

Research Article

Electronic Medical Record Entity Recognition via Machine Reading Comprehension and Biaffine

Jun Cao , **Xian Zhou** , **Wangping Xiong** , **Ming Yang, Jianqiang Du, Yanyun Yang, and Tianci Li**

School of Computer, Jiangxi University of Traditional Chinese Medicine, Nanchang, China

Correspondence should be addressed to Wangping Xiong; 20030730@jxutcm.edu.cn

Received 6 July 2021; Revised 16 September 2021; Accepted 28 September 2021; Published 19 October 2021

Academic Editor: Waqas Haider Bangyal

Copyright © 2021 Jun Cao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The entity recognition of Chinese electronic medical record is of great significance to medical decision-making. The main process of entity recognition is sequence tagging, which has problems such as nested entity and boundary prediction. In this paper, we proposed a NER method called Bert-MRC-Biaffine, which formulates the NER as an MRC task. The approach of the machine reading comprehension framework is to introduce prior knowledge, the query about entities. The biaffine mechanism scores pair start and end tokens in a sentence so that the model is able to predict named entities accurately. The proposed method outperforms from the electronic medical record dataset, called CCKS2017 data, and the TCM dataset. We also remove components to evaluate the contribution of individual components of our model. Experiments on two datasets demonstrate the effectiveness of our model.

1. Introduction

With the widespread of EMRs, research and application of natural language processing in medical field has attracted much attention [1]. In the medical field, medical entities mainly include symptoms, diseases, drugs, treatments, and body parts. These are an important part of establishing a medical knowledge base. However, the data in the EMR cannot be used directly; we need to adopt the NER to extract information.

Early NER system has been traditionally approached by the hidden Markov model (HMM), maximum entropy model, and conditional random field (CRF) [2–4], which performs well on identifying entities in clinical EMRs. However, these approaches rely heavily on handcrafted features and task-specific resources. In addition, the support vector machine (SVM) [5] is well adapted to the field. What counts is such task-specific knowledge is costly to develop [6], making sequence labeling models difficult to adapt to new tasks or new domains.

In the past few years, neural network-based approaches have become popular in many NLP tasks. Recently,

recurrent neural networks (RNN) [7], together with its variants such as long short-term memory (LSTM) [8, 9] and gated recurrent unit (GRU) [10], have shown great success in modeling sequential EMRs. Based on this method, a large number of studies have focused on splicing additional features to improve the recognition effect of the model, such as splicing part of speech features [11] and adding attention mechanism [12]. Similarly, several CNN-based neural network models have been proposed to solve sequence labeling tasks in EMRs like [13, 14], achieving competitive performance against traditional models.

These systems have some limitations even that can be flexibly used in EMRs. There is space for the improvement in the annotation in a specific field, such as relying on a large-scale annotation data, as well as nested named entity recognition. The latter causes the incorrect prediction of entity boundary. For example, “心肺检查(cardiopulmonary examination)” has the aid of body parts to describe medical examination entities, and it is easy to label as body parts, resulting in a decrease in performance.

To handle the issues, we introduced the MRC into the span extraction based on the BERT [15] framework and

biaffine mechanism. The latter aims, such as some optimization algorithms [16, 17], to improve the ability of the model to classify. We first apply the BERT encoder to encode the sentence and questions. Then, we combine two representations and feed them into the transformer to model interactive information of each sentence. On top of the transformer, we apply two separate FFNNs to represent the start/end of the spans and finally to a biaffine classifier. We conduct extensive experiments on CCKS 2017 dataset and TCM dataset. Our model is able to improve performance when plugging with different NER models.

The main contributions of our work are as follows:

- (i) We propose a neural model based on MRC with biaffine mechanism to extract entities facts from sentences, where the entities could extract from real word data.
- (ii) Our model could consider the nested entity problem through MRC mechanism. We formalize the task as a question answering task: each entity type is characterized by a question answer template, and entities are extracted by answering template questions.
- (iii) We conduct analyses to understand why our approach performs so well and how different model parts affect final performance in other dataset.

2. Related Work

Most of previous studies on electronic medical record entity recognition primarily focus on English clinical texts. Various machine learning methods have obtained significant performance on English EMRs. Compared with English EMR, Chinese EMR faces more obstacles, which may due to the following reasons: fuzzy word boundary and complex composition forms in all kinds of tense and few open access Chinese EMR corpus and the size is usually small. Until recently, related work of Chinese EMR is developing rapidly, boosting the performance of the models [18].

By giving a sentence with annotated entities, tradition machine learning will focus on the construction and selection of different features, which directly affect the effect of NER. Liang et al. [19] integrated the sentence category classifier from SVM- and CRF-based Chinese EMR to recognize drug names. Their approach achieves F1 of 0.93 and 0.91 in traditional Chinese and Western medicine drug names, respectively. Zhang and Li [20] study the influence of multifeature combination such as part of speech feature, keyword feature, and dictionary feature on CRF sequence tagging. Lei et al. [5] investigated the effects of different types of machine learning models including CRF, SSVM, SVM, and ME in the admission records.

There are also relevant studies on Chinese EMRs' entity recognition using deep learning methods [21]. Yang et al. [22] investigated EMR label specification in English, developed a detailed EMR Chinese naming entity label specification, and completed the EMR Chinese foundation work. Wu et al. [23] applied a deep neural network on the NER task in Chinese electronic medical record with only word

embeddings, achieving the F1-score of 0.92. In addition, Huang et al. [24] and Wei et al. [25], both of them, proposed a bidirectional LSTM-CRF model for Chinese EMRs' entity recognition.

As we all know, pretraining and fine-tuning of entity recognition is the application trend. We proposed a NER method for Chinese EMRs based on pretraining and used as feature extractors; MRC introduces prior knowledge about entities and biaffine as the classifier. In addition, we also conducted ablation experiments on the TCM dataset to verify the effect of model components on the experimental results and the robustness of the overall model.

3. Methodology

3.1. Task Definition. The goal of NER is to classify each token or word in sentence and assign it to a corresponding label. Inspired by MRC [26–28] and the dependency parsing model [29], we formulate the task as an MRC task and thus convert the dataset style into a set, which contains context, query, and answer. Context is an input sentence, query is a query sentence predefined according to the dataset, and the answer is the gold entity span.

3.2. Model Details. We used the pretrained model as our model backbone. It employed the BERT encoder to present the input sentence and feed the output into two FFNNs and to a biaffine classifier. Figure 1 shows an overview of the model implemented in the MRC framework. Firstly, given an input sentence of n word token, the encoder module is responsible to map each token to feature information, which will feed into interaction module. BERT is a multilayer bidirectional transformer-based language representation model. For each word, it is used to learn deep representations, and it has been proven greatly effective in many downstream tasks. Specifically, it is composed of a stack of identical transformer blocks. We denote the block as T . The detailed operations of word representation are as follows:

$$\begin{aligned} h_0 &= M_1 W_1 + M_2 W_2 + W_p, \\ h_i &= T(h_{i-1}), \quad i \in [1, N], \end{aligned} \quad (1)$$

where $M_{1,2}$ is the matrix of one-hot vectors of indices in the input sentence, W_i is the words embedding matrix, W_s is the segment embedding matrix, W_p is the positional embedding matrix, where p represents the position index, h_i is the hidden state vector, the last representation of input sentence at i th layer, and N is the numbers of blocks.

The approach of the MRC framework is to introduce prior knowledge, the query about entities. Table 1 shows an overview of the architecture. The answers, obtained by the model based on questions, are the entities that should be labeled. We utilized the MRC task span extraction, which introduces prior knowledge in the span extraction and constructs a query set for all entity labels in the dataset. The span extracted by the model is the answer to the question and the entity that should be tagged. Normally, query is much shorter than the sentence text; the query input is

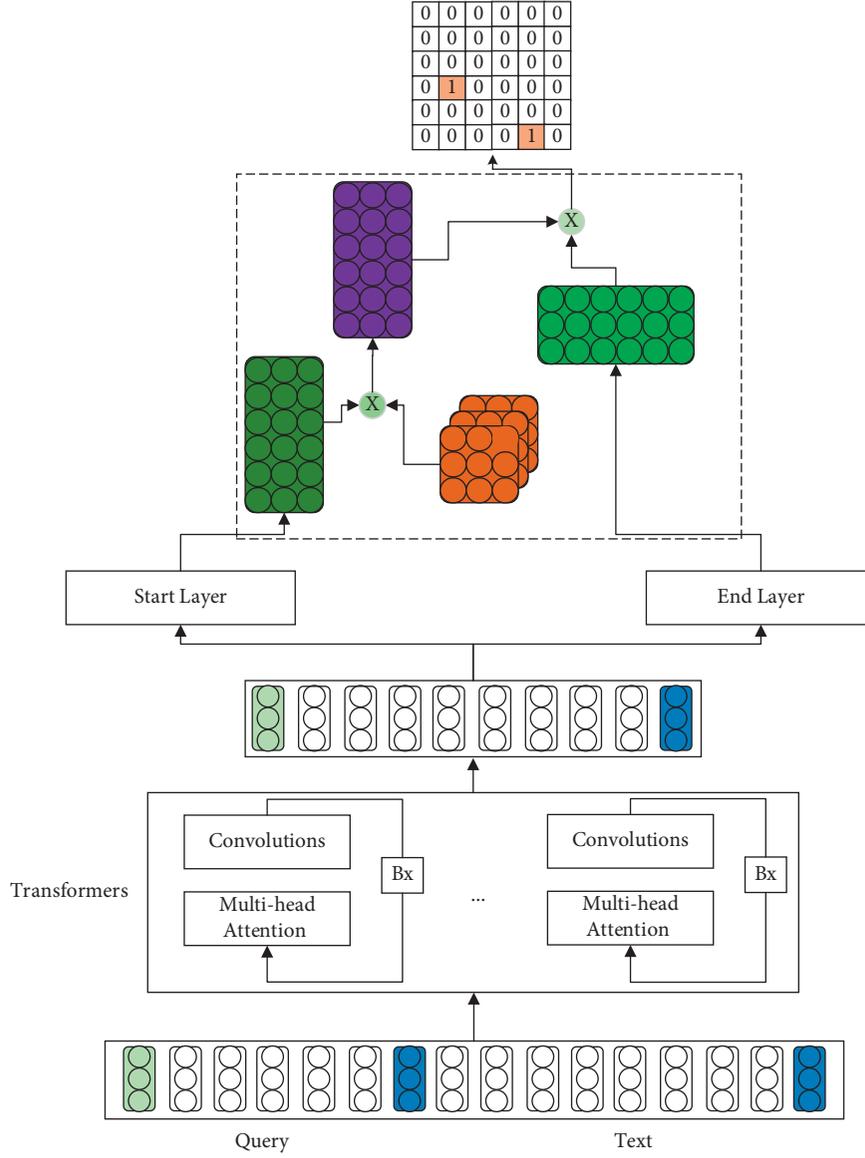


FIGURE 1: The network architectures of our model.

regarded as a whole to obtain problem encoding through the transformer. The detailed operations of question representation are as follows:

$$\begin{aligned}
 h_0^Q &= M_1^Q W_t^Q + M_2^Q W_s^Q + W_p^Q, \\
 h_i^Q &= T(h(i-1)^Q), \quad i \in [1, N], \\
 Q &= h_N^Q,
 \end{aligned} \tag{2}$$

where $M_{1,2}^Q$, W_t^Q , W_s^Q , W_p^Q , and h_i^Q are the same as introduction to the BERT encoder section and the difference represents the question text. Q is the finally obtained question information representation. For exploring the deep relationship between the text and the question, the attention mechanism is designed to extract relevant information. Note that, in our work the interaction may be performed multiple times. The operations of interaction are as follows:

$$\begin{aligned}
 S_i &= F(h_i, Q), \\
 w_i &= \exp S_i, \\
 O &= \sum w_i p_i,
 \end{aligned} \tag{3}$$

where F is the function for calculating the relevance of the sentence to the question, w_i is the relevance weight, and O is the hidden layer representation after interaction.

The strategy for choosing the span in the MRC framework is usually to construct two binary classifiers. One is used to predict whether the token is a start index, and the other is used to predict whether the token is an end index. In this study, we adopted the biaffine mechanism, which allows the token at the head and tail of the entity to interact with information, thus solving the problem of nested entity. After obtaining the hidden layer representation from the interaction module, we apply two separate FFNNs to create different representations for the start and end of spans. It

TABLE 1: Overview of the span extraction task via MRC.

MRC Span extraction	
Text sentence	The patient was in good health and accidentally fell five hours ago; then, the patient felt pain and went to the hospital; the X-ray showed that the left intertrochanter of patient is fractured; the fluid infusion was provided to patient for symptomatic treatment
Corpus	What is the body part in the EMR?
Query	
Answer/label	Left intertrochanter
entity	

allows the model to learn to identify the start or end of the spans separately. Finally, we employ a biaffine mechanism over the sentence to create a tensor. We apply the biaffine operator:

$$\begin{aligned}
 h_i^s &= FFNNs(O_i), \\
 h_i^e &= FFNNe(O_i), \\
 B(s, e) &= h_i^s \wedge T U h_i^e + W(h_i^s \oplus h_i^e) + b,
 \end{aligned} \tag{4}$$

where U is a $d \times c \times d$ tensor, W is a $2d \times c$ matrix and a bias vector b , respectively, and d is the size of output layers of FFNNs, while c is the number of entity classes. Finally, we feed the output vectors $B(s, e)$ of the biaffine layer into softmax for the entity scores and assign each span a NER category:

$$\begin{aligned}
 y_i &= \arg \max B(s, e), \\
 \text{score}(i_c) &= \log \sum \exp B(s, e).
 \end{aligned} \tag{5}$$

In context X , there could be multiple entities of the entity span. This means that multiple entity label indexes can be predicted from one input sequence. At the training time, we optimize the case of named entity recognition with the following loss function:

$$L = \frac{1}{N} \sum \log P(y_i | \text{score}(i_c)). \tag{6}$$

4. Experiments

In this section, we conduct several experiments to show effectiveness of our framework. We evaluate the performance of our model on clinical NER datasets, by plugging it into three base NER models.

4.1. Dataset. We conducted experiments on the dataset of CCKS2017 EMR evaluation. There are mainly 400 EMRs, including five types of entities: body parts (BODY), disease symptoms (SIGNS), medical checks (CHECK), diseases (DISEASE), and treatments (TREATMENT). We follow the standard train/dev/test splits and use both the train set and dev set for training, with more detailed statistics of the dataset listed in Table 2.

Each entity type is associated with a type-specific question generated by the templates. We utilized the annotation guideline as references to construct queries, thus conducted experiments on five types of medical entities. Table 3 lists the items of queries we constructed.

The original corpus text is split into a series of Chinese characters, each of which is tagged with a BIO format. B-tag indicates the begin of the entity, I-tag indicates inside the entity, and O-tag indicates that the character is outside the entity. To meet the input requirements of the model, the data should be preprocessed as {context, query, answer} format. When predicting the boundary of the entity, we should predict the entity type. According to the dataset, there are a total of 11 labels to be predicted. We report Precision, Recall, and F1 scores for all evaluations. The named entity is considered correct when both boundary and category are predicted correctly.

4.2. Settings and Results. We compare the performance with three baseline systems: BiLSTM-CRF, the combination of BiLSTM with CRF, BERT, the language representation model for task fine tuning, and BERT-BiLSTM-CRF, the BiLSTM-CRF substructure is connected after the BERT layer. The results are presented in Table 4.

We use a unified setting for all of experiments. Table 5 shows hyperparameters for our experiments. Table 4 presents the result for the entity identification task (Precision, Recall, and F1). Finally, in the last column of Table 4, we reported an additional F1 measurement, which is the average F1 performance of the five entities.

As shown in Table 4 and Figure 2(a), body part, disease symptoms, check, and disease entity labeling achieved good performances in all methods and that of treatment achieved poor performance. The number and average length of three types of entities, including body, check, and symptom, are far more than those of the other two types of entities, including check and disease, respectively. The performances are because the model prefers to overfitting when a small scale of data is trained. In addition, our system achieved an F1 score of 92.8% on the corpus. It outperforms the results of the comparison by large margin.

As shown in Table 4 and Figure 2(b), the performance of the proposed method is slightly better than against models. Our model introduces prior knowledge to enhance language representation, which breaks the limitation that the knowledge required for span extraction is limited to context and acquires more domain-related knowledge for inference prediction. It explains from the theoretical level that the F1 of the model can reach a good performance of 92.8% even on a small-scale corpus, and the F1 of various entities is greater than that of conventional model and the BERT fine-tuning method. This indicates that the Bert-MRC-Biaffine method proposed in this paper is effective.

TABLE 2: The statistical results of different entities in the dataset.

Type	Train	Validate	Test
Disease symptoms	6484	730	614
Body parts	8492	922	855
Medical checks	7987	836	723
Diseases	853	119	86
Treatments	515	109	88

TABLE 3: Set of queries about different entities.

Query	Entity
What is the treatment method for diseases	TREATMEN
What is body part in EMR	BODY
What is the symptom of the disease	SIGN
What is the medical examination item in EMR	CHECK
What is the disease entity in EMR	DISEASE

TABLE 4: The result for the entity recognition task (Precision, Recall, and F1).

Methods	Evaluation	Entity categories					OVERALL
		BODY	CHECK	DISEASE	SIGNS	TREATMENT	
BiLSTM-CRF	P	85.87	94.51	85.51	90.03	71.91	89.15
	R	83.16	95.3	67.05	98.53	74.42	89.94
	F1	84.49	94.9	75.16	94.09	73.14	89.54
BERT-BiLSTM-CRF	P	87.82	89.50	91.11	89.40	63.16	86.43
	R	85.16	93.83	73.21	97.92	69.23	88.32
	F1	86.47	91.61	81.19	93.47	66.06	87.36
BERT-Softmax	P	86.38	87.92	88.70	85.43	62.48	84.47
	R	83.68	86.39	72.09	93.71	67.13	86.04
	F1	85.01	87.15	79.54	89.38	64.72	85.25
BERT-MRC-Biaffine	P	91.68	96.84	94.20	95.33	75.93	91.93
	R	89.06	93.91	81.32	96.40	86.82	93.68
	F1	90.35	95.35	87.29	95.86	81.01	92.80

As shown in Table 4 and Figure 2(a), the F1 of body parts in our model is a few percentage points on an average than that in other methods. Although the number of body parts' entities is the largest, the prediction performances are poor in other methods. The reason is that there are more short words in this dataset, and the model tagging recognition is more inclined to short words. When long words are labeled, only part of words will be taken as the labeling result. This is easy to cause boundary errors. For example, other entities should be described with the help of body parts, such as cardiopulmonary examination, reflex neck test, Achilles knee tendon reflex, and knee effusion. Non-MRC methods will identify these nested entities as body parts. The experimental results show that MRC, which obtains the labeling entity in the form of answering questions, can solve the problem of entity nesting without considering the influence of the number of long and short entity words on model labeling.

To verify the robustness of the model, we conducted the same model comparison experiments on the TCM dataset, which is a small dataset consisting of 5 categories: formula, syndrome, herb, symptom, and disease.

The experimental results in Table 6 show that our proposed model achieves the highest F1 value of 83.30, indicating that it can be effectively applied to other datasets with good robustness and generality.

In addition, to evaluate the contribution of individual components of our model, we remove part components and use TCM dataset for evaluation.

We replace the biaffine classifier with a softmax layer that is frequently used in models for produce the outputs. When we replace the biaffine of our model with a softmax layer, the performance drops by 0.84 percentage points, as shown in Table 7. The performance difference shows the support of a biaffine module and confirms it is an important factor for the outperformance.

We ablate MRC module and as expected, the model performance drops by a number of 3.7 percentage point in Table 7. It shows that prior knowledge introduced by MRC is one of the most important factors for the outperformance.

We remove both MRC and biaffine, and the model results show a decrease of 4.32 percentage points in Table 7. It proves the validity of the two structures for our proposed model.

TABLE 5: The statistical results of different entities in the dataset.

Parameter	Value
Epoch	100
Learning rate	0.001
Dropout	0.5
Clip	5
Batch_size	20
Bert size	1024
FFNN size	150
Optimizer	Adam

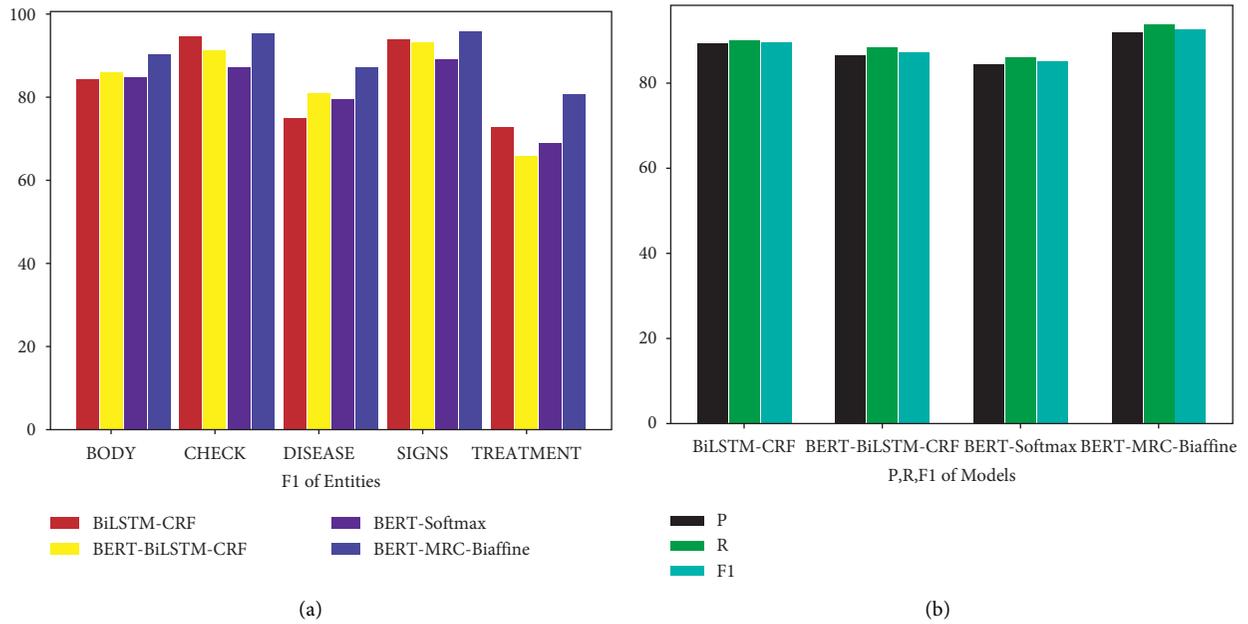


FIGURE 2: The performance comparison for entities and models. (a) F1 of entities. (b) P, R, and F1 of models.

TABLE 6: The result for the TCM dataset.

Methods	P	R	F1
BiLSTM-CRF	82.11	77.23	79.59
BERT-BiLSTM-CRF	81.67	74.24	77.78
BERT-Softmax	83.74	77.70	78.98
BERT-MRC-Biaffine	87.02	80.68	83.30

TABLE 7: The performance between full model and removed component models in TCM dataset.

Methods	P	R	F1	
BERT-MRC-Biaffine	87.02	80.68	83.3	—
BERT-MRC(-Biaffine)	86.40	78.86	82.46	+0.84
BERT-Biaffine(-MRC)	83.95	78.10	79.6	+3.7
BERT(-MRC-Biaffine)	83.74	77.70	78.98	+4.32

5. Conclusions

In this paper, we presented a neural architecture for medical entity recognition. MRC introduces entity-related prior knowledge into the named entity recognition model, which can achieve better performance in the context of small-scale corpus. The biaffine mechanism allows the tokens at the head and the end of the entity sequence to exchange information. The experimental results have been improved to a certain extent. It is an effective decoding method for the NER model.

There are several potential directions for future work. We aim to improve model performance by exploring multitask learning approaches to combine more information such as POS jointly trains with NER task to enhance the representations learn in our model. In addition, we will propose several different ways as query to verify whether it will greatly affect the performance of the model.

Abbreviations

EMR:	Electronic medical record
NER:	Named entity recognition
MRC:	Machine reading comprehension
TCM:	Traditional Chinese medicine
CCKS2017:	China Conference on Knowledge Graph and Semantic Computing in 2017
POS:	Part of Speech.

Data Availability

The datasets used in this study are adopted from the Chinese EMR named entity recognition task in China Conference on Knowledge Graph and Semantic Computing in 2017 (<http://www.ccks2017.cn/>), but restrictions apply to the availability of these data, which were used under license for the current study, so are not publicly available. The TCM dataset used and analysed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (nos. 82160955, 61762051, and 82060826), the Jiangxi Natural Science Foundation (nos. 20192BAB205094 and 20202BAB202019), the National Key R&D Plan (no. 2019YFC1712301), and the Science and technology research project of Jiangxi Provincial Department of Education (no. GJJ190683).

References

- [1] B. Ji, S. Li, J. Yu et al., "Research on Chinese medical named entity recognition based on collaborative cooperation of multiple neural network models," *Journal of Biomedical Informatics*, vol. 104, Article ID 103395, 2020.
- [2] Y. Li, S. L. Gorman, and N. Elhadad, "Section classification in clinical notes using supervised Hidden Markov Model," in *Proceedings of the 1st ACM International Health Informatics Symposium*, November 2010.
- [3] W. Li, D. Zhao, B. Li, X. Peng, and J. Liu, "Combining CRF and rule based medical named entity recognition," *Application Research of Computers*, vol. 32, no. 4, pp. 1082–1086, 2015.
- [4] Ö. Uzuner, B. R. South, S. Shen, and S. L. DuVall, "2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text," *Journal of the American Medical Informatics Association JAMIA*, vol. 18, no. 5, pp. 552–556, 2011.
- [5] J. Lei, B. Tang, X. Lu, K. Gao, M. Jiang, and H. Xu, "A comprehensive study of named entity recognition in Chinese clinical text," *Journal of the American Medical Informatics Association: JAMIA*, vol. 21, no. 5, pp. 808–814, 2014.
- [6] X. Ma and F. Xia, "Unsupervised dependency parsing with transferring distribution via parallel guidance and entropy regularization," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*, Baltimore, USA, June 2014.
- [7] Y. Xia, J. Zheng, and Y. Zhao, "Named entity recognition of electronic medical," *Records Based on Deep Learningelectronic Technology*, vol. 31, no. 11, 2018.
- [8] L. Wan, Y. Luo, and Z. Li, "The recognition of naming entity of Bi-LSTM Chinese electronic medical records based on the joint training of Chinese characters and words," *China Digital Medicine*, vol. 14, no. 2, pp. 54–56, 2019.
- [9] A. N. Jagannatha and H. Yu, "Structured prediction models for RNN based sequence labeling in clinical text," *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, vol. 2016, pp. 856–865, 2016.
- [10] K. Cho, B. Van Merriënboer, and C. Gulcehre, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*, Doha, Qatar, October 2014.
- [11] F. Yang, Y. Zhang, L. Qin, and J. Li, "A research progress of clinical named entity recognition from Chinese electronic medical records," *Chinese Digital Medicine*, vol. 15, no. 2, pp. 9–12, 2020.
- [12] Q. Zhang, X. Liu, and J. Fu, "Neural networks incorporating dictionaries for Chinese word segmentation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.
- [13] M. Yin, C. Mou, K. Xiong, and J. Ren, "Chinese clinical named entity recognition with radical-level feature and self-attention mechanism," *Journal of Biomedical Informatics*, vol. 98, Article ID 103289, 2019.
- [14] M. Gao, Q. Xiao, S. Wu, and K. Deng, "An Attention-based ID-CNNs-CRF model for named entity recognition on clinical electronic medical records," in *Proceedings of the Artificial Neural Networks and Machine Learning—ICANN 2019: Workshop and Special Sessions*, Munich, Germany, September 2019.
- [15] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.
- [16] W. H. Bangyal, J. Ahmad, H. T. Rauf, and R. Shakir, "Evolving artificial neural networks using opposition based particle swarm optimization neural network for data classification," in *Proceedings of the 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, November 2018.

- [17] W. H. Bangyal, J. Ahmad, and H. T. Rauf, "Optimization of neural network using improved bat algorithm for data classification," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 4, pp. 670–681, 2019.
- [18] Y. Zhang, X. Wang, Z. Hou, and J. Li, "Clinical named entity recognition from chinese electronic health records via machine learning methods," *Preprint*, vol. 6, no. 4, p. e50, 2018.
- [19] J. Liang, X. Xian, X. He et al., "A novel approach towards medical entity recognition in Chinese clinical text," *Journal of Healthcare Engineering*, vol. 2017, Article ID 4898963, 16 pages, 2017.
- [20] X. Zhang and Z. Li, "Chinese electronic medical record named entity recognition based on multi-feature fusion," *Software Guide*, vol. 16, no. 2, pp. 128–131, 2017.
- [21] L. Gong, Z. Zhang, and S. Chen, "Clinical named entity recognition from Chinese electronic medical records based on deep learning pretraining," *Journal of Healthcare Engineering*, vol. 2020, no. 2, 8 pages, Article ID 8829219, 2020.
- [22] J. Yang, Y. Guan, and B. He, "Chinese electronic medical record named entity and entity relationship corpus construction," *Journal of Software*, vol. 27, no. 11, pp. 2725–2746, 2016.
- [23] Y. Wu, J. Min, and J. Lei, "Named entity recognition in Chinese clinical text using deep neural network," *Studies in Health Technology and Informatics*, vol. 216, pp. 624–628, 2015.
- [24] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," arXiv preprint arXiv:1508.01991, 2015.
- [25] Q. Wei, T. Chen, R. Xu, Y. He, and L. Gui, "Disease named entity recognition by combining conditional random fields and bidirectional recurrent neural networks," *Database The Journal of Biological Databases and Curation*, vol. 2016, 2016.
- [26] S. Liu, X. Zhang, S. Zhang, H. Wang, and W. Zhang, "Neural machine reading comprehension: methods and trends," *Applied Sciences*, vol. 9, no. 18, p. 3698, 2019.
- [27] S. Wang and J. Jiang, "Machine comprehension using match-Lstm and answer pointer," arXiv preprint arXiv:1608.07905, 2016.
- [28] X. Li, J. Feng, Y. Meng, Q. Han, F. Wu, and J. Li, *A Unified MRC Framework for Named Entity Recognition*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2020.
- [29] J. Yu, B. Bohnet, and M. Poesio, "Named entity recognition as dependency parsing," in *Proceedings of the 58th Annual Conference of the Association for Computational Linguistics*, ACL, July 2020.