Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 10, No. 3, 2023

TCM disease diagnosis based on convolutional cyclic neural network algorithm

Fanpeng Kong, Mingchuan Zhang

Henan University of Science and Technology, Luoyang 471000, China

Abstract: The unique diagnosis and treatment mode of traditional Chinese medicine provides a lot of diagnostic basis with reference value for modern medicine, which is welcomed by people all over the world and has attracted extensive attention of medical researchers. However, the unique diagnosis and treatment method of traditional Chinese medicine also brings difficulties to the dissemination of traditional Chinese medicine. The diagnosis method of traditional Chinese medicine is difficult to objectively and quantitatively express, and the diagnosis process is also closely related to the subjective experience of traditional Chinese medicine doctors. , lacking established standards. In view of the above-mentioned difficulties in the modernization of traditional Chinese medicine, the development of modern information science and technology has brought a turning point to the intelligentization of disease differentiation in traditional Chinese medicine. The core of traditional Chinese medicine diagnosis and treatment is diagnosis and classification. How to convert the difficult to describe disease diagnosis and treatment process of traditional Chinese medicine into language that can be distinguished and understood by ordinary doctors, and build an auxiliary diagnosis and treatment model to provide reference for Chinese doctors to see a doctor is the main problem to be solved in this paper.

Keywords: TCM diagnosis of diseases; Recurrent Neural Network; Convolutional Neural Network.

1. Introduction

TCM disease differentiation is one of the directions of future medical development, and the combination of TCM disease differentiation and TCM syndrome differentiation will be the general trend in the future. However, the traditional Chinese medicine diagnosis and treatment methods have not kept up with the trend of the times and hindered the development of Chinese medicine. With the development of modern information science and technology, Chinese medicine has developed in the auxiliary diagnosis and treatment. The auxiliary diagnosis and treatment model is still relatively rare. TCM disease differentiation is a mapping relationship from TCM symptoms to TCM syndromes, which conforms to the overall architecture of neural networks. Through the convolution layer in the convolutional neural network, the TCM symptom features with rich background relationships are extracted, and the extracted TCM symptom sequence features are input into the recurrent neural network. The extracted TCM symptoms are associated with TCM disease categories, so as to realize TCM disease identification and provide a reference for TCM doctors to diagnose and treat.

2. Application Scenarios of TCM Disease Diagnosis

The use of auxiliary diagnosis and treatment technology to diagnose diseases is an inseparable model in modern society. When entering a hospital, you will always come into contact with various medical devices first. Doctors can obtain various indicators of the patient's body through the diagnosis of medical devices. This allows for a more accurate diagnosis of the patient's condition. Auxiliary diagnosis and treatment in TCM is equally important. From patient information and symptoms, through the TCM disease identification system model, the patient's disease category can be directly derived. This diagnosis and treatment model can not only reduce the workload of TCM doctors during diagnosis, but also help medical resources. Physicians in deficient areas can make diagnoses, thereby alleviating the problem of uneven distribution of medical resources.

3. Recurrent Neural Network Algorithm

In 1986, Michaell Jordan proposed a neural network framework, a new recurrent network based on the theory of distributed parallel processing. This neural network is called a simple recurrent network [1]. This simple neural network is regarded as a the most basic network architecture of recurrent neural network. Recurrent Neural Network [2,3,4] (Recurrent Neural Network, RNN) is fully connected between layers without nodes. The input and the previous hidden state are combined into a vector containing the information of the current and previous input, and then the new hidden state or network memory is output through the activation function. The hidden state here can also be called a time step. Only one-time step can proceed to the next one, so the input time series problem also needs to be considered. Recurrent neural networks have been widely used in various fields related to sequence data. The basic structure of the recurrent neural network is shown in Figure 1 (the picture comes from the network).

The recurrent neural network also contains a three-layer basic structure [5,6,7]. Through the loop connection of the hidden layer, the network state of the previous moment and the current network state can be transmitted to the next moment, which is an extension in the time dimension. The input of the recurrent neural network is mainly the forward propagation of the RNN. The hidden layer of the recurrent neural network mainly receives data in two directions, one is to receive the network state data at the previous moment, and the other is the newly input data of the input layer. Therefore, The RNN will have dependencies on the state of all previous moments [8]. At the same time, the network can map the relationship between the input sequence data and the output sequence, and can also have multiple mapping relationships, one-to-many correspondence mode. In this way of training data, the output value is affected by the previous input and subsequent input, which will make the result more accurate [9].

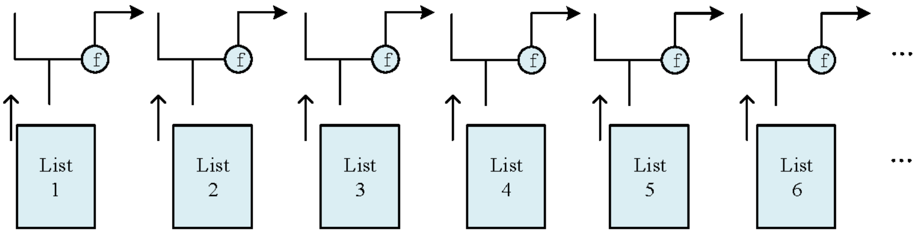

Fig. 1 Recurrent Neural Network Structure

Convolutional Neural Networks (CNN), which is also a very mature and classic neural network architecture, is widely used in various fields of life, mainly including target detection and image vision. The field of natural language processing. As early as the 1960s, influenced by the process of the human brain's visual information being transmitted to the brain, the early research basis of convolutional networks, that is, the receptive field theory, was proposed [10]. In various large-scale image field competitions, CNN was used to identify the high accuracy rate won the championship. So far, the convolutional neural network has become one of the popular networks. In recent years, convolutional neural networks have been applied in the fields of natural language processing such as text classification and feature recognition. Its characteristics can reduce the complexity of the network, keep the model stable to a certain extent, and be easy to optimize and train.

In text processing [11], convolutional neural networks can learn local responses from temporal or spatial data, but lack the ability to learn sequential related sequences; recurrent neural networks build data processing models based on sequences, but it is difficult to fully extract features. Therefore, in order to take full advantage of the advantages of convolutional neural networks [12] and recurrent neural networks, this chapter proposes an algorithm C-RNN that combines convolutional neural networks and recurrent neural networks to be applied in TCM disease differentiation. The output of one layer of the convolutional layer of the convolutional neural network is input into the cyclic neural network. The convolutional neural network is constructed on the basis of the pre-trained word vectors. In order to better improve the sequence correlation by using the advantages of the two algorithms Learning ability, the feature map structure of convolutional neural network is organized into sequence features as input to recurrent neural network [13]. When constructing a recurrent neural network in this way, each sentence is first converted into a continuous window feature, which is used to help decompose the changing factors in the sentence. Before inputting the medical data of traditional Chinese medicine diagnosis into the neural network, the input based on the sequence is selected, which does not rely on grammatical analysis.

4. Convolutional Recurrent Neural Network Algorithm

In this chapter, Convolutional Neural Networks and Recurrent Neural Networks are stacked in a unified framework for text language modeling [14,15]. The combination of Convolutional Neural Networks and Recurrent Neural Networks has been used in computer vision tasks, with applications in image captioning, speech recognition translation, tracking stock price fluctuations, image labeling, and more. Most of these models use classical convolutional or recurrent neural networks in artificial neural networks, or transfer the output feature vector of a fully connected layer convolutional network into a recurrent neural network as an input vector [16]. In this chapter, through the convolutional layer in CNN, the TCM symptom features with rich background relationships are extracted, and the extracted TCM symptom sequence features are input into the recurrent neural network [17]. The extracted TCM symptoms are associated with TCM disease categories. Therefore, the network architecture in this section enables recurrent neural networks to learn long-term dependencies from high-order sequential features. In the C-RNN algorithm [18], the convolutional neural network is directly built on word sequences instead of syntactic parse trees. Figure 2 is a schematic diagram of the basic structure of the C-RNN algorithm [19].

Before training the data, it is necessary to encode the data, which is easy for machine recognition. The Chinese data is complex, the meaning of words is difficult to understand, and there is no obvious separation mark between words. The jieba word segmentation tool is the word segmentation tool used in this chapter. A dictionary is created, which contains more than 20,000 Chinese words. The trie tree is generated based on this dictionary. When scanning the dictionary, the number of occurrences of each word is counted. This process is called word graph scanning, which is equivalent to each word having a prefix, which can greatly speed up the search for words, and also supports the detection of traditional characters. The main research work includes separation of vocabulary, abbreviated tagging of parts of speech and extraction of keywords. After Chinese word segmentation, the data needs to be represented by a vector, and the Skip-Gram mode in Word2Vec is used to convert the word into a vector that the neural network can recognize.

To detect and extract features from different positions on TCM case text data sequences, one-dimensional convolution contains a filter vector. is the dimensional word vector for the th word in each TCM text sequence? Let denote the sequence of input vectors, where is the length of the sequence. Indicates the length of the filter. When extracting the features of TCM text data, the convolution calculation of the filter is represented by a vector. For each position in the TCM symptom sequence, a vector window will be generated, and it also contains a continuous TCM symptom word vector, which is expressed as formula (1).

$$w_j = [x_j, x_j, \cdots, x_{j+k-1}] \tag{1}$$

where the row-vector concatenation of the text is represented by a comma. Using the filter and the vector convolution window (k-grams), the feature vector in the TCM case text can be extracted, and an effective feature map of TCM symptoms to disease classes can be generated; the feature map of the window vector of each TCM feature element is as follows:

$$c_i = f(w_j \circ m + b) \tag{2}$$

In formula (2), represents the multiplication of element features. represents the bias term, and is a curve transfer function of type. The C-RNN model uses multiple filter tools to extract features from TCM case texts, and generates multiple feature maps, which rearrange the features to represent each text feature window.

$$W = [c_1; c_2; \cdots; c_n] \tag{3}$$

In formula (3), the connection of column vectors is represented by semicolons, and the mapping feature collected by the th filter is Each row in is a new feature representation of a sequence window vector at, generated by filters. The new feature vector sequence window representation is then input into the RNN, and for each sequence, the RNN input is specified, due to discontinuous feature selection. An RNN network model is stacked behind the CNN, and the pooling layer operation in the CNN is removed.

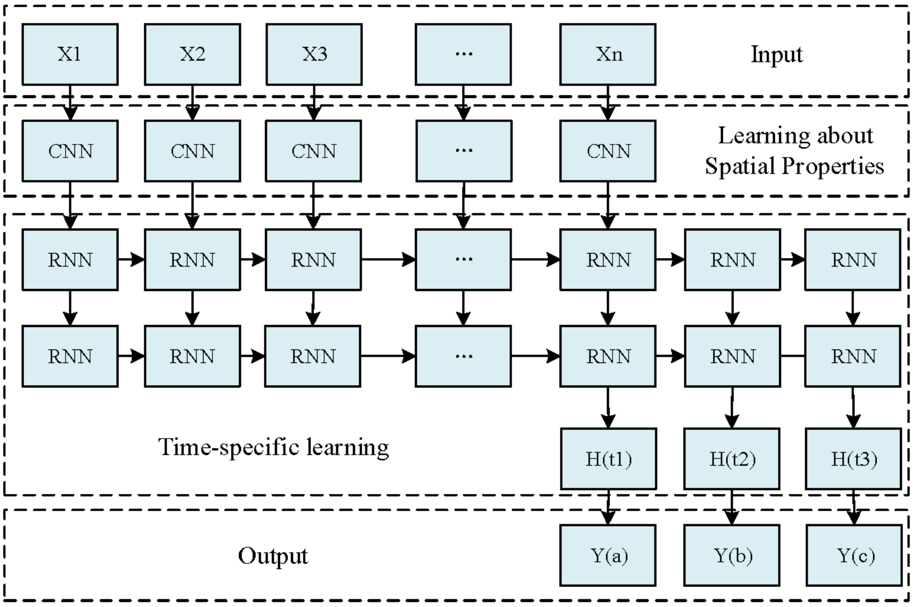

Fig. 2 Convolutional Recurrent Neural Network Architecture

This is achieved based on previous hidden states and inputs from the current stage. The gate of forgetting, input and output. All gates together determine the current memory and current hidden state.

The LSTM architecture has a series of repeated modules for each time series. At each time step, the output of the text sequence module is controlled by a set of gates in , which is a function of the hidden state of the previously input TCM text feature vector and the time step input of the currently input text feature vector . Also includes input gates, output gates, and forget gates, which together determine the current memory cell and the current hidden state. Represents the memory dimension in LSTM, all vectors in this architecture share the same dimension. The RNN transfer function is defined as follows formula (4).

$$\begin{cases} i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ q_t = \tanh(W_q \cdot [h_{t-1}, x_t] + b_q) \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ c_t = f_t \odot c_{t-1} + i_t \odot q_t \\ h_t = o_t \odot \tanh(c_t) \end{cases} \tag{4}$$

In the above formula, is the sigmoid function, and its output is [0,1], which represents the product of the elements. tanh is a hyperbolic tangent function, the output is [1, 1], to understand the structure of the algorithm, it can be regarded as the degree to which the control information is discarded. The information here represents the information of the old storage unit, and the control is stored in the current How much new information is contained in the cell is determined by the control based on how much the memory cell outputs. LSTM is specially designed for time series data and can be used to learn long-term dependencies, so LSTM based on convolutional layers is chosen to learn the mapping relationship on such high-level feature sequences of TCM case texts. When processing text, after outputting the hidden state of the last time step of LSTM, this state is used to represent the text document of traditional Chinese medicine, and the function SoftMax layer is added to the neural network architecture, and the error training of the entire model is performed by minimizing the cross entropy. Given a training sample of TCM symptoms and a true label of TCM disease category, where is the predicted possible label of the sample, and its evaluation probability of evaluating each label, the error is defined as formula (5).

$$L(x^{(i)}, y^{(i)}) = \sum_{j=1}^{k} 1\left\{y^{(i)} = j\right\} \log(\tilde{y}_j^{(i)}) \tag{5}$$

where 1{condition} is an indicator that 1{condition is true} = 1, otherwise 1{condition is false} = 0. Model parameter learning is performed using stochastic gradient descent

(SGD), and the network is trained with the RMSprop optimization algorithm.

At the beginning of training, a fixed length is used to represent the maximum length of case symptoms in the TCM symptom dataset. Because the data transmission in the C-RNN model requires a fixed length of the input data, a fixed value is set for the data length of the data set. The length of some data will be longer or shorter than this fixed length. Short text data needs to use Special symbols are supplemented into text sequences of the same length to represent unknown text data. For long text sequence data, redundant length data sequences need to be cut off. Word2vec vector is an initialization word vector tool. It will first select 20 TCM text datasets for pre-training, and the dimension of the vector is set to 300. When the unknown word vector is initialized, the uniform distribution in the range of [-0.25, 0.25] is selected. In the process of training the TCM text data set, the representation of the word vector and the parameters of the network model will be adjusted accordingly.

5. Experimental process and result analysis

Both convolutional neural network and recurrent neural network belong to the classical algorithm network of artificial neural network, but both networks have their own advantages and disadvantages. Convolutional neural network has better feature extraction ability, which can search from global search to Chinese medicine text data set The cyclic neural network searches for the features of TCM texts in order, which will make it easier to miss the important symptoms of TCM. Therefore, the convolutional layer of the convolutional neural network is selected to extract the features of TCM texts, and then the extracted text vectors are input. In the recurrent neural network, training is carried out to realize the diagnosis of diseases in traditional Chinese medicine.

5.1. Experimental dataset

The original medical record data set of TCM is extracted from the medical books "Traditional Chinese Medicine Surgery", "Traditional Chinese Medicine Internal Science", "Traditional Chinese Medicine Pediatrics" and other series of books with a total of 8270 cases. There are 11 kinds of diseases in organs, emotions, diseases of qi, blood and body fluid, diseases of head, body and body, diseases of skin and mucous membranes, reproductive diseases, diseases of eyes, ears, nose, throat and mouth, tumors and cancer and other diseases. Disease category. Each symptom and condition type have its own variable name, as shown in Table 1.

Table 1. Dataset

Num	Symptom	Syndromes
1	Photophobia, tearing, astringent pain, red tongue and yellow coating, fast pulse.etc	Eye, Ear, Nose, Throat and Mouth Diseases
2	Sputum smear, concentrate or culture, mostly positive for Mycobacterium tuberculosis. etc	viscera diseases
3	Thirsty and drinking more, eating more and easily hungry, frequent urination, and body weight loss. etc	Qi, blood and body fluid diseases
4	The tongue is pale and fat, the coating is white and the pulse is sinking. etc	head body disease

In the process of syndrome differentiation using the ABC-BP neural network algorithm, it is necessary to train each piece of data and find the mapping relationship between TCM symptoms and TCM syndrome types.

The experimental data set should be divided into two parts. Eighty percent is used for data training. In order to ensure the accuracy of the experiment, the remaining part is used for testing. The datasets used for training and testing are all processed datasets.

5.2. Evaluation method

On the TCM data set, the convolutional cyclic neural network algorithm, the convolutional neural network algorithm and the cyclic neural network algorithm (where RNN is the traditional cyclic neural network) are respectively trained on the data set experimentally, and the obtained training results are compared with the actual ones. The expected value is compared, and the accuracy and loss rate are calculated. The calculation formula is shown below.

$$Accuracy = \frac{|TRD|}{|ERD|} \tag{6}$$

$$Lossrate = \frac{|LOSS|}{|TRD|} \tag{7}$$

Among them, $|TRD|$ represents the number of correct TCM disease categories obtained in the experimental training; $|LOSS|$ represents the loss of precision during the experimental training process, and $|ERD|$ represents the expected number of TCM disease types in the experimental training data. The obtained algorithm training model is used for experimental testing.

5.3. Experimental training result graph

The strengths and weaknesses of the C-RNN algorithm learning are evaluated using the TCM disease identification dataset. The experimental results show that the C-RNN model can achieve excellent results under multiple benchmarks compared with a wide range of neural network models. It can be seen from the results that the cyclic neural network C-RNN model improves the accuracy of disease identification, improves the algorithm efficiency, and reduces the training loss in the process of TCM disease identification. Combining Convolutional Neural Networks and Recurrent Neural Networks outperforms separate multi-layer CNN models and traditional RNN models, suggesting that C-RNNs can learn high-level representations of long-term dependencies better than other models. In clinical practice, it can help TCM doctors to conduct auxiliary diagnosis and treatment, and to a certain extent promote the intelligent development of TCM diagnosis and treatment.

After the comparison and analysis between the algorithms, the comparison of the experimental results of the three algorithms is shown in Figure 3-left. During the experiment, the loss rate of the algorithm model is calculated, which can reflect the performance of the model and measure the practicability of the model prediction. The comparison of the loss rate of the three algorithms is shown in Figure 3-right.

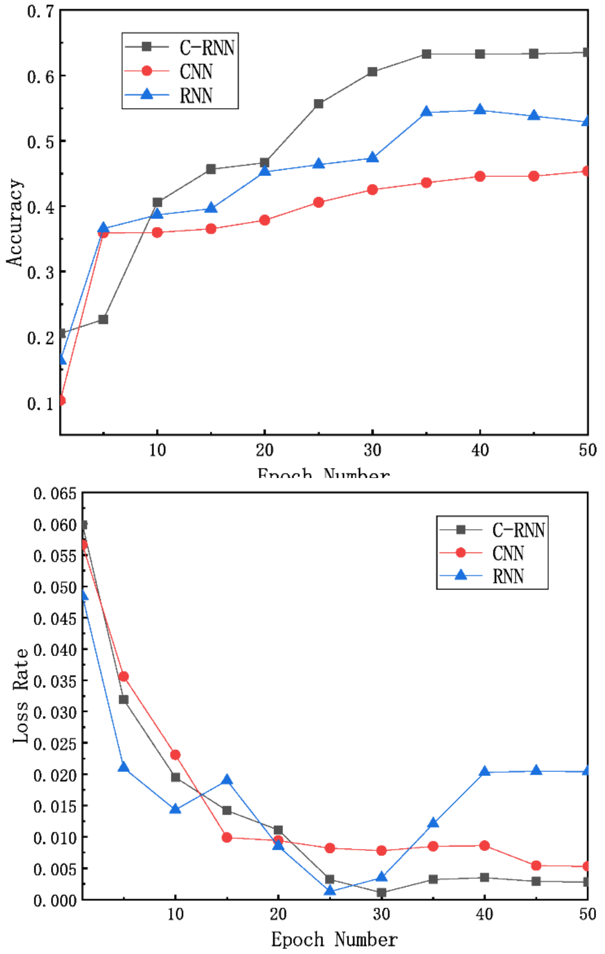

Fig. 3 Accuracy comparison chart and loss rate comparison chart

5.4. Experimental results

The strengths and weaknesses of the C-RNN algorithm learning are evaluated using the TCM disease identification dataset. The experimental results show that the C-RNN model can achieve excellent results under multiple benchmarks compared with a wide range of neural network models. It can be seen from the results that the cyclic neural network C-RNN model improves the accuracy of disease identification, improves the algorithm efficiency, and reduces the training loss in the process of TCM disease identification. Combining Convolutional Neural Networks and Recurrent Neural Networks outperforms separate multi-layer CNN models and traditional RNN models, suggesting that C-RNNs can learn high-level representations of long-term dependencies better than other models. In clinical practice, it can help TCM doctors to conduct auxiliary diagnosis and treatment, and to a certain extent promote the intelligent development of TCM diagnosis and treatment.

6. Conclusion

This paper mainly studies the algorithm selection of TCM disease differentiation, and designs the TCM syndrome differentiation standard and the architecture design of the recurrent neural network in the process of TCM disease differentiation. Aiming at the spatial characteristics of the convolutional neural network and the time sequence characteristics of the recurrent neural network, the algorithm in this chapter combines the two to solve the diagnosis problem of TCM disease differentiation. Finally, the simulation experiment of the algorithm is carried out, and the results show that the method has a high evolution rate and high convergence accuracy. This paper proposes to apply the improved C-RNN algorithm to TCM disease differentiation, and compares it with the traditional RNN and CNN algorithms respectively. Compared with the traditional neural network method, the time complexity of this method has been significantly improved, especially the C-RNN algorithm has good robustness in solving high-dimensional functions, which is similar to the traditional complex traditional Chinese medicine disease differentiation method. than has a good fit.

References

[1] M. Chen, W.-Z. Li, L. Qian, et al. Ming chen 2020 next poi recommendation based on location interest mining with recurrent neural networks. Journal of Computer Science and Technology (English version), vol. 35, no. 3, pp. 603–616, 5 2020.

[2] Y. Yu, X. Si, C. Hu, et al. A review of recurrent neural networks: Lstm cells and network architectures. Neural computation, vol. 31, no. 7, pp. 1235–1270, 2019.

[3] C. P. Khatri, N. Parikh, S. N. Solanki, et al. Snippet extractor: Recurrent neural networks for text summarization at industry scale. 2020.

[4] B. Li, Y. He.A feature-extraction-based lightweight convolutional and recurrent neural networks adaptive computing model for container terminal liner handling volume forecasting. Discrete Dynamics in Nature and Society, vol. 2021, no. 1, pp. 1–17, 2021.

[5] S. E. Okel, F. Sommen, E. Selmanaj, et al. Tissue-border detection in volumetric laser endomicroscopy using bi-directional gated recurrent neural networks. 2021.

[6] N. Padmaja, P. Balasubramaniam. Finite-time passification offractional-order recurrent neural networks withproportional delay andimpulses: anlmi approach. 2021.

[7] S. Jain, C. Bruckmann, C. Mcdougall. Nft appraisal prediction: Utilizing search trends, public market data,linear regression and recurrent neural networks. 2022.

[8] F. M. Awan, R. Minerva, N. Crespi.Using noise pollution data for traffic prediction in smart cities: Experiments based on lstm recurrent neural networks. IEEE Sensors Journal, vol. PP, no. 99, pp. 1–1, 2021.

[9] Y. Li, X. Wang. Almost periodic solutions in distribution of clifford-valued stochastic recurrent neural networks with time-varying delays. Chaos, Solitons Fractals, vol. 153, 2021.

[10] G. J. Pan, Y. P. Liu. Research and discussion on academic thoughts of diagnosis and treatment of phlegm disease in tcm during the ming dynasty. China Journal of Traditional Chinese Medicine and Pharmacy, pp. 494– 497, 2009.

[11] C. J. Zhang, Z. S. Yan, L. Q. Song. An algorithm of mining association rules of treatment information dealing with tcm nephrosis. Journal of Natural Science of Heilongjiang University, 2005.

[12] J. P. Yang, H. Zhao, Y. Z. Du, et al. Study on quantitative diagnosis model of tcm syndromes of post-stroke depression based on combination of disease and syndrome. Medicine, vol. 100, no. 12, p. e25041, 2021.

[13] S. Adavanne, P. Pertil, T. Virtanen. Sound event detection using spatial features and convolutional recurrent neural network. IEEE, 2017.

[14] C. C. Kao, W. Wang, M. Sun, et al. R-crnn: Region-based convolutional recurrent neural network for audio event detection. Interspeech 2018, 2018.

[15] N. K. Kim, K. K. Hong. Polyphonic sound event detection based on residual convolutional recurrent neural network with semi-supervised loss function. IEEE Access, vol. PP, no. 99, pp. 1–1, 2021.

[16] A. Onan. Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification. 2022.

[17] K. Kobayashi, S. Matsushita, N. Shimizu, et al. Automated detection of mouse scratching behaviour using convolutional recurrent neural network. Scientific Reports, vol. 11, no. 1, p. 658, 2021.

[18] J. Yi, J. Park. Hypergraph convolutional recurrent neural network. 2020.

[19] Z. Cui, K. Henrickson, R. Ke, et al. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. IEEE Transactions on Intelligent Transportation Systems, vol. PP, no. 99, 2019.