

Received September 3, 2019, accepted September 16, 2019, date of publication September 23, 2019, date of current version October 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2942937

Intelligent Analysis of Medical Big Data Based on Deep Learning

HANQING SUN¹, ZHENG LIU¹, GUIZHI WANG¹, WEIMIN LIAN¹, AND JUN MA^{1,2} 

¹School of Information Engineering, Henan University of Animal Husbandry and Economy, Zhengzhou 450000, China

²School of Mechanical Engineering, Dalian University of Technology, Dalian 116000, China

Corresponding author: Jun Ma (majundalian@mail.dlut.edu.cn)

This work was supported in part by the Henan University of Animal Husbandry and Economy of Scientific Research and Innovation Team under Grant 2018KYTD19, and in part by the Scientific Research Projects of Higher Education Institutions in Henan Province under Grant 18A510011.

ABSTRACT With the wide application of computer technology, medical health data has also increased dramatically, and data-driven medical big data analysis methods have emerged as the times require, providing assistance for intelligent identification of medical health. However, due to the mixed medical big data format, many incomplete records, and a lot of noise, it is still difficult to analyze medical big data. Traditional machine learning methods can't effectively mine the rich information contained in medical big data, while deep learning builds a hierarchical model by simulating the human brain. It has powerful automatic feature extraction, complex model construction and efficient feature expression, and more important. It is a deep learning method that extracts features from the bottom to the top level from the original medical image data. Therefore, this paper constructs a data analysis model based on deep learning for medical images and transcripts, and is used for intelligent identification and diagnosis of diseases. The model uses massive medical big data to select and optimize model parameters, and automatically learns the pathological analysis process of doctors or medical researchers through the model, and finally intelligently conducts disease judgment and effective decision based on the analysis results of medical big data. The experimental results show that the method can analyze the medical big data, and can realize the early diagnosis of the disease. At the same time, it can analyze the physical health status according to the patient's physical examination records and predict the risk of a certain disease in the future. Greatly reduce the work pressure of doctors or medical researchers and improve their work efficiency.


INDEX TERMS Medical big data, analysis, deep learning, intelligent recognition.

I. INTRODUCTION

With the development of the economy and the advancement of computer technology and medical information processing technology, the application of medical big data recording systems in hospitals has become more and more extensive [1], [2]. The variety of medical big data is a collection of data for patient contact, examination, and diagnosis in time series. From the format point of view, it is mainly medical record and medical imaging inspection records, specifically current and past diagnostic records, laboratory tests, clinical records, prescriptions and medical examination images (mainly X-ray, CT, B-ultrasound and nuclear magnetic Resonance) et al [3]. Medical health records are the most important part of medical big data analysis [4], [5]. For a specific disease, doctors can

analyze the disease according to the existing medical big data, and can diagnose the disease in advance to prevent and control the disease before the onset of the disease, accurately predict the disease development, and make more high-risk patients. Active treatment prescriptions make patients more confident in treatment and rehabilitation, and always help to achieve the best therapeutic effect [6].

In the analysis of medical big data, first of all, it is necessary to record and collect images based on the medical examination indicators of the patient. Then, the difference between the normal image and the image of the patient is analyzed for disease analysis and early diagnosis. Many traditional medical big data analysis algorithms have been studied, including logistic regression [7], random forest algorithm [8], and support vector machine [9] and so on. With the development of deep learning, the application of this method in data mining [10], computer vision [11], [12] and natural language

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

processing [13], [14] has achieved continuous success. And it has become difficult to extract features by the preferred method of the task. It extracts deep and abstract features from the data and captures long-range dependencies in the data in an efficient manner, enabling efficient analysis of both image and text data. With the increasing number of medical images and transcript data, deep learning methods have better performance in big data analysis than traditional methods. And they also require less time and computational resources in data preprocessing and feature extraction [15], [16]. In recent years, many research teams have tried to apply deep learning methods in medical big data analysis [17], [18]. Table 1 shows the deep learning applications in medical big data analysis. Schirmer *et al.* [19] first analyzes the brain state signal and proposes a method to improve brain state decoding. However, the algorithm does not have a good analysis effect on the brain state signal, and there is a lot of noise information. Lu *et al.* [20] designed a deep learning architecture to classify data analysis results in multiple categories. Due to the large number of deep learning framework layers, the time performance of the algorithm is low. Chaddad *et al.* [21] proposed the use of both supervised and unsupervised learning algorithms to learn the characteristics of deep convolution images, so as to analyze the medical images. But the algorithm can only process medical image information, and cannot process medical text information. Cebul *et al.* [22] uses data analysis through a deep learning network and generalized estimation equations to evaluate the effectiveness of diabetes care based on medical transcript data. However, the assumptions of the generalized estimating equation are more stringent, and the model is not portable. Hripsak and Albers [23] demonstrates that medical big data can provide unprecedented amounts of information for clinical research. However, there is no way to effectively analyze medical data. Amarasingham *et al.* [24] proposes an automated data analysis model that analyzes data from hospitalized patients with heart failure and accurately predicts the probability of prehospitalization and death in the next 30 days. This method focuses on the probabilistic predictions that follow, but uses a simpler method to analyze the data. Ho *et al.* [25] uses the tensor decomposition method to analyze the test indicators and physical health status of each patient. However, the connection between images cannot be analyzed, which may result in misjudgment of the recognition result. Wei and Denny [26] studied a cognitive method based on natural language processing for patient disease state analysis and recognition. However, this method cannot analyze the relationship between the parts of the image and the nuances in the image, nor can it handle arbitrarily sized text records. Winterstein *et al.* [27] combined text data mining methods with CNN networks to analyze clinically recorded data, and then to explore the characteristics of coronary artery disease, and to predict the risk of coronary artery disease. However, this method requires a number of words in the input text. If more characters are input, a misjudgment may occur. Khalifa and Meystre [28] uses natural language processing methods to analyze clinically

recorded data. However, this method cannot process medical image data, and the time performance does not meet the real-time requirements.

Regardless of the analysis of medical image data or the analysis of transcript data, the existing deep learning model cannot analyze the ties of the various parts of the image and the nuances thereof, nor can it handle arbitrarily sized transcripts. The existing deep learning model is less robust. And the model design needs to be modified when the disease characteristics change, which consumes a lot of manpower and time, limits the versatility and adaptability of the model. And it is difficult to implement in clinical applications.

Based on the above problems, this paper constructs different neural network structures for data analysis of the most common medical images and transcripts in medical big data. For medical image data, the image data is first matrixed, and the correlation coefficients between different regions of interest in the brain are calculated. Then, the deep learning model of AutoEncoder which is most suitable for this data is designed to extract the correlation features between different regions of interest to achieve accurate and intelligent pre-diagnosis. For medical text data, the 3D convolutional neural network is used to extract the temporal features between the internal features of the data and different data. At the same time, the spatial pyramid structure is used to make the model can process the input of any length to realize the intelligent prediction of the disease.

Specifically, the technical contributions of our paper can be concluded as follows:

First: This paper elaborates and summarizes the difficulties, research status and research methods of deep learning in medical data recognition. The related deep learning methods and models are studied. And the construction of the network provides a reference value.

Second: Different neural network structures are designed for the most common medical images and text records in medical big data. Data analysis of medical images and transcript data is carried out to realize intelligent recognition and prediction of diseases.

The rest of our paper was organized as follows. Theory of deep learning was introduced in Section II. Section III described related theoretical derivation of the proposed system. Experimental results and analysis were discussed in detail in Section IV. Finally, Section V concluded the whole paper.

II. DEEP LEARNING IN MEDICAL DATA ANALYSIS

A. DIFFICULTIES IN MEDICAL DATA ANALYSIS WITH DEEP LEARNING

Because of the outstanding achievements in deep learning in the fields of speech recognition, image semantic segmentation, scene recognition, and target detection, it has also enabled researchers in medical data analysis to see the dawn. However, there are still some challenges in the field

TABLE 1. Applications of deep learning in biological and medical big data analysis.

Category	Time	Problem	Model
Disease diagnosis	2014,2015	AD/MCI classification	SAE/DBN
	2013,2014	Cancer / Tumor Diagnosis	CNN /DBN+SVM
	2011,2014	Diagnosis of brain diseases	SAE/CNN/DBN
	2015	Nuclear cataract classification	CNN+SVM
	2014	Diagnosis of chronic gastritis	DBN
Medical image processing	2013,2014	Automatic segmentation of medical images	Stacked ISA/CNN
	2013	Image key point discovery	CNN/DBN/CNN
	2015	MRI image reconstruction	SAE
Medical big data modeling	2014,2015	Brain development, brain circuit modeling	DBN/CNN
	2013	Emotional analysis	DBN
Protein structure prediction	2014,2015	Secondary structure prediction	SAE/DBN
	2012,2014	Tertiary structure prediction	DBN/3DNN/GSN
	2015	Template-based structural prediction	DBN
Sequencing data processing	2015	DNA labeling	DNN
	2014,2015	RNA variable shear analysis	SAE+DNN
Expression profile analysis	2015	Non-coding RNA discovery	lncRNA-MFDL
	2014,2015	Subtype identification based on expression profiling	DBN/SAE
Other	2015	Drug design and discovery	DNN
	2014	Protein Model Quality Assessment	SAE

of medical data analysis, which are reflected in the following points:

1) DIMENSIONS OF MEDICAL DATA

With the replacement of medical electronic devices, it is possible that a patient will get much data. For image data in medical data, deep learning is mostly to train some two-dimensional plane images. So it is necessary to sample the three-dimensional medical images and convert them into two-dimensional image information, which will not only reduce the picture. The resolution also loses a lot of image information, but there are some good ways to preserve the original image. For example, Roth *et al.* [29] proposed a 2.5D sampling method, which is mainly in the second when dimensioning, the sampled images are no longer trained separately, but three spatially orthogonal images are combined into one RGB image, and the original image information is retained as much as possible. Many researchers have designed a 3D deep learning model to directly analyze the original image, such as Brosch *et al.* [30], Dou *et al.* [31] and others. This eliminates the cumbersome sampling steps and the whole process is straightforward.

2) NUMBER OF MEDICAL DATA

The reason why deep learning can excel in many fields is because a large amount of learning data, the features obtained through these learning data have stronger expressive ability than the features extracted by manual methods, so that a better effect can be obtained. Therefore, this requires that the deep learning model must have enough learning data sets. But in the field of medical data analysis, the training data is seriously insufficient, which requires researchers to solve: how to quickly obtain a large amount of tag data and how to

use the limited data to train the deep learning model and get satisfactory results.

3) UNSUPERVISED TRAINING AND SUPERVISED TRAINING

For most tasks that apply deep learning algorithms, such as scene detection, image segmentation, object recognition, etc., all use supervised training methods for research. But there are also a few work focused on unsupervised training methods, such as images coding, image representation, etc. Due to the inherent shortcomings of medical data, many data cannot use the deep learning model, so the unsupervised training algorithm is more reasonable for the analysis of medical images. The model of unsupervised deep learning currently used has a Boltzmann machine RBM. This algorithm model can directly learn good feature representations from the unlabeled training dataset. The disadvantage is that these features do not necessarily have the best classification effect. van Tulder and de Bruijne [32] used the deep convolution Boltzmann model to train directly from the training data set, and then used the training convolution kernel to extract the training data, and merged these features with the features obtained by the discriminant model. The final result is better than the classification result of the separate generation discriminant model.

For the above questions, the researchers spent a lot of time and effort, but also made some good progress, such as van Grinsven *et al.* [33] tried to dynamically select the wrong false negative samples during the training process. The data is used as the next classification data to improve the learning effect of the deep learning algorithm in medical data. In the training process of the deep learning algorithm, the optimization of the model weight coefficient requires multiple iterations, at each time. In the iteration, a random

subsample data set selected from the training samples is used to train the network model, and the loss function is minimized by the gradient descent algorithm.

B. DEEP LEARNING METHOD IN MEDICAL DATA ANALYSIS

Medical big data is a multi-modal, complex data that continues to grow rapidly and contains a wealth of information. The challenges associated with medical big data include how to quickly and accurately collect and obtain medical health data, how to efficiently use high-speed networks to deliver medical health data reliable and efficient transmission, how to use artificial intelligence related machine learning and deep learning techniques to extract useful information from medical big data, and how to develop intelligent applications for the majority of medical staff and ordinary people.

The deep learning method is developed from the artificial neural network model. The raw data is layer-by-layer abstracted by combining multiple nonlinear processing layers, and different levels of abstract features are obtained from the data and used for classification prediction. Compared with traditional machine learning methods, it has the following three characteristics:

a. Deep model architecture. The multi-layer structure of the deep learning model is very similar to the animal’s visual processing system. Compared with other shallow models, such as support vector machine, deep learning models have more hidden layers and contain more nonlinear transformations, which makes the ability of deep learning to fit complex models greatly enhanced.

b. Multi-layer data feature representation. The deep learning model takes the original form of the data as input, and then uses the output of the current layer as the input of the next layer, stacking layer by layer, thereby summarizing the higher-level feature representation, so as to describe the complex data structure.

c. Unsupervised learning. The deep learning model adds unsupervised learning process to the training, and obtains a good initial value of the model through pre-training, which can effectively improve the training effect. In addition, the training of unlabeled data also increases the scale of available data.

Deep learning itself contains a large number of algorithmic models, which itself refers to a multi-layered complex neural network layer, which is now defined as a deep algorithmic structure with two or more layers, so deep learning is the name of a multi-layered algorithm. At present, deep learning is divided into the following two ways in the research and analysis of medical data:

(1) Self-training model

Deep learning the analysis of medical data includes algorithmic models including restricted Boltzmann machines (RBM), deep belief networks (DBN), deep Boltzmann machines (DBM), convolutional neural networks (CNN), etc. These models are all trained through a large amount of data, and the process is to tuned many parameters, such as learning

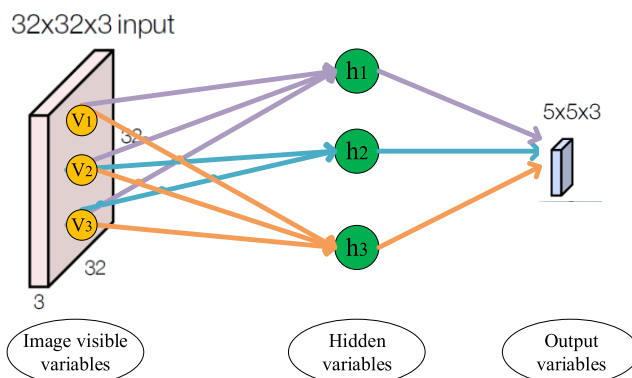


FIGURE 1. RBM drawn as a Markov network.

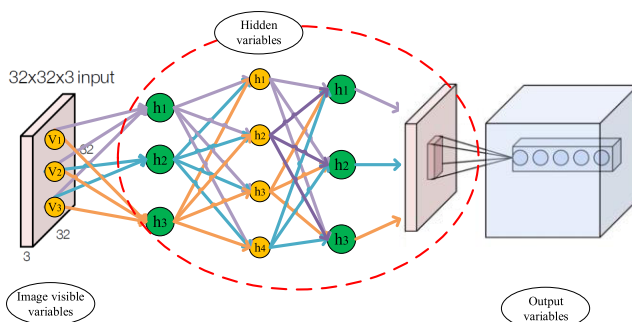


FIGURE 2. Deep belief network model.

rate, momentum coefficient, weight attenuation coefficient and so on.

RBM is a typical generation model in deep learning, including a layer of visible variables and single-layer hidden variables. RBM is essentially an undirected probability graph model that can be stacked to form a deeper model with no connections allowed between any of the visible or hidden layers [34]. Figure 1 is a binary graph structure for RBM. The model is described, one layer of the figure is the visible layer, and the other layer is the hidden layer.

Figure 2 is a schematic diagram of the general model of the deep belief network. As can be seen from the model, the top two layers are hidden layers and the bottom layer is the visible layer. An undirected connection is taken between the two hidden layers, and a directed connection is taken between the hidden layer and the visible layer, and the arrow points to the v layer closest to the data.

Figure 3 is a schematic diagram of the structure of the DBM general model. The model is the same as the DBN model in Figure 2, consisting of a bottom visible layer and a top two hidden layers, with connections only between adjacent layer units, but between the visible layer and the hidden layer is none connect to the connection instead of the directed connection.

CNN is a feedforward neural network. The general convolutional neural network structure diagram is shown in Figure 4. It consists of a convolutional layer, a pooled sampling layer, and a fully connected layer. The medical image can be used as the original input of the bottom layer

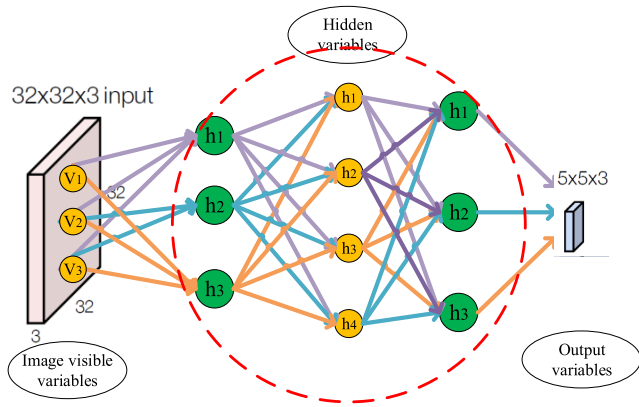


FIGURE 3. Deep Boltzmann machine model.

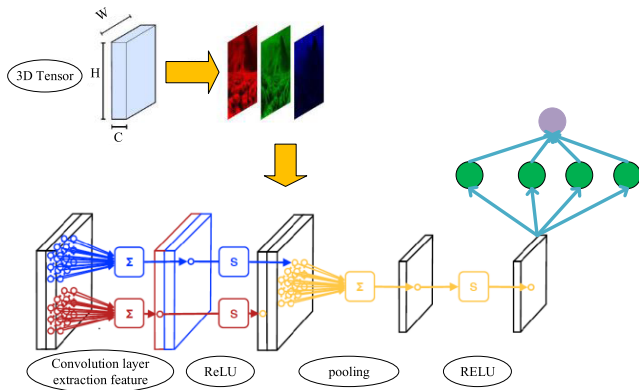


FIGURE 4. Example of convolution neural network.

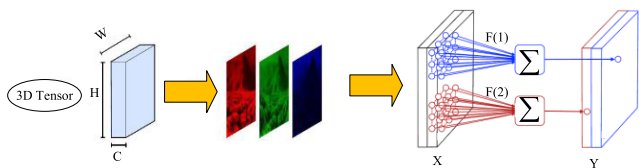


FIGURE 5. Two-dimensional convolution calculation.

of the network, and then transmitted to the next layer in turn. Each layer extracts the most significant feature of the image data through a convolution kernel, and the output result is the classification recognition result of the image. As shown in Figure 5, the convolutional layer convolves the target image through the convolution kernel to generate a feature map to realize local feature perception and feature extraction.

(2) Migration learning and fine-tuning

In the field of medical data analysis, obtaining a comprehensive annotation data set like ImageNet [35] is a huge challenge. Therefore, when the data is scarce, there are several ways to choose:

a. Migration learning, this method generally uses a network model trained from the data set, such as AlexNet [36], CaffeNet [37], GoogleNet [38], etc. And then they extract features from the network layer and use it to train a separate pattern classifier. For example, Shameer *et al.* [39] used the pre-trained convolutional neural network model as a feature generation container to classify lung nodule data.

b. Fine-tuning, when there is a small data set and the classification task is very important. A suggested method is to use a pre-trained deep learning model to initialize the classification network of this data set, and then reduce the learning rate to supervise the training of the entire network layer.

At present, many papers have achieved good results using the above schemes. Roth *et al.* [29] applied convolutional neural networks to improve the detection of CT image data by existing computer-aided detection systems. The nodule detection model generates all candidate image data in the classification task, and also proposes a 2.5D image generation algorithm, which makes the training model obtain higher dimensional image data features, which has a great influence on the obtained experimental results. Finally, a convolutional neural network is trained using the processed image data, and an amazing accuracy is obtained. The sensitivity of lesion detection across the database has increased by 13%–34%. Dou *et al.* [31] proposed a more complex deep learning model to directly train three-dimensional raw data, which ensures that the entire algorithm can learn more useful high-dimensional features. And the author further improves detection performance, reduced computational cost, and some modifications to the initial detection. Compared with the traditional sliding window algorithm, the whole method reduces the calculation, speeds up the detection process, and achieves a high sensitivity of 93.16%. Ma *et al.* [40] first generated positive and negative samples of the data set, and used a self-designed 13-layer depth network for preliminary screening. The positive sample generated by the deep network model was used as the training set of the next depth model. The rejection of some unqualified negative samples is achieved, making the remaining training samples more representative. Li *et al.* [41] extracted the mid-high-dimensional features from the convolutional neural network model trained by natural images, and used the extracted features to classify them. They used these features to achieve good results in target recognition, scenario analysis and classification recognition. Eberl *et al.* [42] used a theoretically more reliable method. They built a deep learning model and then trained the parameters of the depth model through AutoEncoder. The advantage of this is that a large amount of unlabeled data can be used, pre-trained by AutoEncoder. The model, with a small amount of labeled data training, can achieve good results, and achieved a correct rate of 91.4% in the 1417 Basel cell carcinoma CT medical image data test.

III. MEDICAL BIG DATA ANALYSIS MODEL BASED ON DEEP LEARNING

Literature [42] constructed a deep learning model and then trained the parameters of the depth model through AutoEncoder. The advantage of this is that a large amount of unmarked data can be used, pre-trained by AutoEncoder. The model is trained with a small amount of marker data and can achieve good results. Literature [29] proposed a 2.5D image generation algorithm to enable the training model to

obtain higher dimensional image data features. And it has a great impact on the experimental results obtained. Under the guidance of the above work, this paper proposes different network models for different forms of medical data. Aiming at medical image data, a deep learning model based on stacked AutoEncoder is proposed, which uses time series data and correlation coefficient data of medical images for data analysis. For medical text data, a 3D convolution kernel is used to extract information between the phrase internal information and the phrase of the input data, and then combined with the SPP to identify the medical text data.

**A. MEDICAL IMAGE DATA ANALYSIS MODE
BASED ON AUTOENCODER**

This paper builds a deep learning model based on stacked AutoEncoder, and uses time series data and correlation coefficient data of medical images for data analysis.

The parameters of the AutoEncoder model designed in this paper mainly include cost function and gradient. Firstly, the cost function is introduced. The cost function in this paper is mainly composed of three parts, namely the mean square error, the sparseness penalty and the weight attenuation. The mean square error describes the difference between the observed value and the expected value. Given a data set $(x^1, y^1), \dots, (x^m, y^m)$ with m training samples, where x^i represents one piece of data entered, y^i represents its label, then the mean square error of the definition of data sample (x^i, y^i) is as follows:

$$J(W, b, x, y) = (1/2)||h_{W,b}(x^i) - y^i||^2 \tag{1}$$

Therefore, for all m training samples, the mean square error is:

$$J(W, b) = \frac{1}{m} \sum \frac{1}{2} ||h_{W,b}(x^i) - y^i||^2 \tag{2}$$

Secondly, sparse penalty terms are added to the calculation of this network. A lot of research has confirmed that only a small part of the brain's various regions of interest are active at the same time or when performing the same task, which means that the brain's activities are sparse. The brain's MRI image is a true reflection of the brain's active state over a period of time. Therefore, the time series data and correlation coefficient data obtained during data preprocessing are also sparse. Therefore, the AutoEncoder model must be increased the sparsity penalty. The AutoEncoder model is very similar to the cognitive process of the brain, so the number of hidden layer nodes that the entire neural network activates in each training is limited. Use x^i to indicate the output of the j node when the input data is $a_j^2(x^i)$ in the second layer of the network, and

$$\rho_j = \frac{1}{m} a_j^2(x^i) \tag{3}$$

The sign ρ_j indicates that the output of this point affects all m samples. In addition, the sparsity parameter ρ in the sparse penalty is normally set to a value slightly larger than but very close to 0 to ensure that the entire AutoEncoder network is

sparse. In this paper, ρ_j is initialized to ρ . If ρ is 0.01, most of the hidden layer nodes will produce a value greater than 0 to make the average activation of the entire network close to 0.01.

In this paper, KL divergence is used to ensure that ρ_j is close to ρ . Therefore, the sparsity penalty term based on KL divergence is:

$$\sum KL(\rho_j||\rho) = \sum \rho \log \frac{\rho}{\rho_j} + (1 - \rho) \log \log \frac{1 - \rho}{1 - \rho_j} \tag{4}$$

Finally, in order to prevent over-fitting problems in this network, we must add weight attenuation items. The weight attenuation term controls the entire network structure by reducing the range of weights to prevent over-fitting of the input data. The weight attenuation parameter λ can describe the magnitude of this function in the cost function. If λ is too small, the weight decay term has too little effect on the cost function, and over-fitting is prone to occur. Conversely, if λ is too large, under-fitting is likely to occur, and neither of these should occur. In this paper, we will first set a smaller value of λ , and then adjust the size of the parameter λ according to the feedback of the experimental results. The formula of the weight attenuation term is:

$$J(W, b) = \frac{\lambda}{2} \sum \sum \sum (W_{ji}^l)^2 \tag{5}$$

After synthesizing all the elements in the cost function, the resulting function formula is as follows:

$$J_{sparse}(W, b) = \frac{1}{m} \sum \frac{1}{2} ||h_{W,b}(x^i) - y^i||^2 + \beta \sum KL(\rho_j||\rho) + \frac{\lambda}{2} \sum \sum \sum (W_{ji}^l)^2 \tag{6}$$

Gradient is also an important parameter of deep neural networks. And it plays a major role in the process of adjusting and optimizing the network. In the method of calculating the gradient, a back propagation algorithm is generally used to calculate the gradient of the cost function for each of the previous layers. The backpropagation algorithm is a supervised learning method that adjusts the parameters of the entire network by comparing the difference between the expected value and the true output. The gradient formula is as follows:

$$\frac{\partial J_{sparse}(W, b)}{\partial W_{ij}^l} = \frac{1}{m} \sum \frac{\partial J(W, b, x^i, y^i)}{\partial W_{ij}^l} + \lambda W_{ij}^l \tag{7}$$

$$\frac{\partial J_{sparse}(W, b)}{\partial b_{ij}^l} = \frac{1}{m} \sum \frac{\partial J(W, b, x^i, y^i)}{\partial b_{ij}^l} \tag{8}$$

In the training process of the neural network model, the error can also be transmitted by the back propagation algorithm. The error of the last layer is as follows:

$$\delta_i^{n_l} = \frac{\partial \frac{||y - h_{w,b}(x)||^2}{2}}{\partial z_i^{n_l}} \tag{9}$$

In the forward propagation process, the error of the first layer can be calculated by the error of the latter layer.

The error calculation formula for the first layer is as follows:

$$\delta_i^l = (\sum W_{ij}^l \delta_i^{l+1} + \beta(-\frac{\rho}{\rho_j} + \frac{1-\rho}{1-\rho_j}))f'(z^l) \quad (10)$$

After obtaining the cost function and the gradient and the expression and transfer method of the error, the entire network can be trained using the L-BFGS algorithm. Firstly train each layer of network separately, then use the output of each layer as the input of the next layer of network, complete the pre-training process of the entire AutoEncoder model, and get the activation value a^m of the last layer for subsequent medical big data. Analyze the foundation.

After the AutoEncoder model completes the feature extraction of the input, it is necessary to classify all the data according to the feature, and the classification task is the softmax layer, which can be regarded as the extension of the logistic regression for multi-category classification conditions. The cost function of the softmax layer is as follows:

$$J_{\text{softmax}}(\theta) = -\frac{1}{m} \sum \sum \{y^j = j\} \log p(y^j = j|x^i, \theta) + \frac{\lambda}{2} \sum \sum \theta_{ij}^2 \quad (11)$$

where k is the number of labels, $\frac{\lambda}{2} \sum \sum \theta_{ij}^2$ is the weight decay term in a softmax regression, and $p(y^j = j|x^i, \theta)$ is the probability to classify x^i as j . The expression $p(y^j = j|x^i, \theta)$ is:

$$p(y^j = j|x^i, \theta) = \frac{e^{\theta_j^T x^i}}{\sum e_i^T x^i} \quad (12)$$

The network training process is mainly divided into two steps, namely a pre-training step and a subsequent fine-tuning step. Pre-training trains each layer of the network through unsupervised greedy methods, and then uses the tagging information of the data to improve model performance through a supervised fine-tuning step.

It is stipulated that the AutoEncoder model has a total of n_l layers, and the pre-training of each layer is not independent of each other. In order to get the optimal parameters for each layer in the AutoEncoder model, this paper uses the following three steps. First, you need to set the weight parameter (W, b) of the entire network to a value close to 0, then optimize the parameters by minimizing the cost function, and finally calculate the output activation vector for each layer and input it to the next layer. If this is the last layer, use it as input to the following softmax classification function. The specific details of each step are described below.

In the pre-training process, all the steps are aimed at minimizing the cost function and then obtaining a better set of parameters for subsequent training. A better set of weighting parameters can make the whole network structure closer to the optimal solution in the ideal state, and these parameters need to start from a random number slightly larger than 0, so that the whole network can converge more easily.

After setting the initial value of each parameter, it is necessary to use a nonlinear optimization algorithm to optimize the cost function parameters and find the parameter set when

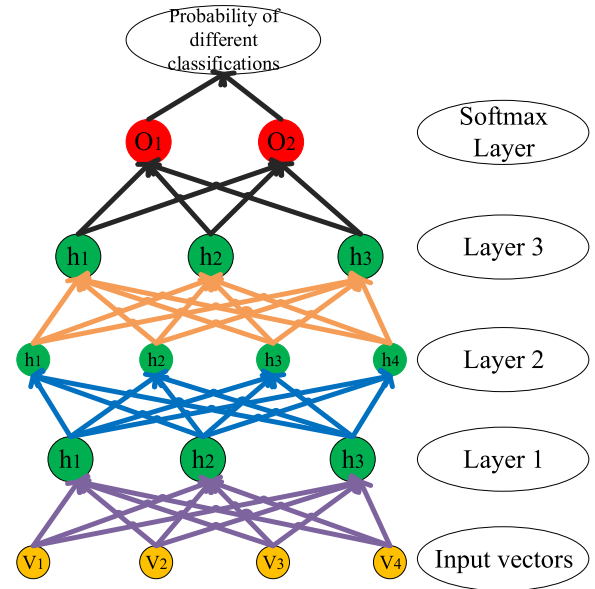


FIGURE 6. AutoEncoder deep learning architecture.

the cost function is the smallest. Among all nonlinear optimization algorithms, the Limited L-BFGS algorithm is one of the best methods for solving nonlinear problems. The L-BFGS algorithm is optimized by the BFGS algorithm, and they all belong to the quasi-Newton algorithm. In the BFGS algorithm, the partial derivative of the cost function and the second-order partial derivative must be calculated, so it takes a lot of time and computational resources. Newton's algorithm is usually not applied to the optimization of the objective function in practice. In contrast, the quasi-Newton algorithm can overcome this by calculating the approximate inverse Hessian matrix. As an improved version of the BFGS algorithm, the L-BFGS algorithm consumes less computation time, computing power, and memory resources than the BFGS algorithm. It only needs to use a small number of vectors to represent the matrix, without the need to calculate and store the entire inverse Hessian matrix, so the L-BFGS algorithm is the most efficient way to minimize the cost function. To minimize $J_{\text{softmax}}()$, the gradient of equation (11) is as follows:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum \{x^i \{y^j = j\} - p(y^j = j|x^i, \theta)\} + \lambda \theta_j \quad (13)$$

After that, the L-BFGS algorithm is used again to optimize the cost function. In the case of a given cost function and gradient, the activation vector of the last layer of the original pre-trained network will be the input of the softmax classification layer. The entire AutoEncoder model architecture is shown in Figure 6.

Fine-tuning is to improve the overall performance of the entire repeating AutoEncoder network in Figure 6. The pre-training process is a feature of greedily looking for input data, while the fine-tuning step actually slightly adjusts the boundary features of different categories. The method of

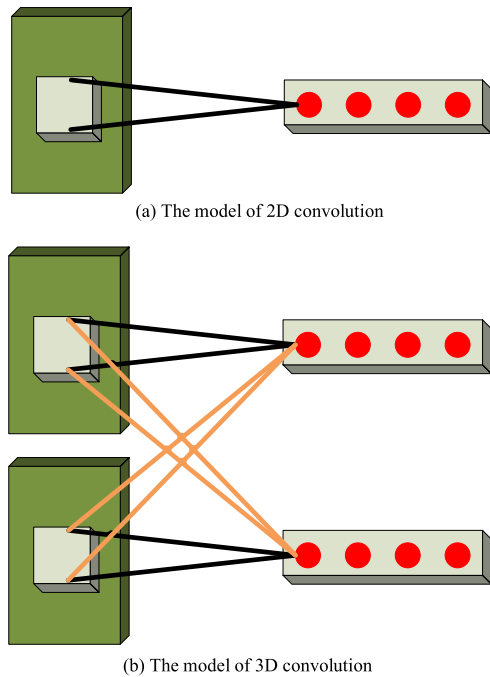


FIGURE 7. 3D convolution vs. 2D convolution.

fine-tuning is similar to the process of minimizing the cost function described earlier. In this step, the L-BFGS method is also used to minimize the difference between the expected output and the actual output in the softmax model.

B. MEDICAL TEXT DATA ANALYSIS MODEL BASED ON 3D CONVOLUTIONAL NEURAL NETWORK AND SPP

The 3D convolutional neural network solves the shortcomings of 2D convolutional neural networks that can only process two-dimensional static data. It has great advantages in processing tasks such as video processing, natural language processing, and motion recognition with time information. Therefore, this article uses a 3D convolution kernel to extract information between the phrase internal information and the phrase of the input data. And regardless of the size of the input data, the same network structure can be used for processing, which greatly expands the scope of application of the model. The difference between its main structure and 2D convolution is shown in Figure 7.

In Figure 7, the left input of the 2D convolution structure is an image, while the input to the left of the 3D convolution structure is multiple images, and the right side is the feature map generated after convolution. As can be seen from the figure, in the 2D convolutional neural network, the input image is scanned by using a fixed-size convolution kernel, the value after each convolution is calculated, and then the feature map is synthesized. In the 3D convolutional neural network, it is necessary to preserve the variation characteristics of different images over time on the premise of extracting image features. Therefore, firstly, the data of the same position in successive images is combined into a three-dimensional matrix, and then this three-dimensional matrix is

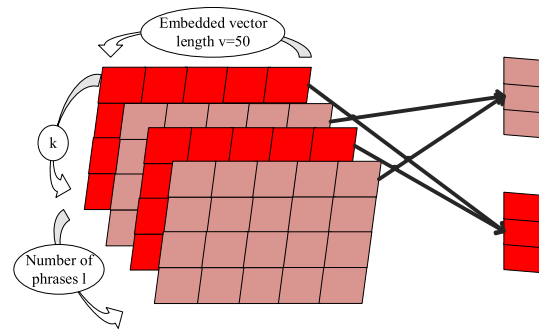


FIGURE 8. 3D convolution operation diagram.

convolved using a three-dimensional convolution kernel, and feature values are generated to be combined to form a feature map. Since a feature map contains information of a plurality of images at the same time, the three-dimensional data can be processed and the time series information in the input data can be extracted.

In the 3D convolutional layer, the input data and the convolution kernel are both a three-dimensional matrix. For the convenience of description, the size of the input data defined herein is $X_p \in R^{l \times k \times v}$, where l is the number of phrases into which the patient’s input data is divided. Among them, k is the length of a phrase, that is, the number of embedded vectors contained in a phrase, and v is the length of each word2vec vector. In this paper, the vector length v is 50. In the process of dividing each patient’s data into sentences as a phrase, there is only one vector between each phrase. For example, if the length k of the phrase is set to 3, the vector of 1, 2, and 3 is the first phrase. The 2, 3, 4 vector is the 2nd phrase, the 3, 4, 5 vector is the 3rd phrase, and so on, so the input data length T and l of each patient is $l = T - k + 1$.

In the convolution process, in order to maintain the integrity of the content contained in a word2vec vector, no convolution operation is performed on the dimension of v , only convolution is performed on the dimensions of l and k , and the dimension of l contains the sequence information between different records that are not adjacent. The dimension of k contains the time information between adjacent records, so convolution in the dimensions of l and k can extract the internals of the input data. Regardless of whether it is sequential information between adjacent data or non-adjacent data, combined with the information inside each data represented by word2vec vector, the 3D convolutional neural network constructed in this paper can completely analyze the characteristics of the input data. And it makes an accurate analysis.

A schematic diagram of a specific convolution operation is shown in Figure 8. In the figure, this paper assumes that the convolution width of each convolution kernel in the k direction is h , the convolution width in the l direction is m , and there are N convolution kernels in total, then the convolution in the case of input p the size of the feature map is $X_p \in R^{(k-h+1) \times (l-m+1) \times l \times N}$.

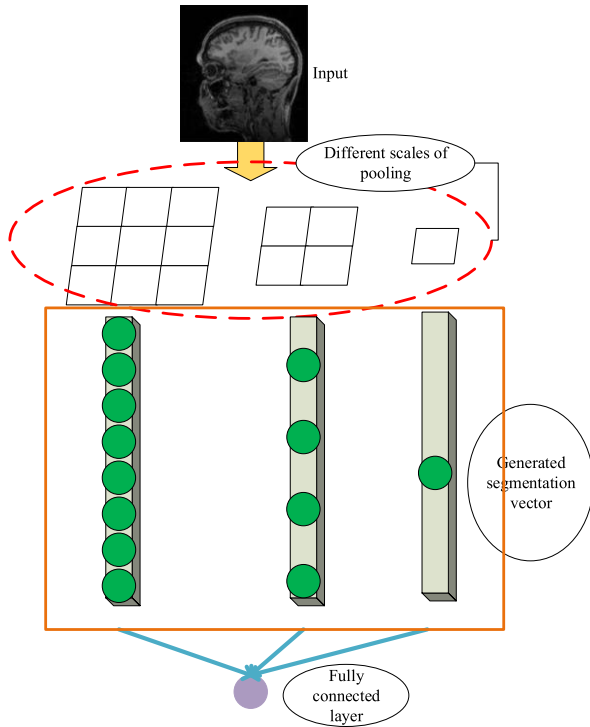


FIGURE 9. Space pyramid pooling structure.

The space pyramid pooling solves the problem of different size of feature maps. Under the premise of retaining the main features of feature maps, it also extracts features from different sizes of receptive fields. It is the best solution for processing input data of different scales now. Applied to any deep learning model, deep learning plays a huge role in irregular data processing. The algorithm structure of the spatial pyramid pooling is shown in Figure 9. The main idea is that the input image of any scale is divided into $n \times n$ regions, and then each region is pooled to obtain a pooled vector, and then n is set to other values. Re-divide the input image into $n \times n$ regions, and then get a pooled vector, and finally combine all the vectors into a fixed-length vector as the input of the fully connected layer.

In the 3D maximum pooling layer, pooling operations are improved based on spatial pyramid pooling, but are quite different from the original spatial pyramid pooling. In this layer, there are mainly two-step pooling operations, namely, a pooling operation inside each phrase, and a pooling operation between phrases, that is, phrases. For the pooling operation inside the phrase, the most representative features of the adjacent phrases can be extracted, and the pooling operation between the phrases can extract the most representative features in the successive time series. In order to ensure that the network can handle input of any size at the same time, this paper also adopts the advantage of pyramid pooling, that is, pooling different scales of input for different sizes to ensure that they can output the same size feature map, specific 3D the maximum pooling operation is shown in Figure 10.

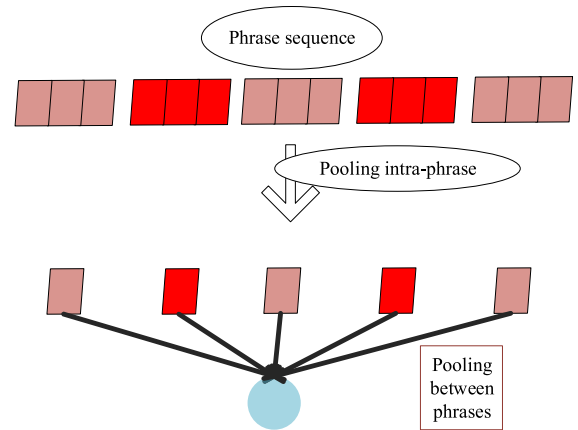


FIGURE 10. Maximum pooling operation.

In this paper, the size of all input records is divided into several approximate ranges, and zero padding is performed for input data that is less than this length range. For example, when the length k of the set phrase is 3, the data is divided into 4 length segments, which are 98, 142, 186, and 246 respectively. If the input matrix data length is less than 98, it is filled with zeros into an input of length 98, and the input length is greater than 98. Filled with a length of 142 input, input length greater than 142 less than 186, filled with an input of length 186, input length greater than 186 less than 246, filled with an input of length 246. Then input these lengths into the neural network to perform pooling operations in different scales, so that all data, regardless of the input length, can be changed to the same size after passing through the 3D maximum pooling layer. In this way, the subsequent fully connected layer can directly connect all the feature data for classification operation.

This paper proposes a medical text data analysis model based on 3D convolutional neural network and SPP. The model can analyze and accurately predict the patient's disease risk in the next 0 days, 90 days, and 180 days after inputting medical text data of any length. The model consists of two 3D convolutional layers, two common pooling layers, one pyramid structure pooling layer and one fully connected layer. The whole network structure is shown in Figure 11.

In this model, the paper still uses the input matrix data length of 98, the length k of each phrase is 3, and the word2vec embedded vector length is 50 as an example to analyze the change of data size after each layer operation. Since the input data size is $98 \times 3 \times 50$, after the first 3D convolutional layer passes N_1 convolution kernels, the output feature map size is $N_1 \times 94 \times 2 \times 1$. After passing through the first pooling layer, the data size becomes $N_1 \times 47 \times 1 \times 1$. Then, after passing through a 3D convolutional layer of N_2 convolution kernels, the generated feature map size is $N_2 \times 46 \times 1 \times 1$. After passing through a pooling layer, the obtained data size is $N_2 \times 12 \times 1 \times 1$. After that, the data will be input into the fully connected layer through a pyramid pooling structure. This ensures that different sizes of data obtained

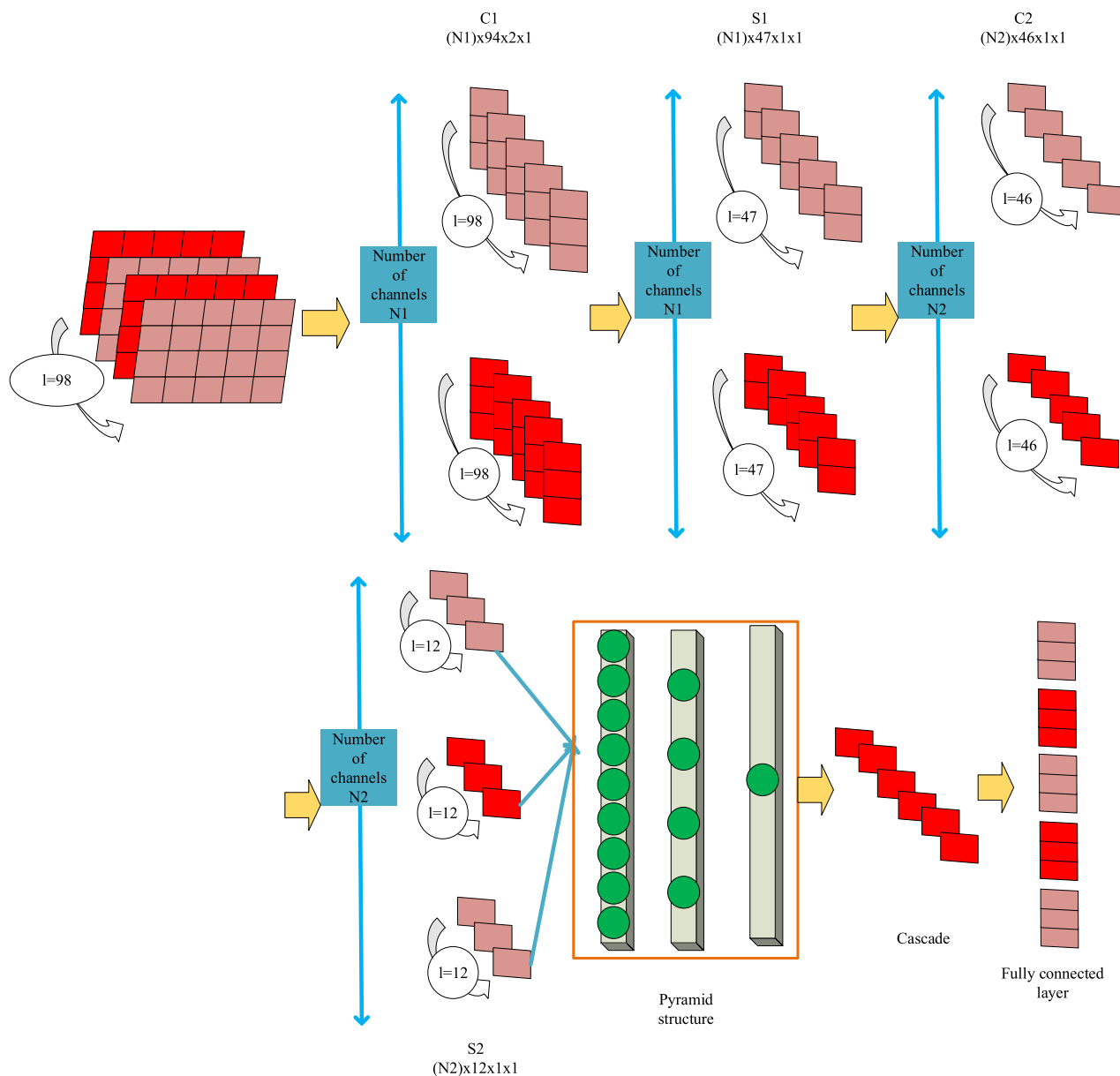


FIGURE 11. Medical text data analysis model based on 3D convolutional neural network and SPP.

by convolution and pooling operations can be pooled when inputting data of any size.

More critically, using pooling operations of different scales, the most important feature information can be extracted in different time ranges, that is, the network structure can extract the correlation of medical events in different time scales. In this paper, the maximum pooling range of 6 is used, and the range is 4 maximum pooling and pure full-join operation. Therefore, when the input data length is 98, the obtained feature sequence sizes are 2, 3, and 12 respectively. Therefore, the output of the pyramid pooling structure is $N_2 \times 17 \times 1 \times 1$. Finally, the data is integrated into a one-dimensional vector through the fully connected layer, and the softmax classification algorithm is used to predict the disease risk of the analyzed results.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL DATA DESCRIPTION

In order to verify the validity of the proposed method for medical image data, this paper selects the brain MRI medical image data set as medical image data. This paper selects 91 MCI data samples and 79 normal control data samples. Subjects in the normal control group ranged in age from 63.2 to 88.3, and MCI patients ranged in age from 66.5 to 87.3. There was no significant difference in age and gender between the two groups. All subjects were from the database of ADNI. In this paper, the patient’s electronic medical record data is selected as the medical text. The electronic medical record data of the patients in the experiment are from the IBM T.J. Watson Research Center, which includes more than 4 years of data for 319,650 patients [43].

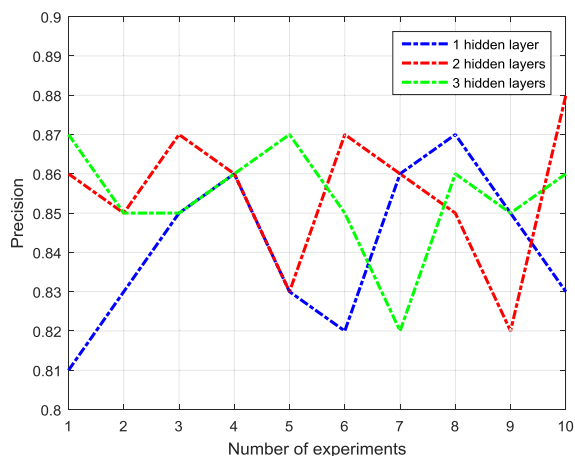


FIGURE 12. Different number of hidden layers of AutoEncoder model test accuracy.

Due to the limited sample data, the model designed in this paper may have over-fitting problems. In order to reduce the impact of this problem on the experimental results, this paper uses the k-fold cross-validation method. In this experiment, k is set to 10. All subjects' samples were randomly divided into 10 equal-sized subsample sets, one sample set was used as a validation set to verify the classification effect of the r model, and the remaining 9 subsample sets were used as training sets. Next, the cross-validation process is repeated 10 times, and each subsample set is used once as a validation set. After completing these steps, the overfitting is well resolved.

B. PARAMETER OPTIMIZATION OF AUTOENCODER MODEL

In this experiment, the optimal AutoEncoder network structure was found by 10-fold cross-validation of all training data. First, in this paper, an AutoEncoder network with different numbers of hidden layers is established, and then repeated experiments are performed for each network structure. In the experiment, it is found that after the number of experiments exceeds 10 times, the average value of the accuracy will gradually stabilize and there will be no large fluctuations. Therefore, this paper repeats 10 classification experiments for different network structures, and compares the input when the input is the correlation coefficient data. The accuracy of the test, if the average accuracy is higher, means that the classification effect is better, as shown in Figure 12. As can be seen from Figure 12, the AutoEncoder network with 3 hidden layers achieved the highest classification accuracy in 6 of all 10 experiments. In addition, the network average classification accuracy of the two hidden layers is 86.47%, while the average accuracy of the network of one hidden layer is 83.53%, and the average network accuracy of the three hidden layers is 84.70%. The network architecture with 3 hidden layers has the highest classification accuracy in the cross-validation test. Therefore, this paper selects 3 hidden layers of AutoEncoder network structure to establish the analysis model of medical image data.

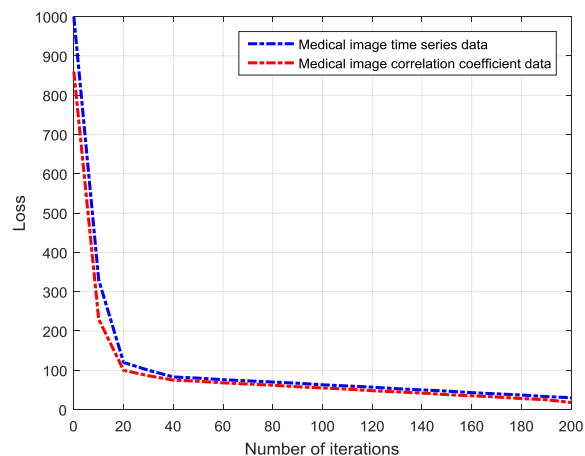


FIGURE 13. Loss reduction process for time series data and correlation coefficient data.

Choosing the right number of hidden layer nodes is very important for training the AutoEncoder model. In most documents, the data of hidden layer nodes is usually determined empirically. If this number is too small, the network will not be able to fit complex data; if this number is too large, it will increase training time and cause over-fitting problems. In this article, we use an AutoEncoder network with 3 hidden layers, each with 200 nodes. In order to express the advantages of using correlation coefficient data, this paper also compares the convergence of the loss function of time series data and correlation coefficient data of medical images. Figure 13 shows the variation of the training error loss with the number of iterations during the fine-tuning phase. It can be seen from Figure 13 that the training error is rapidly reduced at the beginning of the training under the condition of two kinds of input data; after about 50 iterations, the convergence speed of the training error gradually slows down and tends to be horizontal. At the same time, for the input time series data, after 100 iterations, the training error finally stabilizes at about 5×10^{-3} , and when the input is the correlation coefficient data, the training error is stabilized at about 1×10^{-3} after about 50 iterations.

C. PERFORMANCE BASED ON AUTOENCODER MEDICAL IMAGE DATA ANALYSIS

In the experiments in this paper, medical image sequence data and correlation coefficient data were placed in different machine learning models for training. The models are Linear Discriminant Analysis (LDA), Logistic Regression (LR), Support Vector Machine (SVM), and the AutoEncoder network used in this paper. It also has the models of literature [40], literature [41] and literature [42].

Since the accuracy of the classification is not the same when the data is randomly arranged, this paper repeats the experiment for classification 10 times, and 10-fold cross-validation is performed in each experiment. The mean and standard deviation of all experimental results were then calculated to compare the performance of the different models.

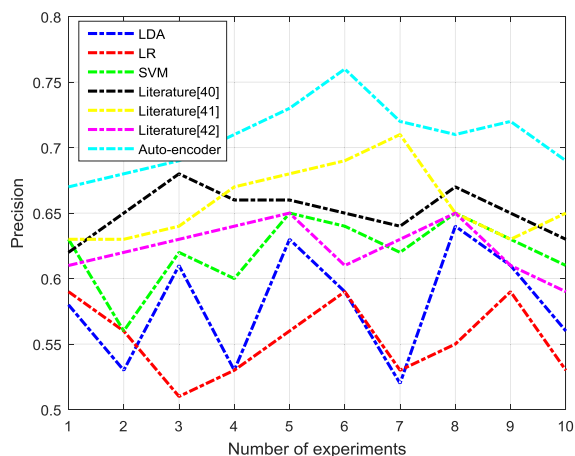


FIGURE 14. The input is the test accuracy of the time series data.

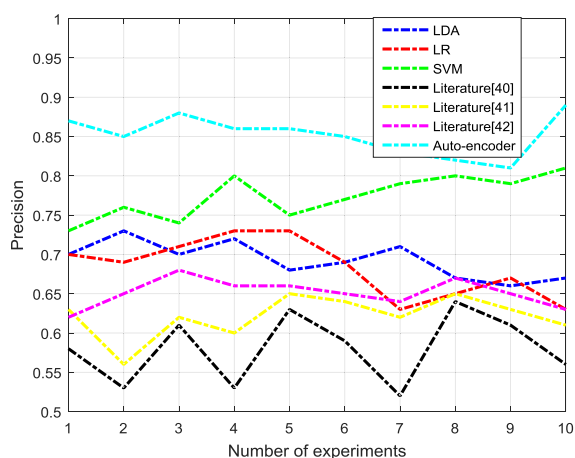


FIGURE 15. The input is the test accuracy of the correlation coefficient data.

In order to more intuitively represent the difference in data analysis accuracy of different algorithms, this paper draws the fluctuations of the experiment with the accuracy under different input data and different models, as shown in Figure 14 and Figure 15. Figure 14 shows the accuracy of each algorithm as a function of the number of experiments when the input data is medical image time series data. As can be seen from the figure, the classification accuracy of AutoEncoder is relatively high. When the input is correlation coefficient data, the test accuracy of each algorithm in all experimental sequences is as shown in Figure 15. The curves and data on the graph show that the accuracy of each algorithm has been greatly improved after the input of the medical image time series data becomes the correlation coefficient data, but the data analysis accuracy of the AutoEncoder network model is significantly higher than other models. This also shows that the AutoEncoder model is much more stable than other models that correspond to other input data.

In order to further verify the data analysis ability of the proposed model, the data information analyzed by the model is applied to the intelligent identification and prediction of diseases. By comparing the ROC curves of different models

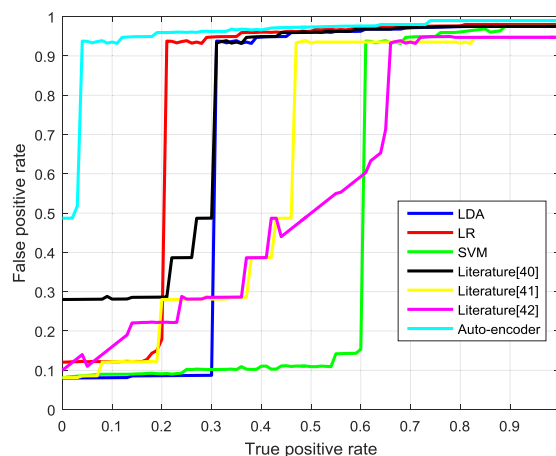


FIGURE 16. Enter the time series data for each model ROC curve.

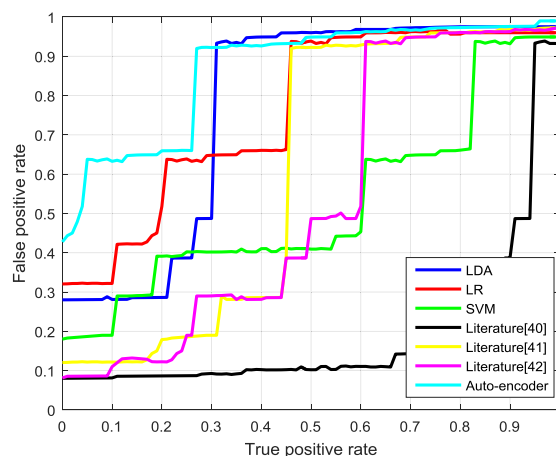


FIGURE 17. Enter the correlation coefficient data for each model ROC curve.

TABLE 2. AUC values for different models corresponding to different data.

	Medical image time series data	Medical image correlation coefficient data
LDA	0.585	0.712
LR	0.532	0.767
SVM	0.521	0.768
Literature [40]	0.532	0.793
Literature [41]	0.567	0.812
Literature [42]	0.578	0.832
AutoEncoder	0.619	0.913

in medical image time series data and correlation coefficient data, the pros and cons of their data analysis ability and prediction effect are compared. If the ROC curve is closer to the upper left corner, the closer the area under the curve (AUC) is to 1, the better the classification effect of the model, and vice versa. Figure 16 and Figure 17 are ROC plots based on experimental results, and Table 2 calculates their corresponding AUC values.

As can be seen from Figure 16, the ROC curve of the AutoEncoder model is mostly at the top left of the other curves. The ROC curves for LR and SVM linger around the

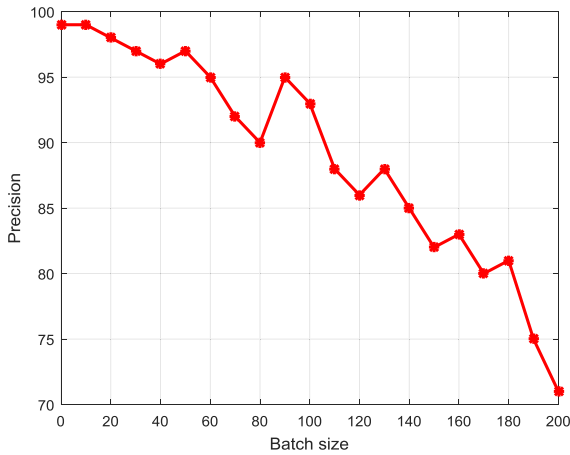


FIGURE 18. The accuracy of data analysis accuracy of heart failure varies with batch size.

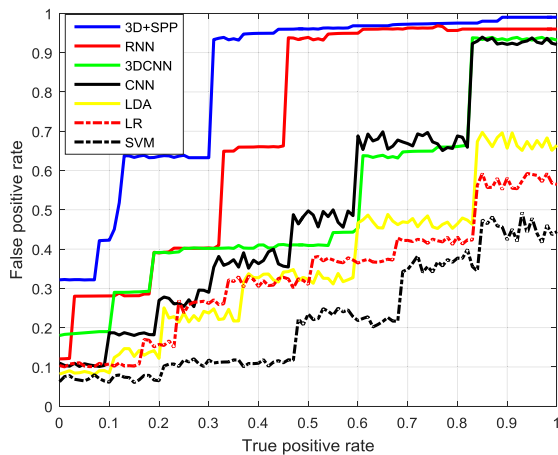


FIGURE 19. ROC curve for different models.

diagonal, while the ROC curve for LDA is slightly higher than the diagonal. It can also be seen from the AUC value calculated in Table 2 that the AUC value of AutoEncoder is 0.619, which is the best under the condition that the input is time series data, the lowest LR and SVM, basically about 0.5, the AUC value of LDA. Between AutoEncoder and LR, SVM, it is 0.585. It can be seen from Figure 17 that when the input is the correlation coefficient data, the ROC curve of all models is obviously biased to the upper left compared with the input time series data, and the classification effect of AutoEncoder is still significantly better than other models. Therefore, regardless of the input data, the data analysis effect of the AutoEncoder model constructed in this paper is much better than the comparison model.

D. PARAMETER OPTIMIZATION BASED ON 3D CONVOLUTIONAL NEURAL NETWORK AND SPP MODEL

In order to analyze the influence of the model structure of the fine-tuning 3D CNN-SPP on the data analysis effect, different batch sizes, different network depths and different phrase lengths of input data are set.

First of all, in the experiment, the range of the batch size can be varied from 1 to several hundred. Therefore, the

TABLE 3. Network data analysis results of different layers.

Model	Number of convolution layers	Data analysis accuracy
3D CNN-SPP	2	0.891
3D CNN-SPP	3	0.875
3D CNN-SPP	4	0.882

TABLE 4. Data analysis effect of network model under different phrase lengths.

Model	Number of phrase length	Data analysis accuracy
3D CNN-SPP	1	0.868
3D CNN-SPP	2	0.878
3D CNN-SPP	3	0.892
3D CNN-SPP	4	0.878

paper sets the batch size from 1 to 200, and the data analysis accuracy of the disease is shown in Figure 19. As can be seen from the figure, the data analysis accuracy is very high when the batch size is small, and the data analysis accuracy is the highest when the batch size is 3. Therefore, in practical applications, it is necessary to comprehensively evaluate according to the requirements of training accuracy and time, and select an appropriate batch size.

For the 3D CNN-SPP model in this paper, different model depths have a great impact on the data analysis of disease. This paper sets up a network structure with 2, 3, and 4 convolution layers, as shown in Table 3. As can be seen from Table 3, when the network contains a 2-layer convolution structure, the data analysis accuracy of heart failure is higher than that of 3 or 4 convolutional layers. The 3D CNN-SPP model designed in this paper has the best data analysis accuracy when it has 2 convolution layers. At the same time, in the experiment, the phrase lengths $k = 1, 2, 3, 4$ were set respectively, and the data analysis effects under the same model were compared. The experimental results are shown in Table 4.

As can be seen from the table, when the phrase length $k = 3$, the data analysis achieves the highest accuracy and the advantages are obvious. It is shown that when the length of the phrase is 3, the model of this paper achieves the best data analysis effect, which means that for the input medical text data, feature extraction with 3 as the time span is the best choice.

E. PERFORMANCE OF MEDICAL TEXT DATA ANALYSIS BASED ON 3D CONVOLUTIONAL NEURAL NETWORK AND SPP

In order to evaluate the advantages and disadvantages of 3D convolutional neural network and SPP combined medical text data analysis models, some excellent models were selected, including 3D convolutional neural network (3D CNN), common convolutional neural network (CNN), logistic regression (LR), support vector machine (SVM), random forest (RF) algorithms, and RNN model.

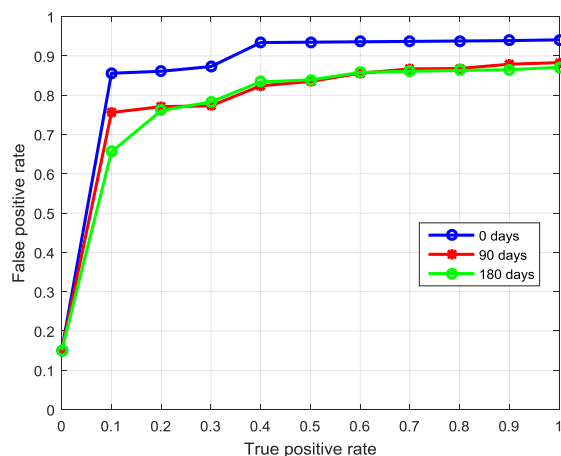


FIGURE 20. Recognition accuracy rate and AUC score with the number of convolution flow curve.

In order to further verify the data analysis ability of the proposed model, the ROC curves of the medical model time series data and correlation coefficient data are drawn by different models to compare their data analysis ability and prediction effect. If the ROC curve is closer to the upper left corner, that is, the area under the curve (area under the curve, AUC) is closer to that, and Figure 19 is a ROC graph plotted based on the experimental results.

As can be seen from Figure 19, the ROC curve of the medical text data analysis model based on 3D convolutional neural network and SPP is mostly located at the upper left of other curves and is closer to the upper left corner. The ROC curve of the RNN model is slightly lower than the model proposed in this paper, indicating that the model proposed in this paper is better. The ROC curves of 3DCNN and CNN are not much different in the diagonal, while the ROC curves of LR and SVM are the lowest. Therefore, regardless of the input data, the data analysis results of the medical text data analysis model based on 3D convolutional neural network and SPP constructed in this paper are far better than the comparison model.

It is also important to be able to perform intelligent recognition prediction in the task of data analysis based on the patient's electronic medical record data. Among them, the longer the time to analyze and predict from the medical record data, the earlier the measures can be taken to control the disease development, and the early diagnosis and treatment plan is very important to improve the cure rate of the disease. In order to analyze the data analysis ability of the model in this paper, the heart failure prediction is performed at different times in advance, and the network is compared at these three different time nodes. The data analysis capabilities of the model and the corresponding prediction accuracy are shown in Figure 20.

The ROC curve can comprehensively evaluate the data analysis ability and intelligent prediction effect of the model under different conditions. It can be seen from Figure 20 that the ROC curve at 0 days ahead is completely at the upper left

of the other curves, indicating that its data analysis ability and intelligent prediction effect are much better than 90 days and 180 days earlier. Explain that the deep prediction is based on the patient's electronic medical record records for intelligent prediction. It also proves the data analysis ability of the model for medical text data.

V. CONCLUSION

The sharp increase in medical big data puts higher demands on medical big data processing methods, and its complex and diverse formats increase the difficulty of its analysis. Among the existing data analysis methods, deep learning is undoubtedly the most effective one. According to the two main data formats of medical image data and medical text data, this paper designs the corresponding deep learning model separately, and realizes the intelligent identification of accurate early diagnosis and risk prediction for specific diseases. Firstly, the AutoEncoder deep learning model is designed, and it can pre-train the network in advance and reduce the consumption of computing time and computing resources. Therefore, the method can be easily extended to other medical image data analysis and processing, which is of great significance for improving the accuracy of disease diagnosis. Secondly, a deep learning model combining 3D convolutional neural network and spatial pyramid pooling is designed. The 3D convolution structure in the model can preserve the timing characteristics of different data while extracting the internal features of the data, while the spatial pyramid pooling structure can process input data of arbitrary length, thus making effective data analysis for the future of the patient in intelligent identification and control of disease risk.

REFERENCES

- [1] I. Tatsuya, "Information processing unit," *J. Mater. Process. Technol.*, vol. 3, no. 8, pp. 165–196, 2018.
- [2] D. Sumathi and P. Poongodi, "Secure medical information processing in cloud: Trust with swarm based scheduling," *J. Med. Imag. Health Informat.*, vol. 6, no. 7, pp. 1636–1640, Nov. 2016.
- [3] G. S. Birkhead, M. Klompas, and N. R. Shah, "Uses of electronic health records for public health surveillance to advance public health," *Annu. Rev. Public Health*, vol. 36, no. 1, pp. 345–359, Mar. 2015.
- [4] L. C. Gurrin, J. J. Kurinczuk, and P. R. Burton, "Bayesian statistics in medical research: An intuitive alternative to conventional data analysis," *J. Eval. Clin. Pract.*, vol. 6, no. 2, pp. 193–204, May 2010.
- [5] A. M. Hassan, Y. F. Hassan, and M. H. Kholief, "A deep classification system for medical data analysis," *J. Med. Imag. Health Informat.*, vol. 8, no. 2, pp. 250–256, Feb. 2018.
- [6] K. Lu, "Number of imputations needed to stabilize estimated treatment difference in longitudinal data analysis," *Stat. Methods Med. Res.*, vol. 26, no. 2, pp. 674–690, Oct. 2017.
- [7] V. Taslimitehrani, G. Dong, N. L. Pereira, M. Panahiazar, and J. Pathak, "Developing EHR-driven heart failure risk prediction models using CPXR(Log) with the probabilistic loss function," *J. Biomed. Informat.*, vol. 60, pp. 260–269, Apr. 2016.
- [8] Z. P. Liu, L. Y. Wu, Y. Wang, X. S. Zhang, and L. Chen, "Prediction of protein-RNA binding sites by a random forest method with combined features," *Bioinformatics*, vol. 26, no. 13, pp. 1616–1622, 2010.
- [9] J. H. Chen and J. Z. Lin, "Developing an SVM based risk hedging prediction model for construction material suppliers," *Autom. Construct.*, vol. 19, no. 6, pp. 702–708, Oct. 2010.

- [10] P. Włodarczak, J. Soar, and M. Ally, "Multimedia data mining using deep learning," in *Proc. 5th Int. Conf. Digit. Inf. Process. Commun.*, Oct. 2015, pp. 190–196.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [12] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1725–1732.
- [13] A. Lazaridis, I. Mporas, and T. Ganchev, "Improving phone duration modelling using support vector regression fusion," *Speech Commun.*, vol. 53, no. 1, pp. 85–97, Jan. 2011.
- [14] Y. Cui, S. Ahmad, and J. Hawkins, "Continuous online sequence learning with an unsupervised neural network model," *Neural Comput.*, vol. 28, no. 11, pp. 2474–2504, 2016.
- [15] A. Ogawa and T. Hori, "Error detection and accuracy estimation in automatic speech recognition using deep bidirectional recurrent neural networks," *Speech Commun.*, vol. 89, pp. 70–83, May 2017.
- [16] T. G. Lid, G. E. Eide, and I. Dalen, "Can routine information from electronic patient records predict a future diagnosis of alcohol use disorder?" *Scandin. J. Primary Health Care*, vol. 34, no. 3, pp. 215–223, Jul. 2016.
- [17] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, "Deep learning-based image segmentation on multimodal medical imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 162–169, Mar. 2019.
- [18] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.
- [19] R. T. Schirmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Aug. 2017.
- [20] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, "Multimodal and multiscale deep neural networks for the early diagnosis of alzheimer's disease using structural MR and FDG-PET images," *Sci. Rep.*, vol. 8, no. 1, Apr. 2018, Art. no. 5697.
- [21] F. A. Chaddad, C. Desrosiers, and T. Niazi, "Deep radiomic analysis of MRI related to alzheimer's disease," *IEEE Access*, vol. 6, pp. 58213–58221, 2018.
- [22] R. D. Cebul, T. E. Love, A. K. Jain, and C. J. Hebert, "Electronic health records and quality of diabetes care," *New England J. Med.*, vol. 365, no. 9, pp. 825–833, Sep. 2011.
- [23] G. Hripcsak and D. J. Albers, "Next-generation phenotyping of electronic health records," *J. Amer. Med. Informat. Assoc. Jamia*, vol. 20, no. 1, pp. 117–121, Sep. 2013.
- [24] R. Amarasingham, B. J. Moore, Y. P. Tabak, M. H. Drazner, C. A. Clark, S. Zhang, W. G. Reed, T. S. Swanson, Y. Ma, and E. A. Halm, "An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data," *Med. Care*, vol. 48, no. 11, pp. 981–988, 2010.
- [25] J. C. Ho, J. Ghosh, S. R. Steinhubl, W. F. Stewart, J. C. Denny, B. A. Malin, and J. Sun, "Limestone: High-throughput candidate phenotype generation via tensor factorization," *J. Biomed. Informat.*, vol. 52, pp. 199–211, Dec. 2014.
- [26] W. Q. Wei and J. C. Denny, "Extracting research-quality phenotypes from electronic health records to support precision medicine," *Genome Med.*, vol. 7, no. 1, Apr. 2015, Art. no. 41.
- [27] A. G. Winterstein, P. Kubilis, S. Bird, R. M. Cooper-DeHoff, G. A. Nichols, and J. A. Delaney, "Misclassification in assessment of diabetogenic risk using electronic health records," *Pharmacoepidemiol. Drug Saf.*, vol. 23, no. 8, pp. 875–881, Aug. 2014.
- [28] A. Khalifa and S. Meystre, "Adapting existing natural language processing resources for cardiovascular risk factors identification in clinical notes," *J. Biomed. Informat.*, vol. 58, pp. S128–S132, Dec. 2015.
- [29] H. R. Roth, L. Lu, A. Seff, K. M. Cherry, J. H. S. Wang, J. Liu, E. Turkbey, and R. M. Summers, "A New 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations," *Med. Image. Comput. Assist. Interv.*, vol. 17, no. 1, pp. 520–527, 2014.
- [30] T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1229–1239, May 2016.
- [31] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V. C. Mok, L. Shi, and P.-A. Heng, "Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1182–1195, May 2016.
- [32] G. van Tulder and M. de Bruijne, "Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted Boltzmann machines," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1262–1272, May 2016.
- [33] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.
- [34] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [exploratory DSP]," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 145–154, Jan. 2011.
- [35] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [37] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, Nov. 2014, pp. 675–678.
- [38] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.
- [39] K. Shameer, M. A. Badgeley, R. Miotto, B. S. Glicksberg, J. W. Morgan, and J. T. Dudley, "Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams," *Briefings Bioinf.*, vol. 18, no. 1, pp. 105–124, Feb. 2016.
- [40] J. Ma, Y. Yang, D. Huo, Z. Wang, X. Zhai, J. Chen, H. Sun, W. An, J. Jie, and P. Yang, "LincRNA-RoR/miR-145 promote invasion and metastasis in triple-negative breast cancer via targeting MUC1," *Biochem. Biophys. Res. Commun.*, vol. 500, no. 3, pp. 614–620, Jun. 2018.
- [41] Z. Li, J. Li, A. Ren, R. Cai, C. Ding, X. Qian, J. Draper, B. Yuan, J. Tang, Q. Qiu, and Y. Wang, "HEIF: Highly efficient stochastic computing-based inference framework for deep neural networks," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 38, no. 8, pp. 1543–1556, Aug. 2019.
- [42] M. Eberl, D. Mangelberger, J. B. Swanson, M. E. Verhaegen, P. W. Harms, M. L. Frohm, A. A. Dlugosz, and S. Y. Wong, "Tumor architecture and notch signaling modulate drug response in basal cell carcinoma," *Cancer Cell*, vol. 33, no. 2, pp. 229–243, Feb. 2018.
- [43] F. Alfaro-Almagro et al., "Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank," *Neuroimage*, vol. 166, pp. 400–424, Feb. 2018.



HANQING SUN graduated from Zhengzhou University, in 2009. He is currently a Teacher with the Henan University of Animal Husbandry and Economy. His research interests include information processing and data analysis.



ZHENG LIU graduated from Zhengzhou University, in 2009. He is currently a Teacher with the Henan University of Animal Husbandry and Economy. His research interests include information processing and data analysis.



GUIZHI WANG graduated from Zhengzhou University, in 2008. She is currently a Teacher with the Henan University of Animal Husbandry and Economy. Her research interest includes data analysis.



JUN MA graduated from the Shandong University of Technology, in 2017. He is currently pursuing the master's degree with the School of Mechanical Engineering, Dalian University of Technology. His research interests include reliability analysis of complex systems and fault diagnosis.

...



WEIMIN LIAN graduated from Zhengzhou University, in 2003. He is currently a Teacher with the Henan University of Animal Husbandry and Economy. His research interests include software engineering and data analysis.