Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 19(5): 5134–5152.
DOI: 10.3934/mbe.2022240
Received: 17 December 2021
Revised: 24 February 2022
Accepted: 15 March 2022
Published: 18 March 2022

Research article

Span-based model for overlapping entity recognition and multi-relations classification in the food domain

Mengqi Zhang[1,2], Lei Ma[1,2], Yanzhao Ren[3], Ganggang Zhang[4] and Xinliang Liu[1,2,*]

[1] School of E-business and Logistics, Beijing Technology and Business University, Beijing 100048, China

[2] National Engineering Laboratory for Agri-product Quality Traceability, Beijing Technology and Business University, Beijing 100048, China

[3] School of Computer Science and Engineering, Beijing Technology and Business University, Beijing 100048, China

[4] Digital Campus Construction Center, Capital Normal University, Beijing 100048, China

* **Correspondence:** Email: liuxinl@btbu.edu.cn.

Abstract: Information extraction (IE) is an important part of the entire knowledge graph lifecycle. In the food domain, extracting information such as ingredient and cooking method from Chinese recipes is crucial to safety risk analysis and identification of ingredient. In comparison with English, due to the complex structure, the richness of information in word combination, and lack of tense, Chinese IE is much more challenging. This dilemma is particularly prominent in the food domain with high-density knowledge, imprecise syntactic structure. However, existing IE methods focus only on the features of entities in a sentence, such as context and position, and ignore features of the entity itself and the influence of self attributes on prediction of inter entity relationship. To solve the problems of overlapping entity recognition and multi-relations classification in the food domain, we propose a span-based model known as SpIE for IE. The SpIE uses the span representation for each possible candidate entity to capture span-level features, which transforms named entity recognition (NER) into a classification mission. Besides, SpIE feeds extra information about the entity into the relation classification (RC) model by considering the effect of entity's attributes (both the entity mention and entity type) on the relationship between entity pairs. We apply SpIE on two datasets and observe that SpIE significantly outperforms the previous neural approaches due to capture the feature of overlapping entity and entity attributes, and it remains very competitive in general IE.

Keywords: information extraction; span-based approach; overlapping entity recognition; category marker; multi-relations classification; entity attributes

1. Introduction

Considering the food field, the number of electronic texts has increased owing to the rapid growth of computer technology and the response to changes in people's dietary habits. Texts in the food field contain a number of precious knowledge. For example, considering the literature, we can learn what microorganisms are used for in food preservation. Moreover, regarding recipes, we can know how a dish is prepared, etc. This study is a completion of the project of the 'Construction of Knowledge Graph for Safety Risk Analysis and Identification of Food Ingredient for the Beijing 2022 Olympic and Paralympic Winter Games'. We consider Chinese dishes as examples to show more details. Chinese dishes are rich in ingredients and complex in preparation, and the recipes contain a wealth of information. The dish name may even contain information about ingredients and cooking methods. An entity can be the subject or object of multiple entity pairs in the same sentence. The aim of IE is to extract triple facts from a given sentence, and Figure 1 provides an example from the Recipe-IE dataset. The triple facts are in the form of (subject, relation, object) or (head entity, relation, tail entity). However, manually annotating useful information from a number of texts requires an enormous amount of labor and is time-consuming [1]. Deep learning (DL), a branch of machine learning (ML), supports more effective extraction of entities and their relations because of its ability to learn representations of data via multiple processing layers [2]. Deep learning-based IE is a process of extracting structured knowledge from unstructured text. This is usually composed of two subtasks: 1) recognizing named entity from raw unstructured text and predicting which pre-defined entity category they belong to [3]; 2) identifying the relationship between pairs of entities mentioned and classifying them into predefined relation categories [4].

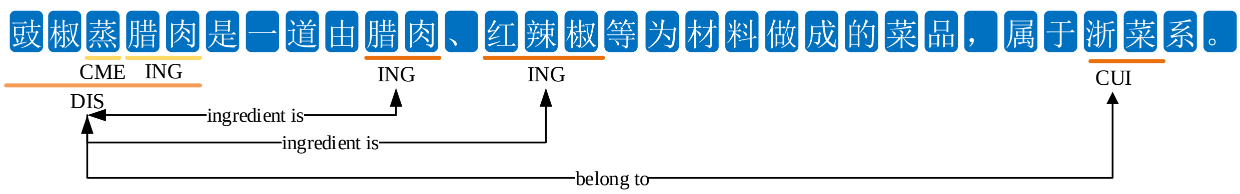


Figure 1. An example of entities, relations and sentence from the Recipe-IE dataset. Given an input sentence "豉椒蒸腊肉是一道由腊肉、红辣椒等为材料做成的菜品，属于浙菜系" (SSauteed Preserved Pork with Black Bean Sauce chili is a dish made of SSauteed Preserved Pork and red pepper, belonging to Zhejiang cuisine). The IE system is aimed to extract that 豉椒蒸腊肉(SSauteed Preserved Pork with Black Bean Sauce chili), 腊肉(SSauteed Preserved Pork), 红辣椒(red pepper), 浙菜(Zhejiang cuisine) are entities of type 菜名(DIS), 原材料(ING), 原材料(ING) 菜系(CUI). The ingredient_is and belong_to are relation types. The spans of red and orange underscores represent a pair of overlapping entity. A line with an arrow connects two entities and there is a relationship between them.

Conventional approaches train one model to extract entities and another model to determine the relationship between them. This is known as the pipeline approach. This method is easy to implement, and the two models have high flexibility. From 2014 (Miwa and Sasaji [5]), the most promising end-to-end approach has been used to jointly model these two tasks. The two tasks make predictions in a unified structural framework or share representations. Although joint models can be used to mitigate error propagation issues, it does not mean that it always has better performance than the pipeline method. One main argument about why the joint model is superior to the pipeline model is that its capable of capturing the interactions between entities and relations. However, in 2019, Zhong et al. [6] used a pipeline approach for the IE task and discovered that entity information is significant for relation prediction, while they have not found that the relation model can improve the entity model. In addition, the two tasks need different formats and features to predict entities and relations, while the joint model shares the encoder, which cannot better learn task-specific features.

In this study, a pipeline model known as SpIE is presented to handle IE from complex Chinese texts. The SpIE models all possible spans and uses span representation to perform entity recognition on all the spans in parallel. The fine-tuned span representation (with added entity mentions and type information) was subsequently used to perform the relation classification on the pair of recognized entities. The contributions of our study are summarized below.

- This study offers a novel approach for complex information extraction from unstructured Chinese texts, using a suitable language model to generate span representation for overlapping entity recognition and multi-relations classification. Considering the Beijing 2022 Olympic and Paralympic Winter Games, we established a new dataset known as Recipe-IE.
- The SpIE considers the impact of its own entity attributes and innovatively feeds the fusion features of entity mentions and types into the RC task.
- The experimental results on the publicly DuIE dataset and the self-constructed Recipe-IE dataset demonstrate that SpIE markedly outperforms the traditional neural approaches for complex IE and remains competitive in normal IE tasks.

2. Related work

Information extraction is primarily oriented towards open linked data and extracts usable triples (knowledge units) through automated techniques. When a sentence with well-labeled entities and relations are provided, the goal of IE is to obtain triples from the sentence. These triples are the basis for a series of high-quality factual representations to further build the upper-level model. They also support various downstream natural language processing tasks, such as query answering [7] and recommendation systems [8]. Recently, the construction of knowledge graphs in vertical fields [9–13], such as the biomedical domain and the field of urban traffic, has shown great application prospects and commercial value.

Manually annotating knowledge from various unstructured texts based on predefined rules [14] is labor-intensive, expensive, and difficult to adapt to data changes. Several existing approaches use traditional ML-based models, such as conditional random fields (CRF) and hidden Markov models (HMM), to handle this task, which rely on hand-crafted rules. These ML-based models are too inflexible to scale into a new dataset. To overcome this shortcoming, DL-based approaches have been proposed to capture features automatically because of their ability to learn representations [15]. When raw data are fed to the machine, it can automatically discover the potential representations needed for classification or detection [16]. We divided the literature on the application of DL-based approaches to the IE task. In this study, we classify the existing studies on IE into four major categories: table-filling, multi-task learning, sequence labeling, and span-based approaches.

2.1. Table-filling

Miwa et al. [5] and Gupta et al. [17] transformed NER and RC tasks into a simple table-filling problem and modeled these two tasks jointly through table representation. The main difference between the two is that the former used a history-based entity and relation table, and the latter labeled each word pair using a bidirectional recurrent neural network (Bi-CNN).

2.2. Multi-task learning

This family of models is essentially the same as the pipeline method. Bekoulis et al. [18] modeled the NER and RC tasks as multi-head selection problems, which did not require external NLP tools. The contextualized representation of each word was generated from a long short-term memory (LSTM) layer. Miwa and Bansal [19] used a Bi-LSTM to tag the entities using a begin inside-last outside unit (BI-LOU) scheme and a bidirectional tree-structured RNN model for relation extraction. The attention mechanism can also be introduced in joint models. Furthermore, Nguyen and Verspoor [20] adopted a pipeline approach. A Bi-LSTM-CRF-based model for NER and a bi-affine attention mechanism was presented for RC. The aforementioned approaches relied heavily on the LSTM and its variants.

2.3. Sequence labeling

Sequence labeling requires additional features, which is essentially a distributed representation for input. Zheng et al. [21] proposed an end-to-end model with tagging scheme for IE and experimental results showed the effectiveness of their method. Ma et al. [22] introduced a NN architecture using the combination of LSTM, CNN and CRF, which required no feature engineering or data processing. Cao et al. [23] and Lv et al. [24] adopted a Bi-LSTM-CRF model to extract information from biomedical and software engineering context respectively. The difference between the two is that the former based on the BIOES (Begin, Inside, Outside, End, Single) scheme with an additional M tag to indicate the relation type that the entity took part in whereas the later extract information from multi-source software knowledge. Extra knowledge has also been used for IE: Zhou et al. [25] introduced an attention-based Bi-LSTM model to acquire the semantic information from sentences without NLP tools. The attention layer was used to generate a weight vector, and then merged each word-level feature into a sentence-level feature vector in order. Li et al. [26] proposed a multi-grained information and external linguistic knowledge for Chinese information extraction. Polysemy and ambiguity will be avoided by introducing external linguistic knowledge. Sequence labeling can also be used for key-phrase extraction [27] from scholarly documents, which is similar to IE.

2.4. Span-based approaches

The above approaches tackle both entity and relation extraction as sequence labeling tasks which are processed by token-level models with the begin inside outside (BIO) scheme or its variant. However, this task is still challenging because the structure of Chinese texts is more complex, and the texts are rich in information. For example, “腊肉” in entities “豉椒蒸腊肉” (an subject) and “腊肉” (an object) has the same token representation, this situation with overlapping entity cannot be handled well. Here, we propose a span-based model that is beneficial for covering overlapping entity by providing different representations for all spans. The idea of using span representations to replace the token-level is not entirely new because it has been studied in neural coreference resolution [28]. Recently, the idea of span representation has been improved by the DYGIE model [29]. It introduced a general architecture for IE tasks and captured the interaction of spans by using dynamically constructed span graphs to update the span representation. The DYGIE++ (Wadden et al.) [30] replaced LSTM encoding with BERT-based representations to capture the semantic information within and across sentences. Different from DYGIE++, PURE [6] used separate encoders for both NER and RC model, and took the entity labels gotten from NER model as the input feature. PURE reported better performance than previous work. However, this approach is not suitable for Chinese and ignores the role of entity attributes in predicting the relation between two entities. We propose an improved span-based model for Chinese IE in the food domain. In contrast with PURE, 1) we use the Chinese pre-trained BERT [31] model known as "BERT-wwm-ext1" [32] to generate the raw token representations. A suitable pre-trained language model can deliver further gains. 2) We use entity attributes (fusion information of entity mention and type) as the input feature for the RC task. 3) We used a unique representation for each span in RC model regardless of which entity pairs it belongs to in one sentence, due to the certainty of the entity. In addition, SpIE uses a feed-forward neural network (FFNN) and a softmax layer for entity and relation classification. The results of the experiment show that SpIE markedly outperforms the baseline approaches.

3. Materials and methods

3.1. Overview of the proposed framework

This study presents a novel framework known as SpIE to solve the problem of overlapping entity recognition and multi-relations classification. Figure 2 displays the workflow of our framework which is divided into three modules. 1) Regarding the embedding on the original corpus, BERT-wwm-ext1 [32] was used to obtain the token representation, and span representation generated from the token representation. 2) Considering the entity recognition model, the model is used to detect both the boundaries and types of entities. 3) Regarding the relation classification model, the model is used to classify each pair of entities into predefined relation types.

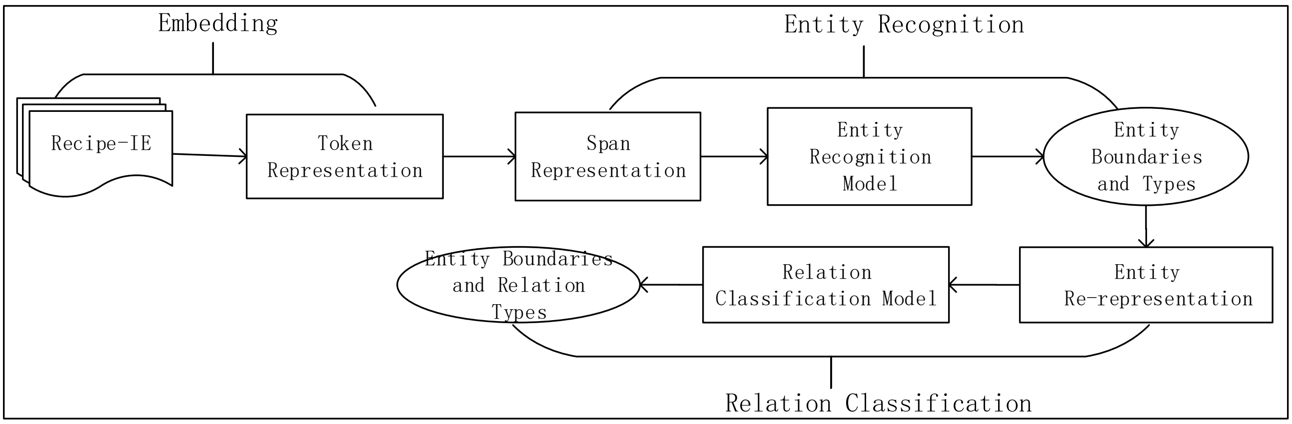


Figure 2. The workflow of our model. We use rectangles to indicate specific operations or processes, arrows to indicate the workflow of the system, and elliptical boxes indicated the results of entity recognition and relation classification.

Figure 3 illustrates the SpIE architecture. Figure 3(a) is an entity model for recognizing the named entity, and Fihure 3(b) is a relation model for classifying each span pair into predefined relation types. In the subsequent subsections, we explain the main components of the SpIE in detail.

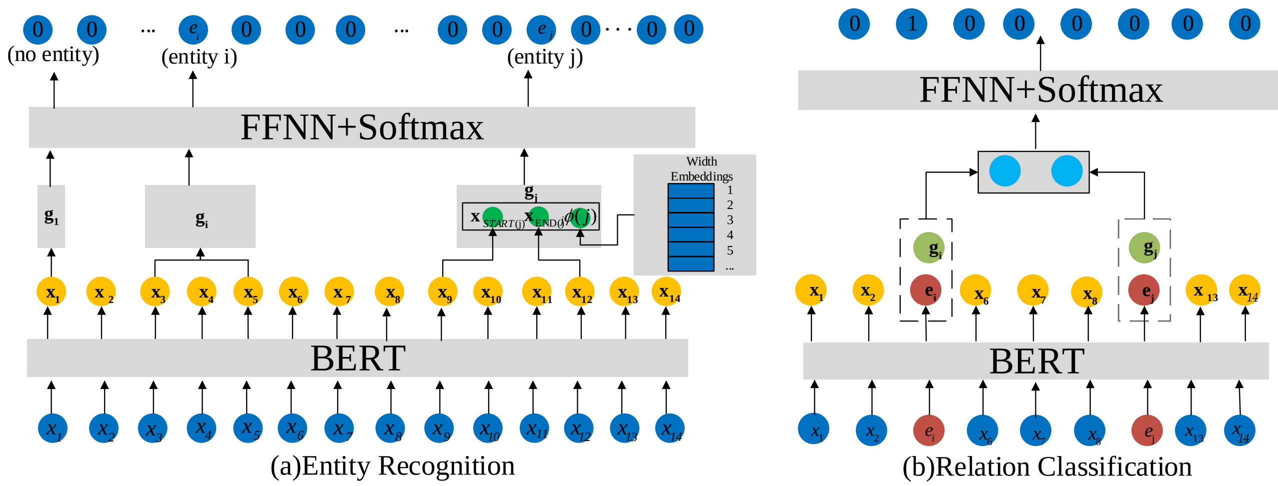


Figure 3. The overall architecture of the SpIE model.

3.2. *Different part of the framework*

3.2.1. Span representation generation

The input is a sentence $X = x_1, x_2, ..., x_n$ with n tokens. Because of the remarkable performance of the recent transformer networks, this study adopts BERT-wwm-ext1 [32] to obtain the contextualized embedding $\mathbf{x_n}$. Our model constructs all the possible spans $S = s_1, s_2, ..., s_m$, where $m = \frac{n(n+1)}{2}$. Following Zhong's model [6], s_i is defined as all the tokens from $START(i)$ to $END(i)$ for $1 \leqslant i \leqslant$ m. This indicates that $START(i)$ and $END(i)$ denote the start and end indices of s_i, respectively. The spans are sorted based on the order of $START(i)$, and spans with the same start index are sorted by $END(i)$. Given a span $s_i \in S$, s_i's representation $\mathbf{g_i}$ is obtained by concatenating the vector representation of the start, end, and span width features:

$$\mathbf{g_i} = [\mathbf{x}_{START(i)}; \mathbf{x}_{END(i)}; \phi(i)] \tag{3.1}$$

Considering Eq (3.1), $\mathbf{x}_{START(i)}$ and $\mathbf{x}_{END(i)}$ are the head and tail token embedding of s_i, respectively. Moreover, $\phi(i)$ is the embedding of the span width feature.

3.2.2. Named entity recognition

Considering this step, the task is to predict the best entity category for each candidate span. Regarding each s_i, we predict an entity category $y_e(s_i) \in \varepsilon \cup NONE$. The output of the task is $Y_E = (s_i, e)$, where $s_i \in S, e \in \varepsilon \cup NONE$. We use ε to represent a set of predefined entity categories. $y_e(s_i) = NONE$ indicates that s_i does not belong to any predefined entity category. Considering each span s_i, SpIE uses a two-layer FFNN with ReLU activation to compute a vector of entity category scores in parallel. Thereafter, the softmax function was applied to the entity category score to obtain a distribution over the entity categories. A pair of overlapping entity belongs to different spans with different representations. These are difficult to define in the traditional token-level model.

Considering span i, the representation $\mathbf{g_i}$ is used to predict the entity category $e \in \varepsilon$:

$$score_i^{ner} = FFNN_{ner}(\mathbf{g_i}) \tag{3.2}$$

$$p_i^{ner} = softmax(score_i^{ner}) \tag{3.3}$$

The output sizes of $FFNN_{ner}$ and p^{ner} are equal to the number of NER classes. Considering each span s_i, the predicted entity category corresponds to the span's highest entity category score. If a span is assigned to an entity category other than NONE, it will determine whether there is a relationship between it and other spans.

3.2.3. Relation classification

The goal of the RC model is to predict the best relation category between the entity pair (s_i, s_j), where s_i and s_j are gold entities ($y_e(\cdot) \in \varepsilon, \cdot$ can be s_i or s_j) and are not overlapping entity. The study only considers binary relations, where (s_i, s_j) and (s_j, s_i) are considered to be different. Within a sentence, we consider each ordered pair of spans. The spans were obtained from the NER model. The output of the RC is $Y_R = (s_i, s_j, r)$, where $s_i, s_j \in S, r \in \mathcal{R} \cup NONE$, and $\mathcal{R}$ are the sets of predefined relation categories. $r = NONE$ indicates that there is no relationship between the pair of candidate spans. A previous study [33] re-used span representation to predict their relations. However, other studies [34, 35] have validated that the original span representation only captures the contextual information around each entity, which ignores the dependencies between each other. Furthermore, other studies [36, 37] proved that extra information was vital to RC. Based on this idea, we introduce a pair of spans with an entity label. The experimental results demonstrate that it is easier to determine the relationship between a pair of spans.

Considering the relation model, we use category markers to replace entity mentions to obtain the representation of the label. We denote the adjusted X by X^*, and Figure 4 provides an example.

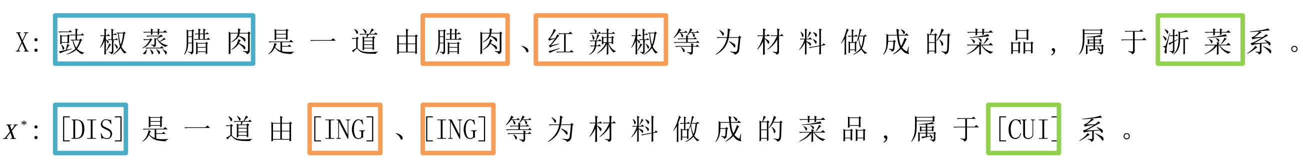


Figure 4. An example of replace entity mentions with type information.

X^* is fed into the pre-trained language model to obtain the representation of entities' type $\mathbf{e_i}$, which incorporates the contextual information. Thereafter, we concatenate $\mathbf{e_i}$ with $\mathbf{g_i}$ (generated from the NER model) to represent the entity mention with the type information. $\hat{\mathbf{g}}_i = [\mathbf{e_i}; \mathbf{g_i}]$ is a new presentation of the span. Regarding a pair of spans (s_i, s_j), we compute an ordered pair embedding $\mathbf{r}_{i,j} = [\hat{\mathbf{g}}_i; \hat{\mathbf{g}}_j]$. Considering (s_i, s_j), the representation $\mathbf{r_{i,j}}$ is used to predict the relation category $r \in \mathcal{R}$:

$$score_{i,j}^{re} = FFNN_{re}(\mathbf{r_{i,j}}) \tag{3.4}$$

$$p_{i,j}^{re} = softmax(score_{i,j}^{re}) \tag{3.5}$$

The output sizes of $FFNN_{re}$ and $p_{i,j}^{re}$ are equal to the number of RE classes. Regarding each pair s_i, s_j, the predicted relation category corresponds to the pair's highest relation category score.

3.2.4. Loss

In this study, we use supervised training on both the NER and RC models. When a set of sentences, well-labeled entities, and relations are given, we use the cross-entropy loss for these two models:

$$\mathcal{L}_{ner} = -\sum_{s_i \in S} \log P_{ner}(e_i^* \mid s_i) \tag{3.6}$$

$$\mathcal{L}_{re} = -\sum_{s_i, s_j \in S', s_i \neq s_j} \log P_{re}(r_{i,j}^* \mid s_i, s_j) \tag{3.7}$$

Considering Eq (3.6), $\mathcal{L}_{ner}$ represents the cross-entropy over the entity category including none, and e_i^* represents the correct entity category for s_i. Regarding the relation model, we only consider the correct entities $S \subset S'$. Moreover, considering Eq (3.6), $\mathcal{L}_{re}$ represents the cross-entropy over the relation classes, including none, and $r_{i,j}^*$ represents the correct relation category for the ordered pair of spans (s_i, s_j).

3.3. Dataset preparetion and processing

This study aims to provide a high-performance model to automatically extract entities and relations from unstructured Chinese texts to construct and expand the knowledge graph. We construct a new dataset to describe recipe information. We build a web crawler in Python using two libraries (requests and Beautiful Soup) to collect the experimental data from Baidu Baike (*https://baike.baidu.com/item/*豉椒蒸腊肉). These data are complex samples of high-density knowledge for the Beijing 2022 Olympic and Paralympic Winter Games.

Sixteen annotators and two supervisors annotated the corpus using the corpus annotation system developed by the National Engineering Laboratory for Agri-product Quality Traceability. In order to ensure the correctness of the annotation, annotators were trained before annotating, and they annotated back-to-back in pairs. When the annotation data is inconsistent, the supervisor annotated again. Finally, the correct annotated corpus make up the Recipe-IE dataset and there are 2318 sample data obtained from 750 recipes that are semantically segmented for periods '。'. There are some distinctive features in the recipes corpus. 1) Information about the ingredients and cooking methods may be contained in the dish name, which will lead to overlapping entity. This feature is shown in Figure 1. An underlined span indicates an entity. The spans underlined in red and orange form a pair of overlapping entity. 2) Considering a sentence, a subject corresponds to multiple objects, and the relationship between any two spans belongs to different relation categories. Regarding Figure 1, the line with the arrow connects two spans, and there is a relation between them. Each triple describes the relevant dish's knowledge. The study developed seven types of dish entities: 菜名(dish name), 菜系(cuisine), 原材料(ingredient), 工艺(cooking method), 营养物质(nutrient), 适合人群 (people), 功效(efficacious) and seven relation categories: 属于(belong_to), 原料是 (ingredient_is), 烹饪方

法(cook_by), 含有(contain), 适合(suitable_for), 功效(efficacy_is), 相克(overcome). The distribution of the Recipe-IE dataset is shown in Figure 5, and Table 1 displays the statistics for all entities and relations.

There are 683 instances of sentences with overlapping entity in the Recipe dataset, representing 29.45% of the total. Over 27% of the sentences contain over one triple, and the maximum length of the gold entity is 11.

(a) **(b)**

Figure 5. Figure (a),(b) are the type distribution of entity and relation separately.

Table 1. Statistics of all Chinese dishes entities and relations.

Type of Entity	Count	Type of Relation	Count
dish name(DIS)	1155	belong_to	857
cuisine(CUI)	857	ingredient_is	740
ingredient(ING)	3037	cook_by	86
cooking method(CME)	348	contain	859
nutrient(NUT)	859	suitable_for	98
people(PEO)	98	efficacy_is	592
efficacious(EFF)	592	overcome	365

4. Results

4.1. Experimental setup

We experimented on the Recipe-IE dataset and part of the publicly available DuIE dataset. We followed Luan's pre-processing steps and randomly divided the dataset into training, dev, and test sets. There was no duplication between the sets. The training set was used for model training in the training phase for the model training stage. The dev set validates the trained model and adjusts the hyper-parameters to optimize the model. The model was evaluated using the test set. Table 2 presents the information for the two datasets.

The BERT-wwm-ext1 [32] model was used as a sentence encoder, which was pre-trained in the Chinese language. We used Pytorch with a GPU to complete the module. The context window size was

300 with the two-layer FFNN with 150 hidden units. We trained the model with the Adam optimizer and a warmup ratio of 0.1. In addition, we trained the model for 10 epochs with a dropout of 0.2, a learning rate of 2e-5, and a batch size of 16.

Table 2. The sample distribution of and the average length of each sample.

Dataset	$\mid \varepsilon \mid$	$\mid \mathcal{R} \mid$	Sentence		
			Train	Dev	Test
Recipe-IE	7	7	1623	463	233
DuIE	7	7	1731	216	200

4.2. Evaluation metrics

This study follows the standard evaluation protocol and measures the F1-score, precision (P), and recall (R) as the evaluation metrics. The F1, P, and R are calculated as follows:

$$P = \frac{TP}{TP + FP} \tag{4.1}$$

$$R = \frac{TP}{TP + FN} \tag{4.2}$$

$$F1 = \frac{2 \times P \times R}{P + R} \tag{4.3}$$

Considering the above equations, false negative (FN), false positive (FP), true positive (TP), and true negative (TN) indicate incorrect negative prediction, incorrect positive prediction, correct positive prediction, and correct negative prediction, respectively.

Bekoulis et al. [38] proposed three types of evaluation, namely: 1) *Strict*: an entity is correct only if both the entity boundaries and type are correct, 2) *Boundaries*: only make sure the entity boundaries are correct and the type is not considered, 3) *Relaxed*: for a multi-token entity, only ensure at least one correct type is assigned to the tokens comprising the entity. We use *Strict* to evaluate the IE task. For named entity recognition, an entity is considered to be correctly predicted only if both span boundaries and the type match the truth. Considering the relation extraction, a predicted relation is considered as a correct prediction only if both the boundaries of the two spans and the predicted relation type are correct. In fact, the representation used for relationship prediction has fused the type information, and the correct span boundaries can ensure that the span type is also correct.

4.3. Experimental results

The performance of the SpIE is benchmarked by comparing it with the latest models in the entity relation extraction task. There are two types of baseline models. 1) The basic Bi-LSTM-CRF model with different input features (both different embedding and tagging schemes) are defined as token-level models. Cao et al. [23] adopted an NN with an M-tag and an original tagging scheme for entities and relation extraction. Gao et al. [37] used a character-and word-attention-enhanced NN for information extraction. 2) Dixit et al. [33] used a span-based , which use Bi-LSTM and ELMo for raw token embedding. PURE [6] is a span-based model which uses bert-base-uncased as the base encoder and directly uses entity label as the input feature for the RC model. These two models are used as a baseline to verify whether different language models and input features affect the experimental results.

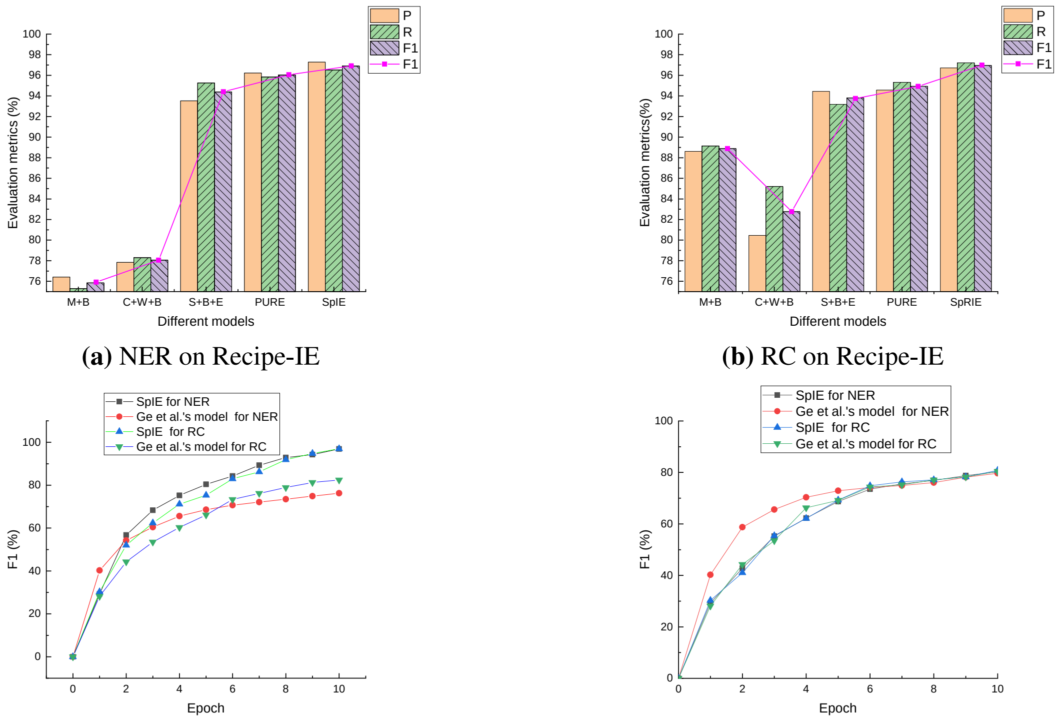
We conducted our experiment using SpIE and other baseline models on the self-constructed Recipe-IE dataset. The SpIE performed markedly better than the other models in both the NER and RC. Table 3 and Figure 6 show the experimental results in detail. Comparing SpIE to token-level models, we conclude that performing non-span-based models will struggle when dealing with complex Chinese samples, such as overlapping entity and multi-relations. These models take a token representation as the input. However, traditional token-level models have an inherent limitation because the representation is fixed and cannot model overlapping entity. Comparing SpIE to other span-based models, we observe that the performance of these two models is close to that of our model and higher than that of those token-lever models, in which PURE is closer to that of our model in NER task. We come to the conclusions： 1) The span-based representation is more useful for overlapping entity recognition. 2) The key to achieving this result is that the encoder, using whole word masking as the core mechanism, captures the interactions between tokens. A suitable pre-trained language model can deliver further gains. 3) The SpIE model fully considers the impact of the entity type on relation extraction, and the inserted entity type information highlights the subject, object, and type. Considering Figures 6(a),(b) compared to the token-level models on the Recipe-IE dataset, our model achieves a F1 improvement of +21.05 and +18.83% for entity recognition and +8.09 and +14.20% for relation classification. Compared to other span-level models, our model achieves a F1 improvement of +2.51 and +1.05% for entity recognition and +3.15 and +2.02% for relation classification.

Table 3. The predicted results of different methods on IE.

Methods	Entity recognition			Relation classification		
	P(%)	R(%)	F1(%)	P(%)	R(%)	F1(%)
M tag+ Bi-LSTM(M+B) [23]	76.41	75.29	75.85	88.62	89.14	88.88
Char+word+ Bi-LSTM(C+W+B) [37]	77.84	78.29	78.06	80.46	85.21	82.77
Char+word+ CNN(C+W+B) [39]	75.92	77.41	76.25	84.37	82.39	83.37
Span + Bi-LSTM +ELMo(S+B +E) [33]	93.53	95.27	94.39	94.44	93.19	93.81
PURE [6]	96.23	95.84	96.03	94.57	95.32	94.94
Our model (SpIE)	97.28	96.52	96.90	96.72	97.21	96.96
SpIE (overlapping entities only)	96.23	95.47	95.85	-	-	-
SpIE (flat entities only)	98.37	97.04	97.70	-	-	-

To future study the capability of the proposed SpIE in extracting general triple facts, we conduct extended experiments on DuIE dataset. Since Ge et al. [39] mixed the characters and words feature for Chinese IE and their study has demonstrated this method has reported a state-of-the-art performance on DuIE datasets. We evaluated the performance of our model and Ge et al.'s model on both Recipe-IE and DuIE datasets. After 10 epochs, due to the interaction between raw tokens being fully captured, our model achieved the same performance as Ge et al.'s model, while their model reported poor performance on Recipe-IE datasets. Table 3, Figures 6(c),(d) describe the specific performance of these two models on these two datasets, which demonstrates that our SpIE remains competitive in the normal IE task.

The SpIE model transforms both NER and RC into a typical multi-label classification task. Table 4 shows the results of each classification task. To further investigate the capabilities of the SpIE on IE, we statistically analyzed the details for each category, and the experimental results are shown in Table 4. Our model performs well on all entity and relation categories, especially on the dish name, ingredient, ingredient_is, and efficacy_is. These categories have a better performance because they have more structural features. Specifically, the position of the dish name is relatively fixed and consecutive ingredients are linked by ','.



(a) NER on Recipe-IE

(b) RC on Recipe-IE

(c) Results on Recipe-IE

(d) Results on DuIE

Figure 6. NER and RC results on different datasets.

Table 4. F1-scores of each entity and relation category.

Entity(label)	P(%)	R(%)	F1(%)	Relation(label)	P(%)	R(%)	F1(%)
DIS	98.12	98.34	98.23	belong_to	96.94	95.92	96.43
CUI	97.65	97.81	97.74	ingredient_is	98.14	97.22	97.68
ING	98.71	97.94	98.32	cook_by	95.18	94.57	94.87
CME	94.96	93.63	94.29	contain	96.54	97.49	97.01
NUT	95.02	95.87	95.44	suitable_for	93.40	95.04	94.21
PEO	92.15	94.49	93.31	efficacy_is	96.45	97.60	97.02
EFF	96.98	95.86	96.42	overcome	94.57	95.68	95.12

5. Discussion

In this section, we demonstrate why this model performed well.

5.1. Span pruning

A number of spans under consideration require significant computational overhead and cause memory issues. We set the maximum length of the span to seven and only consider spans that are entirely within a sentence. Our model creates a representation for the span instead of a token, and the maximum length we set can capture over 99.90% of the entities. Table 5 shows the statistics of the effects of different policies on span generation. When the span's max length changes from one to seven, we obtain a 3.3 × speedup, and the F1 score increases from 76.14 to 96.90%. However, when all the possible spans are considered, the time consumption reaches 7.4 × speedup, and F1 only increases by 0.21%. "Policy" indicates the max length of the span permitted. "Coverage" denotes the rate of coverage of entity, "F1" and "Speed" are the F1 score for NER and time consumption under different span generation policies, respectively. The trend of the fold in Figure 7 shows the phenomenon more intuitively. Considering L = 7 to all the spans, the time consumption exhibits an extraordinarily steep increase, whereas there is no significant difference in the F1-score. Considering entity coverage, F1-score, and time consumption, we set the maximum span length to seven.

Table 5. Statistics on the effects of different policy for span generation. We measured speed on a single NVIDIA GeForce 10,606 GB.

Policy	Coverage (%)	F1	Speed (sent/s)
All spans	100.00	97.11	1755.71
L = 7	99.90	96.90	237.64
L = 4	95.98	90.15	194.14
L = 1	46.29	76.14	72.14

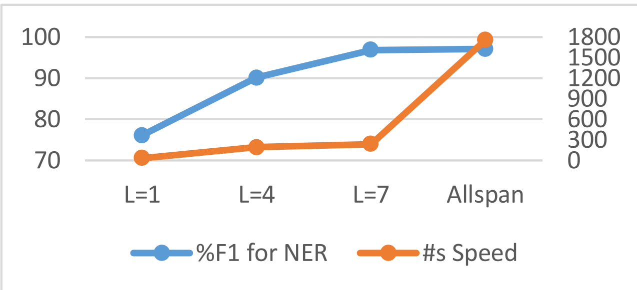


Figure 7. The trends of the impact.

We provide additional analysis of the performances on overlapping entities and flat entities separately. SpIE's performance on "overlapping entity only" and "flat entity only" are similar to it on the whole dataset, and the ability to recognize overlapping entity is slightly lower than that of flat entity. While in Table 5, when we select L = 1, we treat the overlapping entities as flat entities, which reports poor performance. The span-based method achieves significant performance for overlapping entity recognition at the cost of speed.

5.2. Category marker

We argue that both entity mentions and types are important in determining the relationship between two entities. To train the relation model, a category marker (entity type information) was inserted into the original sentence. The presentation of $\mathbf{r}_{i,j} = [\hat{\mathbf{g}}_i; \hat{\mathbf{g}}_j]$ is the input of the relation model, and $\hat{\mathbf{g}}$ contains both entity mention and type information. To validate the importance of entity category markers, we present four options (illustrated in Figure 8) for the input of the RC model.

No Marker: The NO MARKER relation model references Luan [29] and Dixit's [33] relation model. The input of the relation model is $[\mathbf{g}_i; \mathbf{g}_j; \mathbf{g}_i \circ \mathbf{g}_j]$, where $\mathbf{g}_i$ and $\mathbf{g}_j$ are the span representations obtained from the entity model ($\mathbf{g}_i, \mathbf{g}_j \neq NONE$) and $\mathbf{g}_i \circ \mathbf{g}_j$ refers to the element-wise product to concatenate the hidden representation.

M Marker: We use [S], [/S], [O], [/O] (the marker without entity type information) at the input layer and use each entity's start marker to represent the entity.

T Mareker: The T markers e_i and e_j are used to replace the entity mentions s_i,s_j at the input layer. The representations $\mathbf{e_i}$ and $\mathbf{e_j}$ have both context and type information.

M+T Marker: As described in Section 3.3, we use $\hat{\mathbf{g}}_i = [\mathbf{e_i}; \mathbf{g_i}]$ to describe the gold entity s_i, and $\hat{\mathbf{g}}_i$ contains context, entity mention and type information.

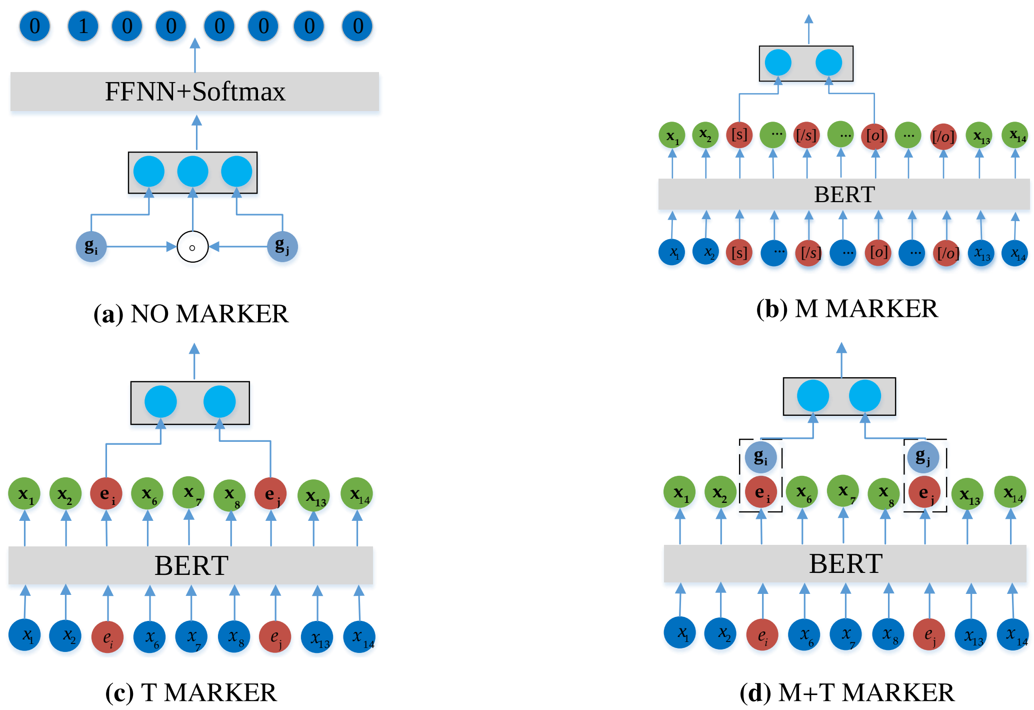


Figure 8. RC model with different input features.

Furthermore, we conducted a comparison experiment on the Recipe-IE dataset. The relation of the F1-score with different input features is shown in Table 6. Type information is the most important feature to relation classification, and the M+T Marker achieves a F1 improvement of score for 1.72% than T Marker.

Table 6. Statistics on the F1 of different input features.

Feature	F1(%)
No Maeker	90.85
M Marker	92.77
T Marker	95.24
M+T Marker	96.96

Table 7. Model setting ablation on the development set.

Feature	F1 for NER	F1 for RC
SpIE	96.90	96.96
+ LSTM layer	97.17 (0.27↗)	-
− FFNN	93.35 (3.55↘)	-
$+\hat{\mathbf{g}}_i \circ \hat{\mathbf{g}}_j$	-	97.35 (0.39↗)

5.3. Ablation test

To thoroughly analyze the effectiveness of the various components of SpIE, including the strategy of embedding, modeling with FFNN, and learning with correlation features, we added an LSTM layer above the BERT, and the performance improved slightly. However, this difference is small. This indicates that the in-domain pre-training language model can effectively capture full contextual representation. Thereafter, we skipped the FFNN, directly connected the BERT layer to the softmax layer, and observed that the performance dropped significantly. This indicates that BERT is not sufficient to classify spans into predefined entity categories. When we replaced $[\hat{\mathbf{g}}_i; \hat{\mathbf{g}}_j]$ with $[\hat{\mathbf{g}}_i; \hat{\mathbf{g}}_j; \hat{\mathbf{g}}_i \circ \hat{\mathbf{g}}_j]$ for the RC task, the performance increased slightly. However, the time consumption reaches $1.2\times$ speedup. In other word, using their element-wise multiplication is not enough to capture the interaction characteristics between this two spans. It comes to the same conclusion with Section 5.2 that No Marker reported the poorest performance. Table 7 lists the performance of the ablation test.

6. Conclusions

This study proposes SpIE as a general information extraction framework for overlapping entity recognition and multi-relation classification from Chinese texts in the food domain. Moreover, SpIE tackles information extraction as two classification tasks through a span-based representation. The span-based model provides different representations for all the spans, which makes handling overlapping entity possible. The connection features of entity attributes fully considers the influence of entity mentions and types on the relationship between entity pairs. The comparative experiments and ablation tests indicate that i) span representation is effective for overlapping entity detection; ii) both entity mentions and type information are vital for relation classification; iii) SpIE outperforms previous neural approaches in complex IE tasks mentioned in this study and is still competitive in general IE tasks. In the future, we plan to extend SpIE to coarse-grained (document level) text extraction and demonstrate its effectiveness in more complex IE tasks.

Acknowledgments

This project was supported by Beijing Science and Technology Planning Project [Grant Numbers Z191100008619007].

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Y. Lin, S. Shen, Z. Liu, H. Luan, M. Sun, Neural relation extraction with selective attention over instances, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, **1** (2016), 2124–2133. https://doi.org/10.18653/v1/P16-1200
2. M. Nickel, K. Murphy, V. Tresp, E. Gabrilovich, A review of relational machine learning for knowledge graphs, *Proc. IEEE*, **104** (2016), 11–33. https://doi.org/10.1109/JPROC.2015.2483592
3. D. Nadeau, S. Sekine, A survey of named entity recognition and classification, *Lingvisticae Investigationes*, **30** (2007), 3–26. https://doi.org/10.1075/li.30.1.03nad
4. N. Bach, S. Badaskar, A review of relation extraction, 2007. Available from: https://www.cs.cmu.edu/ nbach/papers/A-survey-on-Relation-Extraction.pdf.
5. M. Miwa, Y. Sasaki, Modeling joint entity and relation extraction with table representation, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (2014), 1858–1869. https://doi.org/10.3115/v1/D14-1200
6. 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
7. R. D. Semmel, J. Mayfield, Automated query formulation using an entity-relationship conceptual schema, *J. Intell. Inf. Syst.*, **8** (1997), 267–290. https://doi.org/10.1023/A:1008629825231
8. J. A. P. Sacenti, R. Fileto, R. Willrich, Knowledge graph summarization impacts on movie recommendations, *J. Intell. Inf. Syst.*, **58** (2022), 43–66. https://doi.org/10.1007/s10844-021-00650-z
9. Y. Li, Y. Chen, I. Podlubny, Stability of fractional-order nonlinear dynamic systems: Lyapunov direct method and generalized MittagLeffler stability, *Comput. Math. Appl.*, **59** (2010), 1810–1821. https://doi.org/10.1016/j.camwa.2009.08.019
10. H. Zhou, S. Ning, Z. Liu, C. Lang, Z. Liu, B. Lei, Knowledge-enhanced biomedical named entity recognition and normalization: application to proteins and genes, *BMC Bioinf.*, **21** (2020). https://doi.org/10.1186/s12859-020-3375-3
11. Y. Huang, B. Niu, An improved method for web text affective cognition computing based on knowledge graph, *Comput. Mater. Continua*, **59** (2019), 1–14. https://doi.org/10.32604/cmc.2019.06032

12. J. Tan, Q. Qiu, W. Guo, T. Li, Research on the construction of a knowledge graph and knowledge reasoning model in the field of urban traffic, *Sustainability*, **13** (2021), 3191. https://doi.org/10.3390/su13063191
13. M. Habibi, L. Weber, M. Neves, D. L. Wiegandt, U. Leser, Deep learning with word embeddings improves biomedical named entity recognition, *Bioinformatics*, **33** (2017), i37–i48. https://doi.org/10.1093/bioinformatics/btx228
14. C. Friedman, P. O. Alderson, J. H. M. Austin, J. J. Cimino, S. B. Johnson, A general natural-language text processor for clinical radiology, *J. Am. Med. Inf. Assoc.*, **1** (1994), 161–174. https://doi.org/10.1136/jamia.1994.95236146
15. J. Li, A. Sun, J. Han, C. Li, A survey on deep learning for named entity recognition, *IEEE Trans. Knowl. Data Eng.*, **34** (2022), 50–70. https://doi.org/10.1109/TKDE.2020.2981314
16. Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature*, **521** (2015), 436–444. https://doi.org/10.1038/nature14539
17. P. Gupta, H. Schütze, B. Andrassy, Table filling multi-task recurrent neural network for joint entity and relation extraction, in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, (2016), 2537–2547. Available from: https://aclanthology.org/C16-1239.pdf.
18. G. Bekoulis, J. Deleu, T. Demeester, C. Develder, Joint entity recognition and relation extraction as a multi-head selection problem, *Expert Syst. Appl.*, **114** (2018), 34–45. https://doi.org/10.1016/j.eswa.2018.07.032
19. M. Miwa, M. Bansal, End-to-end relation extraction using LSTMs on sequences and tree structures, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, **1** (2016), 1105–1116. https://doi.org/10.18653/v1/P16-1105
20. D. Q. Nguyen, K. Verspoor, End-to-end neural relation extraction using deep biaffine attention, in *Advances in Information Retrieval*, **11437** (2019), 729–738. https://doi.org/10.1007/978-3-030-15712-8_47
21. S. Zheng, F. Wang, H. Bao, Y. Hao, P. Zhou, B. Xu, Joint extraction of entities and relations based on a novel tagging scheme, in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, **1** (2017), 1227–1236. https://doi.org/10.18653/v1/P17-1113
22. X. Ma, E. Hovy, End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, **1** (2016), 1064–1074. https://doi.org/10.18653/v1/P16-1101
23. M. Cao, Z. Yang, L. Ling, H. Lin, J. Wang, Joint drug entity and relations extraction based on neural network, *J. Comput. Res. Dev.*, **56** (2019), 1432–1440. https://crad.ict.ac.cn/EN/Y2019/V56/I7/1432
24. W. Lv, Z. Liao, S. Liu, Y. Zhang, MEIM: A multi-source software knowledge entity extraction integration model, *Comput. Mater. Continua*, **66** (2021), 1027–1042. https://doi.org/10.32604/cmc.2020.012478

25. P. Zhou, W. Shi, J. Tian, Z. Qi, B. Li, H. Hao, et al., Attention-based bidirectional long short-term memory networks for relation classification, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, **2** (2016), 207–212. https://doi.org/10.18653/v1/P16-2034
26. Z. Li, N. Ding, Z. Liu, H. Zheng, Y. Shen, Chinese relation extraction with multi-grained information and external linguistic knowledge, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019), 4377–4386. https://doi.org/10.18653/v1/P19-1430
27. R. Alzaidy, C. Caragea, C. L. Giles, Bi-lstm-crf sequence labeling for keyphrase extraction from scholarly documents, in *The World Wide Web Conference*, New York, (2019), 25512557. https://doi.org/10.1145/3308558.3313642
28. K. Lee, L. He, M. Lewis, L. Zettlemoyer, End-to-end neural coreference resolution, in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, (2017), 188–197. https://doi.org/10.18653/v1/D17-1018
29. Y. Luan, D. Wadden, L. He, A. Shah, M. Ostendorf, H. Hajishirzi, A general framework for information extraction using dynamic span graphs, in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, **1** (2019), 3036–3046. https://doi.org/10.18653/v1/N19-1308
30. D. Wadden, U. Wennberg, Y. Luan, H. Hajishirzi, Entity, relation, and event extraction with contextualized span representations, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, (2019), 5784–5789. https://doi.org/10.18653/v1/D19-1585
31. 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
32. Y. Cui, W. Che, T. Liu, B. Qin, Z. Yang, Pre-training with whole word masking for Chinese BERT, *IEEE/ACM Trans. Audio*, **29** (2021), 3504–3514. https://doi.org/10.1109/TASLP.2021.3124365
33. K. Dixit, Y. Al-Onaizan, Span-level model for relation extraction, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019), 5308–5314. https://doi.org/10.18653/v1/P19-1525
34. L. Baldini Soares, N. FitzGerald, J. Ling, T. Kwiatkowski, Matching the blanks: Distributional similarity for relation learning, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019), 2895–2905. https://doi.org/10.18653/v1/P19-1279
35. Z. Zhang, X. Han, Z. Liu, X. Jiang, M. Sun, Q. Liu, ERNIE: Enhanced language representation with informative entities, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019), 1441–1451. https://doi.org/10.18653/v1/P19-1139
36. H. Peng, T. Gao, X. Han, Y. Lin, P. Li, Z. Liu, et al., Learning from context or names? An empirical study on neural relation extraction, in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (2020), 3661–3672. 10.18653/v1/2020.emnlp-main.298

37. Y. Gao, Y. Wang, P. Wang, L. Gu, Medical named entity extraction from chinese resident admit notes using character and word attention-enhanced neural network, *Int. J. Environ. Res. Public Health*, **17** (2020), 1614. https://doi.org/10.3390/ijerph17051614

38. G. Bekoulis, J. Deleu, T. Demeester, C. Develder, Adversarial training for multi-context joint entity and relation extraction, in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, (2018), 2830–2836. https://doi.org/10.18653/v1/D18-1307

39. J. Ge, S. Li, F. Yiqiu, Joint extraction method of Chinese entity relationship based on mixture of characters and words, *Appl. Res. Comput.*, **38** (2021), 2619–2623. https://doi.org/10.19734/j.issn.1001-3695.2021.01.0006