett.*, vol. 29, no. 19, pp. 1667–1670, Oct. 1, 2017.

[22] H. Boughrara, M. Chtourou, C. B. Amar, and L. Chen, ''Facial expression recognition based on a mlp neural network using constructive training algorithm,'' *Multimedia Tools Appl.*, vol. 75, no. 2, pp. 709–731, 2016.

[23] X. Qu, W. Wang, K. Lu, and J. Zhou, "Data augmentation and directional feature maps extraction for in-air handwritten Chinese character recognition based on convolutional neural network," *Pattern Recognit. Lett.*, vol. 111, pp. 9–15, Aug. 2018.

[24] X. Wang, C. Chen, Y. Cheng, and Z. J. Wang, "Zero-shot image classification based on deep feature extraction," *IEEE Trans. Cogn. Develop. Syst.*, vol. 10, no. 2, pp. 432–444, Jun. 2018.

[25] H. Lu, X. Fu, C. Liu, L.-G. Li, Y.-X. He, and N.-W. Li, "Cultivated land information extraction in UAV imagery based on deep convolutional neural network and transfer learning," *J. Mountain Sci.*, vol. 14, no. 4, pp. 731–741, 2017.

[26] B. Liu, X. Yu, P. Zhang, A. Yu, Q. Fu, and X. Wei, "Supervised deep feature extraction for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1909–1921, Apr. 2018.

[27] D. T. Nguyen, K. W. Kim, H. G. Hong, J. H. Koo, M. C. Kim, and K. R. Park, "Gender recognition from human-body images using visible-light and thermal camera videos based on a convolutional neural network for image feature extraction," *Sensors*, vol. 17, no. 3, pp. 637–648, 2017.

[28] B. Yong, G. Zhang, H. Chen, and Q. Zhou, "Intelligent monitor system based on cloud and convolutional neural networks," *J. Supercomput.*, vol. 73, no. 7, pp. 3260–3276, 2018.

WEIDONG LIU graduated from the Guangzhou University of Traditional Chinese Medicine, in 2012, majoring in orthopedics. His research interests include the application of wearing equipment and data mining in medical treatment.

CAIXIA QIN graduated from the Guangzhou University of Traditional Chinese Medicine, in 2012. She is currently with Shenzhen Luohu Hospital Group Luohu People's Hospital. She is good at using TCM syndrome differentiation method for disease treatment.

KUN GAO received the Ph.D. degree from Southern Medical University. He is currently with the Shenzhen Hospital of Traditional Chinese Medicine. He is good at the research of orthopedic diseases treated by traditional Chinese medicine.

HENG LI graduated from the Guangzhou University of Traditional Chinese Medicine. He is currently a Doctor of medicine with the Shenzhen Hospital of Traditional Chinese Medicine. He is good at the research of orthopedic diseases treated by traditional Chinese medicine.

ZUEN QIN received the bachelor's degree from the Guangxi University of Chinese Medicine. He is currently with Hechi Traditional Chinese Medicine Hospital. He is good in integrated traditional Chinese and western medicine in the treatment of orthopedic diseases, specializing in orthopedic surgery, especially spinal, osteopathy, joint, and other diseases diagnosis and treatment.

YAFEI CAO graduated from the Guangzhou University of Traditional Chinese Medicine. He is currently a Professor, the Chief Physician, and a Doctor of medicine. He is good at the diagnosis and treatment of orthopedic diseases and has some research on the medical treatment under the Internet and big data.

WEN SI received the Ph.D. degree from Shanghai University, in 2011. He was a Postdoctoral Research Associate with Fudan University's Rehabilitation Medicine, Huashan Hospital. He joined Shanghai Business School, in 2011, where he is currently an Associate Professor with the Internet of things (IoT) Engineering Department. His research interests include biomedical engineering and the Internet of things technologies. His experiences mainly focused in technical scheme about human motion detects, including the tri-axial force on foot–ground interface and human-body model of movement, which may enhance the effect of recovery training.