

A review of Chinese named entity recognition

Jieren Cheng^{1,2}, Jingxin Liu^{1,*}, Xinbin Xu¹, Dongwan Xia¹, Le Liu¹, Victor S. Sheng³

¹School of Compute Science and Cyberspace Security
Hainan University, Haikou, 570228, China

[email: 2210489897@qq.com,291108957@qq.com,839119912@qq.com,1220474455@qq.com]

²Hainan Blockchain Technology Engineering Research Center
Hainan University, Haikou, 570228, China
[e-mail: cjr22@163.com]

³Department of Computer Science Texas Tech University
TX, 79409, American
[e-mail:victor.sheng@ttu.edu]

*Corresponding Author: Jingxin Liu

*Received April 5, 2021; revised May 11, 2021; accepted May 16, 2021;
published June 30, 2021*

Abstract

Named Entity Recognition (NER) is used to identify entity nouns in the corpus such as Location, Person and Organization, etc. NER is also an important basic of research in various natural language fields. The processing of Chinese NER has some unique difficulties, for example, there is no obvious segmentation boundary between each Chinese character in a Chinese sentence. The Chinese NER task is often combined with Chinese word segmentation, and so on. In response to these problems, we summarize the recognition methods of Chinese NER. In this review, we first introduce the sequence labeling system and evaluation metrics of NER. Then, we divide Chinese NER methods into rule-based methods, statistics-based machine learning methods and deep learning-based methods. Subsequently, we analyze in detail the model framework based on deep learning and the typical Chinese NER methods. Finally, we put forward the current challenges and future research directions of Chinese NER technology.

Keywords: Chinese word segmentation, Deep learning, Machine learning, Model framework, Named entity recognition.

1. Introduction

The purpose of Named Entity Recognition (NER) is to identify named entities such as person names, place names, and organizational structure names in the corpus, that is to identify entities with specific meanings in the text [1]. NER is also called "entity recognition", "entity extraction" and "entity segmentation", etc. It is the basis of many research questions in the field of Natural Language Processing such as relation extraction [2], event extraction [3], knowledge graph [4], machine translation [5], question answering system [6, 7] and so on.

The history of NER. NER was formally proposed as a clear concept at the MUC-6 meeting [8]. The conference pointed out that the named entity task is composed of three branches: entity recognition, numerical expression recognition and time expression recognition, besides all solutions to the NER tasks were based on rule matching. In the subsequent MUC-7 conference [9], although most named entity recognition still uses rule-based methods, there also have some attempts at named entity recognition based on statistical methods, such as Hidden Markov Model (HMM) [10], Maximum Entropy Model (MEM) [11] and so on. At the CoNLL-2003 conference [12], four types of named entities were identified and classified, Person (PER), Organization (ORG), Location (LOC), and other miscellaneous entities. The conference mainly used machine learning methods. Before the Bakeoff-2006 conference, most NER methods were based on English text as the main research object. In this conference, the NER methods of Chinese text became the main research object.

As shown in Fig. 1, it describes the development trend of Chinese NER. According to this trend, we mainly analyze and summarize the methods of Chinese NER from three dimensions: (1) Rule-based methods, which require domain experts to make a lot of annotations on entities, and its portability is poor. (2) Statistics-based machine learning methods use statistical algorithms and artificially set characteristics to train the model. (3) Based on the deep learning methods, the input data are used to automatically obtain the entity label and entity category through the deep learning model. In this review, we briefly introduce (1), (2), and summarize (3) in detail. Finally, through a comparative analysis of typical deep learning methods, we put forward the current challenges and future directions of Chinese NER.

The motivation for this review. In recent years, with the success in various fields, Chinese NER technology has attracted much attention. When the Chinese NER methods are applied to the field of Chinese social media, external knowledge or joint training model is usually used to improve the recognition of small annotation corpus [13]. Li Weiyan et al. [14] applied Chinese NER technology to the medical field, which adopts automatic recognition and extraction to the medical text entities. Zhou Jie et al. [15] proposed an approach to automatically construct a Chinese NER corpus from Chinese Wikipedia. Although NER methods have been proposed decades ago, most NER tasks are based on English text, and there are relatively few surveys and reviews on Chinese NER. Li Jing et al. [16] made a systematic summary of NER, and carefully analyzed NER methods from traditional approaches and deep learning approaches, but did not provide a system summary for Chinese NER approaches, and Chinese NER technology still faces many challenges.

Contributions of this review. In this paper, we summarize Chinese NER methods including rule-based methods, statistics-based machine learning methods, and deep learning-based methods. At the same time, we analyze the model framework based on deep learning and typical Chinese NER methods and applications. In addition, we put forward the current challenges and the future development directions in Chinese NER.

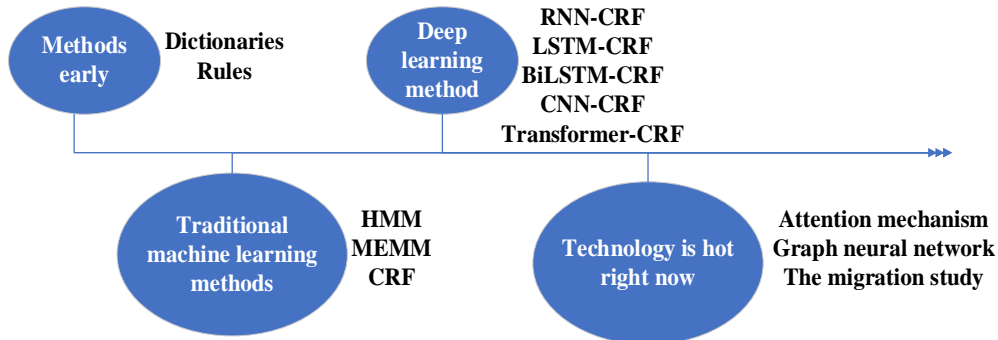


Fig. 1. Development Trend of Chinese Named Entity Recognition Method

2. Background

2.1 Definition of NER

Named entity recognition [1] is to extract entities with specific meanings from text information. In academic terms, the entities involved in NER generally refer to three major categories and seven subcategories. The three major categories include Entity category, Time category, and Number category, and the seven subcategories contain Person, Location, Organization, Time, Percentage, Currency, Date. In practical applications, it is generally only necessary to identify the Person, Location, Organization, Time, and Date. The specific process, enter a sentence as: "张三在海南大学上学". After the sentence is recognized by the named entity, "张三" and "海南大学" in the sentence will be identified. The label of "张三" is the Person (PER), and the label of "海南大学" is the Organization (ORG).

2.2 Sequence tag system

Solving the problem of sequence labeling is considered as the key of NER [17]. There are many different annotation modes that can be used for different datasets. In general, the common annotation methods are BIO, BIOES, BMEWO, etc. At present, the BIOES is the most common named entity annotation mode. In some areas where entities are denser, the BIOES mode is selected to better identify these entities. The more complex the annotation system, the higher the accuracy, but the corresponding training time increase. Therefore, the appropriate annotation system should be chosen according to the actual situation [18].

2.3 Evaluation metrics

For the development of a NER system, a comprehensive evaluation of the system is essential and vital to its development. Based on the evaluation metrics, the definition of data in NER is as follows: TP means that the correct entity sample is identified in NER, which is actually a correct sample; TN means that the wrong entity sample was identified in the NER, which is actually the wrong sample; FP means that the correct entity sample is identified in NER, but it is actually the wrong sample; FN identifies the wrong entity sample in NER, but it is actually the correct sample.

According to the above four types of data, the used evaluation indicators are mainly accuracy, recall, and F-value to evaluate NER tasks. Their definitions are as follows,

Accuracy rate refers to the number of truly correct samples among the samples predicted to be correct in the NER process;

$$P = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

Recall rate refers to how many of the actual correct samples in NER are predicted to be correct samples;

$$R = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

F1 value is a measurement method of the harmonic average of precision and recall. Usually, when the parameter α is equal to 1, it is F1.

$$F1 = \frac{2 \times P \times R}{R + P} \times 100\% \quad (3)$$

Generally, the results of NER are shown that a higher accuracy rate leads to a lower recall rate; a higher recall rate causes a lower accuracy rate; Therefore, the F1 value is selected as the reconciliation standard. High recall rate is paid more attention to NER, but high accuracy rate is more focused on information retrieval [19].

3. Chinese NER methods

3.1 Rule-based recognition methods

Rule-based recognition methods [20] are the earliest mainstream methods to appear in Chinese NER, and they are constructed under the existing rule system. The rule-based NER methods rely on a large number of language experts to formulate rule templates with punctuation marks, keywords, etc. [21]. With the development of technology, Wang et al. [22] used a sublingual mechanism to propose a new Chinese name recognition method.

3.2 Statistics-based machine learning recognition methods

With the rise of machine learning in the field of Chinese Natural Language Processing at the beginning of the 21st century, the research on Chinese NER also turned to a combination of statistics and machine learning [23]. These methods mark few samples through artificially setting features, then use statistical algorithms and artificially set features to train the model. Based on these characteristics, commonly used statistical models for sequence labeling include: MEM [11], HMM [10], Conditional Random Field (CRF) [24, 25], etc. Hu Hongping et al. [26] used CRF as a Chinese NER model to compare the effect of two different levels of models based on character-level and word-level. Zhou Jie et al. [15] combined multi-party information rules with a supervised named entity classifier to identify named entity types in Chinese Wikipedia articles, and used entity linking methods to identify ambiguous named entity types. In addition, a method for selecting a labeled corpus based on core article extension is proposed, which can automatically adapt to the field of test data, so as to obtain a better NER model as a training corpus.

3.3 Recognition methods based on deep learning

In recent years, with the rapid development of neural networks, methods of deep learning made major breakthroughs in the field of image recognition [27, 28], speech recognition [29] and natural language processing [30, 31]. In NER, the appearance of BiLSTM-CRF [32] opened

the prelude of the deep learning era of NER. Its appearance makes the model more concise and more robust, and becomes the baseline for deep learning to solve the problem of NER. According to the process framework, the Chinese NER model of deep learning can be divided into embedding presentation layer, sequence modeling layer, and label decoding layer. In this section, we introduce the model framework and typical methods of Chinese NER based on deep learning.

3.3.1 Model framework

The embedding representation layer [33] is the first step of the deep learning NER process, and its function is to convert textual information into a vector representation that can be recognized by the computer. After obtaining the representation of these input vectors, the input vectors are passed to the sequence modeling layer [34]. The sentence features of each sentence are extracted and transmitted to the label decoding layer [35]. The entity labels are predicted, and the corresponding input is generated the tag sequence.

Embedding presentation layer. Due to the particularity of Chinese, the basic unit of input is a Chinese character. It can be represented word-level embedding representation [36] and character-level embedding representation [37].

Adopting word-level embedding representation, the pre-trained word vector is generally used, which can well represent the input word. In the medical field, because BERT requires a large number of Wikipedia as a corpus for training, it will lead to a slower training speed, and its training greatly depends on computing resources. But the actual medical field unlikely to provide expensive computing resources for actual operations. For this problem, Zhu Jiayi et al. [38] collected and sorted out a medical corpus of about 36 million characters, utilized CCKS2019 data as a testset to perform NER tasks with the generated character-level vectors and word-level vectors. Based on the medical corpus, Word2vec is used for training, which can be deployed in relatively low-configuration devices. It is faster, lighter, and more valuable in the actual medical field. However, word-level embedding representation will cause errors in NER, due to errors in Chinese word segmentation.

Adopting character-level embedding representation can avoid errors caused by word segmentation. However, the disadvantage is that the semantic information existing between adjacent characters is not used and the word segmentation boundary is unknown. Therefore, adding word-level semantic information and features on the basis of character-level representation became the mainstream embedding representation methods for Chinese NER [39]. Ye Na et al. [40] combined word representation and character representation to propose a Chinese NER model that combines character and word vectors. In order to solve the problem of strong correlation between adjacent characters, Zhang Naixin et al. [41] proposed a new dynamic embedding method, which uses the attention mechanism to combine the character and word vector features at the embedding layer.

Sequence modeling layer. It uses the existing neural network model to model the input sequence. The main neural network models are network models based on RNN and its variants [42, 43], models based on CNN [44, 45], and models based on Transformer [46, 47].

The neural network model based on RNN and its variants. In order to solve the problems in the field of Chinese biomedical NER, Li et al. [48] proposed a model based on RNN (WCP-RNN), which combines the input representation of characters and word vectors to obtain orthographic and lexical semantic features. Yang Yaosheng et al. [49] proposed a group tagging method for Chinese NER based on adversarial training ideas, which makes full use of the noisy sequence labels of multiple annotators. This method uses two Bi-LSTMs to represent the general information and specific information of the annotators, then encodes at the LSTM

network layer and finally passes to the decoding layer to obtain Chinese labels. When network coding is based on the model of RNN and its variants, the main disadvantage is the speed problem caused by the RNN unable to parallelize network coding.

The neural network model based on CNN. Wang Chunqi et al. [50] proposed a gated CNN-based framework, which is used different datasets for experiments on simplified Chinese NER and traditional Chinese NER. The experimental results reflect the advanced performance of the model. Chen Hui et al. [51] proposed a CNN-based NER model, the Gated Relational Network (GRN). This model is simple and effective, which introduces gated layer to build connection between random two words, and uses gated mechanism to merge global features for all words. GRN has a better performance in capturing long-distance information compared with the ordinary CNN model. Network coding based on the CNN model can model sentences in parallel and has higher efficiency, but it is difficult to deal with long-distance dependence.

The network model based on Transformer. Li Xiaonan et al. [46] proposed a FLAT(Flat-Lattice Transformer) structure to be applied to Chinese NER. This model relies on the powerful functions of Transformer and carefully designed specific location codes to fully utilize lattice information and has efficient parallelism. Xue Mengge et al. [47] presented a Transformer Encoder Extension (PLTE), which not only models all characters and matching words at the same time in batch processing, but also introduces a porous mechanism to enhance the ability of local modeling. Guo Xiaoran et al. [52] proposed a character-level Chinese NER approach based on Transformer encoder, which combines direction information with character vectors at the embedding layer. At the same time, the Transformer encoder is introduced to further obtain the relationship characteristics between words. Compared with RNN models coding, Transformer-based models have efficient parallelism; Compared with CNN, it can capture long-distance features more effectively. However, many current experimental studies have shown that the effect of Transformer is far inferior to the BiLSTM model at dealing with Chinese NER tasks. The main reason is that length of Chinese sentences divided into individual sentences according to characters is generally short, and the advantages of long-distance modeling of the Transformer encoder cannot be used.

Label decoding layer. After obtaining the contextual feature representation, Label decoding layer predicts the entity label to generate the corresponding output label sequence. Conditional Random Fields [53-55] are currently recognized label decoders used in deep learning for NER tasks. The main reason is that CRF considers the interdependence between tags on the basis of text information modeling, so as to get a better solution. If there is no CRF layer, the tags are independent of each other. In order to identify a large number of informal writing entities in Chinese social media, Dong Chuanhai et al. [56] presented a multi-channel LSTM-CRF model based on out-of-domain annotation data, which utilizes different channels to share the same character embedding to improve NER performance. At the same time, choosing CRF as the decoder helps to enhance the recall rate. In the field of e-medicine, Liu Kaixin et al. [57] added four features to Chinese clinical NER based on CRF, and constructed a medical dictionary to capture the features. The research results show that these features are beneficial to the recognition of named entities to varying degrees in Chinese clinical NER tasks. The above is the model framework of Chinese NER based on deep learning. Fig. 2 is based on the BiLSTM-CRF model framework, which is the most common architecture in the current Chinese NER model.

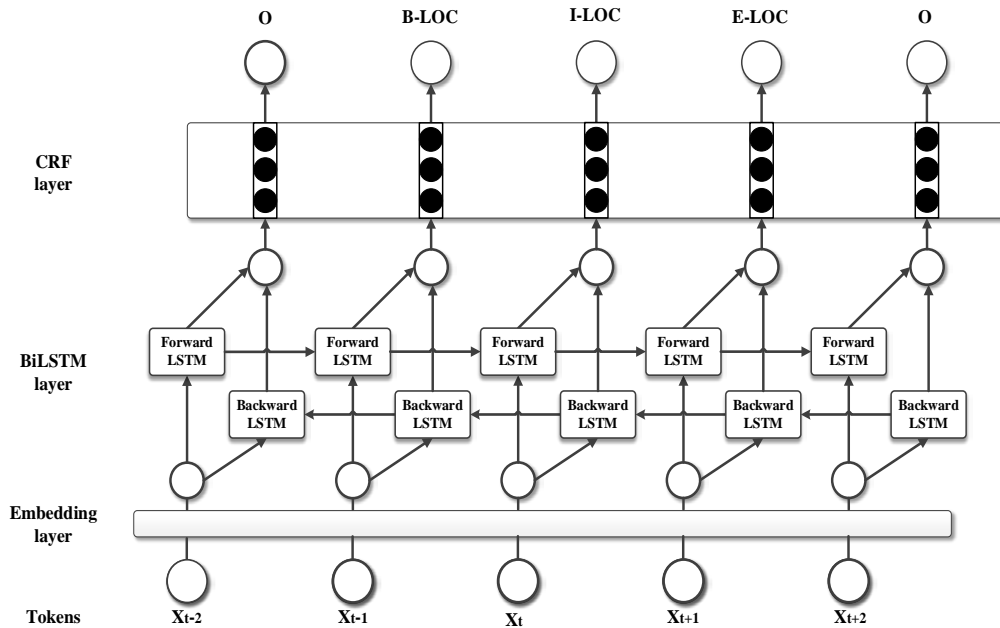


Fig. 2. Chinese NER model based on BiLSTM-CRF

3.3.2 Typical recognition methods

Chinese NER methods combined with dictionary. Due to the unique difficulties of Chinese characters the coupling between Chinese NER and Word Segmentation is very strong, and the character-level representation methods cannot make good use of the information of Chinese words. Therefore, the mainstream methods for improving the accuracy of Chinese NER are to add dictionary information to character-level sentences in recent years. The general idea is to build a structure with a large amount of word information on the character-level sentence through vocabulary matching, then encode this structure to obtain the sentence representation with the adding word information, then use CRF to decode the tag sequence to realize Chinese NER. At present, the overall improvement methods for adding vocabulary information to Chinese NER can be divided into two categories: One type is dynamic improvement in the sequence modeling layer. The other is to improve in the embedded presentation layer.

The dynamic improvement methods in the sequence modeling layer. They combine vocabulary information into the model by transforming the sequence modeling layer. Zhang Yue et al. [42] first proposed a Lattice-LSTM model, which opened the prelude to the dynamic improvement method in the sequence modeling layer in Chinese NER. This model adds an extra word-level LSTM cell between non-adjacent words in the character-level model. Its advantage is that it cleverly uses the information between words and sequences to eliminate ambiguity. In the medical field, Zhao et al. [43] first presented an adversarial training-based method (AT-Lattice-LSTM-CRF) suitable for Chinese clinical NER. This model utilizes LSTM coding to make the model consider the word and character information in a balanced manner, so as to make full use of the clinical entity information of the electronic health record. The experimental results show that the addition of noise to the training of the Adversarial Training (AT) not only enhances the robustness of the neural network approaches, but also enhances the effect of the model. Gui Tao et al. [44] proposed an LR-CNN model. This model introduces the Rethinking mechanism to merge vocabulary and uses high-level features to guide the weight distribution of the lower level, which can better solve the problem of word

boundary conflicts. Gui Tao et al. [58] also introduced a dictionary-based GNN model with global semantics to alleviate the lexical ambiguity of Chinese NER. This model combines the semantic information of characters, potential words and the entire sentence through multiple graph interactions to effectively solve the problem of word ambiguity. Li Xiaonan et al. [46] presented the Flat-Lattice-Transformer model, which transforms the lattice structure into a flat structure of a set of spans, so that each span has a character or potential word in the lattice structure corresponding to it. This model uses Transformer-specific position coding to utilize lattice information, and has efficient parallelism, and can directly model long-distance dependency. The experimental results show that the model runs very efficiently and surpasses other models that combine dictionary in performance. The performance improvement on large datasets is particularly obvious.

The improvement methods in the embedding presentation layer. Through the transformation of the embedding layer, the vocabulary information is integrated into the model. Liu Wei et al. [59] presented a new word-character vector embedding model (WC-LSTM), which adds the word information to the beginning or end characters of a word, so as to effectively utilize word boundary information and reduce the impact of Chinese Word Segmentation errors. Ma Ruotian et al. [60] proposed a Soft-Lexicon method, which cleverly combines word information and word boundary information into the embedding presentation layer of the model. The model starts from the constructed features and accurately restores the matching results of the dictionary. In the constructed features, each word can be represented by a corresponding word vector, and there is no problem of missing information and word segmentation error propagation.

Nested NER methods. Typically, the task of NER does not consider the problem of nested entities, but in the actual text, there are many nested entities to cause each entity to correspond to multiple labels. For Nested NER, some researchers still regard them as common sequence labeling methods to complete. Ju Meizhi et al. [61] dynamically stacked Flat NER layers to identify nested entity, and proposed a dynamic hierarchical model. The model divides each nested named entity into multiple layers to recognize, and passes the information obtained to the next layer of entity recognition after each layer is completed, and so on. From the experimental results, it can be inferred that in this method, the use of internal entities greatly facilitates the detection of external entities. Jana Straková et al. [62] connected multiple tags of nested entities into a multi-tag, and proposed a neural architecture for Nested NER, which allows multiple neural network tags to be constructed in an enhanced BILOU coding scheme. Mohammad Golam Sohrab et al. [63] proposed a simple neural network model. The model first lists all possible subsequences as potential entities, and then classifies them through a deep neural network. Zheng Changmeng et al. [64] proposed a Nested NER boundary-aware neural model, which uses a sequence labeling model to detect boundaries to accurately locate entities. Joseph Fisher et al. [65] proposed a new Nested NER neural network structure, which predicts the internal relationship between entities. This network structure does not do enumeration or boundary prediction, but uses two adjacent entities in the upper layer to do prediction of the entity relationship between them to reduce the possibility of sub-sequences.

NER methods based on transfer learning. Due to NER models based on deep learning usually require large-scale labeled data to better train the model. When the label data are insufficient, the deep learning model cannot fully learn the hidden features of the data, which greatly reduces the performance of the Chinese NER model based on deep learning. At the same time, the task of Chinese NER is mostly used in fields where information is specialized, and the correlation between the various fields is not large, and the portability is not high [66]. Therefore, it is difficult to graft existing label data and deep learning models into resource-

poor fields. In the face of the above problems, transfer learning [67] according to virtues of its small dependence on data and labels, relaxation of independent and identical distribution constraints became the main option to solve resource-poor NER. The NER methods based on transfer learning uses a large amount of label data and pre-trained models in the source domain to improve the learning performance of the target domain, which can focus on migrating part of the parameter or feature representation of the source domain model to the target domain model without using additional alignment information, and realizes the task of cross-domain Chinese NER migration. Sheng Jiabao et al. [35] proposed a transfer learning model that combines character vectors and word vectors. This model symmetrically converts low-resource data into high-resource data to improve the performance of the deep learning model in less annotated. Peng Dunlu et al. [68] presented a deep learning model (TL-NER) combined with transfer learning technology for limited labeled data, which can be applied to a small number of signed data and a large number of unlabeled text data to complete the task of Chinese NER. In the field of electronic medical records, Dong Xishuang et al. [69] combined the multi-task BiLSTM model with transfer learning and proposed a new transfer learning model. The model obtains the potential knowledge from the Chinese corpus in the general domain, and applies it to the task of NER for mining Chinese medical terms. The experimental evaluation results of real datasets show that this method can be used as a solution to enhance the performance of NER under limited data. With the influence of Generative Adversarial Networks (GAN) boom in recent years, the introduction of GAN into migration learning has become the pursuit of the majority of Chinese NER researchers. For the Chinese NER task with a small amount of annotation data, the Chinese word segmentation task can be used to help complete it. However, Chinese word segmentation neither retains the specific information of the word, nor exploits the word boundary information. In response to this problem, Cao Pengfei et al. [70] proposed a new adversarial transfer learning framework, which can use the shared word boundary features of the two tasks in Chinese NER and Chinese Word Segmentation and prevent loss of specific information.

Table 1. Summary of Chinese NER combined with dictionary

Work	Innovation	Network encode layer	Tag decode layer	Performance(F-score)
Zhang et al. ^[42]	Lexicon, Lattice Structure	LSTM	CRF	OntoNotes: 73.88% MSRA NER: 93.18% Weibo NER: 58.79% Chinese Resume: 94.46%
Gui et al. ^[44]	Lexicon Rethinking	CNN	CRF	OntoNotes: 74.45% MSRA NER: 93.71% Weibo NER: 59.92% Chinese Resume: 95.11%
Gui et al. ^[58]	Lexicon	GNN	CRF	OntoNotes: 74.89% MSRA NER: 93.46% Weibo NER: 60.21% Chinese Resume: 95.37%

Li et al. ^[46]	Converting Lattice into Flat Structure	Transformer	CRF	OntoNotes: 76.45% MSRA NER: 94.12% Weibo NER: 60.32% Chinese Resume: 95.45%
Liu et al. ^[59]	Self-attention	LSTM	CRF	OntoNotes: 74.43% MSRA NER: 93.36% Weibo NER: 57.51% Chinese Resume: 94.96%
Ma et al. ^[60]	SoftLexicon	LSTM	CRF	OntoNotes: 75.64% MSRA NER: 93.66% Weibo NER: 61.42% Chinese Resume: 95.53%

4. Challenges and future directions of Chinese NER

Through the description of Chapter 3, we have a detailed understanding of the Chinese NER methods. The rule-based Chinese NER methods are simple to understand, however it requires a large number of rules to be formulated manually. These methods have poor scalability and portability, which are very difficult to deal with many types of complex NER tasks. The statistics-based machine learning methods mainly use the original corpus for training. The labeling of the corpus does not require a lot of linguistic knowledge, but only needs to use the corpus of the new field for training. However, the statistics-based methods do not have the accuracy of linguistic experts, so the performance of NER is often not very good, and it needs to rely on a large number of artificially labeled sample data. Compared with the previous two methods, the Chinese NER methods based on deep learning do not require a lot of artificial features. As long as the word vectors, character vectors and dictionary features are combined, better results can be achieved. Therefore, the deep learning-based methods are more suitable for the task of Chinese NER. NER as an important sub-task in Chinese text information extraction [71-73]. It is also an important part in the field of Chinese Natural Language Processing [74-76]. It has been used in social multimedia [77-79], bio-medicine [80-82], medical treatment [83-85] and other fields. Due to the particularity of Chinese characters, some Chinese NER methods based on deep learning still have some problems. In response to these problems, many researchers proposed many improvements, but there are still some shortcomings. In this section, we propose the challenges and future directions of Chinese NER methods.

4.1 Challenges

Chinese NER methods combined with dictionary. This kind of method is a common method to enhance the effect of Chinese NER, but there are still some problems, as shown in [Table 1](#).

For the dynamic improvement method of the sequence modeling layer. In the Lattice-LSTM model proposed by Zhang yue et al, due to the particularity of the Lattice structure, the words inside the character cannot receive the information of word, and the model is only designed for LSTM, which has a problem of model migration. At the same time, if the model structure is dynamically changed according to the input needs, LSTM is a cyclic structure, which will result in a slower running speed. In the LR-CNN model proposed by Gui Tao et al, due to the introduction of the Rethinking mechanism, the model requires multiple iterations,

which reduces the running speed. In addition, the model is designed for CNN, which is difficult to model long-distance dependence, and there is also a problem of model migration. In the dictionary-based GNN model proposed by Gui Tao et al, because the model requires LSTM as the underlying encoder to encode the inductive bias of the sequence, the Chinese NER task is transformed into a node classification task, so capturing the directed acyclic Lattice structure, which makes the model structure more complicated. Li Xiaonan et al. proposed the Flat-Lattice-Transformer model, which performs better than the previous model in the four datasets, and surpasses the baseline model in performance. However, the model uses BERT for pre-trained and Transformer Model for encoding, which leads to higher complexity of the model structure.

For the improvement method in the embedding layer. In the word character vector representation model (WC-LSTM) proposed by Liu Wei et al, although the effectiveness of the model on the four datasets is more effective than previous Lattice model, the effect is still not very good when faced with recognizing new words.

Nested NER methods. These methods refer to the situation where another entity is nested in one entity, and one token corresponds to multiple entities. At present, the methods for nested recognition are still immature, and there are many challenges. The dynamic hierarchical model proposed by Ju Meizhi et al. takes full advantage of the internal information of the entity to encourage external entity recognition in an end-to-end manner. But the disadvantage of this method is that if the first-layer prediction is wrong, the probability of error transmission will be very large. At the same time, there is no parallel training, resulting in longer training time. The BILOU scheme proposed by Jana Straková et al. is indeed simple and effective, but the disadvantage is that the number of tags increases exponentially and the distribution of tags is too sparse. For these problems, many researchers proposed the idea of entity classification to solve them, but there are still many problems. In the method of Mohammad Golam Sohrab et al. listing all potential subsequences, it would be very complicated to list all the subsequences for a longer sequence, and it is necessary to consider how to reduce the complexity. In addition, there are many negative cases, so we need to consider how to reduce the negative cases. Zheng Changmeng et al. proposed a method of predicting the boundary, which reduces the complexity and the number of negative examples to a certain extent, but from the experimental results, the improvement effect is not very high. The Nested NER neural network framework was proposed by Joseph Fisher et al. Experiments show that this method has a better effect than the boundary prediction method, but the improvement effect is not obvious. Although the idea of entity classification tried its best to solve the problem of too many negative examples and too high complexity, it still did not achieve a good effect, and the calculation method was very complicated. Meanwhile the length of the sub-sequence is limited by a manually set threshold, so some of them cannot be captured if they exceed the length.

NER methods based on transfer learning. Although many researchers made many explorations in this field, there are still many problems in the current research progress. Sheng Jiabao et al. proposed a transfer learning model that combines character and word vectors. This model uses a combination of characters and words to find the maximum path of Chinese words. The experimental results show that although the performance of NER tasks on low-resource datasets is improved under this model, more redundancy will be generated and the time complexity is relatively high. In the TL-NER model proposed by Peng Dunlu et al, the model has a good recognition effect in experiments, but relatively few comparative experiments have been done. In addition, it was not tested in recognized Chinese NER datasets such as Chinese Resume NER, Sighan NER and Weibo NER, and the conclusions obtained are relatively thin. The adversarial transfer learning framework based on Chinese NER

proposed by Cao Pengfei et al. improved from the experimental results, but relatively few comparative experiments have been done. Besides, this framework was not tested in more Chinese NER datasets. The current field migration of Chinese NER has been changed from many aspects, but further optimization and exploration are needed.

4.2 Future direction

Chinese Nested NER. According to the survey of a large number of Chinese text information, we find that the probability of nested entities appearing in real text is still quite large. Regarding the methods of Nested NER, whether in Chinese or English, the effect is not pretty good, and the F1 value on each dataset does not exceed 80%. In response to this problem, most of the current domestic and foreign researches are still at a relatively preliminary stage. In the future, dealing with Chinese Nested NER, we can consider as much as possible to use the information of the internal and external entities of the nested entity to obtain more fine-grained semantic information from the underlying text to achieve a deeper text understanding.

Resource-scarce Chinese NER. Methods such as transfer learning, adversarial learning and other methods are fully utilized to solve the problems of NER in resource-poor fields and reduce the workload of manual labeling, which are also the focus of recent research. Although transfer learning as a potential knowledge transfer and data expansion solution, can perform feature conversion between different datasets to improve the performance of cross-application Chinese NER, but this feature conversion is also conditional. Transfer learning made changes to the field transfer of Chinese NER from many aspects, however, it is a big challenge and further optimization and exploration are needed. In the future Chinese NER researches, facing with insufficient corpus scale, we can consider combining the attention mechanism, graph neural network and transfer learning technologies to solve the problem of Chinese NER with scarce resources.

Flexible and non-standardized word formation. Chinese and English are fundamentally different in basic forms. Many special nouns in English are distinguished in capital letters, while Chinese cannot be distinguished in special forms. Due to the popularity of the current network language, crawling network text information became an increasingly common means of acquiring knowledge. When performing Chinese NER in the face of this type of text information, due to the flexibility of network language and the complexity and variety of Chinese word formation, it is difficult to distinguish it by traditional deep learning models alone. In the future, we can consider to introduce transfer learning and attention mechanisms into the model to distinguish many complex and flexible new nouns.

Chinese NER combined with dictionary. Due to the particularity of Chinese, the embedding layer can not accurately express Chinese semantics by using character vectors or word vectors alone. The majority of Chinese NER researchers consider to introduce dictionary into the Chinese NER model to better obtain word boundary information and avoid the impact of Chinese word segmentation errors on Chinese NER. In the previous chapter, we described Chinese NER methods combined with dictionary, which have its own advantages and disadvantages. In the future, the two methods can be combined, combining character vectors with dictionary information at the embedding presentation layer, and combining with the dictionary at the sequence modeling layer to better obtain Chinese semantic information.

5. Conclusion

This paper reviews the research results of Chinese NER and provides a comprehensive interpretation for researchers in the field of Chinese NER. This review includes the research background of Chinese NER, sequence labeling system, evaluation metrics, model framework based on the deep learning, the research results of Chinese NER, current challenges and future study directions. First, we introduce the definition, development history, sequence labeling system and evaluation metrics of NER. Then, we divide Chinese NER approaches into rule-based approaches, statistics-based machine learning approaches, and deep learning-based approaches. Subsequently we compare and analyze the advantages and disadvantages of typical Chinese NER methods. Finally, we further compare and summarize the current challenges and future research directions in Chinese NER. We hope this paper can provide a good reference for the research of Chinese NER.

Acknowledgement

This work was supported by the Major science and technology project of Hainan Province(Grant No.ZDKJ2020012), Hainan Provincial Natural Science Foundation of China (Grant Nos. 2019RC098 and 620MS021), Key Research and Development Program of Hainan Province (Grant No.ZDYF2020040) and National Natural Science Foundation of China (NSFC) (Grant No.61762033).

References

- [1] D. Nadeau, S. Sekine “A survey of named entity recognition and classification,” *Linguisticae Investigationes*, vol. 30, no. 1, pp. 3-26, 2007. [Article\(CrossRef Link\)](#)
- [2] L. Yin, X. Meng, J. Li and J. Sun, “Relation extraction for massive news texts,” *Computers, Materials & Continua*, vol. 60, no. 1, pp. 275–286, 2019. [Article\(CrossRef Link\)](#)
- [3] X. Ma, Y. Lu, Y. Lu, Z. Pei and J. Liu, “Biomedical event extraction using a new error detection learning approach based on neural network,” *Computers, Materials & Continua*, vol. 63, no. 2, pp. 923–941, 2020. [Article\(CrossRef Link\)](#)
- [4] H. Zhou, T. Shen, X. Liu, Y. Zhang, P. Guo et al., “Survey of Knowledge Graph Approaches and Applications,” *Journal on Artificial Intelligence*, vol. 2, no. 2, pp. 89-101, 2020. [Article\(CrossRef Link\)](#)
- [5] J. Qiu, Y. Liu, Y. Chai, Y. Si, S. Su et al., “Dependency-based local attention approach to neural machine translation,” *Computers, Materials & Continua*, vol. 59, no. 2, pp. 547–562, 2019. [Article\(CrossRef Link\)](#)
- [6] Y. Sharma, S. Gupta, “Deep learning approaches for question answering system,” *Procedia computer science*, vol. 132, pp. 785-794, 2018. [Article\(CrossRef Link\)](#)
- [7] Z. Dou, X. Wang, S. Shi, and Z. Tu, “Exploiting Deep Representations for Natural Language Processing,” *Neurocomputing*, vol. 386, pp. 1-7, 2020. [Article\(CrossRef Link\)](#)
- [8] R. Grishman, B. Sundheim, “Message understanding conference-6: A brief history,” in *Proc. of the 16th International Conference on Computational Linguistics*, vol. 1, pp. 466-471, 1996. [Article\(CrossRef Link\)](#)
- [9] N. Chinchor, P. Robinson, “MUC-7 named entity task definition,” in *Proc. of the 7th Conference on Message Understanding*, vol. 29, pp. 1-21, 1997. [Article\(CrossRef Link\)](#)
- [10] G. Zhou, J. Su, “Named Entity Recognition using an HMM-based Chunk Tagger,” in *Proc. of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 473-480, July, 2002. [Article\(CrossRef Link\)](#)

- [11] M. Fresko, B. Rosenfeld, and R. Feldman, "A hybrid approach to NER by MEMM and manual rules," in *Proc. of the 14th ACM international conference on Information and knowledge management*, pp. 361-362, 2005. [Article\(CrossRef Link\)](#)
- [12] E. Sang, F. Meulder, "Introduction to the conll-2003 shared task: Language-independent named entity recognition," in *Proc. of NAACL-HLT*, pp. 142-147, 2003. [Article\(CrossRef Link\)](#)
- [13] M. Li, F. Kong, "Combined self-attention mechanism for named entity recognition in social media," *Journal of Tsinghua University(Science and Technology)*, vol. 59, no. 6, pp. 461-467, 2019. [Article\(CrossRef Link\)](#)
- [14] W. Li, W. Song, X. Jia, J. Yang, Q. Wang, Y. Lei, K. Huang, J. Li, and T. Yang, "Drug Specification Named Entity Recognition Base on BiLSTM-CRF Model," in *Proc. of 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*. *IEEE*, vol. 2, pp. 429-433, 2019. [Article\(CrossRef Link\)](#)
- [15] J. Zhou, B. Li, and G. Cheng, "Automatically building large-scale named entity recognition corpora from Chinese Wikipedia," *Frontier of Information and Electronic Engineering*, vol. 16, no. 11, pp. 940-956, 2015. [Article\(CrossRef Link\)](#)
- [16] J. Li, A. Sun, J. Han, and C. Li, "A Survey on Deep Learning for Named Entity Recognition," *IEEE Transactions on Knowledge and Data Engineering*, 2020. [Article\(CrossRef Link\)](#)
- [17] W. Wei, Z. Wang, X. Mao, G. Zhou, P. Zhou, and S. Jiang, "Position-aware self-attention based neural sequence labeling," *Pattern Recognition*, vol. 110, 2020. [Article\(CrossRef Link\)](#)
- [18] S. Pradhan, A. Moschitti, N. Xue, O. Uryupina, and Y. Zhang, "CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes," in *Proc. of Joint Conference on EMNLP and CoNLL-Shared Task*, pp. 1-40, 2012. [Article\(CrossRef Link\)](#)
- [19] E. Yan, Y. Ding, S. Milojević, and C. Sugimoto, "Topics in dynamic research communities: An exploratory study for the field of information retrieval," *Journal of Informetrics*, vol. 6, no. 1, pp. 140-153, 2012. [Article\(CrossRef Link\)](#)
- [20] S. Zhang, N. Elhadad, "Unsupervised biomedical named entity recognition: Experiments with clinical and biological texts," *Journal of biomedical informatics*, vol. 46, no. 6, pp. 1088-1098, 2013. [Article\(CrossRef Link\)](#)
- [21] Y. Wen, C. Fan, G. Chen, X. Chen, and M. Chen, "A Survey on Named Entity Recognition," *Communications, Signal Processing, and Systems*, pp. 1803-1810, 2020. [Article\(CrossRef Link\)](#)
- [22] L. Wang, W. Li, and C. Chang, "Recognizing unregistered names for Mandarin word identification," in *Proc. of International Conference on Computation Linguistics*, vol. 4, pp. 1239-1243, 1992. [Article\(CrossRef Link\)](#)
- [23] Y. Zhang, T. Zhang, "Research on Me-based Chinese NER model," in *Proc. of 2008 International Conference on Machine Learning and Cybernetics*, *IEEE*, vol. 5, pp. 2597-2602, 2008. [Article\(CrossRef Link\)](#)
- [24] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. of 18th International Conference on Machine Learning*, 2001. [Article\(CrossRef Link\)](#)
- [25] Z. Xu, X. Qian, Y. Zhang, Y. Zhou, "CRF-based Hybrid Model for Word Segmentation, NER and even POS Tagging," in *Proc. of Sixth SIGHAN Workshop on Chinese Language Processing*, pp. 167-170, 2008. [Article\(CrossRef Link\)](#)
- [26] H. Hu, H. Zhang, "Chinese Named Entity Recognition with CRFs: Two Levels," in *Proc. of International Conference on Computational Intelligence & Security*. *IEEE Computer Society*, vol. 2, pp. 1-6, 2008. [Article\(CrossRef Link\)](#)
- [27] J. Cheng, Y. Yang, X. Tang, N. Xiong, Y. Zhang and F. Lei, "Generative Adversarial Networks: A Literature Review," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 12, pp. 4625-4647, 2020. [Article\(CrossRef Link\)](#)
- [28] X. Zhang, S. Zhou, J. Fang and Y. Ni, "Pattern recognition of construction bidding system based on image processing," *Computer Systems Science and Engineering*, vol. 35, no.4, pp. 247-256, 2020. [Article\(CrossRef Link\)](#)

- [29] S. Xu, D. Qu, and X. Long, "An Adaptation Method in Noise Mismatch Conditions for DNN-based Speech Enhancement," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 10, pp. 4930-4951, 2018. [Article\(CrossRef Link\)](#)
- [30] Y. Lin, H. Lei, J. Wu, X. Li, "An Empirical Study on Sentiment Classification of Chinese Review using Word Embedding," *Computer Science*, 2015. [Article\(CrossRef Link\)](#)
- [31] S. Hwang, J. Hong, and Y. Nam, "Towards Effective Entity Extraction of Scientific Documents using Discriminative Linguistic Features," *KSII Transactions on Internet & Information Systems*, vol. 13, no. 3, pp. 1639-1658, 2019. [Article\(CrossRef Link\)](#)
- [32] Z. Dai, X. Wang, P. Ni, Y. Li, G. Li, and X. Bai, "Named Entity Recognition Using BERT BiLSTM CRF for Chinese Electronic Health Records," in *Proc. of 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, IEEE, pp. 1-5, 2019. [Article\(CrossRef Link\)](#)
- [33] Y. Shen, H. Yun, Z. Lipton, Y. Kronrod, and A. Anandkumar, "Deep Active Learning for Named Entity Recognition," in *Proc. of the 2nd Workshop on Representation Learning for NLP*, pp. 252-256, 2017. [Article\(CrossRef Link\)](#)
- [34] R. Zhang, W. Lu, S. Wang, X Peng, R. Yu, and Y. Gao, "Chinese clinical named entity recognition based on stacked neural network," *Concurrency and Computation Practice and Experience*, 2020. [Article\(CrossRef Link\)](#)
- [35] A. Akbik, D. Blythe, and R. Vollgraf, "Contextual string embeddings for sequence labeling," in *Proc. of the 27th international conference on computational linguistics*, pp. 1638-1649, 2018. [Article\(CrossRef Link\)](#)
- [36] R. Yin, Q. Wang, R. Li, P. Li, B. Wang, "Multi-Granularity Chinese Word Embedding," in *Proc. of Conference on Empirical Methods in Natural Language Processing*. pp.981-986, 2016. [Article\(CrossRef Link\)](#)
- [37] Y. Jin, J. Xie, W. Guo, C. Luo, D. Wu, and R. Wang, "LSTM-CRF Neural Network with Gated Self Attention for Chinese NER," *IEEE Access*, vol. 7, pp. 136694-136703, 2019. [Article\(CrossRef Link\)](#)
- [38] J. Zhu, P. Ni, Y. Li, J. Peng, Z. Dai, G. Li, and X. Bai, "An Word2vec based on Chinese Medical Knowledge," in *Proc. of 2019 IEEE International Conference on Big Data*. IEEE, pp. 6263-6265, 2019. [Article\(CrossRef Link\)](#)
- [39] Q. Zhao, D. Wang, J. Li, and F. Akhtar, "Exploiting the concept level feature for enhanced name entity recognition in Chinese EMRs," *The Journal of Supercomputing*, vol. 76, no. 8, pp. 6399-6420, 2020. [Article\(CrossRef Link\)](#)
- [40] N. Ye, X. Qin, L. Dong, X. Zhang, and K. Sun, "Chinese Named Entity Recognition Based on Character-Word Vector Fusion," *Wireless Communications and Mobile Computing*, vol. 2020, no. 3, pp. 1-7, 2020. [Article\(CrossRef Link\)](#)
- [41] N. Zhang, F. Li, G. Xu, W. Zhang, and H. Yu, "Chinese NER Using Dynamic Meta-Embeddings," *IEEE Access*, vol. 7, pp. 64450-64459, 2019. [Article\(CrossRef Link\)](#)
- [42] Y. Zhang, J. Yang, "Chinese NER Using Lattice LSTM," in *Proc. of ACL*, vol. 1, pp. 1554-1564, 2018. [Article\(CrossRef Link\)](#)
- [43] S. Zhao, Z. Cai, H. Chen, Y. Wang, F. Liu, and A. Liu, "Adversarial training based lattice LSTM for Chinese clinical named entity recognition," *Journal of Biomedical Informatics*, vol. 99, 2019. [Article\(CrossRef Link\)](#)
- [44] T. Gui, R. Ma, Q. Zhang, L. Zhao, Y. Jiang, and X. Huang, "CNN-Based Chinese NER with Lexicon Rethinking," in *Proc. of IJCAI*, pp. 4982-4988, 2019. [Article\(CrossRef Link\)](#)
- [45] Y. Zhu, G. Wang, and B. Karlsson, "CAN-NER: Convolutional Attention Network for Chinese Named Entity Recognition," in *Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 3384-3393, 2019. [Article\(CrossRef Link\)](#)
- [46] X. Li, H. Yan, X. Qiu, and X. Huang, "FLAT: Chinese NER Using Flat-Lattice Transformer," in *Proc. of ACL*, pp. 6838-6842, 2020. [Article\(CrossRef Link\)](#)
- [47] M. Xue, B. Yu, T. Liu, Y. Zhang, E. Meng, and B. Wang, "Porous Lattice-based Transformer Encoder for Chinese NER," in *Proc. of ACL*, 2019. [Article\(CrossRef Link\)](#)

- [48] J. Li, S. Zhao, J. Yang, Z. Huang, B. Liu, S. Chen, H. Pan, and Q. Wang, "WCP-RNN: a novel RNN-based approach for Bio-NER in Chinese EMRs," *Journal of Supercomputing*, vol. 76, no. 3, pp. 1450-1467, 2020. [Article\(CrossRef Link\)](#)
- [49] Y. Yang, M. Zhang, W. Chen, W. Zhang, H. Wang, and Min Zhang, "Adversarial Learning for Chinese NER from Crowd Annotations," in *Proc. of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018. [Article\(CrossRef Link\)](#)
- [50] C. Wang, W. Chen, and B. Xu, "Named Entity Recognition with Gated Convolutional Neural Networks," *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Springer*, vol. 10565, pp. 110-121, 2017. [Article\(CrossRef Link\)](#)
- [51] H. Chen, Z. Lin, G. Ding, J. Lou, Y. Zhang, and B. Karlsson, "GRN: Gated relation network to enhance convolutional neural network for named entity recognition," in *Proc. of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 6236-6243, 2019. [Article\(CrossRef Link\)](#)
- [52] X. Guo, P. Luo, T. Wang, and W. Wang, "Chinese Named Entity Recognition based on Transformer Encoder and BiLSTM," *Design Engineering*, pp. 68-80, 2020. [Article\(CrossRef Link\)](#)
- [53] Z. Hao, H. Wang, R. Cai, and W. Wen, "Product named entity recognition for Chinese query questions based on a skip-chain CRF model," *Neural Computing and Applications*, vol. 23, no. 2, pp. 371-379, 2013. [Article\(CrossRef Link\)](#)
- [54] M. Zhu, D. Zhen, "Chinese Named Entity Recognition for Clothing Knowledge Graph Construction," in *Proc. of IOP Conference Series: Materials Science and Engineering*, vol. 646, no. 1, pp. 012043-012049, 2019. [Article\(CrossRef Link\)](#)
- [55] Z. Zhao, Z. Chen, J. Liu, Y. Huang, X. Gao, F. Di, L. Li, and X. Ji, "Chinese named entity recognition in power domain based on Bi-LSTM-CRF," in *Proc. of the 2nd International Conference on Artificial Intelligence and Pattern Recognition*, pp. 176-180, 2019. [Article\(CrossRef Link\)](#)
- [56] C. Dong, H. Wu, J. Zhang, and C. Zong, "Multichannel LSTM-CRF for Named Entity Recognition in Chinese Social Media," *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Springer, Cham*, vol. 10565, pp. 197-208, 2017. [Article\(CrossRef Link\)](#)
- [57] K. Liu, Q. Hu, J. Liu, C. Xing, "Named Entity Recognition in Chinese Electronic Medical Records Based on CRF," in *Proc. of 2017 14th Web Information Systems and Applications Conference (WISA)*, pp. 105-110, 2017. [Article\(CrossRef Link\)](#)
- [58] T. Gui, Y. Zou, Q. Zhang, M. Peng, J. Fu, Z. Wei, and X. Huang, "A Lexicon-Based Graph Neural Network for Chinese NER," in *Proc. of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp.1040-1050, 2019. [Article\(CrossRef Link\)](#)
- [59] W. Liu, T. Xu, Q. Xu, J. Song, and Y. Zu, "An Encoding Strategy Based Word-Character LSTM for Chinese NER," in *Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, vol. 1, pp. 2379-2389, 2019. [Article\(CrossRef Link\)](#)
- [60] M. Peng, R. Ma, Q. Zhang, and X. Huang, "Simplify the usage of lexicon in Chinese NER," in *Proc. of ACL*, pp. 5951-5960, 2020. [Article\(CrossRef Link\)](#)
- [61] M. Ju, M. Miwa, and S. Ananiadou, "A neural layered model for nested named entity recognition," in *Proc. of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, vol. 1, pp. 1446-1459, 2018. [Article\(CrossRef Link\)](#)
- [62] J. Straková, M. Straka, and J. Hajič, "Neural architectures for nested NER through linearization," in *Proc. of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 5326-5331, 2019. [Article\(CrossRef Link\)](#)
- [63] M. Sohrab, M. Miwa, "Deep exhaustive model for nested named entity recognition," in *Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 2843-2849, 2018. [Article\(CrossRef Link\)](#)
- [64] C. Zheng, Y. Cai, J. Xu, H. Leung, and G. Xu, "A boundary-aware neural model for nested named entity recognition," in *Proc. of EMNLP-IJCNLP*, pp. 357-366, 2019. [Article\(CrossRef Link\)](#)

- [65] J. Fisher, A. Vlachos, “Merge and label: A novel neural network architecture for nested NER,” in *Proc. of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 5840-5850, 2019. [Article\(CrossRef Link\)](#)
- [66] S. Chen, X. Ouyang, “Overview of Named Entity Recognition Technology,” *Radio Communications Technology*, vol. 46, no. 3, pp. 251-260, 2020.
- [67] L. Yao, H. Huang, K. Wang, S. Chen, and Q. Xiong, “Fine-Grained Mechanical Chinese Named Entity Recognition Based on ALBERT-AttBiLSTM-CRF and Transfer Learning,” *Symmetry*, vol. 12, no. 12, 2020. [Article\(CrossRef Link\)](#)
- [68] D. Peng, Y. Wang, C. Liu, and Z. Chen, “TL-NER: A Transfer Learning Model for Chinese Named Entity Recognition,” *Information Systems Frontiers*, vol. 22, no. 6, pp. 1291-1304, 2020. [Article\(CrossRef Link\)](#)
- [69] X. Dong, S. Chowdhury, L. Qian, X. Li, Y. Guan, J. Yang, and Q. Yu, “Deep learning for named entity recognition on Chinese electronic medical records: Combining deep transfer learning with multitask bi-directional LSTM RNN,” *PloS one*, vol. 14, no. 5, pp. 1-15, 2019. [Article\(CrossRef Link\)](#)
- [70] P. Cao, Y. Chen, K. Liu, J. Zhao, and S. Liu, “Adversarial Transfer Learning for Chinese Named Entity Recognition with Self-Attention Mechanism,” in *Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 182–192, 2018. [Article\(CrossRef Link\)](#)
- [71] S. Sun, Z. Dai, X. Xi, X. Shan, and B. Wang, “Power Fault Preplan Text Information Extraction Based on NLP,” in *Proc. of 2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*, IEEE, pp. 617-621, 2018. [Article\(CrossRef Link\)](#)
- [72] J. Chen, H. Hou, and J. Gao, “Inside Importance Factors of Graph-Based Keyword Extraction on Chinese Short Text,” *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 19, no. 5, pp. 1-15, 2020. [Article\(CrossRef Link\)](#)
- [73] F. Ren, S. Yuan, and F. Gao, “Extraction of Transitional Relations in Healthcare Processes from Chinese Medical Text based on Deep Learning,” in *Proc. of the 2019 4th International Conference*, pp. 56-60, 2019. [Article\(CrossRef Link\)](#)
- [74] Q. Yin, S. Wang, Y. Miao, and D. Xin, “Chinese Natural Language Processing Based on Semantic Structure Tree,” in *Proc. of International Conference on Computer Science & Applications. IEEE*, pp.130-134, 2015. [Article\(CrossRef Link\)](#)
- [75] W. Jiang, Y. Wang, and Y. Tang, “SDTCNs: A Symmetric Double Temporal Convolutional Network for Chinese NER,” in *Proc. of Wireless Algorithms, Systems, and Applications, 15th International Conference, WASA 2020, Qingdao, China, September 13–15, 2020, Proceedings, Part I*, pp.194-205, 2020. [Article\(CrossRef Link\)](#)
- [76] C. Yu, S. Wang, and J. Guo, “Learning Chinese Word Segmentation Based on Bidirectional GRU-CRF and CNN Network Model,” *International journal of technology and human interaction*, vol. 15, no. 3, pp. 47-62, 2019. [Article\(CrossRef Link\)](#)
- [77] J. Xu, H. He, X. Sun, X. Ren, and S. Li, “Cross-Domain and Semisupervised Named Entity Recognition in Chinese Social Media,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 11, pp. 2142-2152, 2018. [Article\(CrossRef Link\)](#)
- [78] N. Peng, M. Dredze, “Named entity recognition for chinese social media with jointly trained embeddings,” in *Proc. of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 548-554, 2015. [Article\(CrossRef Link\)](#)
- [79] B. Wang, Y. Chai, and S. Xing, “Attention-based Recurrent Neural Model for Named Entity Recognition in Chinese Social Media,” in *Proc. of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, pp. 291-296, 2019. [Article\(CrossRef Link\)](#)
- [80] Y. Zhang, X. Wang, Z. Hou, and J. Li, “Clinical named entity recognition from Chinese electronic health records via machine learning methods,” *JMIR medical informatics*, vol. 6, no. 4, pp. 242-254, 2018. [Article\(CrossRef Link\)](#)
- [81] L. Luo, N. Li, S. Li, Z. Yang, H. Lin, “DUTIR at the CCKS-2018 Task1: A Neural Network Ensemble Approach for Chinese Clinical Named Entity Recognition,” in *Proc. of the Evaluation Tasks at the China Conference on Knowledge Graph and Semantic Computing (CCKS-Tasks 2018)*, pp.7-12, 2018. [Article\(CrossRef Link\)](#)

- [82] 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*, vol. 2016, pp. 1-8, 2016. [Article\(CrossRef Link\)](#)
- [83] Y. Chen, C. Zhou, T. Li, H. Wu, X. Zhao, K. Ye, and J. Liao, "Named entity recognition from Chinese adverse drug event reports with lexical feature based BiLSTM-CRF and tri-training," *Journal of biomedical informatics*, vol. 96, 2019. [Article\(CrossRef Link\)](#)
- [84] R. Zhang, Y. Gao, R. Yu, R. Wang, and W. Lu, "Medical Named Entity Recognition Based on Overlapping Neural Networks," *Procedia Computer Science*, vol. 174, pp. 27-31, 2020. [Article\(CrossRef Link\)](#)
- [85] Y. Li, G. Du, Y. Xiang, S. Li, L. Ma, D. Shao, X. Wang, and H. Chen, "Towards Chinese Clinical Named Entity Recognition by Dynamic embedding using Domain-specific knowledge," *Journal of Biomedical Informatics*, vol. 106, 2020. [Article\(CrossRef Link\)](#)



Jieren Cheng received the Ph.D. degree from School of Computer, National University of Defense Technology in 2010. Now he is a professor and Ph.D supervisor at Hainan University, and the member of CCF. His research interests include cloud computing, artificial intelligence, network security and intelligent transportation, etc.



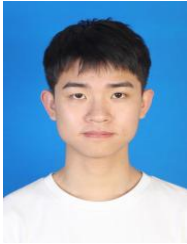
Jingxin Liu received the B.S. degree in computer science and technology from Bengbu College, Bengbu, China, in 2019. He is currently working toward the M.S. degree in software engineering at Hainan University, Hainan, China. His research interests focus on named entity recognition and deep learning.



Xinbin Xu received the B.E. degree in Hunan University of Science and Technology, Hunan, China, in 2020. He is currently working toward the M.S. degree in cyberspace security at Hainan University, Hainan, China. His research interests focus on blockchain consensus mechanism and machine learning.



Dongwan Xia received the B.E. degree in Henan University of Economics and Law, Zhengzhou, China, in 2019. She is currently working toward the M.S. degree in Hainan University, Haikou, China. Her research interests focus on natural language processing and deep learning.



Le Liu received the B.S. degree in plant protection from Hainan University, Haikou, China, in 2019. He is currently working toward the M.S. degree in computer science and technology at Hainan University, Hainan, China. His research interests focus on transfer learning.



Victor S. Sheng received the Ph.D. degree in computer science from the University of Western Ontario. He is an Assistant Professor with the Department of Computer Science, University of Central Arkansas, Conway, AR, USA, and the founding Director of the Data Analytics Laboratory. His research interests include data mining, machine learning, and related applications.