Automated Classification of Textual SPECT Diagnostic Reports with TextCNN Model

Chengcheng Han 1,2, Qiang Lin 1,2*, Zhengxing Man 1,2, Yongchun Cao 1,2, and Haijun Wang 3

1School of Mathematics and Computer Science, Northwest Minzu University; No. 1, Xibei Xincun Rd. Lanzhou city, Gansu Province, China 730030
2Key Laboratory of Streaming Computing and Applications, Northwest Minzu University
3Department of Nuclear Medicine, Gansu Provincial Hospital; No. 204, Donggang Xilu Rd. Lanzhou city, Gansu Province, China 730000

2307115582@qq.com, qianglin@xbmu.edu.cn, wwhhh@q.com, 408690991@qq.com, 1718315929@qq.com

*Correspondence: qianglin@xbmu.edu.cn; Tel.: +86 - 13909405358

Abstract. The SPECT diagnostic text contains several aspects of the patient’s personal information, image description, and suggested results. In order to construct a diagnostic model of nuclear medical text, a classification method of nuclear medical text based on deep learning was proposed. TextCNN was applied to propose the classification method of diseases. A set of real nuclear medical text data was used to verify the proposed method, the experimental results show that the proposed method has a good classification effect.

1. Introduction

Medical imaging plays key role in diagnosis and treatment of a variety of diseases in modern clinical medicine. It has become an important technical means to display both the structural and functional variants of the body in a non-invasive way. When a solid cancer has invaded into the bone, the patient would develop bone metastasis which can be captured by functional imaging, such as SPECT imaging. In order to automatically diagnose disease with textural SPECT diagnostic reports, a classification method of nuclear medical text based on deep learning was proposed. TextCNN was applied to propose the classification method of diseases. A set of real nuclear medical text data was used to verify the proposed method, the experimental results show that the proposed method is workable for the text classification task.

2. Related work

It is impossible for doctors to exist everywhere in real life, and Anand [1] proposed a diagnostic method for the same set of symptoms that patients deal with by using current computing power and information technology. Keke [2] proposed a new model to describe the new relationship generated by the cloud-based PCT system, which can help patients reduce the impact of cloud-based self-diagnosis. However, this self-diagnosis has a certain error compared with the doctor’s diagnosis. Bourkache [3]
proposed an image analysis algorithm for the diagnosis of breast cancer, which can accurately distinguish between tumor mammograms and healthy breast x-rays under the background of big data.

At present, the field of SPECT nuclear medicine is a research hotspot in the academic circle. In order to explore the possibility of the application of computer X-ray imaging system in nuclear medical imaging, Najeeb [4] proposed an automatic segmentation image processing method and an image analysis method based on the segmentation image center. Ming [5] summarized the treatment, interpretation and diagnostic performance of nuclear medicine and molecular imaging in two unique and underdiagnosed cardiomyopathy i.e., cardiac amyloidosis and cardiac sarcoidosis.

3. Materials and methods
In this section, the used data of SPECT diagnostic reports and the developed diagnostic model based on Text-CNN will be elaborated.

3.1. Dataset
The used diagnostic reports were collected in the process of diagnosing bone metastasis and related diseases in Department of Nuclear Medicine, Gansu Provincial Hospital from Jan. 2018 to Dec. 2019. For the examination of a patient, two SPECT images (i.e., the anterior and posterior) and a corresponding diagnostic report in text format are recorded. Figure 1 demonstrates an example of SPECT examination, with a male patient diagnosed with bone metastasis in ribs.

![Figure 1. An example of SPECT examination. a) The imaging results; and b) Textual diagnostic report](image)

For the textual diagnostic report, we can see that a diagnostic report text can be divided into two main parts, i.e., descriptive information of lesions and diagnosis solutions.

Our objective is to propose a text classification method based on deep learning by studying the diagnosis description and results, and to accurately construct a nuclear medicine auxiliary diagnosis model. Table 1 outlines the used data of textual diagnostic reports in this work.

| Diseases             | Bone metastasis | Arthritis | Degenerative changes | Normal |
|----------------------|-----------------|-----------|----------------------|--------|
| Number               | 748             | 1021      | 532                  | 820    |

3.2. Data preprocessing
It can be seen from the description of the diagnosis report. Text preprocessing usually includes the process of medical text acquisition, word segmentation and text cleaning.

The content of nuclear medicine text is described in natural language. Lesion characterization is the key words found from a large number of nuclear medicine texts that can describe the characteristics of lesions, including position, shape, level and state. Given any SPECT case, you can always find
position, shape, level and state, and Classification information (or at least some of them). The first four are the description of the lesion itself, and the last is the description of the disease classification. Based on this, the lesions can be formally expressed as five tuples as follows:

\[ RL = (P, S, L, T, C) \]  

where \( P, S, L, T, \) and \( C \) represents lesion position (Position), shape (Shape), level (Level), state (sTate), and disease Classification (Class), respectively.

3.3. Text-CNN based model

The TextCNN proposed by Kim [6] is a three-layer convolutional neural network based on multiple convolution kernels, it's an improvement from CNN. Then, through embedding layer, the words are mapped into a low dimensional space, which is essentially a feature extractor, encoding semantic features in the specified dimensions. Then extract the largest feature through the convolutional pooling layer, and finally use softmax to output the probability value. As shown in Figure 3, the input layer is a 9*6 word vector matrix \( X \). Among them, 9 represents the number of words in each sentence, and 6 represents the dimension of the corresponding word vector. Two convolution kernels are used to perform convolution operations on word vectors, and the three convolution kernels are 2*6 and 3*6 respectively. In this paper, Tanh is selected as the activation function, which can well solve this problem.

![Figure 2. The architecture of TextCNN network.](image)

4. Results

In this section, the developed diagnosis model will be evaluated with a group of real-world data as shown in Table 1.

4.1 Experimental setup

Accuracy, as an evaluation index in this experiment, refers to the closeness between the test result and the true value. In order to facilitate the understanding of the following definitions, Table 2 shows the used evaluation metrics in this paper.

| Matrix | Acc  | Prec  | Rec  | F1               |
|--------|------|-------|------|------------------|
| Definition | \((TP+TN)/(TP+FP+TN+FN)\) | \(TP/(TP+FP)\) | \(TP/(TP+FN)\) | \(2(Pre \times Re)/(Pre + Re)\) |

The experimental hyperparameters determine the quality of an experimental result. Table 3 describes the main parameter configuration of TextCNN.
Table 3. Hyperparameters of the defined model

| Hyperparameters   | Implication                              | Value             |
|-------------------|------------------------------------------|-------------------|
| filter_sizes      | Convolution kernel size                  | Determined by experimental setting |
| num_filters       | Number of convolution kernels            | Determined by experimental setting |
| dropout_keep_prob | Dropout rate                             | 0.5               |
| embedding_dim     | The embedding dimension                   | 128               |
| batch_size        | Number of samples                        | 64                |
| num_epoch         | num_epoch of model                       | 200               |

4.2. Experimental results

This experiment uses TextCNN to classify text. In order to better observe the influence of filter_sizes and num_filters on the experiment. Three groups of experiments are [1, 2, 3], [2, 3, 4], [3, 4, 5]. Table 4 gives the experimental results. It can be seen that filter_sizes is set to [3, 4, 5], and the value of Acc is higher.

Table 4. Influence of Filter_sizes on experiment

| Plan  | Filter_sizes | Acc    |
|-------|--------------|--------|
| plan 1 | [1, 2, 3]    | 0.934968 |
| plan 2 | [2, 3, 4]    | 0.934968 |
| plan 3 | [3, 4, 5]    | 0.940298 |

After the size of the convolution kernel is determined, the number of num_filters also determines the quality of the experiment. It can be seen from Table 5 that when num_filters is 128, better results can be achieved.

Table 5. The impact of Num Filters on experiments

| Plan  | Num_filters | Acc    |
|-------|-------------|--------|
| plan 1 | 32          | 0.939232 |
| plan 2 | 64          | 0.938166 |
| plan 3 | 128         | 0.940298 |

In order to make the output result more accurate, test the influence of different activation functions on the experimental results. It can be seen from Table 6 that the Acc values of ReLU and Tanh are the same, both are 0.940298, but the loss value of Tanh is less, so Tanh is chosen. The activation function in the TextCNN article is ReLU. When a relatively large gradient passes through the ReLU neuron, the neuron will be invalid, which will affect the experimental results. This article chooses Tanh as the activation function, which can solve this problem well.

Table 6. The result of the experiment without activating the function

| Plan  | Activation function | Acc    | Loss    |
|-------|---------------------|--------|---------|
| plan 1 | ReLu                | 0.940198 | 0.179256 |
| plan 2 | Sigmoid             | 0.934968 | 0.174531 |
| plan 3 | softmax             | 0.934968 | 0.167223 |
| plan 4 | Tanh                | 0.940298 | 0.170199 |

The pooling layer extracts key features from all the features after convolution, which mainly includes two pooling methods: max_pool and avg_pool. Table 7 shows the accuracy and loss value of each pooling method. It can be seen from the table that the loss value of avg_pool is smaller, indicating that the pooling method can obtain more accurate features. This is because TextCNN only uses max-pooling in the pooling layer, only considers the maximum feature value of each feature map, and does not consider other factors. Therefore, some important information may be lost. It can be seen from Table 7 that the pooling method of avg_pool has a better classification performance.
Table 7. Experimental results of different pooling methods

| pooling methods | Acc    | Loss     |
|-----------------|--------|----------|
| max_pool        | 0.934968 | 0.174531 |
| avg_pool        | 0.934968 | 0.169175 |

The choice of Adam algorithm and dropout rate will also affect the experimental results. Therefore, Table 8 shows the parameter comparison of the two methods. It can be seen from the table that when the parameters of the Adam algorithm are set to 1e-4 and the dropout rate is 0.5, the classification result can be optimal.

Table 8. Experimental results of Adam and drop rate

| (Adam parameter, Dropout rate) | Acc | Loss |
|-------------------------------|-----|------|
| (1e-3, 0.4)                   | 0.93| 0.19 |
| (1e-4, 0.5)                   | 0.94| 0.16 |
| (1e-5, 0.6)                   | 0.93| 0.20 |

Through the above experiments, the algorithm achieves good results under the settings of this experiment. It can be seen that the normal class are all correctly classified, followed by the classification of disease bone metastasis has a good effect. In order to understand correctly the confusion matrix, take Bone transfer as an example, 205 pieces of data were predicted correctly, but 2 pieces were predicted as arthritis, and 14 pieces were predicted as debatable. From this point of view, the error between arthritis and degenerative changes in the confusion matrix is relatively large. This is because, according to a medical analysis, degenerative changes are a special type of arthritis. For the actual diagnosis text, most Humans undergo degenerative changes in their bones. Arthritis mostly occurs in the joints of the human body, so the situation in Figure 3 will appear.

![Confusion matrix](image)

**Figure 3. Confusion matrix**

Table 9 gives the evaluation indicators of each disease, from which it can be seen that the classification effect of bone metastasis is relatively good, with F1 of 0.95.

Table 9. Evaluation index of each disease

| Evaluation metrics | Prec | Rec | F1   | Acc  |
|--------------------|------|-----|------|------|
| Normal             | 1.00 | 1.00| 1.00 | 0.94 |
| Bone               | 0.92 | 0.99| 0.95 |      |
| Arthritis          | 0.93 | 0.92| 0.92 |      |
| Degenerative       | 0.90 | 0.79| 0.84 |      |
In order to make it easier to observe the relationship between accuracy and loss value, Figure 4 shows a graph of the two changes with the number of iterations. In the first 10 iterations, the accuracy of the model increases greatly, and then gradually stabilizes. It can also be seen from the figure that the loss curve gradually stabilizes as the number of epochs increases, and drops to about 0.18.

![Figure 4. Curves of loss and accuracy](image)

4.3. Discussions
In general, the accuracy of the TextCNN algorithm is 0.94, and the loss value is 0.18, which can obtain high accuracy, provide help for medical diagnosis and prediction, and realize real human-computer interaction. Bone metastasis is the result of pathological changes in other cancers. For the actual diagnosis text, some of the cases are bone metastases occurring throughout the body. There are many locations of the disease, and the main data used in the experiment is also bone metastasis, so the evaluation indexes obtained are relatively high. Arthritis has actual reference conditions. The legs support human bones to ensure basic postures such as standing and walking. The legs are bound to be affected over time, and various evaluation indicators are relatively high.

5. Conclusions
Focusing on the classification of diseases in nuclear medical texts, in this work, we developed a nuclear medical text classifier based on the classic TextCNN model. First, the pre-processing process of converting nuclear medical text data into TextCNN model is proposed. Second, different strategies are introduced in the standard TextCNN model to construct the classifier by comparing different parameters. Last, a set of real nuclear medical text data is used to evaluate the proposed classifier. The experimental results showed that the average Acc index obtained was not less than 94%.

Funding:
This research was funded by Northwest Minzu University for Central University basic Scientific Research Operating Expenses Special Fund to Support the Graduate Program Yxm2020101; the National Natural Science Foundation of China, grant number 61562075; the Gansu Provincial First-class Discipline Program of Northwest Minzu University, grant number 11080305; and the Program for Innovative Research Team of SEAC, grant number [2018] 98.

References
[1] P. Anand, S. Chatterjee, I. Agarwal, Shantanu, A. Anupam, K. A. Jain, Intelligent Clinical Diagnosis System, 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Bengaluru, India, 2019: 1-5.
[2] K. Gai, M. Qiu, S. Jayaraman, L. Tao, Ontology-Based Knowledge Representation for Secure Self-Diagnosis in Patient-Centered Teleheath with Cloud Systems, 2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing, New York, NY, 2015: 98-103.
[3] N. Bourkache, M. Laghrrouch, S. Sidhom, Gabor Filter Algorithm for medical image processing: evolution in Big Data context, 2020 International Multi-Conference on: Organization of Knowledge and Advanced Technologies (OCTA), Tunis, Tunisia, 2020: 1-4.

[4] A. M. Najeeb, A. M. Osman, A. Al-Tijani, A. Abdelrahman, Implementation of computed radiography in nuclear medicine imaging, 2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), Khartoum, 2018: 1-4.

[5] M Yang, R Arsanjani, M C. Roarke1, Advanced Nuclear Medicine and Molecular Imaging in the Diagnosis of Cardiomyopathy. American Journal of Roentgenology. 2020, 215(5): 1208-1217.

[6] Y. Kim, Convolutional neural networks for sentence classification. 2014, arXiv: 1408.5882. [Online]. Available: https://arxiv.org/abs/1408.5882.