



*Research article*

# Entity relationship extraction from Chinese electronic medical records based on feature augmentation and cascade binary tagging framework

Xiaoqing Lu<sup>1</sup>, Jijun Tong<sup>2,\*</sup> and Shudong Xia<sup>3</sup>

<sup>1</sup> School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China

<sup>2</sup> School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou, China

<sup>3</sup> The Fourth Affiliated Hospital Zhejiang University School of Medicine, Jinhua, China

\* **Correspondence:** Email: [jijuntong@zstu.edu.cn](mailto:jijuntong@zstu.edu.cn).

**Abstract:** Extracting entity relations from unstructured Chinese electronic medical records is an important task in medical information extraction. However, Chinese electronic medical records mostly have document-level volumes, and existing models are either unable to handle long text sequences or exhibit poor performance. This paper proposes a neural network based on feature augmentation and cascade binary tagging framework. First, we utilize a pre-trained model to tokenize the original text and obtain word embedding vectors. Second, the word vectors are fed into the feature augmentation network and fused with the original features and position features. Finally, the cascade binary tagging decoder generates the results. In the current work, we built a Chinese document-level electronic medical record dataset named VSCMeD, which contains 595 real electronic medical records from vascular surgery patients. The experimental results show that the model achieves a precision of 87.82% and recall of 88.47%. It is also verified on another Chinese medical dataset CMeIE-V2 that the model achieves a precision of 54.51% and recall of 48.63%.

**Keywords:** deep learning; neural networks; natural language processing; Chinese electronic medical records; biomedical information processing

---

## 1. Introduction

With the progress of information technology and the internet, the utilization of electronic medical records (EMRs) has skyrocketed worldwide. In the United States, this percentage rose drastically from

10% to almost 96% within a decade (2008 to 2017). Similarly, in China, the increase reached 85% [1]. Electronic medical records, serving as a crucial source of clinical data, encompass rich medical information, including patient's chief complaints, medical history, examination items and results, medication details, treatment plans, and more. Extracting such valuable information from unstructured text to construct triplet data constitutes a pivotal prerequisite for subsequent related research tasks, such as disease risk prediction [2,3], clinical knowledge bases [4,5], and assisted diagnosis [6,7]. In the past, the extraction of entity relations from electronic medical records was a manually intensive task that was not only time-consuming but also relied highly on physician's experience, leading to potential omissions and errors. Therefore, efficient and accurate methods for electronic medical record entity relationship extraction are of immense importance.

Until now, the extraction of entity relations from Chinese electronic medical records has continued to pose significant challenges. First, there is a scarcity of high-quality medical datasets in the Chinese domain compared to English EMRs, posing a limitation for comprehensive research. Second, Chinese semantic expression exhibits exceptional complexity. On one hand, with the same entity, Chinese can potentially convey vastly different meanings in varying contexts. On the other hand, Chinese has few strict entity boundaries, such as spaces between words, which complicates accurate extraction. Third, medical text tends to be longer and more information-dense than ordinary text, leading to an increased difficulty in extracting triplets as their quantity within the text increases.

To address the aforementioned challenges, we constructed a modest-scale document-level dataset, named VSCMeD, utilizing real Chinese electronic medical records. Furthermore, we introduced a network model based on a feature augmentation and cascade binary tagging framework (FA-CBT). This model divides the task into two parts: subject decoding and relation-object joint decoding. Initially, to enhance the encoding of lengthy electronic medical records, we incorporated a pre-trained model that is specifically tailored for Chinese natural language processing tasks. The subject classifier then predicts all potential subjects within the given sequence. Subsequently, our model combines subject features, sequential features, and semantic features through a dedicated feature augmentation layer. This comprehensive, fused information is then processed by the relation-specialized object decoder, ultimately yielding the predictive outcome. This approach aims to enhance the precision and efficiency of entity relation extraction from complex Chinese electronic medical records.

Our main contributions can be summarized as follows:

1) We built a real Chinese document-level electronic medical record dataset named VSCMeD, including genuine EMRs from over 500 vascular surgery patients. The maximum length of the data entries exceeded 1,000 characters.

2) We proposed a network model based on feature augmentation and cascade binary tagging framework (FA-CBT), for feature deterioration in lengthy texts and entity overlapping. Experimental results on VSCMeD demonstrated that the highest achieved precision is 87.82%, with a recall rate of 88.47%.

## 2. Related works

Relation extraction has become one of the most important tasks in natural language processing due to its significant implications. In recent years, a series of studies on relation extraction have been published.

The relation extraction task is often decomposed into a pipeline of two independent tasks: entity recognition (NER), which extracts all entities from the text, and relationship classification (RC), which

predicts whether a relation exists between the extracted entities [8–13]. Most studies have used traditional methods like conditional random fields (CRF) [14], recurrent neural networks (RNN), and RNN-based improvements like long short-term memory (LSTM) [15]. Wei et al. [16] pioneered a triplet extraction approach based on a cascade binary tagging framework. Although the method is still a two-step pipeline, it treats relation extraction as a mapping from subject to object and relation and can handle entity overlapping issues. This paper also adopts this tagging and decoding framework.

The two models in the pipeline approach are independent of each other, ignoring potential correlations between the tasks, and can cause error propagation issues. Therefore, recent researchers have started to view relation extraction as a complete task, known as the joint approach, which directly extracts entities and relations between them from text without determining entity types. At the same time, pre-trained BERT [17]-derived models have also been widely adopted, and many pre-training models targeting the medical domain have emerged, such as PubMedBERT [18], ClinicalBERT [19], and MedBERT [20], which have greatly facilitated the development of medical natural language processing. Wang et al. [21] designed an entity-aware self-attention layer based on BERT, showing great performance on standard benchmarks. Shang et al. [22] viewed relation extraction as a table-filling task, building a matrix for each relation and using four types of tags to extract subjects and objects under specific relations, which can even handle nested entities. Lu et al. [23] proposed a unified framework for general information extraction, enabling unified modeling of tasks including entity extraction, relation extraction, event extraction, sentiment analysis, etc., with good transferability and generalization between different tasks. However, the pipeline approach is not necessarily inferior to the joint approach. Zhong and Chen [24] proposed a very simple span-based pipelined approach for entity and relation extraction and established new state-of-the-art results on standard benchmarks.

Specifically in the medical information field, relevant foreign research started earlier, and there are high-quality datasets like i2b2 [25], CDR [26], MIMIC [27,28], etc. Li et al. [29] proposed a novel edge-oriented graph neural network based on document structure and external knowledge called SKEoG, which achieved an F1 score of 70.7 on the CDR dataset. Chen et al. [30] combined BERT with a one-dimensional convolutional neural network (1d-CNN) to fine-tune the pre-trained model for relation extraction, achieving F1 scores of 70.85 on the i2b2 dataset and 71.56 on the CDR dataset. Sun et al. [31] constructed document-level syntactic dependency graphs and encoded them using a graph convolution network (GCN), and then extracted relations by using weighted context features and long-range dependency features, achieving an F1 score of 73.3 on the CDR dataset.

Due to the lack of Chinese medical datasets, domestic related research is relatively scarce. Medical institutions often use internal data for research, which is very difficult to obtain. The China Health Information Processing Conference (CHIP) initiated the Chinese biomedical language understanding evaluation (CBLUE [32]) and released the Chinese medical named entity recognition dataset (CMeEE) and the Chinese medical entity relation extraction dataset (CMeIE), attracting more researchers to Chinese information extraction. Chang et al. [33] constructed a Chinese dataset consisting of stroke and diabetes electronic medical records and validated their model's performance in extracting overlapping triplets on this dataset and CMeIE. Pang et al. [34] constructed a specific relation attention-guided graph neural networks (SRAGNNs) model to jointly extract entities and their relations in Chinese EMRs. They demonstrated improved performance for Chinese medical entity and relation extraction on two Chinese datasets. Zhang et al. [35] utilized GCN to capture inter-sentence dependency and studied the impact of different pruning operations on performance. Ye et al. [36] proposed a method of subsequence and distant supervision based active learning which was annotated

by selecting information-rich subsequences as a sampling unit. They achieved the best performance on the CMeIE dataset.

### 3. Method

In this section, we describe our relation extraction approach, which involves extracting triplets— $\langle \text{subject, relation, object} \rangle$ —from electronic medical records based on predefined relationships. We adopt the cascade binary tagging framework. In prior studies, a single subject and its corresponding triplet were randomly selected from a given data entry as a training sample. In our approach, we include all triplets from each data entry, which are distributed together for training, thereby reducing the number of epochs and shortening the overall training time. Moreover, we performed pre-decoding feature processing and augmenting, elevating the performance of the approach. The method can be summarized into two steps: extracting subjects, then extracting relations and objects based on the subjects. It can be denoted as:

$$f_{relation}(\text{subject}) \rightarrow \text{object} \tag{1}$$

Taking the sentence from Figure 1 as an example, “With a medical history of hypertension for five years, the individual self-administers one Compound Hypotensive Tablet per day.” The model initially predicts the potential subject as “hypertension” and subsequently forecasts possible objects under predefined relationships. This process results in the formation of the triplet  $\langle \text{hypertension, pharmacological treatment for the disease, compound hypotensive tablet} \rangle$ .

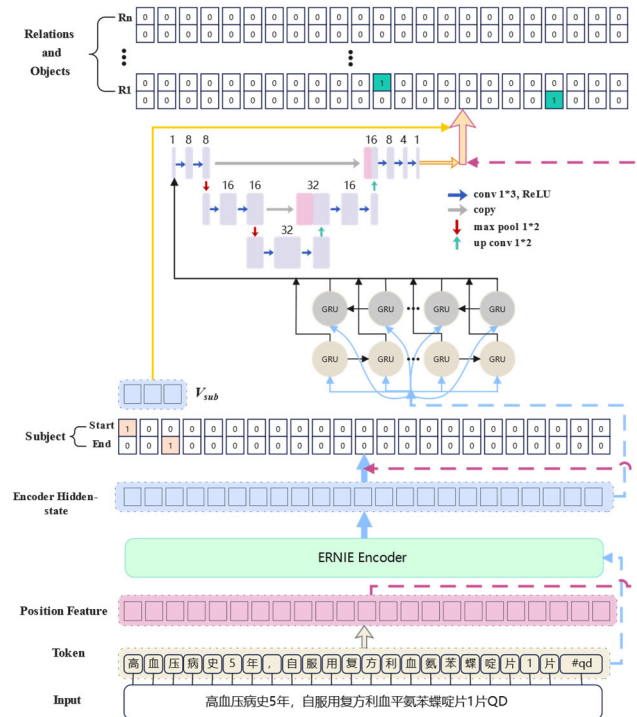


Figure 1. The overall model architecture of FA-CBT.

The overall model architecture is shown in Figure 1, and the process is described in detail in the following sections.

### 3.1. Encoder

The encoder is used to map the input tokens into a low-dimensional vector space, which is the foundation of natural language processing (NLP). In the past, we often used Word2Vec [37] to implement this step, but the resulting word vectors were static and performed poorly on specific tasks. The emergence of pretrained models, especially BERT [17], has brought milestone changes to the NLP field and been rapidly and widely applied to various NLP tasks. However, BERT and its derived pretrained models can encode at most 512 tokens, while real Chinese EMRs often exceed this length. Due to the uniqueness of the text, simple splitting cannot be applied. ERNIE [38] models the lexical, syntactic, and semantic information in the text. Similar to BERT, it employs a multi-layered transformer as its fundamental encoder. However, it enhances pre-trained language models with prior knowledge, avoiding the direct addition of knowledge embeddings. Moreover, it introduces a multi-stage knowledge masking strategy, integrating phrase-level and entity-level knowledge into language representations, and supporting up to 2048 input tokens. And ERNIE 3.0 outperformed the state-of-the-art models on 54 Chinese NLP tasks. Therefore, we choose ernie-3.0-base-zh as the encoder layer, as shown in Eq (2):

$$H = ERNIE(t_1, t_2, \dots, t_n) = (h_1, h_2, \dots, h_n) \quad (2)$$

where  $(t_1, t_2, \dots, t_n)$  represents the input token sequence and  $(h_1, h_2, \dots, h_n)$  represents the feature vector of the corresponding token.

### 3.2. Feature augmentation layer

Gate recurrent unit (GRU) [39] is a variant of LSTM with fewer parameters and a simpler structure, enabling greater efficiency. It can effectively capture long-distance dependencies, making it well-suited for relation extraction from electronic medical records. The inputs to the GRU unit comprise the current time step input vector  $x_t$  and previous time step hidden state  $h_{t-1}$ . The implementations of a GRU unit are as follows:

$$z_t = \text{sigmoid}(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \quad (3)$$

$$r_t = \text{sigmoid}(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \quad (4)$$

$$\tilde{h}_t = \tanh(W_{xh}x_t + W_{hh}(r_t \odot h_{t-1}) + b_h) \quad (5)$$

$$h_t = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{t-1} \quad (6)$$

where  $W_{xz}$ ,  $W_{xr}$ , and  $W_{xh}$  are weight matrices for input vector  $x_t$ ;  $W_{hz}$ ,  $W_{hr}$ , and  $W_{hh}$  are weight matrices for hidden state  $h_{t-1}$ ;  $b_z$ ,  $b_r$ , and  $b_h$  are bias vectors; and  $z$  and  $r$  denote the update gate and reset gate respectively. The output vectors from forward and backward GRU units are concatenated to obtain feature vectors with bidirectional semantics:

$$h_t = \vec{h}_t \oplus \overleftarrow{h}_t \quad (7)$$

UNet [40] has demonstrated astounding performance in medical imaging. Inspired by it, we designed a model for medical text feature augmentation. We perform convolution with 1\*3 kernels and

stride 1 to capture inter-semantic features in sequences and augment long-distance characteristics. The entire model has 13 convolutional layers, with padding to keep input and output feature vector dimensions identical.

### 3.3. Decoder

We utilize two binary classifiers to identify the start and end positions of the subject respectively. For each token, the corresponding positions are two binary classification problems, described as follows:

$$p_i^{start_s} = \text{sigmoid}(W_{start}(x_i + p_i) + b_{start}) \quad (8)$$

$$p_i^{end_s} = \text{sigmoid}(W_{end}(x_i + p_i) + b_{end}) \quad (9)$$

where  $p_i^{start_s}$  and  $p_i^{end_s}$  denote the probabilities of the  $i$ -th token in the input sentence being the start and end positions of the subject respectively,  $W_{start}, W_{end}$  are weight matrices,  $p_i$  is the position feature, and  $b_{start}$  and  $b_{end}$  are biases. If the probability exceeds the set threshold, the token is tagged as 1, otherwise 0. In this paper, the threshold is 0.5. Subject  $s$  can be denoted as:

$$p_\theta(s|x) = \prod_{t \in \{start_s, end_s\}} \prod_{i=1}^L (p_i^t)^{I\{y_i^t=1\}} (1 - p_i^t)^{I\{y_i^t=0\}} \quad (10)$$

where  $L$  denotes sequence length;  $I\{x\}$  indicates 1 if  $x$  is true, otherwise 0;  $start_s, end_s$  represent binary tags for the start and end positions of the subject; and  $\theta$  denotes model parameters. As multiple  $start_s$  and  $end_s$  of subjects will inevitably be detected, we follow the nearest match principle, matching the most recent  $start_s$  and  $end_s$ , marking tokens within the interval as one entity.

The relation-specialized object decoder builds upon previously extracted subjects, performing object extraction given the subject for each relation separately.

$$x'_i = CNN(\overline{GRU(x_i)} \oplus \overline{GRU(i)}) \quad (11)$$

$$p_i^{start_o} = \text{sigmoid}(W_{start}^r(x'_i + v_{sub} + p_i) + b_{start}^r) \quad (12)$$

$$p_i^{end_o} = \text{sigmoid}(W_{end}^r(x'_i + v_{sub} + p_i) + b_{end}^r) \quad (13)$$

where  $GRU(x)$  denotes feeding  $x$  through the GRU layer,  $CNN(x)$  denotes passing  $x$  through the convolutional layers,  $p_i^{start_o}, p_i^{end_o}$  represent the probabilities of the  $i$ -th token being the start and end positions of the object, and  $v_{sub}$  denotes subject feature information. Similarly, object  $o$  can be denoted as:

$$p_{\Phi_r}(o|s, x) = \prod_{t \in \{start_o, end_o\}} \prod_{i=1}^L (p_i^t)^{I\{y_i^t=1\}} (1 - p_i^t)^{I\{y_i^t=0\}} \quad (14)$$

where  $start_o, end_o$  denote binary tags for the start and end positions of the object, and  $\Phi_r$  represents relation-specialized object decoder parameters.

## 4. Experiment

### 4.1. Dataset

#### 4.1.1. VSCMeD

The VSCMeD dataset contains 595 authentic electronic medical records from a Chinese top tertiary hospital's vascular surgery department. The maximum sequence length is 1207, with an average of 581. Each EMR encompasses patients' details, chief complaints, present illness history, past medical history, physical examination findings, and preliminary diagnosis. We defined 8 types of relations and had a professional physician annotate them using the open-source tool doccano. The annotated triplet structure is represented as  $\langle (\text{sub\_s}, \text{sub\_e}), \text{rel\_idx}, (\text{obj\_s}, \text{obj\_e}) \rangle$ , where  $\text{sub\_s}$ ,  $\text{sub\_e}$ ,  $\text{obj\_s}$ , and  $\text{obj\_e}$  represent the start and end positions of the subject and object in the original text, and  $\text{rel\_idx}$  indicates the relationship type. Tables 1 and 2 show the statistics. The dataset is split into training and test sets with the ratio of 8:2 for experiments.

**Table 1.** Statistics of the relation type.

Relation	Meaning	Count
DiSy	symptoms induced by the disease	758
OpDi	surgical treatment for the disease	884
OpDa	date of the surgery	1543
DrDi	pharmacological treatment for the disease	650
ExDa	date of the examination	419
ExDi	examination confirmed the disease	186
ExIt	items included in the examination	2086
ItVa	results of the examination items	2306
Total		8832

**Table 2.** Statistics of triplet quantities in VSCMeD.

Number of Triples in One Sentence	Count
$0 \leq N < 5$	57
$5 \leq N < 10$	98
$10 \leq N < 15$	127
$15 \leq N < 20$	160
$20 \leq N < 25$	94
$25 \leq N$	59
Total	595

#### 4.1.2. CMeIE-V2

CMeIE-V2<sup>1</sup> is the latest version of CMeIE, with increased task difficulty. CMeIE-V2 has a total

<sup>1</sup> <https://tianchi.aliyun.com/dataset/95414>

of 44 relation types, with maximum sequence length of 300. The training, validation, and test sets contain 14339, 3585 and 4482 instances respectively. As the test set remains unannotated, one must tender model predictions to the CBLUE 32 competition to obtain scores. Table 3 summarizes the distribution of relation quantities per instance in the dataset.

**Table 3.** Statistics of triplet quantities in CMeIE-V2.

Number of Triples in One Sentence	Train	Dev	Test
$0 \leq N < 2$	1788	417	UNK
$2 \leq N < 4$	4981	1272	UNK
$4 \leq N < 6$	3500	879	UNK
$6 \leq N < 8$	1953	482	UNK
$8 \leq N$	2117	535	UNK
Total	14339	3585	4482

#### 4.2. Evaluation index

Triplet extraction necessitates identifying the entity span and corresponding relationship type. Hence, a triplet is deemed accurate only when subject, relation and object are correct. We utilize the universally adopted precision (P), recall (R), and micro F1 score (F1) as evaluation metrics. The formulae are as follows:

$$P = \frac{\text{correct}}{\text{recognized}} \quad (15)$$

$$R = \frac{\text{correct}}{\text{triplets}} \quad (16)$$

$$F1 = \frac{2PR}{(P + R)} \quad (17)$$

where *correct*, *recognized*, and *triplets* denote the number of triplets predicated correctly, the total number of triplets predicated, and the total number of triplets labeled respectively.

#### 4.3. Experimental environment and hyperparameters

The experimental environment employed in this work comprised CPU AMD 5900x, GPU RTX3090, Python 3.9.16, PyTorch 1.13.1. In the experiments, the model was trained for 150 epochs with batch size of 4, utilizing the Adam optimizer and dropout ratio of 0.5, GRU hidden size 384. The initial learning rate was set at  $3 \times 10^{-5}$ , linearly decaying to  $1 \times 10^{-8}$ .

#### 4.4. Baseline

We compared FA-CBT with several other methods:

**LSTM-CRF+RC** is a widely adopted classic model for relation extraction, first identifying plausible entities before classifying relations. We employ it as our baseline model, reproducing the



model using ernie-3.0-base-zh as the encoder, with results derived from our own experiments.

**CasRel** 16 is an extraction model grounded in cascade binary tagging framework, first extracting all candidate subject entities before identifying the associated relations and object entities. We reproduced the model utilizing ernie-3.0-base-zh as the encoder, with results derived from our own experiments.

**OneRel** 22 transforms the task into a table-filling task, employing four special tags to determine subject and object boundaries. We reproduced the model using ernie-3.0-base-zh as the encoder, with results derived from our experiments.

**UIE** 23 is a universal information extraction framework enabling unified modeling across entity extraction, relation extraction, event extraction, sentiment analysis, and other tasks, facilitating robust generalization capabilities between different tasks. We reproduced the model utilizing uie-base as the default encoder, with results derived from our experiments.

**BiRTE** 41 represents an enhancement over CasRel, devising a bidirectional extraction framework extracting (*subject, relation, object*) from both directions  $f_r(s) \rightarrow o$  and  $f_r(o) \rightarrow s$ , ameliorating propagation error. We reproduced the architecture leveraging ernie-3.0-base-zh as the encoder, with results stemming from our experiments.

**FA-CBT** is the model devised in this work. It markedly boosts performance through feature augmentation.

#### 4.5. Results

Table 4 exhibits the outcomes across different methods on VSCMeD. The results demonstrated our proposed approach surpasses all models, as evidenced by the data in Table 4.

**Table 4.** Results of the models on VSCMeD.

Method	<i>P</i> (%)	<i>R</i> (%)	<i>F1</i> (%)
LSTM-CRF+RC (baseline)	43.54	27.13	33.43
CasRel	77.98	68.29	72.82
OneRel	31.83	12.86	18.32
UIE	69.10	87.07	77.05
BiRTE	37.21	20.14	26.14
FA-CBT (Our)	<b>87.82</b>	<b>88.47</b>	<b>88.15</b>

To ascertain the impact of the GRU and CNN in the feature augmentation layer, ablation experiments were conducted as well. The results of the ablation studies are delineated in Table 5.

**Table 5.** Results of ablation experiments.

Method	<i>P</i> (%)	<i>R</i> (%)	<i>F1</i> (%)
FA-CBT (Our)	<b>87.82</b>	<b>88.47</b>	<b>88.15</b>
-CNN	83.09	83.98	83.53
-BiGRU	84.12	84.75	84.43
-CNN -BiGRU	77.98	68.29	72.82

Furthermore, to verify the generalization of our method, we also tested it on the CMeIE-V2

dataset, along with testing OneRel and BiRTE which performed poorly on VSCMeD compared against the best CBLUE baseline. The results are shown in Table 6.

**Table 6.** Results of the models on CmeIE-V2.

Method	<i>P</i> (%)	<i>R</i> (%)	<i>F1</i> (%)
CBLUE Best Baseline	51.96	48.06	49.93
OneRel	<b>58.19</b>	48.00	<b>52.61</b>
BiRTE	50.44	<b>51.03</b>	50.73
FA-CBT (Our)	54.51	48.63	51.40

\*To eliminate the errors caused by differences in pre-trained models, all three methods above used the same pre-trained model, Chinese-RoBERTa-wwm-ext-large 42, as the baseline.

## 5. Discussion

We have devised an NLP model for extracting entity relations from Chinese electronic medical records. Compared to baseline models, it achieves over 50% tremendous F1 score improvement on small document-level datasets and surpasses baseline models in both precision and recall on large sentence-level datasets.

Existing research often seeks innovation in entity tagging and decoder to enable one-step extraction of entity relations, as with OneRel 22 and BiRTE 41. Admittedly, these studies have achieved outstanding results on public datasets, but model complexity also increases, making it difficult to learn adequate features from small datasets. As seen in Table 4, the three-dimensional table constructed in OneRel becomes extremely large when learning long sequences, incurring severe performance degradation. BiRTE attempts to improve upon the extraction mode of CasRel but shows no improvement or complementary effect from bidirectional extraction on small datasets. Only the unidirectional extraction CasRel 16 demonstrates decent performance on small datasets. UIE constructs some negative samples during training, with authors affirming through ablation studies that inserting negative samples is effective under small sample settings. In our tests, UIE attained the second highest F1 score on VSCMeD after our model.

On large datasets, as shown in Table 6, OneRel achieves excellent results on short sequence datasets, owing to its unique Rel-Spec Horns Tagging allowing highly accurate boundary identification and thus the highest precision. Meanwhile, the bidirectional extraction framework of BiRTE obtains the highest recall through complementary effects.

The results in Table 5 show that both BiGRU and CNN layers lead to considerable performance improvements. The BiGRU layer benefits the model nicely by extracting contextual features. CNN layer extracts semantic features on top of sequential features, while dampening the negative impact of ineffective features, boosting the model even more.

However, certain limitations persist in this research. As Chen et al. 43 pointed out, all records originate from the same hospital and the relatively consistent writing style may differentially impact different models. We aim to introduce electronic medical records from multiple collaborating hospitals across various departments in follow-up studies, to test model generalization and scalability.

## 6. Conclusions

In this paper, we present a novel approach based on feature augmentation and cascade binary tagging framework to extract relations from Chinese electronic medical records, which is utilized to derive useful information from clinical texts. Experimental results indicate that the proposed model exhibits remarkable generalization ability, delivering impressive outcomes on both small and large datasets, and it is capable of handling long text sequences, making it highly suitable for the practical application scenario in Chinese hospitals. This exploration holds significant practical value for clinical inquiries in China.

### Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

### Acknowledgments

This research was supported by Zhejiang Provincial Natural Science Foundation of China (LQ22F010006); Basic Public Welfare Research Project of Zhejiang Province (LGF20H150006); National Natural Science Foundation of China (No. 81971688); Medical and Public Health Projects in Zhejiang Province (No. 2020386297); Science Foundation of Zhejiang Sci-Tech University (ZSTU) (20032311-Y).

### Conflict of interest

The authors declare there is no conflict of interest.

### References

1. E. Hossain, R. Rajib, N. Higgins, J. Soar, P. D. Barua, A. R. Pisani, et al., Natural language processing in electronic health records in relation to healthcare decision-making: A systematic review, *Comput. Biol. Med.*, **155** (2023), 106649. <https://doi.org/10.1016/j.combiomed.2023.106649>
2. C. A. Nelson, R. Bove, A. J. Butte, S. E. Baranzini, Embedding electronic health records onto a knowledge network recognizes prodromal features of multiple sclerosis and predicts diagnosis, *J. Am. Med. Inf. Assoc.*, **29** (2021), 424–434. <https://doi.org/10.1093/jamia/ocab270>
3. Z. Ning, D. Du, C. Tu, Q. Feng, Y. Zhang, Relation-aware shared representation learning for cancer prognosis analysis with auxiliary clinical variables and incomplete multi-modality data, *IEEE Trans. Med. Imag.*, **41** (2022), 186–198. <https://doi.org/10.1109/TMI.2021.3108802>
4. X. Li, H. Liu, X. Zhao, G. Zhang, C. Xing, Automatic approach for constructing a knowledge graph of knee osteoarthritis in Chinese, *Health Inf. Sci. Syst.*, **8** (2020), 12. <https://doi.org/10.1007/s13755-020-0102-4>

5. F. Liu, M. Liu, M. Li, Y. Xin, D. Gao, J. Wu, et al., Automatic knowledge extraction from Chinese electronic medical records and rheumatoid arthritis knowledge graph construction, *Quant. Imaging Med. Surg.*, **13** (2023), 3873–3890. <https://doi.org/10.21037/qims-22-1158>
6. J. Wu, X. Liu, X. Zhang, Z. He, P. Lv, Master clinical medical knowledge at certificated-doctor-level with deep learning model, *Nat. Commun.*, **9** (2018), 4352. <https://doi.org/10.1038/s41467-018-06799-6>
7. T. Sun, K. Yan, T. Li, X. Lu, Q. Dong, Auxiliary diagnosis of type 2 diabetes complication based on text mining, in *2022 IEEE 5th International Conference on Big Data and Artificial Intelligence (BDAI)*, (2022), 190–194. <https://doi.org/10.1109/BDAI56143.2022.9862667>
8. D. Zeng, K. Liu, S. Lai, G. Zhou, J. Zhao, Relation classification via convolutional deep neural network, in *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, (2014), 2335–2344.
9. G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, C. Dyer, Neural architectures for named entity recognition, in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (2016), 260–270. <https://doi.org/10.18653/v1/N16-1030>
10. X. Shi, Y. Yi, Y. Xiong, B. Tang, Q. Chen, X. Wang, et al., Extracting entities with attributes in clinical text via joint deep learning, *J. Am. Med. Inf. Assoc.*, **26** (2019), 1584–1591. <https://doi.org/10.1093/jamia/ocz158>
11. Q. Wei, Z. Ji, Z. Li, J. Du, J. Wang, J. Xu, et al., A study of deep learning approaches for medication and adverse drug event extraction from clinical text, *J. Am. Med. Inf. Assoc.*, **27** (2019), 13–21. <https://doi.org/10.1093/jamia/ocz063>
12. X. Yang, J. Bian, Y. Gong, W. R. Hogan, Y. Wu, MADEx: A system for detecting medications, adverse drug events, and their relations from clinical notes, *Drug Saf.*, **42** (2019), 123–133. <https://doi.org/10.1007/s40264-018-0761-0>
13. X. Yang, J. Bian, R. Fang, R. I. Bjarnadottir, W. R. Hogan, Y. Wu, Identifying relations of medications with adverse drug events using recurrent convolutional neural networks and gradient boosting, *J. Am. Med. Inf. Assoc.*, **27** (2019), 65–72. <https://doi.org/10.1093/jamia/ocz144>
14. J. D. Lafferty, A. McCallum, F. C. N. Pereira, Conditional random fields: Probabilistic models for segmenting and labeling sequence data, in *Proceedings of the Eighteenth International Conference on Machine Learning*, (2001), 282–89.
15. S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.*, **9** (1997), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
16. Z. Wei, J. Su, Y. Wang, Y. Tian, Y. Chang, A novel cascade binary tagging framework for relational triple extraction, in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, (2020), 1476–1488. <https://doi.org/10.18653/v1/2020.acl-main.136>
17. J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, **1** (2019), 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
18. Y. Gu, R. Tinn, H. Cheng, M. Lucas, N. Usuyama, X. Liu, et al., Domain-specific language model pretraining for biomedical natural language processing, *ACM Trans. Comput. Healthcare*, **3** (2021). <https://doi.org/10.1145/3458754>

19. E. Alsentzer, J. Murphy, W. Boag, W. Weng, D. Jindi, T. Naumann, et al., Publicly available clinical BERT embeddings, in *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, (2019), 72–78. <https://doi.org/10.18653/v1/W19-1909>
20. C. Vasantharajan, K. Z. Tun, H. Thi-Nga, S. Jain, T. Rong, C. E. Siong, MedBERT: A pre-trained language model for biomedical named entity recognition, in *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, (2022), 1482–1488. <https://doi.org/10.23919/APSIPAASC55919.2022.9980157>
21. H. Wang, M. Tan, M. Yu, S. Chang, D. Wang, K. Xu, et al., Extracting multiple-relations in one-pass with pre-trained transformers, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019), 1371–1377. <https://doi.org/10.18653/v1/P19-1132>
22. Y. Shang, H. Huang, X. Mao, OneRel: Joint entity and relation extraction with one module in one step, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **36** (2022), 11285–11293. <https://doi.org/10.1609/aaai.v36i10.21379>
23. Y. Lu, Q. Liu, D. Dai, X. Xiao, H. Lin, X. Han, et al., Unified structure generation for universal information extraction, in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, **1** (2022), 5755–5772. <https://doi.org/10.18653/v1/2022.acl-long.395>
24. Z. Zhong, D. Chen, A frustratingly easy approach for entity and relation extraction, in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (2021), 50–61. <https://doi.org/10.18653/v1/2021.naacl-main.5>
25. W. Sun, A. Rumshisky, O. Uzuner, Annotating temporal information in clinical narratives, *J. Biomed. Inf.*, **46** (2013), S5–S12. <https://doi.org/10.1016/j.jbi.2013.07.004>
26. C. Wei, Y. Peng, R. Leaman, A. P. Davis, C. J. Mattingly, J. Li, et al., Assessing the state of the art in biomedical relation extraction: overview of the biocreative V chemical-disease relation (CDR) task, *Database*, **2016** (2016), baw032. <https://doi.org/10.1093/database/baw032>
27. A. E. Johnson, T. J. Pollard, L. Shen, L. H. Lehman, M. Feng, M. Ghassemi, et al., MIMIC-III, a freely accessible critical care database, *Sci. Data*, **3** (2016), 160035. <https://doi.org/10.1038/sdata.2016.35>
28. A. E. Johnson, L. Bulgarelli, L. Shen, A. Gayles, A. Shammout, S. Horng, et al., MIMIC-IV, a freely accessible electronic health record dataset, *Sci. Data*, **10** (2023). <https://doi.org/10.1038/s41597-022-01899-x>
29. T. Li, Y. Xiong, X. Wang, Q. Chen, B. Tang, Document-level medical relation extraction via edge-oriented graph neural network based on document structure and external knowledge, *BMC Med. Inf. Decis. Making*, **21** (2021), 368. <https://doi.org/10.1186/s12911-021-01733-1>
30. T. Chen, M. Wu, H. Li, A general approach for improving deep learning-based medical relation extraction using a pre-trained model and fine-tuning, *Database*, **2019** (2019), baz116. <https://doi.org/10.1093/database/baz116>
31. Y. Sun, J. Wang, H. Lin, Y. Zhang, Z. Yang, Knowledge guided attention and graph convolutional networks for chemical-disease relation extraction, *IEEE/ACM Trans. Comput. Biol. Bioinf.*, **20** (2023), 489–499. <https://doi.org/10.1109/TCBB.2021.3135844>
32. N. Zhang, M. Chen, Z. Bi, X. Liang, L. Li, X. Shang, et al., CBLUE: A chinese biomedical language understanding evaluation benchmark, in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, **1** (2022), 7888–7915. <https://doi.org/10.18653/v1/2022.acl-long.544>

33. H. Chang, H. Zan, T. Guan, K. Zhang, Z. Sui, Application of cascade binary pointer tagging in joint entity and relation extraction of chinese medical text, *Math. Biosci. Eng.*, **19** (2022), 10656–10672. <https://doi.org/10.3934/mbe.2022498>
34. Y. Pang, X. Qin, Z. Zhang, Specific relation attention-guided graph neural networks for joint entity and relation extraction in Chinese EMR, *Appl. Sci.*, **12** (2022), 8493. <https://doi.org/10.3390/app12178493>
35. Q. Zhang, M. Wu, P. Lv, M. Zhang, L. Lv, Research on Chinese medical entity relation extraction based on syntactic dependency structure information, *Appl. Sci.*, **12** (2022), 9781. <https://doi.org/10.3390/app12199781>
36. Q. Ye, T. Cai, X. Ji, T. Ruan, H. Zheng, Subsequence and distant supervision based active learning for relation extraction of Chinese medical texts, *BMC Med. Inf. Decis. Making*, **23** (2023), 34. <https://doi.org/10.1186/s12911-023-02127-1>
37. T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, preprint, arXiv: 1301.3781.
38. Y. Sun, S. Wang, Y. Li, S. Feng, X. Chen, H. Zhang, et al., Ernie: Enhanced representation through knowledge integration, (2019). <https://doi.org/10.48550/arXiv.1904.09223>
39. J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, preprint, arXiv:1412.3555.
40. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *Med. Image Comput. Comput. Assisted Int.*, **9351** (2015), 234–241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
41. F. Ren, L. Zhang, X. Zhao, S. Yin, S. Liu, B. Li, A simple but effective bidirectional framework for relational triple extraction, in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, (2022), 824–832. <https://doi.org/10.1145/3488560.3498409>
42. Y. Cui, W. Che, T. Liu, B. Qin, Z. Yang, Pre-training with whole word masking for Chinese BERT, *IEEE/ACM Trans. Audio Speech Lang. Process.*, **29** (2021), 3504–3514. <https://doi.org/10.1109/TASLP.2021.3124365>
43. L. Chen, L. Song, Y. Shao, D. Li, K. Ding, Using natural language processing to extract clinically useful information from chinese electronic medical records, *Int. J. Med. Inf.*, **124** (2019), 6–12. <https://doi.org/10.1016/j.ijmedinf.2019.01.004>



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)