

MSFM: Multi-view Semantic Feature Fusion Model for Chinese Named Entity Recognition

Jingxin Liu¹, Jieren Cheng^{1,2,*}, Xin Peng³, Zeli Zhao³, Xiangyan Tang¹, and Victor S. Sheng⁴

¹ School of Computer Science and Technology, Hainan University, Haikou, 570228, China

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

³ School of Cyberspace Security, Hainan University, Haikou, 570228, China

⁴ Department of Computer Science Texas Tech University TX, 79409, American

*Corresponding Author: Jieren Cheng

*Received March 25, 2022; revised May 3, 2022; accepted May 23, 2022;
published June 30, 2022*

Abstract

Named entity recognition (NER) is an important basic task in the field of Natural Language Processing (NLP). Recently deep learning approaches by extracting word segmentation or character features have been proved to be effective for Chinese Named Entity Recognition (CNER). However, since this method of extracting features only focuses on extracting some of the features, it lacks textual information mining from multiple perspectives and dimensions, resulting in the model not being able to fully capture semantic features. To tackle this problem, we propose a novel Multi-view Semantic Feature Fusion Model (MSFM). The proposed model mainly consists of two core components, that is, Multi-view Semantic Feature Fusion Embedding Module (MFEM) and Multi-head Self-Attention Mechanism Module (MSAM). Specifically, the MFEM extracts character features, word boundary features, radical features, and pinyin features of Chinese characters. The acquired font shape, font sound, and font meaning features are fused to enhance the semantic information of Chinese characters with different granularities. Moreover, the MSAM is used to capture the dependencies between characters in a multi-dimensional subspace to better understand the semantic features of the context. Extensive experimental results on four benchmark datasets show that our method improves the overall performance of the CNER model.

Keywords: Chinese Named Entity Recognition, Multi-head Self-Attention Mechanism, Multi-view Feature Fusion Embedding, Natural Language Processing, Semantic Feature Fusion.

1. Introduction

Named Entity Recognition (NER) is used to extract structured information from unstructured text, such as Person, Location, Organization, etc [1]. NER plays an essential role in many downstream tasks of NLP, including information retrieval (IE) [2, 3], question answering (QA) [4-6], entity-relationship extraction [7-9], knowledge graph construction [10, 11], etc. In recent years, with the rapid development of artificial intelligence [12-15], CNER technology [16] has been widely used in general fields, such as network security field [17], social media field [18], medical field [19], etc. There are many methods to improve the performance of CNER from the perspective of feature extraction [20-24]. One of the classic methods is the Lattice-LSTM model proposed by Zhang yue et al. [22], which pioneered the addition of vocabulary information to the Chinese NER model based on character features. In order to solve the problem that the Lattice-LSTM model cannot parallelize batch processing of data, Liu Wei et al. [23] added vocabulary information to the beginning and ending characters of a word in the character embedding layer. To obtain the word boundary information of words to reduce the impact of word segmentation errors on entity recognition. Ma Ruotian et al. [24] also proposed a simple method in the embedding layer to embed vocabulary information into characters, which only needs to make fine adjustments to the character embedding layer to embed characters into vocabulary information, so as to achieve the effect of character-word information combination.

The above innovations of these methods are based on character features, adding more accurate vocabulary information, and enhancing the semantic understanding of the model from the perspective of character and vocabulary features. However, these methods still lack the multi-directional information acquisition of the text. The most fine-grained Chinese is a character. The above methods only focus on the understanding of the character meaning in the feature extraction of the characters, and ignore the strengthening effect of the shape and sound on the understanding of the character meaning. Moreover, none of these methods dig deep into the inter-character relationship of the sentences in the sentence level.

This paper proposes a Multi-view Semantic Feature Fusion Model (MSFM) for Chinese Named Entity Recognition. The core of the model is a Multi-view Semantic Feature Fusion Embedding Module (MFEM) and a Multi-head Self-Attention Mechanism Module (MSAM). In the embedding presentation layer, a MFEM is constructed to fuse character features, word boundary features, radical features, and pinyin features. Capturing character-based information from multiple aspects effectively solves the problem that most CNER models only consider partial features of word or character to capture sentence information while ignoring the integration of sentence semantic information from multiple perspectives. The MSAM is embedded between the network coding layer and the label decoding layer to capture the contextual information contained in the text itself from subspaces of different dimensions to better understand and mine the semantic features contained in the text itself. Experiments on four public datasets demonstrate that the MSFM can improve the overall performance of the CNER model. The most highlighted point is that our method can facilitate the CNER model to enrich the semantic features at the character level and capture the relationship between sentences at the sentence level, so as to avoid the CNER model from affecting the overall recognition performance of the model due to insufficient features.

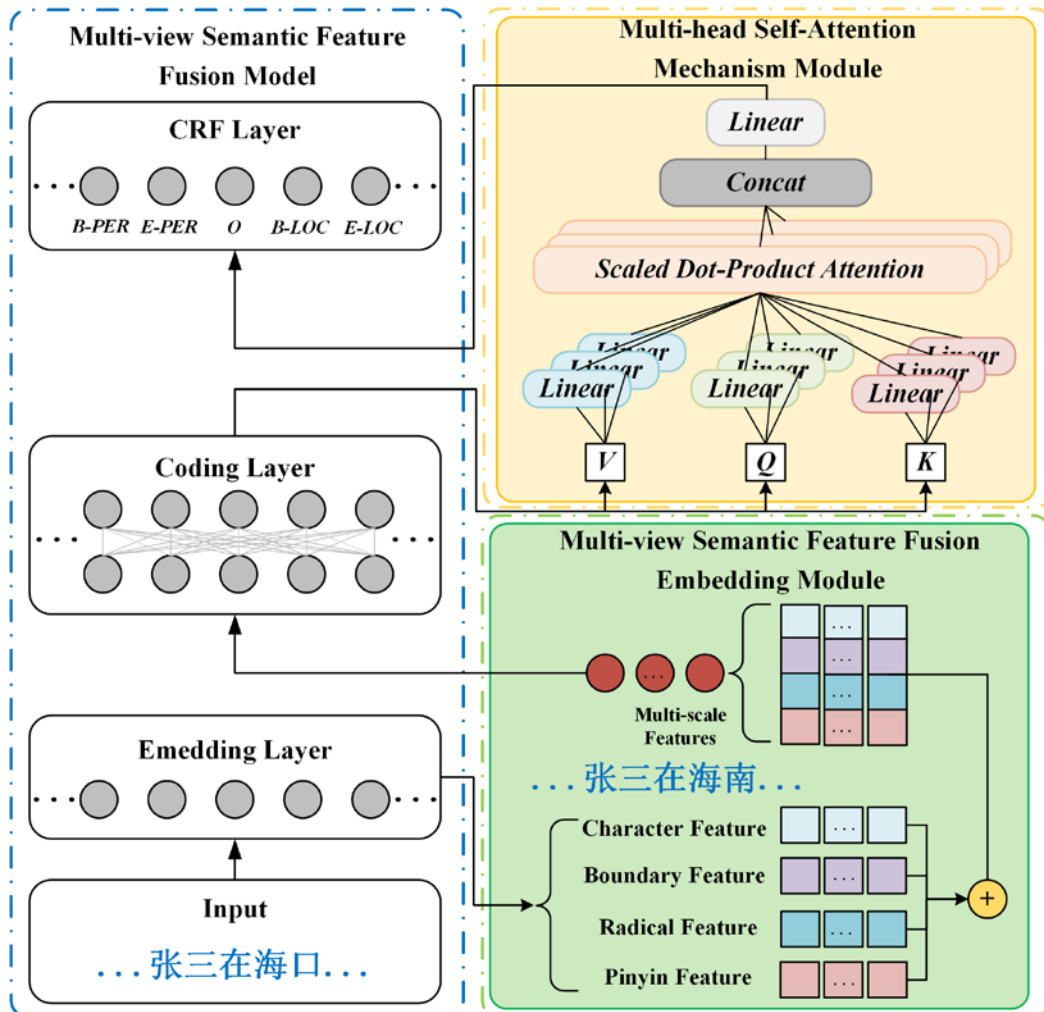


Fig. 1. The overall architecture of Multi-view Semantic Feature Fusion Model for Chinese Named Entity Recognition

The main contributions can be summarized as follows:

- We propose a Multi-view Semantic Feature Fusion Model for Chinese Named Entity Recognition to fully extract the features contained in the text from multiple angles and multiple dimensions to improve the overall recognition performance of the model.
- For the constructed Multi-view Semantic Feature Fusion Embedding Module, we integrate the semantic information of characters from multiple angles to compensate for the insufficient acquisition of semantic features and other issues to improve the performance of model recognition.
- For the embedded Multi-head Self-Attention Mechanism Module, we capture the global dependency of the entire sentence in a multi-dimensional subspace to better understand and extract the semantic features of the text itself, and ultimately improve the recognition performance of the model.
- Experimental results show that our method improves the overall performance of the CNER model, and effectively enhances the ability to capture the characteristics of Chinese text from multiple angles and dimensions.

2. Model

This paper proposes a MSFM for CNER. The overall structure of the model is shown in [Fig. 1](#). The model is mainly divided into four modules, including the multi-view semantic feature fusion embedding module, BiLSTM network coding module, MSAM, and Conditional Random Fields (CRF) tag decoding module. Thereinto in the MFEM, aimed at better understanding the semantic information of the sentence, character features will be integrated from multiple perspectives. In the BiLSTM network coding module, the received multi-view semantic feature fusion vector is subjected to context coding. In the MSAM, the contextual information of the text is captured in a multi-dimensional subspace to fully extract the semantic information contained in the text itself. Finally, in the CRF label decoding module, the label of each character is predicted.

2.1 Multi-view Character Embedding Fusion Module

The construction of MFEM plays a vital role in CNER. The finest granularity in Chinese sentences are Chinese characters. In Chinese, the finest granularity in a sentence is character. Each Chinese character is not completely equal to an independent character. Because the position in each sentence affects its important meaning. Therefore, each Chinese character needs to provide its multi-directional information for the Chinese sentence in which it is located, to ensure the multi-angle understanding of the semantic features of the depth model. The existing embedding methods for CNER are generally divided into three types. The first embedding method is to start from the perspective of a single character [\[25, 26\]](#), mapping each Chinese character in the sentence to the semantic space to form a character vector, and then replace it with a pre-trained word vector. However, its drawback is that the semantic relationship between adjacent Chinese characters is not used, and the word boundary information is ignored. The second embedding method is applied from the perspective of words [\[27\]](#); each sentence is divided into multiple words and mapped to the semantic space to form a word vector, and then replaced with the trained word vector. However, the embedding method is limited to the accuracy of word segmentation, because Chinese cannot be segmented naturally. The third embedding method is the character-word combination [\[28, 29\]](#), which avoids ignoring word boundary information and limiting to the accuracy of word segmentation. However, this method also has less angles of capturing semantic features, and almost fails to interpret Chinese sentences in multiple dimensions.

In order to enable the model to integrate the semantic information of sentences from multiple perspectives, we build a MSFM, as shown in [Fig. 2](#). This module is composed of character features, word boundary features, radical features and pinyin features, which correspond to four parts of information: meaning, word boundary, form, and sound. We define a Chinese sentence S_i whose length is n , and divide each Chinese character into $C_1, C_2, C_3, \dots, C_n$. Among them, each Chinese character C_j is mapped to an embedded vector E_j containing four parts of information, in order to better express the semantic features of the sentence from multiple angles. The fusion process of MFEM is shown in Algorithm 1.

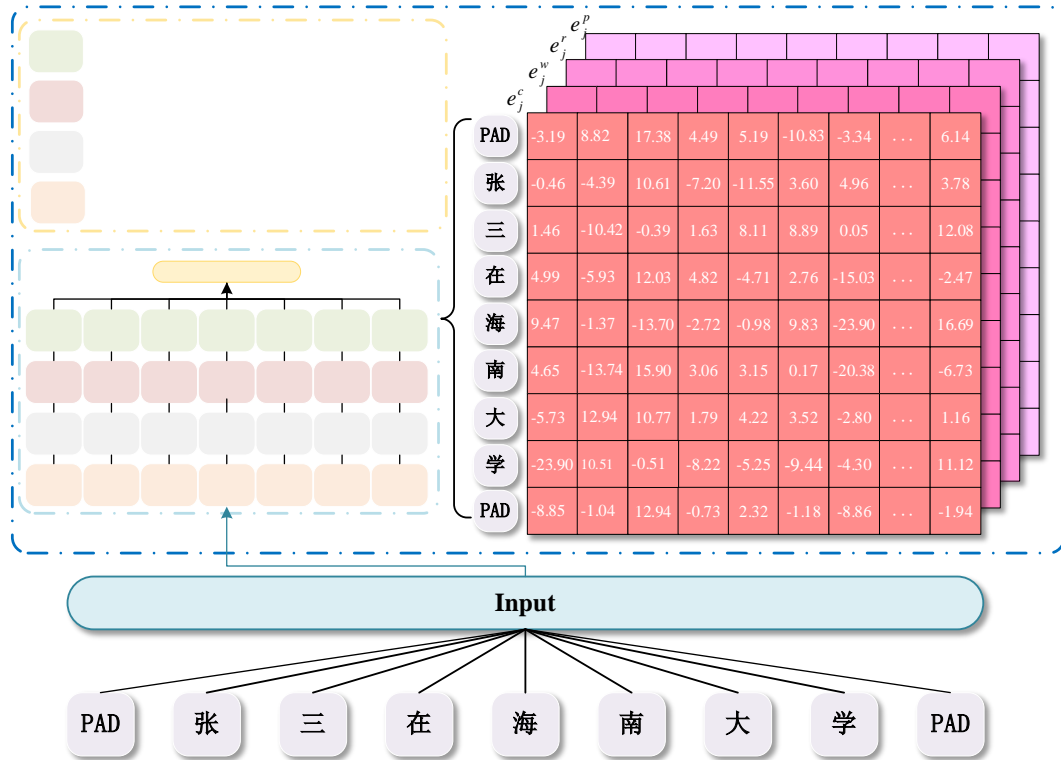


Fig. 2. Multi-view Character Embedding Fusion Module(MCFM)

2.1.1 Character feature embedding

The pre-trained model has been widely used in the field of CNER [30]. We use the classic Word2vec model to train the characters of the Chinese corpus, Gigaword, to form a corresponding set of 50-dimensional character vectors. Then we replace the character vector in the annotation dataset with the 50-dimensional one obtained through being pre-trained, and form the final character vector representation to improve the performance of the deep learning model.

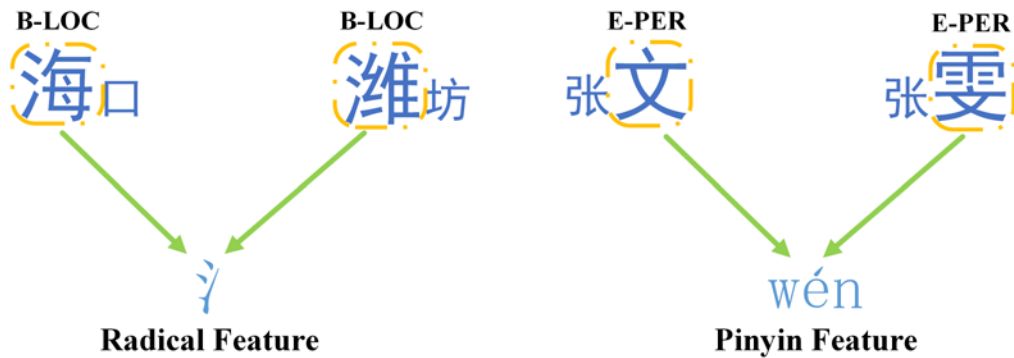


Fig. 3. The relevance of Radical Feature and Pinyin Feature

2.1.2 Word boundary feature embedding

Only use of embedding the character feature will lack the semantic relationship between adjacent characters. Therefore, Peng and Dredze et al. [31] proposed a character feature-based CNER model that we can use soft features of characters to improve the recognition performance of the model. We embed word boundary features in the construction of the MFEM, such as formula (1). e_j^w representing the segmented word boundary label. In the text, we use Jieba word segmentation to attach the word boundary information in a sentence to the characters, in order to form word boundary features, which are used to obtain the connection between adjacent characters.

$$X_j^c = [e_j^c; e_j^w] \quad (1)$$

2.1.3 Radical feature embedding

In terms of the text structure, there is a huge difference between Chinese and English. English is alphabetic writing, while Chinese is hieroglyphic one. Evolving from primitive pictographs, hieroglyphs are essentially texts that depict the shape of objects. Radical refers to any part of a character, including the upper, lower, left, right, inner, and outer parts of it. Radical facilitates us to deeply understand the most basic meaning of Chinese characters. As shown in Fig. 3, if you look at the two words "海口(Hai kou)" and "潍坊(Wei fang)" separately, although they both have "氵", they are not directly related to water, but locations. The label of "海(Hai)" in "海口(Hai kou)" is "B-LOC", and its radical is extracted as "氵". The radical of "潍(Wei)" in "潍坊(Weifang)" is "氵", and its label is also "B-LOC". Among these two characters, one is "left-right structure" and the other is "left-middle-right structure". The characters are distinct in form and structure, but the radicals are the same. Therefore, they are constructed to a certain extent. The implicit connection between them facilitates the model's understanding of semantic information. For another example, "猿猴(apes)", "狸猫(civet cat)", "豹(leopard)", "豺狼(Jackal)" and so on, these words not only have the same "犹", but also contain the meaning of animal in explaining the word. At the same time, their spatial structures are all "left and right" forms, with similar embedded structural feature components, and they all convey implicitly related information from the perspectives of word meaning, font shape, and structural form. Therefore, only using character vectors to represent Chinese characters will lack the essential characteristics of hieroglyphs. Some existing works use Chinese character images as glyph embedding, which is difficult to obtain good results due to its spatial complexity. We embed the radical features in the MFEM, such as formula (2), e_j^r representing the radical of a Chinese character, and the radical information in a sentence is attached to the character to form the radical feature, which can simply capture the glyph features from the characters themselves, in order to enhance semantic understanding.

$$X_j^c = [e_j^c; e_j^r] \quad (2)$$

Algorithm 1 MFEM Fusion Algorithm

```

1:Input: Enter each character ( $C_1, C_2, C_3, \dots, C_n$ ) in each sentence  $S_i$ .
2:Output: Multi-view Semantic Feature Fusion Embedding Vector  $E_j^C$ .
3: Obtain Multi-view Semantic Feature Fusion Embedding Vector.
4:for  $n = 1 \rightarrow N$  do
5:   Capture the character feature in each sentence  $e_j^c$ .
6:   Capture the word boundary feature information in each sentence  $e_j^w$ .
7:   Capture the radical features in each sentence  $e_j^r$ .
8:   Capture the pinyin features in each sentence  $e_j^p$ .
9:end for
10:Generate global context Multi-view Semantic Feature Fusion Embedding  $E_j^C$  by (4).
11:return  $E_j^C$ .

```

2.1.4 Chinese Pinyin feature embedding

Chinese Pinyin is the Latinization of Chinese characters, and a special way to symbolize the pronunciation of Chinese characters. Pinyin consists of 26 letters and 4 tones. Each Chinese character corresponds to a special way of expression, the pronunciation of the Chinese character. Usually, the pronunciation of each Chinese character is composed of initials, vowels and tones. As shown in Fig. 3, "张文(Zhang wen)" and "张雯(Zhang wen)" are both personal names. The pinyin of "文(wen)" is "wén", and the entity tag is "E-PER". The pinyin of the word "雯(wen)" is "wén", and its label is also "E-PER". These two characters are quite different in terms of meaning, form and structure of the characters. But the Pinyin of the two is the same, which builds an inner connection between the two to some extent, which facilitates the model's in-depth understanding of the sound and meaning of the word. We embed pinyin features in the MFEM, such as formula (3), which represents the corresponding pronunciation of a Chinese character. The pinyin information in a sentence is attached to the character to form the pinyin feature, which can simply capture the phonetic feature from the character itself to enhance semantic understanding.

$$X_j^c = [e_j^c; e_j^p] \quad (3)$$

Through the above four feature-extracting methods, the model can capture the semantic features of sentences from multiple perspectives. However, in response to the differences in the embedding dimension of each feature, we cannot directly add various embedding features. So we make two fusion methods [32].

Concat fusion method

As shown in formula (4), the four feature vectors are directly spliced to obtain a new multi-view feature vector to prepare for BiLSTM encoding. Among them, e_j^c represents the feature vector corresponding to a certain Chinese character; e_j^w represents the word boundary feature vector corresponding to the character; e_j^r represents the radical feature vector corresponding to the character, and e_j^p represents the pinyin feature vector corresponding to the character.

E_j^C represents the multi-view semantic feature fusion embedding vector corresponding to the character.

$$E_j^C = e_j^c \oplus e_j^w \oplus e_j^r \oplus e_j^p \quad (4)$$

Concat + Linear fusion method

As shown in formula (5), after we directly splice the four feature vectors, a linear connection layer is added to further fuse the feature vectors.

$$E_j^C = \text{Linear}(e_j^c \oplus e_j^w \oplus e_j^r \oplus e_j^p) \quad (5)$$

In theory, the Concat + Linear fusion method saves a certain amount of calculation space and improves the recognition speed of the model. However, the fusion method may cause the loss of fusion information due to linear splicing, which will negatively affect the recognition of the model. In this article, we use a large number of experimental surfaces Concat fusion methods to achieve the best model performance.

2.2 LSTM encoding module

After the MFEM, the Multi-view Semantic Fusion feature vector is connected into the network coding layer to capture the dependency between adjacent characters. Common network coding layers include Recurrent Neural Network (RNN) [33], Convolutional Neural Network [20], Transformer [34] and so on. We choose the Long and Short-Term Memory Network (LSTM) [35], which is one of the variants of the RNN. The network selectively retains some important information by introducing memory cells and gate structures to solve problems such as gradient disappearance and gradient explosion. But the one-way LSTM can only encode the semantic information of the sentence one-way from front to back, while ignoring the influence of the second half of the information on the first half one. In order to integrate sentence information more comprehensively, in this paper, we adopt a single-layer BiLSTM to capture contextual information, which can obtain the long-distance dependence of the sentence. The specific representation of the hidden state of BiLSTM is in formulas (6), (7), (8), where \vec{H}_j and \overleftarrow{H}_j represent the hidden state of the forward LSTM and the reverse LSTM at position j , respectively. Connecting the two through the Concat connector help to get the hidden state H_j of the BiLSTM at the j position.

$$\vec{H}_j = \text{LSTM}(\vec{H}_{j-1}, E_j^C) \quad (6)$$

$$\overleftarrow{H}_j = \text{LSTM}(\overleftarrow{H}_{j+1}, E_j^C) \quad (7)$$

$$H_j = \text{Concat}(\vec{H}_{j-1}, \overleftarrow{H}_{j+1}) \quad (8)$$

2.3 Multi-Head Self-Attention Mechanism Module

In 2017, the Google machine translation team proposed the Transformer model, which opened the prelude to the application of the Self-Attention Mechanism to text information. Thanks to its outstanding performance in the field of machine translation [36], we find that when the Self-Attention Mechanism processes text information, it gives higher weight to the important text in the sentence, while assigning lower weight to other characters. The essence of the Self-Attention Mechanism is to allocate and select limited information. As shown in formula (9),

the dot product is used to calculate the similarity.

$$Attention(Q, K, V) = Soft\ max(\frac{QK^T}{\sqrt{d_K}})V \quad (9)$$

Among them, Q is the query matrix, K is the key matrix, $\sqrt{d_K}$ prevents the inner product of Q and K from being too large for adjustment. V is the value matrix, and $Attention(Q, K, V)$ represents the calculation result of the attention mechanism.

In this article, we embed the MSAM to make the model fully clarify the dependence between the characters in the sentence. The MSAM is calculated according to formula (10), (11). Under the premise that the parameters are not shared, Q , K , and V are mapped through the parameter matrix, and then the dot product is scaled and reduced, and the process is repeated h times. Finally, the results of each attention $Head_j$ are spliced to obtain $Multi-Head(Q, K, V)$ Capturing the contextual information of sentences in multiple sub-spaces can help to obtain semantic features in multiple dimensions.

$$Head_j = Attention(QW_j^Q, KW_j^K, VW_j^V) \quad (10)$$

$$Multi-Head(Q, K, V) = (Head_j \oplus \dots \oplus Head_h) \quad (11)$$

2.4 CRF label decoding module

The label corresponding to each Chinese character is encoded by BiLSTM at first, and then the contextual information contained in the text is captured through the multi-dimensional subspace of the MSAM, and finally leads to the decoding layer for label prediction. The output of the Softmax function is independent of each other, which can only guarantee the maximum output probability at each moment, but fail to consider the sequence of character labels. However, labeling prediction for named entity recognition needs to consider the dependency between consecutive character labels. Therefore, CRF that can be globally normalized are often used in the NER model to maximize the overall prediction probability of a sentence.

3. Experiments

In order to verify the effectiveness of our proposed method, we use four public corpora's of CNER to experiment with different models by setting different qualification conditions, and discuss the experimental results. We use the classic static word vector pre-trained model. The Word2vec model [30] is used to train the characters of the Chinese corpus Gigaword to form the corresponding character vector set. We choose the baseline model BiLSTM-CRF as the basis for the performance comparison of CNER models, and compare the performance of three CNER models. The first type is the BiLSTM-CRF model that only embedding character features and word boundary features in the embedding layer; The second type is the BiLSTM-CRF model in which the MFEM built by us is embedded in the embedding layer; The third one is to embed the MFEM we built in the embedding layer, and embed the MSAM built between the BiLSTM layer and the CRF layer. In order to control variables, we use a consistent fusion method-Concat, because different fusion methods will change the network structure of the model. As seen from Table 1, Table 2, Fig. 4, we propose achieves the best results in all four datasets, with the highest F1 values. Among them, M1 means only embedding character features + word boundary features (BiLSTM + CRF) model; M2 means character features +

word boundary features + radical features + pinyin features (MFEM + BiLSTM + CRF) model, which is, embedding MFEM for CNER model; M3 represents character feature + word boundary feature + radical feature + pinyin feature (MEFM + BiLSTM + MSAM + CRF) model, which is a MSFM for CNER.

3.1 Datasets

We select four public datasets as the experimental corpus. MSRA NER[37], People Daily NER[32], Chinese Resume NER[22], Weibo NER[38].

The two datasets, MSRA NER and People Daily NER, are mainly derived from official news reports. The use of sentence patterns and grammar are relatively formal and formal and highly official. Chinese Resume NER is mainly derived from the materials of seniors' executives resume. It mainly contains three kinds of entity information, mainly including names of person, position and company with strong logic. Weibo NER comes from social media, and it mainly includes four types of entity information: person names, place names, organization names, and geopolitical entities. Its sentence patterns and words all highlight informal expressions, in random, relatively small data set and serious over-fitting problem. Meanwhile, the proportion of entity tag data in the data set is also quite small, and the results will have some fluctuations, and the model parameters can be adjusted for multiple experiments. **Table 3** shows various information of the dataset.

Table 1. Performance of our method in MSRA and People Daily datasets

Models	A	B	MSRA NER			People Daily NER		
			Precision	Recall	F1	Precision	Recall	F1
M1			84.65	80.68	82.62	83.21	81.07	82.12
M2	√		86.07	83.58	84.81	85.48	85.48	86.17
M3	√	√	86.90	85.46	86.17	82.21	86.66	87.43

Table 2. Performance of our method in the Resume and Weibo datasets

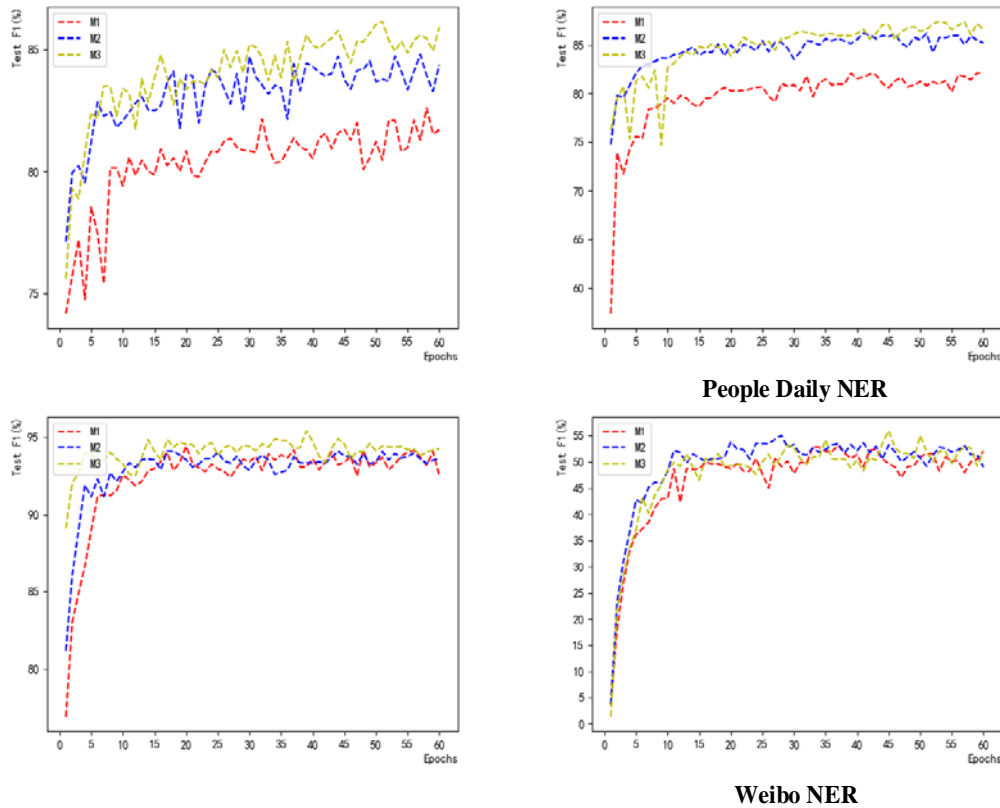
Models	A	B	Chinese Resume NER			Weibo NER		
			Precision	Recall	F1	Precision	Recall	F1
M1			94.79	93.68	94.23	57.84	49.08	53.10
M2	√		94.72	94.66	94.69	61.24	50.00	55.05
M3	√	√	96.08	94.79	95.43	60.75	51.83	55.94

3.2 Hyperparameter settings

In order to verify the overall performance of the radical features and pinyin features, we processed the pre-trained Gigaword word vector to the Chinese character vector according to Word2vec, which helps to improve the performance of the deep learning model. Display style equations should be numbered consecutively, using Arabic numbers in parentheses. We conducted multiple comparison experiments on (BiLSTM + CRF) Model, (MFEM + BiLSTM + CRF) Model and (MFEM + BiLSTM+ MSAM + CRF) Model on four public datasets. In order to control variables, we ensure that other variables and parameter settings remain consistent.

Table 3. Dataset Composition

Dataset	Sentences		
	Train set	Dev set	Test set
MSRA NER	46364	/	4365
People Daily NER	20864	/	4636
Chinese Resume NER	3821	463	477
Weibo NER	1350	270	270

**Fig. 4.** The performance of our method in the four public datasets

In terms of parameter settings, as shown in **Table 4**, in order to obtain the optimal performance of the model of the embodiment of the present application, the Tensorflow 1.13.1 framework and the CPU are both used for training. The specific model parameters are shown in **Table 2**. Character feature embedding dimension is set to 50, Word_boundary feature dimension is set to 20, Radical feature embedding dimension is set to 50; Pinyin feature embedding dimension is set to 50; Droup_keep in Dropout layer is set to 0.5; The number of layers of BiLSTM is set to 1; the Batch size is set to 12, and the learning rate is 0.002; Epochs is set to 60, and Head is set to 8. In addition, gradient truncation technology is used in the model to prevent gradient explosion and Dropout technology to prevent over-fitting. In the gradient truncation technique, set [-5,5] as the gradient truncation range, and finally use the Adam optimizer to update the gradient.

Table 4. Hyper-parameter setting

Items	B_s	Epochs	Lr	Opt	D_k	Lstm_layer	H_d	Head
Value	12	60	0.002	Adam	0.5	1	100	8

3.3 Our performance

Table 1 and **Table 2** show the model performance of MFEM and embedded MSAM on four different public datasets. It can be observed from the table that in the four public datasets, the method we propose possess the highest F1 value, which demonstrated its feasibility.

As shown in **Table 1** and **Fig. 4**, in the MSRA NER dataset, our propose method built a MFEM in 60 rounds of training to increase the F1 value of the static embedding model from 82.26% to 84.81%. Constructing a MSAM Module will boost the MFEM from 84.81% to 86.17%. As shown in **Table 1** and **Fig. 4**, in the People Daily NER dataset, the proposed method built MFEM in 60 rounds of training, And the F1 value of the static embedded model goes up from 82.12% to 86.17%. The MSAM increases the Multi-view Character Embedding Fusion Method from 86.17% to 87.43%. In the MSRA NER and People Daily NER datasets, our method has conrtributed a significant improvement in model recognition. This is because the sentences in these two data sets are mainly derived from official news reports. The sentence structure and grammatical form are more formal and rigorous, which shows that our method is more suitable for more formal language environments.

As shown in **Table 2** and **Fig. 4**, in the Chinese Resume NER dataset, in 60 rounds of training, the F1 value of the Chinese NER method based on MFEM and MSAM reaches 95.43%; The CNER method based on MFEM improves by about 0.74%. Compared to the BiLSTM-CRF model that only embeds character features and word boundary features, it improves by about 1.2%, which shows that the method we proposed has a certain improvement effect on the recognition of the model as a whole. This is because in the Chinese Resume NER corpus, most named entities have certain inherent patterns, such as company names and organization names with certain fixed suffixes, which are easier to identify.

As shown in **Table 2** and **Fig. 4**, in the Weibo NER dataset, our method in 60 rounds of training, the construction of a MFEM and a MSAM are not comparable to the static character embedding model as a whole. The improvement is obvious, the volatility of the three curves is obvious. This is mainly due to the small size of the Weibo NER dataset and fewer entity tags, which leads to the high contingency of model recognition is high. Among the F1 values of the most outstanding model performance, the F1 value of M1, M2 and M3 is 53.10%, 55.05% and 55.94% respectively. Our method still obtains the best results.

According to the results of the above four datasets, we can conclude that the construction of a MFEM and a MSAM has brought more obvious improvements to the CNER model. Most prominently, there are obvious advantages in the two more formal corpora of MSRA NER and People Daily NER, which demonstrate that in the recognition of more formal Chinese texts, capturing semantic features from multiple angles and multiple dimensions are more effective for CNER.

3.4 Result Analysis

Based on the above experimental results, the multi-view semantic feature fusion model used for CNER, only captures sentence features from partial features of words or characters, and improves the overall performance of the CNER model in the meanwhile. Our method is suitable for making up for the insufficient mining of the semantic features contained in itself.

The experimental results can prove its worth. A MFEM is built in the embedding presentation layer. Such fusions as radical features and pinyin features are embedded to integrate with each other from multi-angles, including the font, pronunciation and meaning of the character, so as to compensate for the lack of information captured by a single feature. The semantic information of sentences is integrated from multiple directions to ensure the model feature-extracting more comprehensive. Moreover, the BiLSTM is combined with the MSAM to capture the contextual information contained in the text in different multiple subspaces to better understand and mine the semantic features contained in the text itself. Our method brings more feature-capturing methods and logical association modes to the CNER model from multiple perspectives at the character level and multiple dimensions at the sentence level.

4. Conclusion

This paper proposes a novel MSFM for CNER. The model consists of two core components: MFEM and MSAM. The purpose of MFEM design is to obtain character features, radical features and pinyin features, etc., as well as enhance the meaning, shape, sound and other information of Chinese characters to a large extent, so as to solve the problem of insufficient semantic information acquisition. By introducing MSAM, the model deeply learns the internal dependencies of sentences in multiple different subspaces to capture the internal structural features of sentences. The designed MFEM and used MSAM extract features from multiple angles of characters and effectively integrate multi-dimensional sentence information to improve the recognition performance of the model. Extensive experimental results on multiple datasets show that MSFM can improve the performance of CNER tasks, which proves MSFM can effectively identify entities in complex Chinese sentences from multiple perspectives and multiple dimensions. It is of strong application value in the field of big data mining.

In future work, we hope to use the powerful capabilities of search engines to obtain more effective Chinese texts and expand the capacity of the deep learning model to learn from the corpus, so as to adapt to many non-standardized network text environments.

Acknowledgement

This work was supported by the Major science and technology project of Hainan Province(Grant No. ZDKJ2020012), Key Projects in Hainan Province (Grant No. ZDYF2021GXJS003 and No. ZDYF2020040), National Natural Science Foundation of China (Grant No. 62162024 and No. 62162022), and Graduate Innovation Project (Grant No.Qhys2021-187).

References

- [1] J. Li, A. Sun, J. Han, et al, "A survey on deep learning for named entity recognition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 50-70, Mar. 2020. [Article \(CrossRef Link\)](#)
- [2] Y. Chen, L. Xu, K. Liu, D. Zeng, and J. Zhao, "Event extraction via dynamic multi-pooling convolutional neural networks," in *Proc. of ACL—IJCNLP*, pp. 167–176, Jul. 2015. [Article \(CrossRef Link\)](#)
- [3] P. Virga, S. Khudanpur, "Transliteration of proper names in cross-lingual information retrieval," in *Proc. of ACL*, pp. 57-64, Jul. 2003. [Article \(CrossRef Link\)](#)

- [4] A. Al-Besher, K. Kumar, M. Sangeetha and T. Butsa, “Bert for conversational question answering systems using semantic similarity estimation,” *Computers, Materials & Continua*, vol. 70, no.3, pp. 4763–4780, Oct. 2022.
- [5] X. Yao, B. Van Durme, “Information extraction over structured data: Question answering with freebase,” in *Proc. of ACL*, pp. 956-966, Jun.2014. [Article \(CrossRef Link\)](#)
- [6] D. Molla, M. Van Zaanen, D. Smith, “Named entity recognition for question answering,” in *Proc. of ALTW*, pp. 51-58, Spe. 2006. [Article \(CrossRef Link\)](#)
- [7] J. Liu, L. Gao, S. Guo, et al, “A hybrid deep-learning approach for complex biochemical named entity recognition” *Knowledge-Based Systems*, vol. 221, Jun. 2021. [Article \(CrossRef Link\)](#)
- [8] R. Bunescu, R. Mooney, “A shortest path dependency kernel for relation extraction,” in *Proc. of EMNLP*, pp. 724-731, Oct. 2005. [Article \(CrossRef Link\)](#)
- [9] S. Zhao, M. Hu, Z. Cai, Z., F. Liu, “Dynamic Modeling Cross-Modal Interactions in Two-Phase Prediction for Entity-Relation Extraction,” *IEEE Transactions on Neural Networks and Learning Systems*, Aug. 2021. [Article \(CrossRef Link\)](#)
- [10] Q. He, L. Wu, Y. Yin, et al, “Knowledge-graph augmented word representations for named entity recognition,” in *Proc. of AAAI*, vol. 34(5), pp. 7919-7926, Apr. 2020. [Article \(CrossRef Link\)](#)
- [11] S. Yoo and O. Jeong, “Ep-bot: empathetic chatbot using auto-growing knowledge graph,” *Computers, Materials & Continua*, vol. 67, no.3, pp. 2807–2817, Mar. 2021. [Article \(CrossRef Link\)](#)
- [12] 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, Dec. 2020. [Article \(CrossRef Link\)](#)
- [13] F. Lei, J. Cheng, Y. Yang, X. Tang, V. Sheng and C. Huang, "Improving Heterogeneous Network Knowledge Transfer Based on the Principle of Generative Adversarial," *Electronics*, vol. 10, no. 13, pp.1525, Jun. 2021. [Article \(CrossRef Link\)](#)
- [14] X. Tang, W. Tu, K.Li and J. Cheng, “DFFNet: An IoT-Perceptive Dual Feature Fusion Network for General Real-time Semantic Segmentation,” *Information Sciences*, vol. 565, pp. 326-343, Jul. 2021. [Article \(CrossRef Link\)](#)
- [15] J. Cheng, X. Peng, X. Tang, W. Tu, W. Xu, “MIFNet: A lightweight multiscale information fusion network,” *Int J Intell Syst*, pp. 1- 26. Dec. 2021. [Article \(CrossRef Link\)](#)
- [16] J. Cheng, J. Liu, X. Xu, et al, “A review of Chinese named entity recognition,” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 15, no. 6, pp. 2012-2030, Jun. 2021. [Article \(CrossRef Link\)](#)
- [17] T. Li, Y. Hu, A. Ju and Z. Hu, “Adversarial active learning for named entity recognition in cybersecurity,” *Computers, Materials & Continua*, vol. 66, no.1, pp. 407–420, Oct. 2021. [Article \(CrossRef Link\)](#)
- [18] Y. Nie, Y. Tian, X. Wan, et al, “Named entity recognition for social media texts with semantic augmentation,” in *Proc. of EMNLP*, pp. 1383–1391, Oct. 2020. [Article \(CrossRef Link\)](#)
- [19] B. Ji, R. Liu, S. Li, et al, “A hybrid approach for named entity recognition in Chinese electronic medical record,” *BMC medical informatics and decision making*, vol. 19, no. 2, pp. 149-158, Apr. 2019. [Article \(CrossRef Link\)](#)
- [20] 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, Aug. 2019. [Article \(CrossRef Link\)](#)
- [21] T. Gui, Y. Zou, Q. Zhang, et al, “A lexicon-based graph neural network for chinese ner,” in *Proc. of EMNLP-IJCNLP*, pp. 1040-1050, Nov. 2019. [Article \(CrossRef Link\)](#)
- [22] Y. Zhang and J. Yang, “Chinese ner using lattice lstm,” in *Proc. of ACL*, pp. 1554-1564, Jul. 2018. [Article \(CrossRef Link\)](#)
- [23] W. Liu, T. Xu, Q. Xu, et al, “An Encoding Strategy Based Word-Character,” in *Proc. of NAACL*, pp. 2379-2389, Jun. 2019. [Article \(CrossRef Link\)](#)
- [24] R. Ma, M. Peng, Q. Zhang, et al, “Simplify the usage of lexicon in Chinese NER,” in *Proc. of ACL*, pp. 5951–5960, Oct. 2019. [Article \(CrossRef Link\)](#)

- [25] Q. Zhao, D. Wang, J. Li, and F. Akhtar, "Exploiting the concept level feature for enhanced name entity recognition in Chinese EMRs," *Supercomputing*, vol. 76, no. 8, pp. 6399-6420, 2020. [Article \(CrossRef Link\)](#)
- [26] 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, Sep. 2019. [Article \(CrossRef Link\)](#)
- [27] R. Yin, Q. Wang, R. Li, P. Li, B. Wang, "Multi-Granularity Chinese Word Embedding," in *Proc. of EMNLP-IJCNLP*, pp.981-986, Nov. 2016. [Article \(CrossRef Link\)](#)
- [28] S. Zhao, M. Hu, Z. Cai, Z. Zhang, T. Zhou, F. Liu. "Enhancing Chinese Character Representation With Lattice-Aligned Attention," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1-10, Oct. 2021. [Article \(CrossRef Link\)](#)
- [29] 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, Jul. 2020. [Article \(CrossRef Link\)](#)
- [30] T. Mikolov, K. Chen, G. Corrado, et al, "Efficient Estimation of Word Representations in Vector Space," in *Proc. of NIPS*, Sep. 2013. [Article \(CrossRef Link\)](#)
- [31] N. Peng, M. Dredze, "Improving Named Entity Recognition for Chinese Social Media with Word Segmentation Representation Learning," in *Proc. of ACL*, pp. 149-155, Mar. 2016. [Article \(CrossRef Link\)](#)
- [32] J. Li, K. Meng, "MFE-NER: Multi-feature Fusion Embedding for Chinese Named Entity Recognition," in *Proc. of ICLR*, Sep. 2021. [Article \(CrossRef Link\)](#)
- [33] S. Gajendran, D. Manjula, V. Sugumaran, "Character level and word level embedding with bidirectional LSTM – Dynamic recurrent neural network for biomedical named entity recognition from literature," *Journal of Biomedical Informatics*, vol. 112, no. 445, pp. 103609, Dec. 2020. [Article \(CrossRef Link\)](#)
- [34] X. Li, H. Yan, X. Qiu, and X. Huang, "FLAT: Chinese NER Using Flat-Lattice Transformer," in *Proc. of ACL*, pp. 6838–6842, May. 2020. [Article \(CrossRef Link\)](#)
- [35] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8 pp. 1735-1780, Nov. 1997. [Article \(CrossRef Link\)](#)
- [36] A. Vaswani, N. Shazeer, N. Parmar, et al, "Attention Is All You Need," in *Proc. of NIPS*, Long Beach, CA, USA, Dec. 2017. [Article \(CrossRef Link\)](#)
- [37] G. Levow, "The third international Chinese language processing bakeoff: Word segmentation and named entity recognition," in *Proc. of the Fifth SIGHAN Workshop on Chinese Language Processing*, pp. 108-117, Jul. 2006. [Article \(CrossRef Link\)](#)
- [38] N. Peng, M. Dredze, "Named Entity Recognition for Chinese Social Media with Jointly Trained Embeddings," in *Proc. of EMNLP*, pp. 548-554, Sep. 2015. [Article \(CrossRef Link\)](#)



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.



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.



Xin Peng received the B.E. degree in Southwest University of Science and Technology, Mianyang, China, in 2020. He is currently working toward the M.S. degree in Hainan University, Hainan, China. His research interests focus on data mining and deep learning.



Zeli Zhao majored in data science and big data technology at Hainan University in 2019. He is currently studying at Hainan University in Haikou, Hainan, China. His main research direction is data mining and visualization.



Xiangyan Tang received the M.S. degree from School of Computer, Hunan Agricultural University in 2011. Now she is an associate professor at Hainan University, and currently working toward the Ph.D. degree in computer science and technology at Tianjin University, Tianjin, China. Her research interests include artificial intelligence, network security and intelligent transportation, etc.



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.