

Received December 19, 2019, accepted January 13, 2020, date of publication January 17, 2020, date of current version January 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2967075

Medical Multimedia Big Data Analysis Modeling Based on DBN Algorithm

YING YANG^{ID}

School of Management, Xinxiang Medical University, Xinxiang 453003, China

e-mail: 041121@xxmu.edu.cn

ABSTRACT With the development of medical multimedia analysis methods based on DBN, DBN models have gained the ability to surpass medical experts in the evaluation of multimedia in some clinical examinations. Firstly, based on the existing architecture of the Internet of Things, combined with the actual characteristics of the hospital, the medical multimedia data is accessed from the IoT support platform. Secondly, the medical multimedia data modeling and classification method based on DBN is studied and analyzed. Three network structure models, a deep belief network, a stacking automatic encoder and a convolutional neural network, were introduced and analyzed. The medical multimedia data classification modeling method based on DBN was proposed to further improve the accuracy of medical multimedia data classification. The experimental results show that compared with the traditional feature extraction based neural network classification method, the classification performance is better. Thirdly, the medical state assessment model is constructed based on the multivariate Gaussian distribution theory. To study how to use the multivariate Gaussian distribution theory to design an evaluation model that can evaluate the health status of users efficiently and accurately. Finally, using the MATLAB software platform, through the experiment and simulation of 40 groups of 8×8064 -dimensional physiological big data of 32 volunteers, First, determine the optimal parameters of a set of health assessment models; then use the model to learn the characteristics of physiological parameters; finally, the state assessment model to obtain the health assessment results. The experimental results show that the feature learning model based on convolutional neural network theory can effectively extract the deep features of medical multimedia big data. The health state assessment model based on multivariate Gaussian distribution theory can effectively evaluate the health status of human body.

INDEX TERMS DBN, medical multimedia, neural network, feature extraction, multivariate Gaussian distribution.

I. INTRODUCTION

Medical multimedia big data applications are very broad and cover almost every life process of each of us. With the continuous breakthrough of big data theory, the technology related to large numbers is becoming more mature, and the industries related to large numbers are constantly being formed and improved. Big data will penetrate all aspects of all walks of life, reshape people's lifestyle and seek medical treatment. Big data has gradually penetrated into all aspects of public life and social development and has become an extremely important production factor. The application of big data in the field of medical health big data is also in full swing, which is bound to have a huge impact on it. There are many types

of medical health big data, and the source is complicated. Because it is closely related to people's life processes, data from the medical industry is more important than data from other fields.

One of the most successful cases in the field of medical health is the special project of heart disease research [1]–[3]. Project researchers have long tracked the heart data of a fixed group and then analyzed the data using big data technology. Excavate the causes of heart disease, and finally develop relevant countermeasures based on specific reasons. Using big data technology modeling analysis is to explore key risk characteristics that lead to disease, including body index and personal habits: blood lipids, blood pressure, diabetes, weight, smoking, eating habits and exercise. Literature [4]–[6] clinical data using machine learning to conduct mining analysis, the data includes physiological parameters and clinical

The associate editor coordinating the review of this manuscript and approving it for publication was Kim-Kwang Raymond Choo^{ID}.

diagnosis and treatment records and other data, and then use the multivariate Gaussian model to model these data, and finally the patient's condition is serious Degree of scientific risk assessment. The accuracy of the experimental results has been greatly improved compared with the traditional method, and the prediction error results from 0.91 to 0.69. By establishing a data-driven analysis model for medical multimedia data [7]–[10], it can not only avoid the misdiagnosis or delay of treatment caused by accidental mistakes of doctors and nurses, but also can treat and recover medical institutions. The care plan provides scientific data support, more accurate and efficient diagnosis of some fast-developing conditions, and rapid scientific decision-making. In the big data technology based on artificial intelligence and machine learning, massive data sources are necessary to utilize big data analysis. Only with sufficient data to support the data analysis method can we provide users with a good personalized solution to improve their living habits and lifestyle [11]–[13]. In order to obtain good research results, many research institutes pay great attention to the quality of data in the process of researching genes [14]–[16]. In theoretical research, DBN in time series data analysis and prediction, speech analysis, image recognition many important achievements have been made in other fields. In the literature [17]–[19], many basic concepts related to DBN, the network architecture model and pre-training methods commonly used in DBN are briefly introduced, and the data types suitable for different network models are respectively explained. References [20]–[22] the research focuses on the number of hidden nodes in the model. The experimental results show that a large number of hidden nodes and dense feature extraction units in a single-layer network are essential for implementing high-performance algorithms, and their effects are even comparable to the depth of the algorithm model. DBN In the study of human emotion analysis, four deep belief network (DBN) models are described in the literature [23]–[25], and the robust multimodal features for sentiment classification are studied in an unsupervised manner. The DBN model performs better than the existing sentiment classification method. The paper also proposes a convolution deep confidence network (CDBN) model to learn the key multimodal features of expressing emotions. Using the DBN model to learn features [26]–[29], the feature input recursive neural network for sentiment classification, the classification performance and classification effect reached the level of manual classification, and the efficiency is much higher than the manual classification.

The combination of medical health big data and DBN to create a new model of medical health management has become a new focus in academia. Based on the existing architecture of the Internet of Things, combined with the actual characteristics of the hospital, the medical multimedia data is accessed from the IoT support platform. Secondly, the medical multimedia data modeling and classification method based on DBN is studied and analyzed. The structural model—depth belief network, stack automatic encoder and convolutional neural network are analyzed and introduced, and

the medical multimedia data classification modeling method based on DBN is proposed to further improve the accuracy of medical multimedia data classification. The experimental results prove compared with the traditional neural network classification method based on feature extraction, the classification performance is better. Thirdly, the medical state assessment model is constructed based on the multivariate Gaussian distribution theory. Study how to use the multivariate Gaussian distribution theory to design an evaluation model that can efficiently and accurately assess the health status of users.

The rest of our paper was organized as follows. Analysis of the Collection and DBN of Medical Multimedia Data was introduced in Section II. Section III described Research on medical multimedia big data classification model based on DBN. Experimental results and analysis were discussed in detail in Section IV. Finally, Section V concluded the whole paper.

II. DBN ANALYSIS OF MEDICAL MULTIMEDIA DATA

A. IoT TECHNOLOGY IN THE MEDICAL FIELD

Remote monitoring and home care monitor the physiological parameters of the monitored subjects and perform emergency rescue in the event of an accident. They belong to the telemedicine field and inevitably use WSN technology. Many literatures have introduced concepts such as body area network, sensor Internet of Things or wireless body area sensing Internet of Things. It is actually a wireless network formed by human physiological information, collecting sensors or biosensors implanted in the human body. One dollar of networking needs to be perceived and componentized. According to the existing projects and examples, the remote monitoring and medical data collection framework based on the Internet of Things technology is summarized—the three-level hierarchical structure, which is the sensing layer, the transport layer and the application layer, as shown in Figure 1.

1) PERCEPTUAL LAYER

The physiological signal or the environment information of the monitoring object is obtained mainly through the sensor node. According to the position of the sensor node in the human body, it can be divided into an implanted sensor implanted in the body, a wearable sensor worn on the body, and an ambient sensor that can recognize human behavior around the body class [30]. The sensor node not only monitors continuous physiological signals (such as EEG, ECG, EMG) and discrete physiological signals (such as body temperature, blood pressure), but also monitors the movements and activities of the human body.

2) TRANSPORT LAYER

The transport layer can be connected to the short-distance network Bluetooth, Zigbee, Wi-Fi, or remote network GSM, GPRS, 3G. The primary function is to store, receive, and

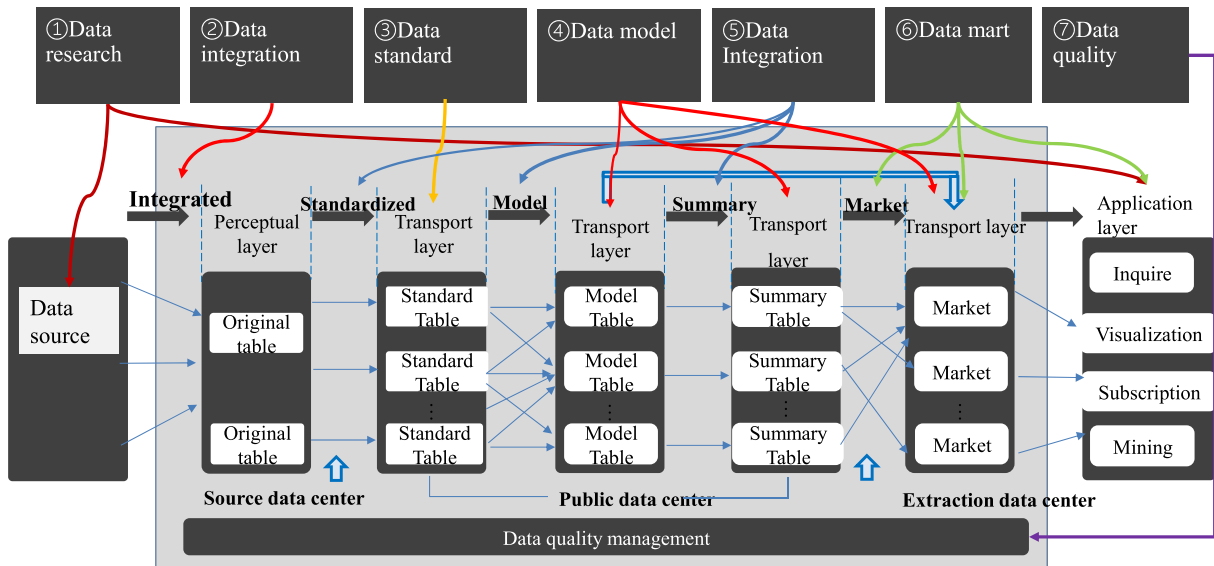


FIGURE 1. General architecture of medical multimedia data collection based on internet of things technology.

analyze data collected from the sensing layer and communicate with external networks.

3) APPLICATION LAYER

Mainly refers to the remote server that provides various application services (such as the storage of physiological parameters or medical data of the monitoring object, the analysis and prediction of the patient’s condition (or health condition) by the expert, and the remote access of the medical data of the monitoring object by the medical staff). And its external network, including: medical database server, medical transaction management server, various dedicated telemedicine analysis devices, various mobile communication terminals (such as mobile phones and PDAs), as well as doctors and experts.

B. DBN ALGORITHM

DBN is to learn the final data feature representation by gradually learning the characteristics of the network through the hidden layer. The learning ideas are summarized as follows:

(1) Due to the complexity of its network, deep network has a multi-hidden layer network structure and has a strong feature learning ability, which can better represent the features obtained by the learned data objects.

(2) Deep networks can be very complicated in training network parameters due to the complexity of their networks. This can use the unsupervised idea to learn the parameters of the network by using “layer-by-layer initialization” training.

In the DBN network, suppose there is a model S, which has n layers (S1,... Sn). The input of the network is I, the output is O, and the calculation process is visually represented as: $I \Rightarrow S1 \Rightarrow S2 \Rightarrow \dots \Rightarrow Sn \rightarrow O$, if the network predicts. It is equal to the last original data I, that is, the input data I has no data information loss after undergoing a series of transformations in the model. The representation of any layer of Si in the network is another valid representation of the original data.

In the automatic deep hierarchical feature learning of DBN, it is assumed that there is a set of input training set I, and the parameters of the network model are adjusted by learning to make the output of the network as close as possible to the input I. In this way, a series of features of the original data can be learned, namely S1,..., Sn. For DBN, the core idea is to stack multiple nonlinear computing layers and use the output of the previous layer as the input to the next layer. Through such a layer-by-layer learning process, a deep representation of the final raw data is obtained.

In addition, the algorithm assumes that the output of the model is equal to its input. In fact, in order to improve the universality of the network and the training speed of the network, it is only necessary to make the reconstructed data of the network close to the real input. In DBN, the basic models are mainly: Auto en-coder and its various encoding networks, first Restricted Boltzmann Machine (RBM, Deep Confidence Network and Convolutional Neural Network, etc.) The core ideas and algorithm flow of each network will be specifically described below.

1) AUTOMATIC CODING

Automatic coding is the simplest network model in DBN. It is a deep network designed by the hierarchical structure of artificial neural networks. In the network, the cost function is constructed with the error of the input data of the network and the actual output data, so as to minimize the objective function training network weight parameter and make the network model reach the optimal state. Based on this idea, through continuous iteration, the deep feature representation of the original data, that is, the characteristics of the original data, is obtained.

Suppose there is now an unmarked training sample set $\{x^{(1)}, x^{(2)}, \dots, x^{(i)}\} \in R^n$. The input signal $x \in [1, 0]^d$ first passes through the autoencoder to a representation of the

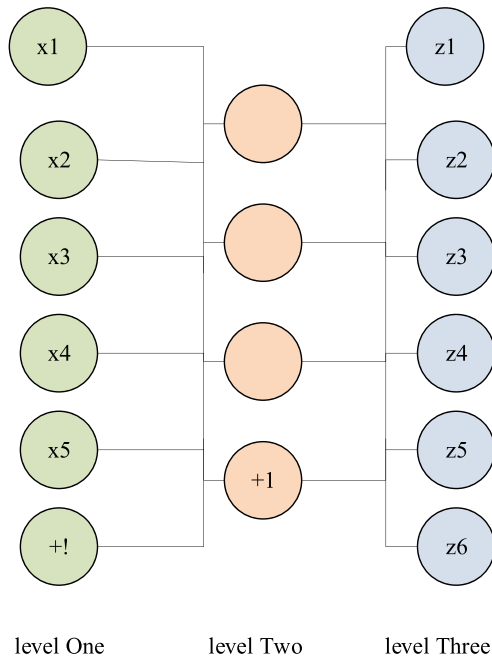


FIGURE 2. Automatic coding structure diagram.

hidden layer. $y \in [1, 0]^{d'}$ can be determined by the following encoding expression:

$$y = s(Wx + b) \tag{1}$$

where S is a nonlinear differentiable function sigmoid, the hidden layer indicates that Y reconstructs a signal z of the same size and size as x by decoding. The same goes through a mapping decoding expression similar to encoding:

$$z = s(W'y + b') \tag{2}$$

The process by which an automatic encoder converts an input space into a new distribution expression by determining the parameters of the encoding function is called encoding. As shown in FIG. 2, the automatic encoder determines the parameters of the decoding function by reducing the reconstruction error, and the process of learning how to reconstruct the output into the input signal space is called decoding, and the encoding parameters of the automatic encoder are also used to construct an input signal. Finally, if the resulting reconstructed signal is similar to the input signal, the hidden layer Y can be seen as another representation of the input, that is, the feature.

Here x is the original input signal, z is the reconstructed signal, and Figure. 2 is a specific reconstruction diagram.

2) STACK AUTOMATIC CODING

The stack self-encoding neural network is a network model composed of a number of automatic encoder positions. In DBN, stack self-encoding neural networks belong to unsupervised neural networks. The output of its lower layer self-encoder in the network acts as the input to the subsequent higher layer self-encoder. The network weight and

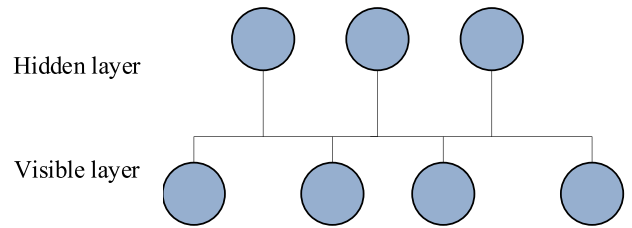


FIGURE 3. RBM structure diagram.

offset parameters corresponding to the k th self-encoder are respectively represented by $W^{(k,1)}, W^{(k,2)}, b^{(k,1)}, b^{(k,2)}$; then, the encoding process of the stack self-encoding neural network trains the step expression algorithm of each layer of network parameters in a front-to-back order:

$$\begin{aligned} \alpha^{(l)} &= f(z^{(l)}) \\ z^{(l+1)} &= W^{(l,1)}\alpha^{(l)} + b^{(l,1)} \end{aligned} \tag{3}$$

Conversely, the decoding operation of each layer of the automatic encoder is performed in a backward-to-forward order, and the algorithm expression is:

$$\begin{aligned} \alpha^{(n)} &= f(z^{(n)}) \\ z^{(l+n+1)} &= W^{(n-l,1)}\alpha^{(l)} + b^{(n-l,1)} \end{aligned} \tag{4}$$

Among them, $a^{(n)}$ is the output value of the highest hidden network, which is the high-order feature of the original data learned by the model, which is a more abstract representation of the data. By using the vector $a^{(n)}$ as the input data for various classifiers, using the features learned by the network for the classification problem can achieve better results than the original data as a classifier input.

3) RESTRICTED BOLTZMANN MACHINE

A restricted Boltzmann network consists of a layer of visible elements and a layer of hidden elements. It essentially belongs to a kind of generated random neural network. The restricted Boltzmann machine network structure is a bipartite graph. There is no connection between the network nodes of the same layer. Only the neurons between the layers can be seen between the layers and the hidden layers. RBM is a neural network algorithm with unsupervised learning. Traditional neural network use generally uses manual extraction of data features, and the network is prone to fall into local optimal solutions. If the back propagation algorithm is used to train the parameters of the network, the classification effect becomes less and less ideal as the number of network layers increases. Therefore, the generated model in the probability map model is introduced into the restricted Boltzmann machine ((RBM) network structure, and the required features are automatically extracted directly from the training set, and the weight of the neural network can be well learned. The problem of manual feature extraction is solved due to personal subjective considerations, etc. The structure diagram of RBM is shown in Figure 3.

The constrained Boltzmann machine is essentially an energy-based neural network model whose energy function between the visible layer variable V and the hidden layer variable h is expressed as:

$$E(v, h : \theta) = - \sum_{ij} W_{ij} v_i h_j - \sum_i b_i v_i - \sum_j a_j h_j \quad (5)$$

The restricted Boltzmann network has two layers, an input layer and a hidden layer. The input layer is also called the visible layer, and the other layer is the hidden layer, which is the feature learning layer. The calculation of the weight of the network is obtained by minimizing the cost function.

4) DEEP TRUST NETWORK

The deep confidence network consists of several RBM network layers superimposed, so the DBN is a Bayesian probability generation model. According to the idea of DBN, the training set data is input into the first layer RBM network, and the calculated output is used as the input of the second layer RBM network to train the network parameters. The DBN is to stack the RBM model by adding layers to improve the network.

DBN is stacked by multiple restricted Boltzmann machines (RBMs) and is an important multi-layer network structure in the DBN network model. Its network training method uses semi-supervised learning algorithm: RBM network weight and offset value initialization Trained by a greedy layer-by-layer unsupervised algorithm, then these weights and offsets are fine-tuned by a supervised backpropagation algorithm. As a DBN model of machine learning, the deep belief network can not only realize autonomous feature learning, but also has more essential characterization of the original data due to the structure of its network.

III. RESEARCH ON MEDICAL MULTIMEDIA BIG DATA CLASSIFICATION MODEL BASED ON DBN

In this section, a deep artificial neural network feature learning model for physiological data is designed using the convolutional neural network in the DBN algorithm. At the same time, the flow of the algorithm will be described in detail.

A. DBN-BASED CLASSIFICATION MODEL

The feature is the use of machine learning to focus on the actual application, the accuracy of the feature directly affects the application effect. The process of feature learning mainly includes training of parameters based on convolutional neural network models and learning of new data features. First, the raw data is preprocessed, and the data of different dimensions are normalized to the same interval. Then, the training set is used to train the parameters of the convolutional neural network in an unsupervised manner, and a set of optimal is obtained through continuous iteration. The parameters enable the characteristics learned by the network to better represent the original data. Finally, the trained network model is used to

learn the feature representation of the new data to prepare for the subsequent medical multimedia big data modeling health assessment.

1) DATA PREPROCESSING

One of the most important steps in the analysis of processed data using DBN algorithms is data preprocessing. In practical applications, when using DBN algorithms to learn data features, pre-processing the original feature data and then performing feature learning can improve the learned feature quality.

The first step in data preprocessing is to normalize the feature data. There are many ways to normalize data, and the specific normalization method usually chooses different normalization methods according to the specific application background of the data. In general, there are several common methods for normalizing feature data:

(1) Simple scaling: After the normalization method re-transforms the feature data of each dimension of the data, the goal is to make the value of the feature data fall within the interval of $[0, 1]$ or $[-1, 1]$. In the data mining, the subsequent predecessor of the feature data is very important, because some algorithms such as epsilon in Pca-whitening default to the feature data is regular and scaled into the specified interval.

(2) Minimum-maximum normalization: The minimum-maximum normalization is a simple linear scaling if min and max represent the minimum and maximum values of the dataset features, respectively. The minimum one maximum normalization is calculated by:

$$v' = \frac{v - \min}{\max - \min} \quad (6)$$

where max is the maximum value of the original set data, and min is the minimum value of the original set data. However, this standardized algorithm also has some disadvantages. When there is new data, it is possible that the changes of the maximum value max and the minimum value min need to be re-discovered the maximum and minimum values and then recalculated. The essence of minimum-maximum normalization is a simple linear transformation of the feature set, thus maintaining the correlation between the transformed data and the original data.

(3) Sample-by-sample mean subtraction: The sample-by-sample mean subtraction algorithm refers to subtracting the statistical average of the data set for each sample in the data set. When the statistics of each dimension of the data set adopted by the algorithm obey the same probability distribution, the data can be preprocessed by the method of sample-by-sample averaging.

(4) Feature Standardization: Feature normalization is intended to have features with zero mean and unit variance in each dimension. This normalization method is widely used in data and preprocessing. In the actual standardized calculation, the detailed process of feature standardization is: first, calculate the mean of the data set in each dimension, and

subtract the average value from the data of each dimension; then, divide by the standard of the data on the dimension difference. The specific calculation method is:

$$v' = \frac{v - u}{\sigma_A} \quad (7)$$

Among them, u, σ_A is the mean and standard deviation, The z-score normalization algorithm is applicable to the case where the maximum and minimum values of the attribute A in the data set cannot be accurately obtained, or the data noise is too large, exceeding the normal value.

When using PCA/ZCA whitening, the input data is first zero-averaged. If $x^{(i)}$ is a feature of the data set, the feature zero-averaging algorithm is:

$$\frac{1}{m} \sum_i x^{(i)} = 0 \quad (8)$$

Data preprocessing is the key to the quality of feature learning. It is suitable for data preprocessing methods to achieve better results even in simple feature learning algorithms. The original data set after data preprocessing can be input into the feature representation of the learning data in the feature learning network.

B. DBN MODEL BASED ON CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network is a kind of deep artificial neural network, because its superiority has become a research hotspot in the field of speech recognition, image recognition and pattern recognition in practical applications. It has many advantages. The weight sharing method of its network structure makes it more similar to the actual working principle of biological neural network, which greatly reduces the number of network weights, reduces the complexity of the network model, and speeds up the network. Calculate the speed. The advantage of the weight sharing method is more prominent when the high-dimensional data is used as the input of the convolutional neural network, and the original data can be used as the input data of the neural network without transformation. This working mechanism can avoid the traditional machine learning algorithm. The process is of tedious feature learning and data reconstruction. Feature learning in the context of health assessment refers to the use of machine learning algorithms to transform raw medical multimedia data into layers to obtain a set of input data processes suitable for use in health assessment models. Typically, features learned from a set of contiguous data are generally statistical features such as the mean, standard deviation, and frequency domain characteristics of the original signal or the normalized signal in the time domain. The physiological timing of big data health status assessment focuses on high-dimensional data learning data characteristics. This paper focuses on the use of convolutional neural network methods to automatically learn the deep feature representation of data from physiological time series big data in an unsupervised learning manner, and then used for health status assessment.

C. HEALTH STATE ASSESSMENT MODEL BASED ON MULTIVARIATE GAUSSIAN DISTRIBUTION

If the random variable x conforms to a Gaussian distribution, its corresponding probability density function can be expressed as:

$$P(x, u, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp - \frac{(x - u)^2}{2\sigma^2} \quad (9)$$

In the unary Gaussian model, a method of calculating the probability corresponding to each feature is taken, and then all the probability values are multiplied to obtain $P(x)$. However, if there are two feature correlation coefficients, and the range of the two features is relatively large, in this case, the ordinary Gaussian probability distribution model usually cannot fit the actual distribution state of the feature. Usually, the multivariate Gaussian distribution model will try to find the difference between the two features at the same time, so it may find a relatively large decision boundary, which leads to the model not fitting the feature distribution well. The Gaussian distribution model is used to fit the feature distribution of high-dimensional medical multimedia big data.

The algorithm flow can be generally divided into the following three steps:

(1) Model establishment, according to the actual application, select the appropriate Gaussian model, and then establish a Gaussian probability distribution function;

(2) Model training, input the characteristic data learned by the network into the network, calculate the parameters related to the Gaussian probability distribution function, and obtain a probability model. Based on the probability distribution of the original input data, establish a health state assessment model in the form of partitioned intervals.;

(3) Model prediction, giving prediction results. For the new data, the probability distribution is calculated according to the trained network model; then, by dividing the probability interval, the health state level to which the feature point belongs is found.

Since most people are in a healthy state in general, only a few cases of medical multimedia data will become abnormal. Based on this idea, the paper designs a method to divide the different feature health state intervals by constructing a health state assessment model by dividing the percentage of feature points and total feature numbers into a certain probability interval. Its specific algorithm is expressed as:

$$\text{if } \begin{cases} 0 < p(x) \leq \varepsilon_0, A \\ \varepsilon_0 < p(x) \leq \varepsilon_1, B \\ \varepsilon_1 < p(x) \leq \varepsilon_2, C \\ \varepsilon_2 < p(x) \leq \varepsilon_3, D \end{cases} \quad (10)$$

where $\varepsilon_n, n = 1, 2, 3 \dots$ is the probability threshold, and the division rule is a percentage of the number of features in a certain probability interval according to the feature point probability $p(x)$, and the smaller the probability of the feature, the health state to which the feature point belongs. The higher is the hazard level, the worse its health status.

Using the K-means algorithm, first cluster the feature data to find the optimal number of cluster centers, and replace the number of health states with the number of cluster centers.

- (1) The first step randomly selects k cluster centroid points;
- (2) In the second step, calculate the distance from each point to the cluster center, and cluster each point into the cluster closest to the point;
- (3) The third step calculates the coordinate average of all points in each cluster, and uses this average as the new cluster center;
- (4) Repeat 2, 3 until the cluster center does not move over a large range or the number of clusters reaches the required level, then the algorithm converges.

Through the clustering algorithm, the optimal number of cluster centers is determined, and the percentage of the number of features in each category and the total number of features is calculated accordingly. According to the percentage, the threshold of the partition probability interval is determined.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL PARAMETER SETTING AND DBN MODEL

Considering that DBN network structure and parameter configuration could not find a reasonable reference, the experiment in this paper can only use multiple methods of comparison. In the Windows 7 64-bit operating system, the experiment was performed using MATLAB 7.8 at a frequency of 3.1 GHz and a memory of 4 GB. First set the DBN parameters Pre-training and Fine-tuning are both 0.05, Mini-batch size is 128. In the experiment, in order to determine the DBN network structure suitable for this experiment, different hidden layer numbers and hidden layer nodes are selected to test, so as to determine the DBN structure suitable for medical multimedia data classification. For the convenience of comparison, DBN in the experiment similar structure is adopted for SAE, CNN is set separately, and the artificial neural network structure used is consistent with the DBN structure in the experiment in terms of the number of input and output neurons, the number of neurons in the hidden layer, and the number of iterations. The transfer function uses the training function, the learning rate is 0.1, the training error is 0.1, the learning rule uses the gradient descent method, and the error back propagation algorithm is used in the study to reduce the weight error.

The first layer 1 contains 12 convolution operations with a feature map layer of size 60×60 and 12 down sampled subsample layers of size 30×30 . The size of the convolution kernel determines the size of each local receptive field, so selecting an appropriate size convolution kernel is critical for DBN feature extraction.

The second layer 2 is similar to the Layer 1 structure, but there are also differences. First, Layer 2 also includes convolution operations and down sampling operations. Second, because of the same operation, Layer 2 also includes

a feature mapping layer and a subsampling layer. Third, the Layer 2 convolution kernel has the same moving step size as Layer 1, which is 1 pixel. But the convolution kernel size in Layer 2 is 3×3 . Secondly, Layer 2 is different from Layer 1 input. Layer 2's convolution operation input is Layer 1 down sampled and has a size of 30×30 .

The third layer 3 has a similar structure to Layer 1 and Layer 2, except that the convolution input of Layer 3 is down sampled by Layer 2 and has a size of 14×14 . The number of convolution kernels is 72, the training parameters are $(3 \times 3 + 1) \times 72$, and the number of network connections $(3 \times 3 + 1) \times 72 \times 12 \times 12$.

The fourth to sixth layers are convolutional layers, and each layer has 120 feature mapping layers, which are all obtained by a convolution kernel of size 3×3 .

The seventh to eighth layers are fully connected layers, wherein the seventh, eighth and ninth layers each have 240, 100, 10 neurons, of which between the seventh and sixth layers, the seventh and eighth layers, the eighth layer and the ninth layer are all fully connected. Finally, the calculation result is input into the Sigmoid function, and a state output is generated.

In the experiments in this paper, the automatic coding model of deep learning is used for verification. Auto Encoder is a neural network that targets the output to maximize the reconstruction of input data. Generally, in an automatic coding network, the input of data is weighted by the network, that is, after the mapping of the activation function, a set of output data y is obtained, and then the weighted inverse mapping of y is performed to reconstruct the input as the target to obtain the data z . Taking the error function of the reconstructed data z and the input data Input as a cost function, the iterative training is used to adjust the weight of the network, so that z is close to the input and the error is minimized. In this training process, the network does not need to use the tag data to calculate the actual output between the data and the tag, but to reconstruct the feature to train the optimal network parameters, so the automatic coding is an unsupervised learning network.. Suppose there is now an unmarked training sample set $x^1, x^2, \dots, x^t \in R^n$. The input signal $x \in [1, 0]^d$ is first passed through an automatic encoder. A representation of the hidden layer is $y \in [1, 0]^{d'}$, whose mapping can be determined by the following encoding expression:

$$y = s(Wx + b) \quad (11)$$

where S is a nonlinear differentiable function sigmoid σ the hidden layer indicates that Y reconstructs a signal z of the same size and size as x by decoding. The same goes through a mapping decoding expression similar to encoding:

$$z = s(W'y + b') \quad (12)$$

In order to reconstruct the error Jw before adjusting the network weight, it is also necessary to converge the original signals of different dimensions, and the z -score normalizes

TABLE 1. Impact of DBN network depth on classification results.

| | Network structure | Acc(%) | Mse | Sen (%) | Spec(%) |
|-----|-------------------------------|--------|--------|---------|---------|
| DBN | 1024-100-150 | 0.8086 | 0.6837 | 0.7667 | 0.7631 |
| | 1024-100-150-200-6 | 0.8464 | 0.6535 | 0.7915 | 0.7636 |
| | 1024-100-150-200-300-6 | 0.8564 | 0.5392 | 0.8067 | 0.8067 |
| | 1024-100-150-200-300-600-6 | 0.8737 | 0.2895 | 0.8434 | 0.8134 |
| | 1024-100-150-200-300-600-1000 | 0.8957 | 0.1638 | 0.7664 | 0.8795 |
| SAE | 1024-100-150 | 0.8137 | 0.6016 | 0.7829 | 0.7568 |
| | 1024-100-150-200-6 | 0.8367 | 0.6138 | 0.8972 | 0.7596 |
| | 1024-100-150-200-300-61 | 0.8734 | 0.5261 | 0.8937 | 0.8196 |
| | 1024-100-150-200-300-600-6 | 0.8794 | 0.3358 | 0.8125 | 0.8284 |
| | 1024-100-150-200-300-800-1000 | 0.8876 | 0.2046 | 0.8337 | 0.8551 |
| ANN | 1-2C-2S-3C-3S-3C-3S- | 0.8864 | 0.5282 | 0.8367 | 0.8035 |
| | SC-SC-SC-10-10-6 | 0.8916 | 0.4834 | 0.8419 | 0.8322 |
| | 1-4C-4S-12C-12S-15C-15S- | 0.9037 | 0.3361 | 0.9237 | 0.8638 |
| | 10C-10C-10C-20-12-6 | 0.9104 | 0.2561 | 0.8972 | 0.8885 |
| | 1-8C-8S-24C-24S-30C-30S- | 0.9118 | 0.2557 | 0.8937 | 0.9282 |

the original signal to the convolutional neural network:

$$X_{norm} = \frac{x - \mu}{std(x)} \tag{13}$$

Mu and std(x) represent the mean and standard deviation of the original signal, respectively. Then, the normalized standard signal is sent to the characteristic of the learning signal of the convolutional neural network, and the relationship between the error JW and the number of iterations is obtained. By analyzing the weight between the original signal X and the reconstructed signal Z, a group can be obtained. Better characteristics of the original signal.

B. COMPARATIVE ANALYSIS OF EXPERIMENTAL RESULTS

Based on different network structure DBN, network depth increase, its modeling ability can be greatly improved, can extract more abstract data feature representation inside the data, further improve the overall Classification performance.

In Table 1, different network depths and number of hidden layer nodes are set to explore the network structure model suitable for medical multimedia data classification. It can be seen from Table 1 that the network structure models most suitable for DBN, SAE and ANN are {1024-100-150-200-300-600-1000fi, 1024-100-150-200-30000-1000-b, 1-12C respectively. -12S-36C-36S-72C-72S-120C-120C-120C-240-100}, when using this type of network structure model, the three classification indicators of Spec Sen and Acc are the highest respectively. However, compared with the three, ANN has the best performance on Acc, but Spec Sen and Mse is slightly lacking. DBN and SAE are basically the same. This is mainly due to the very similar two structural models of DBN and ANN.

As can be seen from Table 2-3, the classification performance is improved compared with DBN and ANN. In different classifications of BENI, MALI, CALC, NCALC, SPIC, NSPIC, DBN and ANN were compared and found that DBN

TABLE 2. Classification sensitivity results of ANN and DBN.

| Category | Spec | |
|----------|--------|--------|
| | ANN | DBN |
| BENI | 0.6741 | 0.8317 |
| MALI | 0.6635 | 0.8548 |
| CALC | 0.6358 | 0.8649 |
| NCALC | 0.6538 | 0.8467 |
| SPIC | 0.6564 | 0.8469 |
| NSPIC | 0.6182 | 0.8359 |

TABLE 3. Classification accuracy results of ANN and DBN.

| Category | Spec | |
|----------|--------|---------|
| | ANN | DBN |
| BENI | 0.7534 | 0.84317 |
| MALI | 0.7635 | 0.8648 |
| CALC | 0.7358 | 0.8249 |
| NCALC | 0.7538 | 0.8865 |
| SPIC | 0.7564 | 0.8272 |
| NSPIC | 0.7182 | 0.8954 |

has a significant improvement over ANN in terms of Classification sensitivity. In terms of Classification accuracy, DBN will also be better than ANN. Higher accuracy. This is mainly because DBN and ANN both use the deep structure network model.

Figure 4 shows the training error rate of the network. As the number of training increases, the overall trend shows a downward trend, among which CNN has the lowest error rate, followed by SAE and DBN, and finally ANN. The classification error rate is of ANN from the whole training trend. The fluctuations are large, the randomness is large, and the overall classification stability is poor. This is due to the randomness of the initial weight of the ANN and the fact that the whole network is easy to fall into the local optimum.

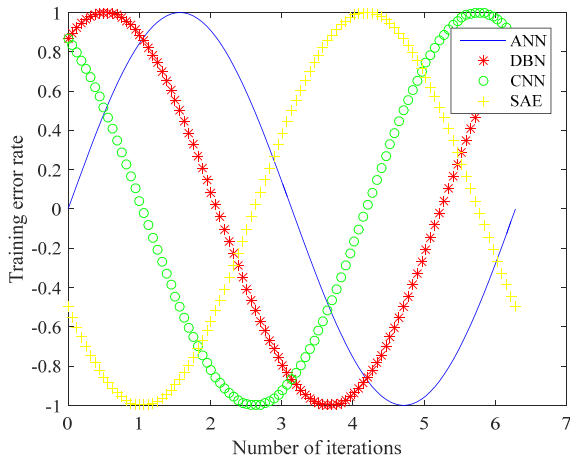


FIGURE 4. Curves of CNN, DBN, SAE and ANN errors with iterations.

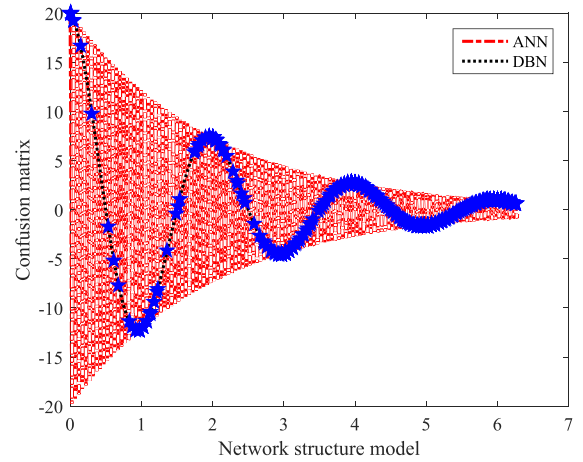


FIGURE 6. DBN, CNN network structure model confusion matrix in medical multimedia big data classification.

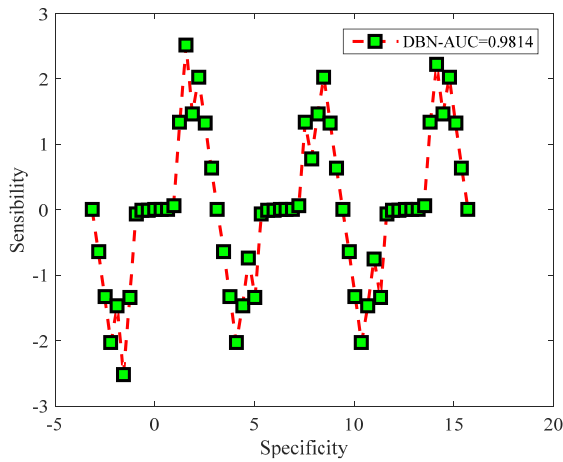


FIGURE 5. DBN network structure classification performance.

Figure 5 shows the curve of the classification performance of the DBN network structure and the area under the curve. It can also be seen that the classification performance of the DBN deep network structure will change with the change of the specification, and its performance peak value will reach 0.9814, which means that the classification performance of the DBN deep network structure is very high in the reasonable state. The DBN big data analysis model is also reasonable in terms of performance.

Figure 6 also shows that ANN and DBN have higher classification accuracy for each category. Therefore, in the MIAS anomaly classification, the DBN network structure model is significantly better than the traditional ANN algorithm.

Figure 7 shows the KAPPA bar, as can be seen from the figure, the yellow part indicates the value of the DBN algorithm. From the value of the y-axis, there will be a certain degree of improvement over the other three methods, which shows that the medical multimedia big data classification modeling based on DBN is better and more stable.

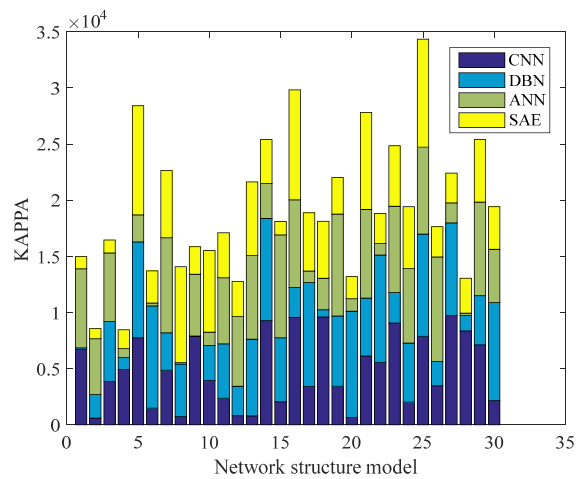


FIGURE 7. ANN, SAE, CNN, DBN network structure model KAPPA value bar.

TABLE 4. Comparison results of ASP of DBN, CNN and CDBN.

| Category | Acc | | |
|----------|--------|--------|--------|
| | SAE | ANN | DBN |
| 1000-200 | 0.8145 | 0.8558 | 0.8824 |
| 1500-250 | 0.8678 | 0.8678 | 0.9121 |
| 2000-300 | 0.8810 | 0.8859 | 0.9324 |
| 2500-350 | 0.8854 | 0.8924 | 0.9482 |
| 3000-400 | 0.8978 | 0.9134 | 0.9495 |
| 4000-500 | 0.8979 | 0.9214 | 0.9558 |

C. MEDICAL MULTIMEDIA DATA CLASSIFICATION BASED ON CONVOLUTION DEPTH BELIEF NETWORK

Table 4-5 shows the comparison of the results of the Acc and Mse experiments in the DBN, CNN and CDBN under the experimental parameter settings. The numbers under the “category” in the table represent the number of X-ray images of the breast tumor in the training set test set. Each result is obtained by 10-fold cross-validation. From the table, it can

TABLE 5. Mse comparison results of DBN, CNN and CDBN.

| Category | Mse | | |
|----------|--------|--------|--------|
| | DBN | CNN | CDBN |
| 1000-200 | 0.6124 | 0.5628 | 0.5238 |
| 1500-250 | 0.6028 | 0.5692 | 0.5112 |
| 2000-300 | 0.5977 | 0.5924 | 0.5021 |
| 2500-350 | 0.5821 | 0.5714 | 0.4918 |
| 3000-400 | 0.5943 | 0.5601 | 0.4805 |
| 4000-500 | 0.5931 | 0.5535 | 0.4348 |

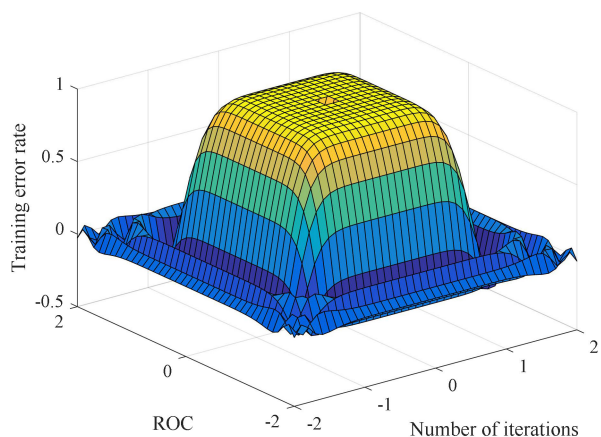


FIGURE 8. DBN training error with the number of iterations.

be found that the performance of CDBN is better than that of CNN. The performance of CNN is better than DBN, and each classification index basically increases with the increase of the number of images in the training set. Improvements have also shown that the DBN structure is highly dependent on big data training samples. Therefore, when the number of training images and the number of test images are 4000-500, respectively, each classification index reaches the best value (no additional training and testing is continued). The number of images is mainly due to the hardware limitations of the experimental conditions. One of the reasons for this is that due to the increased convolution operation and sparse regularization in the CDBN experiment, CDBN can better detect the details in the lung image.

Figures 8 and 9 show the ROC curves and training error rate curves of DBN in training images and test images. It can be seen that DBN has better performance in CT image classification of lungs, with the increase of iteration times. The training error rate was steadily reduced, and there was no over-fitting phenomenon.

D. NETWORK TRAINING AND FEATURE LEARNING VERIFICATION

As shown in Figure 10, as the number of iterations increases, J_w gradually stabilizes after zero, and the reconstructed signal is almost identical to the original signal. When J_w is small enough to approach and a decimal is, the convolutional neural network learns. The resulting feature is another efficient

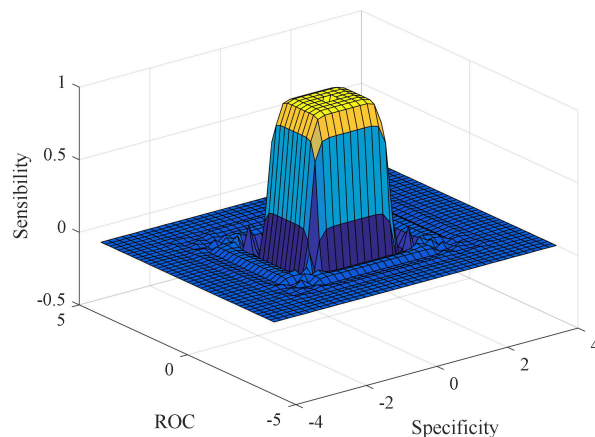


FIGURE 9. DBN network structure classification performance ROC curve.

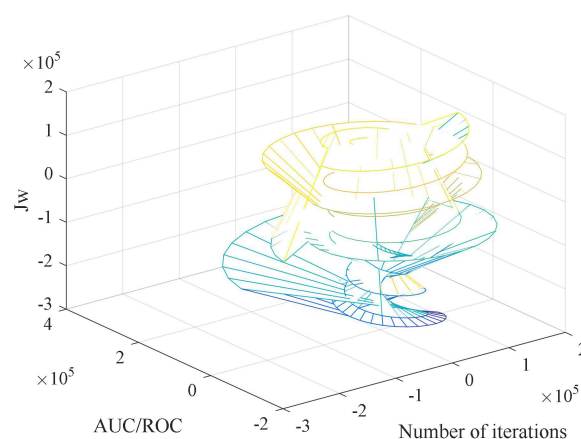
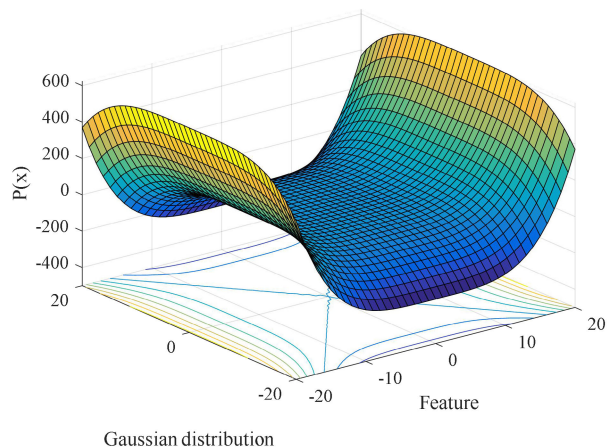


FIGURE 10. Curve relationship between iteration number and J_w .

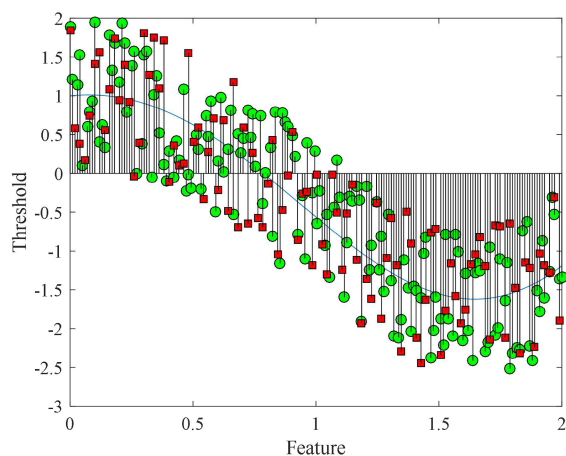
representation of the original signal. When learning the characteristics from the original signal and providing it to the multivariate Gaussian model for health assessment, it is also necessary to consider whether the learned features conform to the Gaussian distribution.

The Gaussian distribution of the physiological signal characteristics was calculated to verify whether it conforms to the Gaussian distribution. As shown in Figure 11, it can be seen intuitively that the medical multimedia data signal characteristics obtained in this experiment conform to the Gaussian distribution.

As shown in Figure 12, by dividing different thresholds, a feature distribution map of different levels of health states of two sets of eight medical multimedia big data signals is obtained. Tests on the dataset show that the proposed method can extract key feature representations of multidimensional medical multimedia data. Then, the feature points of the same probability interval are calculated, and the state of the user's health state is obtained according to the probability of these feature points. In this way, big data-based analysis can help users understand their health status. If analyzed by a doctor or a professional, such data information can be helpful for doctors to diagnose early signs of some diseases. Therefore,



Gaussian distribution

FIGURE 11. Gaussian distribution of features.**FIGURE 12. Feature points with different thresholds.**

the model is also a useful tool to help professionals diagnose potential diseases.

V. CONCLUSION

In the evaluation process of this paper, feature learning is the focus of this paper. Firstly, this paper adopts a multi-layer convolutional network structure to extract the key features of the original data from the input signal. Then, the feature is used as the input of the multivariate Gaussian distribution health state assessment model to evaluate the health status. The model tested eight physiological signals in the data set. The experimental results show that the feature extraction method used in this paper is an effective high-level feature algorithm that can learn data from the original temporal physiological data. It is also a feasible algorithm to use the multivariate Gaussian distribution theory to evaluate the health state. Combined with DBN, the comparative experiments of deep belief network, stack automatic encoder and convolutional neural network three kinds of DBN model medical multimedia database MIAS are studied. The advantages and disadvantages of three network structure models are analyzed. Secondly, in the training network, the idea of automatic coding is used in the training of convolutional

neural networks, which effectively solves the problem that the general convolutional neural network needs to fine-tune the network with tagged data. Finally, this paper uses Gaussian distribution to model the health state, and uses the Gaussian function to calculate the probability values of the network learning, and establish a health level model based on different probability intervals. The experimental results prove that the method used in this paper is an effective method, and the method can also be extended to other types of data.

The method used in this paper is only a convolutional neural network in DBN. Although some other DBN algorithms are mentioned in this paper, the algorithm principle is also described, but there is no other deep trust network, for example. Restricted Boltzmann machines and stack-type automatic coding compare the actual experimental results with convolutional neural networks. Therefore, in the subsequent research, one or more DBN networks can be used to model and analyze the health status of medical multimedia big data.

REFERENCES

- [1] I. Kiral-Kornek, S. Roy, E. Nurse, B. Mashford, P. Karoly, T. Carroll, D. Payne, S. Saha, S. Baldassano, T. O'Brien, D. Grayden, M. Cook, D. Freestone, and S. Harrer, "Epileptic seizure prediction using big data and deep learning: Toward a mobile system," *EBioMedicine*, vol. 27, pp. 103–111, Jan. 2018.
- [2] H. Zhang and Z. Wei, "Risk management of commodity trade business based on DBN and parallel processing of visual multimedia big data," *Multimedia Tools Appl.*, vol. 2019, no. 12, pp. 1–19, 2019.
- [3] Y. Zhao, Q. Dong, H. Chen, A. Iraj, Y. Li, M. Makkie, Z. Kou, and T. Liu, "Constructing fine-granularity functional brain network atlases via deep convolutional autoencoder," *Med. Image Anal.*, vol. 42, pp. 200–211, Dec. 2017.
- [4] Q. Meng, K. Wang, and X. He, "QoE-driven big data management in pervasive edge computing environment," *Big Data Mining Analytics*, vol. 1, no. 3, pp. 222–233, 2018.
- [5] M.-J. Lee, S.-Y. Choi, H.-J. Jeong, and Y.-G. Ha, "A precise image crawling system with image classification based on deep learning," *adv. Sci. Lett.*, vol. 23, no. 3, pp. 1623–1626, Mar. 2017.
- [6] S. Pouyanfar, Y. Tao, H. Tian, S.-C. Chen, and M.-L. Shyu, "Multimodal deep learning based on multiple correspondence analysis for disaster management," *World Wide Web*, vol. 22, no. 5, pp. 1893–1911, Sep. 2019.
- [7] S. Pouyanfar and S.-C. Chen, "Automatic video event detection for imbalance data using enhanced ensemble deep learning," *Int. J. Semantic Comput.*, vol. 11, no. 01, pp. 85–109, Mar. 2017.
- [8] J. F. Torres, A. Troncoso, and I. Koprinska, "Big data solar power forecasting based on DBN and multiple data sources," *Expert Syst.*, vol. 36, no. 4, pp. 12394–12406, 2019.
- [9] P. Zhang, X. Xu, and H. Wang, "Computer-aided lung cancer diagnosis approaches based on DBN," *J. Comput.-Aided Des. Comput. Graph.*, vol. 30, no. 1, pp. 90–106, 2018.
- [10] Z. Yan, Z. Xu, and J. Dai, "The big data analysis on the camera-based face image in surveillance cameras," *Intell. Automat. Soft Comput.*, vol. 24, no. 1, pp. 123–132, Jan. 2018.
- [11] S. Wang and J. Xu, "De novo protein structure prediction by big data and deep learning," *Biophys. J.*, vol. 112, no. 3, p. 55a, Feb. 2017.
- [12] L.-H. Wang, X.-P. Zhao, J.-X. Wu, Y.-Y. Xie, and Y.-H. Zhang, "Motor fault diagnosis based on short-time fourier transform and convolutional neural network," *Chin. J. Mech. Eng.*, vol. 30, no. 6, pp. 1357–1368, Nov. 2017.
- [13] R. Maldonado, T. R. Goodwin, and S. M. Harabagiu, "Active deep learning-based annotation of electroencephalography reports for cohort identification," *Amia Summits Transl. Sci. Proc.*, vol. 2017, pp. 229–238, Jul. 2017.
- [14] Y. Yan, M. Chen, S. Sadiq, and M.-L. Shyu, "Efficient imbalanced multimedia concept retrieval by deep learning on spark clusters," *Int. J. Multimedia Data Eng. Manage.*, vol. 8, no. 1, pp. 1–20, Jan. 2017.

- [15] S. Xiao, Z. Chen, and D. Shuai, "Detecting anomalous emotion through big data from social networks based on a DBN method," *Multimedia Tools Appl.*, vol. 2018, no. 5, pp. 1–22, 2018.
- [16] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: Challenges and approaches," *IEEE Access*, vol. 5, pp. 7776–7797, 2017.
- [17] F. Celesti, A. Celesti, J. Wan, and M. Villari, "Why deep learning is changing the way to approach NGS data processing: A review," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 68–76, 2018.
- [18] P. Kasnesis, C. Z. Patrikakis, and I. S. Venieris, "Changing mobile data analysis through deep learning," *IT Prof.*, vol. 19, no. 3, pp. 17–23, 2017.
- [19] H. K. V. D. Burgh, R. Schmidt, and H. J. Westeneng, "DBN predictions of survival based on MRI in amyotrophic lateral sclerosis," *Neuroimage Clin.*, vol. 13, no. C, pp. 361–369, 2017.
- [20] R. Mehmood, F. Alam, N. N. Albogami, I. Katib, A. Albesri, and S. M. Altowajri, "UTiLearn: A Personalised ubiquitous teaching and learning system for smart societies," *IEEE Access*, vol. 5, pp. 2615–2635, 2017.
- [21] L. Yu, W. Shuai, and M. S. Khan, "A novel deep hybrid recommender system based on auto-encoder with neural collaborative filtering," *Big Data Mining Anal.*, vol. 1, no. 3, pp. 211–221, 2018.
- [22] Z.-J. Yao, J. Bi, and Y.-X. Chen, "Applying deep learning to individual and community health monitoring data: A survey," *Int. J. Autom. Comput.*, vol. 15, no. 6, pp. 643–655, Dec. 2018.
- [23] Z. Rui, R. Yan, and Z. Chen, "Deep learning and its applications to machine health monitoring: A survey," *J. Latex Class Files*, vol. 14, no. 8, pp. 1–14, Aug. 2015.
- [24] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM Model for Particulate Matter (PM_{2.5}) forecasting in smart cities," *Sensors*, vol. 18, no. 7, pp. 2220–2232, Jul. 2018.
- [25] Q. Yu, Y. Shi, J. Sun, Y. Gao, J. Zhu, and Y. Dai, "Crossbar-net: A novel convolutional neural network for kidney tumor segmentation in CT images," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 4060–4074, Aug. 2019.
- [26] X. Ma, Y. Lu, and Y. Lu, "Medical image analysis of phosphorylated protein interaction extraction algorithm based on text mining technology," *Multimedia Tools Appl.*, vol. 2019, no. 3, pp. 1–29, 2019.
- [27] R. Wang, J. Lv, and S. Ma, "A MRI image segmentation method based on medical semaphore calculating in medical multimedia big data environment," *Multimed Tools Appl.*, vol. 77, no. 8, pp. 9995–10015, Apr. 2018.
- [28] W. Zhang and L. Jiang, "Algorithm analysis for big data in education based on depth learning," *Wireless Pers. Commun.*, vol. 102, no. 4, pp. 3111–3119, Oct. 2018.
- [29] H. Liu, X. Mi, and Y. Li, "Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM," *Energy Convers. Manage.*, vol. 159, pp. 54–64, Mar. 2018.
- [30] D. S. Kermany, M. Goldbaum, and W. Cai, "Identifying medical diagnoses and treatable diseases by image-based DBN," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.



YING YANG was born in Henan, China, in 1981. She received the bachelor's degree from the Xi'an University of Finance and Economics, in 2004, and the master's degree from the Wuhan University of Technology, in 2010. Since 2004, she has been working with Xinxiang Medical University. She currently works with Xinxiang Medical University. She has published six articles. Her research interests include hospital engineering management and hospital operation management.

...