

Research Article

Fusion Deep Learning and Machine Learning for Heterogeneous Military Entity Recognition

Hui Li , Lin Yu , Jie Zhang , and Ming Lyu 

School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China

Correspondence should be addressed to Jie Zhang; zhangjie_njust@njust.edu.cn

Received 21 November 2021; Revised 13 December 2021; Accepted 17 December 2021; Published 17 January 2022

Academic Editor: Liqin Shi

Copyright © 2022 Hui Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With respect to the fuzzy boundaries of military heterogeneous entities, this paper improves the entity annotation mechanism for entity with fuzzy boundaries based on related research works. This paper applies a BERT-BiLSTM-CRF model fusing deep learning and machine learning to recognize military entities, and thus, we can construct a smart military knowledge base with these entities. Furthermore, we can explore many military AI applications with the knowledge base and military Internet of Things (MIoT). To verify the performance of the model, we design multiple types of experiments. Experimental results show that the recognition performance of the model keeps improving with the increasing size of the corpus in the multidata source scenario, with the *F*-score increasing from 73.56% to 84.53%. Experimental results of cross-corpus cross-validation show that the more types of entities covered in the training corpus and the richer the representation type, the stronger the generalization ability of the trained model, in which the recall rate of the model trained with the novel random type corpus reaches 74.33% and the *F*-score reaches 76.98%. The results of the multimodel comparison experiments show that the BERT-BiLSTM-CRF model applied in this paper performs well for the recognition of military entities. The longitudinal comparison experimental results show that the *F*-score of the BERT-BiLSTM-CRF model is 18.72%, 11.24%, 9.24%, and 5.07% higher than the four models CRF, LSTM-CRF, BiLSTM-CR, and BERT-CRF, respectively. The cross-sectional comparison experimental results show that the *F*-score of the BERT-BiLSTM-CRF model improved by 6.63%, 7.95%, 3.72%, and 1.81% compared to the Lattice-LSTM-CRF, CNN-BiLSTM-CRF, BERT-BiGRU-CRF, and BERT-IDCNN-CRF models, respectively.

1. Introduction

The US military attaches great importance to the unified management of battlefield resources. In recent years, the US Department of Defense Advanced Research Projects Agency (DAPRA) has issued a number of basic research topics on the unified management of battlefield resources. The topic guides focus on the importance of building the expert knowledge base and knowledge graph for the battlefield resources [1]. They point out the research route and direction of the unified management of battlefield resources for the US military in the future. Relevant guidelines also point out that it is necessary to focus on expanding the coverage of military entity data and improving the quantity and quality of the knowledge base of military entities by combining data from different data sources like the existing data sets of the US Army, military canonical books, reliable professional websites, and

military blogs. In terms of specific applications, the US military has achieved many practical applications; the most famous case is that Palantir applies the intelligence analysis technology and knowledge graphs to assist the US Intelligence Agency in capturing Osama bin Laden and uncovering the Ponzi scheme. Currently, Palantir is working with DAPRA to conduct in-depth research on the application of knowledge graphs and intelligence analysis technology to assist in intelligence gathering, processing, and military resource control. In the area of unified resource management, the PLA Academy of Military Sciences, in conjunction with the Department of Equipment Development, has carried out a number of studies on the management of military resources. In recent years, they have published “Distributed Resource Management Methodology,” “Research on End Resource Management Technology,” and “Research on Resource Management Technology Based on Edge

Computing,” and other related topics. With these recognized military entities, many military apps about military management could be constructed, such as military resource planning, military resource scheduling, and military resource monitoring, according to the topics issued by the DAPRA and the PLA Academy of Military Sciences. Furthermore, the military resource posture could be drawn in real time with the development of the MIIoT technology.

Inspired by the topic guidelines of the US military and the PLA Academy of Military Sciences, we perform battlefield resource entities extraction works from multisource heterogeneous data. The entities, we mainly concern, are computing and storage, battlefield perception, communication networks, weapon platform, and logistics support, as well as integrated mission environment information, combatant, combat agency, combat time, combat location, and event information involved in combat resources. These military entities can be virtualized, and then, various virtual military resource services and military IoT applications can be built in cloud services [2]. With these cloud services, military staffs can carry out studies on the virtual scheduling of military resources, thus improving the utilization of military resources.

Military entity recognition technology is based on general entity recognition technology, which has experienced the development process from dictionary, rule, and machine learning to deep learning [3]. At present, the recognition model based on pretraining language model and deep learning algorithm is the main stream of entity recognition in military and general fields with the increasing computing power of small computers, the maturity of deep learning technology, and the development of pretraining language model [4]. However, it is a challenge work for the military entity recognition with the distinct domain characteristics of military field. Firstly, it is lack of solid and reliable corpus for military entity recognition [5]. Since the data stored in military information systems can only be accessed by the staff with special rights, and fewer AI researchers in military field, since that the quantity and quality of the data sets of military fields cannot be compared with open domain. Secondly, there are few types of corpuses for military entity recognition. Military entities have distinct domain characteristics that multiple terms, multiple abbreviations, multiple types of expressions, multiple nested expressions, and multiple fresh words [3]. As the few types of corpus and data sources, a few characteristics of military entities could be covered with these corpuses, and thus, the military entity recognition on cross data source could not be conducted. Last but not least, it is lack of unified criteria for entity division. A nested entity could be divided into different granularity with different worker, so we need a unified criteria to guide the practice for military entity recognition.

To solve the problems discussed above, three types, abbreviated, scientific or English name, and novel and casual, of military corpuses are constructed, and then, we can explore different research work on military entity recognition with them, such as crossing data source entity recognition. We improve the entity labeling mechanism for entity with fuzzy boundary based on the military entity recognition

works conducted by other researchers. With the constructed corpuses, three kinds of experiments have been carried out, the experiments of applying the BERT-BiLSTM-CRF model to military entity recognition for different size corpus sets, the experiments of applying the BERT-BiLSTM-CRF model to military entity recognition across corpus sets, and the experiments of comparing the performance of multimodel military entity recognition. The experimental results show that, the model used in this paper outperforms the listed models, such as CRF, LSTM-CRF, Lattice-LSTM-CRF, and other models, for military entity recognition. Furthermore, the generalization ability of the model is verified with these experiments; the experimental results provide a reliable reference for researchers to use BERT-BiLSTM-CRF model for military entity recognition.

2. Related Works

Military entity recognition has attracted much attention, and many research works have been conducted. Researchers mainly applied machine learning combined with dictionaries or rules for military recognition at earlier times.

Jiang et al. [6] have proposed a model combined CRF with rules to extract military entities from combat instruments for automated generation of combat orders, which combines external lexical features, grammatical rules of military expressions, and the rule learning capability of CRF. The model achieves an F -value of 75.48% on the corpus constructed by the authors with 300 combat instruments. Feng et al. [7] have proposed a semisupervised military entity recognition method based on CRF for identifying entity military information such as military ranks, military equipment, military facilities, and military institutions. The proposed method takes use of the basic features of military texts to construct a grammatical feature set of military texts and fuses them with the CRF model. Experimental had been carried out on the corpus consisting of combat documents, duty documents, military documents, military online news, military blogs, and military reviews with this model, and the results show that the highest F -value of entity extraction was 90.9%. Shan et al. [8] have proposed a CRF-based military entity recognition method under a small granularity strategy for the military named entities with complex internal nested relationships and inconspicuous grammatical distinctions. This model applies a small granularity strategy, combines it with a CRF model to identify small nondivisible military entities, and finally, the small granular entities are integrated to obtain complete military entities. Experimental carried out on a manually constructed annotated corpus of combat instruments, and the model achieves an F -value of 78% on the corpus.

With the continuous development of deep learning techniques and the increasing computing power of minicomputers, applications of deep learning in the field of military entity recognition are emerging. In [9], a recognition method based on deep neural network models has been studied; the method applies the word vectors and word states as features for weapon name recognition and achieves an F -value of 91.02% on a corpus set constructed with

military website data. Liu et al. [10] apply a BiLSTM-CRF model to identify weaponry equipment names, and this achieves an F -value of 93.88% on the corpus constructed from military documents. Wang et al. [11] propose a character+BiLSTM+CRF model to extract military entity from military corpus, which aims to solve the problem that the complexity of artificial construction features and the inaccuracy of military text segmentation in the traditional methods of military named entity recognition; the experimental results show that the model proposed in this paper outperforms the traditional methods. Other applications for the military entity recognition have been studied, which are based on deep learning, such as recognition models based on a combination of a self-attentive mechanism and BiLSTM-CRF model [12], military entity recognition models based on multineural network collaboration [13], and military entity recognition models based on transfer representation learning [14]. Where the multineural network-based model applies word vectors obtained from BERT pretraining as input and combines with BiLSTM-CRF model for military entity recognition, which achieves an F -value of 84.07% on the corpus constructed based on military websites and military blogs. However, the authors did not consider the sentence contextual features in the training process, which easily led to contextual coreference that could not be effectively processed. Single data source was used, and the scenario of multiple corpus intersection was not considered in the work, which could not verify the generalization ability of the mode. Furthermore, the entity type coverage was relatively small, and the effect of corpus set size on the experimental results and the effect of heterogeneous data on the experimental results were not illustrated in the experimental session. At present, entity recognition based on pretrained models and attention mechanism in the field of generic entity recognition is the mainstream [15–20], which gives important insight into the direction of entity recognition technology development in the military field.

Recognition models based on lexicons, rules, and machine learning are traditional methods, which rely on powerful feature engineering and cost a lot for large-scale applications. Entity recognition models based on pretraining and deep learning do not need to rely on the support of basic feature engineering, and they are the main research direction for future military entity recognition. Some research results for named entity recognition have been achieved with these methods, but there are also many problems, such as no standard corpus to measure the merits of the models, no experimental tests across scenes and corpus, and no comprehensive consideration of all information in the context.

3. Construction of Multisource Heterogeneous Military Corpus

3.1. Analysis of Data Sources and Data Characteristics. Multisource heterogeneous data has the characteristics of wide fields, large span, and rich information. We can take advantage of this and then construct a relatively complete set of military entities for the research work of military entity recognition.

There are two kinds of military text data; one is the nonpublic data, such as combat documents, military documents, military documents, reconnaissance intelligence, military teaching plans, simulation training task scenarios, and simulation logs; the other one is the open-source data, such as military blogs, military news, well-known military websites (such as Phoenix military, Jane’s Defense Weekly, and Hanhe defense), and arms dealer websites. These data can be divided into three categories, according to the specific forms of military entity data: (1) abbreviation type, it is commonly used in military combat documents, such as double 35, which represents double barrel 35 mm artillery; (2) scientific or English name type, many military entities recorded as the forms of scientific or English name in military books, such as 7.62 mm sniper rifle, F-35, and Su-30; (3) novel and casual type, the expression of military entities in network terms is relatively casual, and with many fresh and cool words are used, such as the network terms about j-20, including “Wei Long,” “J-00,” and “door-to-door.” In addition, the entities in these corpora have the characteristics of fuzzy boundary and multinesting.

3.2. Data Preprocessing and Corpus Construction. In this work, we mainly select four kinds of nonpublic data and three kinds of open-source data as the main sources of the experimental corpus. Nonpublic data contains representative military documents, reconnaissance intelligence, simulation training mission scenario, and military books; open-source data contains military blogs, military reviews, and well-known military websites. The data in the selected data sources covers the three types of data discussed above, which can support the experimental needs for cross data scenarios. Most of the nonpublic data is text type, which needs to be electronic, and then, the original text is segmented according to the punctuation marks of “,” “.”, “!”, “;”, and “?” and serialized into data in CSV or TSV format.

We apply crawler to access web data from military blogs, military reviews, and well-known military websites, then extract the text information by text density, and finally, the data is serialized by the proposed method above. The serialized corpus is the original unmarked corpus. We label the original corpus in a word level way [13]. In this way, three kinds of military entity recognition corpus sets, abbreviated, scientific or English name, and novel and casual, three military Entity Recognition Corpus sets, are obtained, covering 12 categories of entity objects including personnel name, military place name, time, military event, military institution, military facility, combat environment, computing and storage, battlefield perception, communication network, weapon platform, and logistics support. The size information of these corpus is shown in Table 1, and the entity information of them is shown in Table 2.

3.3. Data Labeling. There are various forms of military entities in multisource heterogeneous data; in [13], five rules have been proposed to solve the problem for recognizing military entities with fuzzy boundaries, but only partial cases can be solved for the problem with these rules. With this

TABLE 1: 3 types of corpus size information.

Corpus type	Number of sentences	Number of words
Abbreviation type	278149	12655807
Scientific or English name type	340786	14926444
Novel and casual type	686774	47044046
Total	1305709	74626298

problem, we propose some rules considering abbreviations and standardized expressions for the other cases.

Rule 1: numbers are connected with weapons or equipment (or place names, military institution names), and they can be labeled as weapon or equipment entities (place name entities, military institution entities), such as 1130 anti-aircraft gun, 591 highlands, and the 38th army

Rule 2: numbers and length units are connected with the weapon entity, and they can be marked as the weapon entity, such as a 7.62 mm sniper gun

Rule 3: adjectives are connected with weapon entities, and we label them as weapon entities, such as long endurance UAV

Rule 4: the abbreviations of English or Chinese characters of weapon entities are connected with numbers, and they can be marked as weapon entities, such as J-20 and J16

Rule 5: personal names are connected with military institutions, and they can be marked as military institutional entities, such as the “Yang Gensi company”

With these rules proposed in this work and in [13], we can solve the problem that determining the entity boundary in heterogeneous corpus.

Named entities in the military field are characterized by multiple types, multiple professional terms, and less ambiguity [10]. The BIO labeling method is suitable for labeling entities with these characteristics, which is a concise and efficient labeling mechanism. The specific labeling scheme is shown in Table 3.

4. Military Entity Recognition Model

In this work, we apply the BERT-BiLSTM-CRF model to recognize battlefield resource entity recognition from military text. This model uses the word vectors obtained by BERT pretraining as input information and integrates bidirectional LSTM (Long Short-term Memory) and CRF to identify entities from the input information. The model is divided into BERT pretraining language model layer, BiLSTM layer, and the layer, and its structure is illustrated in Figure 1.

Let D denote the military corpus sets, thus $D = \{d_1, d_2, \dots, d_n\}$, where d_i denotes the i th corpus set, $d_i = \{s_{i1}, s_{i2}, \dots, s_{im}\}$, s_{ij} denotes the sentence j in corpus i , $s_{ij} = \{x_{ij1}, x_{ij2}, \dots, x_{ijn}\}$, x_{ijk} denotes the k th word in sentence s_{ij} . At the beginning, the input unit transforms the x_{ijk} to C_k , and then, C_k is transformed to E_k by the transformer encoders; next, E_k goes through the forward LSTM units, $F = \{F_1, F_2, \dots,$

$F_n\}$, and backward LSMT units, $B = \{B_1, B_2, \dots, B_n\}$; then, we get a feature matrix P_k ; at the end, the CRF layer captures the dependencies between adjacent labels according to feature vectors and outputs corresponding labels y_k .

4.1. BERT Pretraining Model. Word vectors applied to military entity recognition can be trained from military corpus with the pretraining model of BERT. This model uses bidirectional transformer network structure to learn semantic feature information of military text context. In particular, two kinds of unsupervised pretraining tasks are developed, which inspired by the idea of cloze filling, to learn more context information from the text during the training process. One is Masked Language Model (MLM), and the other one is Next Sentence Predict (NSP), which are introduced to overcome the unidirectional problem that meets by most word vector generation models. As a result, more fully information can be extracted from the military corpus by the BERT pretraining language model which applies the organic combination of the transformer encoder structure and the unsupervised training task.

4.1.1. Input Unit. The input unit of the BERT model consists of Word Embedding, Segment Encoding, and Positional Encoding of the input sequences. Word features, sentence features, and position features of each word in the input sentence $s_{ij} = \{x_{ij1}, x_{ij2}, \dots, x_{ijn}\}$ should be calculated before they are transformed into the BERT layer. Where the word features are denoted as $(e_{ij1}^w, e_{ij2}^w, \dots, e_{ijn}^w)$, sentence features are denoted as $(e_{ij1}^s, e_{ij2}^s, \dots, e_{ijn}^s)$, and position features are denoted as $(e_{ij1}^p, e_{ij2}^p, \dots, e_{ijn}^p)$. The word feature of the word x_{ijk} is provided by the corresponding word vectors in the vocabulary trained by Google. e_{ijk}^s denotes segment number of the input sequences; e_{ijk}^s takes either 0 or 1. We apply absolute position mode for the position feature; that is, $e_{ijk}^p = k$, where k denotes the position feature of the word at the k th input position.

The word feature, sentence feature, and position feature of the corresponding position in the input sequence are added together to obtain $C_k = e_{ijk}^w + e_{ijk}^s + e_{ijk}^p$, where $C_k \in C$, and $C = (C_1, C_2, \dots, C_n)$. The input feature composition and calculation method are illustrated in Figure 2.

4.1.2. Transformer Encoder Unit. The BERT model consists of a bidirectional transformer encoder network (encoder structure is shown in Figure 3); the structure is illustrated in Figure 4. It takes sequence of $C = (C_1, C_2, \dots, C_n)$ as input, and then, the multilayer transformer encoding unit pretrains the sequence into dynamic word vectors.

The encoder unit includes self-attention network unit, feedforward neural network unit, and basic normalized network unit. The self-attention network unit is used to learn the features of input sequences. The application of the self-attention mechanism helps the BERT to get rid of the problem of long-term dependence on recurrent neural nets, and thus, it can be used to perform parallel computing [21].

TABLE 2: Information on the 3 types of corpus entities.

Entity type	Abbreviation	Scientific or English name	Novel and casual	Total
Personnel name (P)	892	1527	2096	4515
Military place name (L)	8327	4172	20384	32883
Time (T)	1477	3091	4773	9341
Military event (E)	4529	3784	5846	14159
Military establishment (G)	5297	9533	12734	27564
Military installation (F)	2175	1732	2209	6116
Operational environment (S)	1639	873	1194	3706
Computational storage (C)	792	1399	2087	4278
Battlefield perception (R)	3327	1894	2256	7477
Communication network (N)	1079	970	1768	3817
Weapon platform (W)	3096	6718	13895	23709
Logistic service (H)	882	1349	2297	4528
Total	33512	37042	71539	142093

TABLE 3: BIO labeling system for military entity.

Entity category	Entity start	Inside entity
Personnel name (P)	B-P	I-P
Military place name (L)	B-L	I-L
Time (T)	B-T	I-T
Military event (E)	B-E	I-E
Military establishment (G)	B-G	I-G
Military installation (F)	B-F	I-F
Operational environment (S)	B-S	I-S
Computational storage (C)	B-C	I-C
Battlefield perception (R)	B-R	I-R
Communication network (N)	B-N	I-N
Weapon platform (W)	B-W	I-W
Logistic service (H)	B-H	I-H

The principle, the self-attention unit used for learning sentence features, is that it enables each word in a sequence to perform attention operations on each other to capture the input features. The formula for attention calculation is given in equation (1).

$$\text{Attention}(Q, K, V) = \text{soft max} \left(\frac{QK^T}{\sqrt{d_k}} \right) V. \quad (1)$$

In which, Q is the query vector, K is the key vector, V is the value vector, and d_k is the input vector dimension.

It is known that the ability is limited that single attention unit learns input features, so the multihead attention mechanism is applied by the transformer encoder unit. Its working principle is to perform different linear mappings of Q , K , and V and then calculate their attention values, respectively, and then fuse the h attention information obtained.

Equations (2) and (3) are the calculation formulas of multiple attention.

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O, \quad (2)$$

$$\text{head}_i = \text{Attention} \left(QW_i^Q, KW_i^K, VW_i^V \right). \quad (3)$$

Word order of input sequences is not considered for the self-attention unit, and the positional coding unit helps to solve this problem. At time t , the sum of Word Embedding, Positional Encoding, and Segment Encoding is the actual input for the BERT model, which has been discussed in the input unit. We can benefit from the input that the relative position and segment encoding added to ensure that the actual input word vector is different when the same word vector appears in different positions in different sequences.

4.1.3. Unsupervised Training Tasks. Inspired by clotting, the BERT model applies two unsupervised training tasks in the pretraining stage; one is the Masked Language Model, and the other one is the Next Sentence Predict.

In the Masked Language Model task, the model will randomly “remove” 15% of the words in the input sequences and then make the model actively learn the contextual semantic relations of the input sequences from different directions. Through iterative training, the probability of reasoning to get the correct answer is as large as possible, so as to achieve the purpose of learning the text semantics. In the task of Next Sentence Predict, the model randomly selects sentence pairs from the training text, in which the positive and negative samples account for 50%, respectively. Then, the training task is carried out on the training set of sentence pairs, and the BERT model is allowed to judge their correlation, so as to learn the relations between two sentences.

4.2. BiLSTM Layer. The BERT pretraining model provides dynamic word vectors for the whole recognition system, but it is slightly insufficient in learning sentence features.

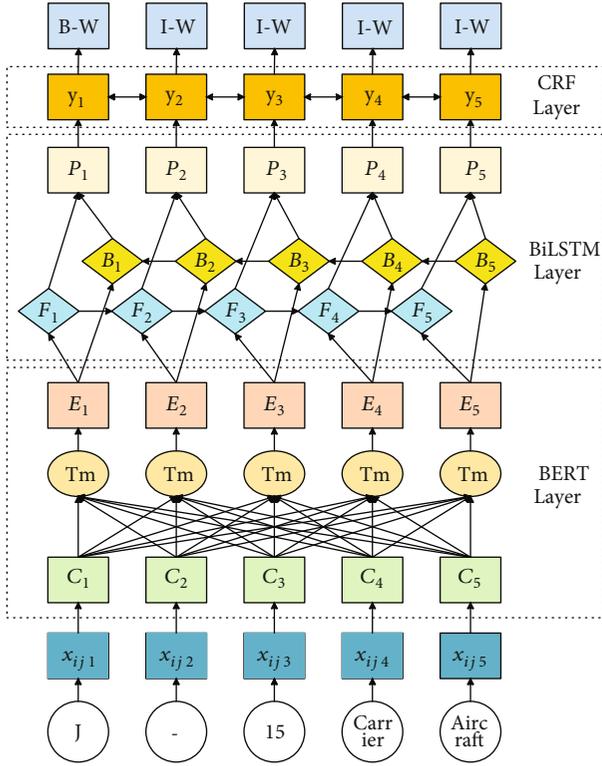


FIGURE 1: Military entity recognition model based on BERT-BiLSTM-CRF model.

Therefore, the BiLSTM layer, illustrated in Figure 2, is introduced to model sentences and learn the features of different input sequences. The input vector $E = (E_1, E_2, \dots, E_n)$ which is the output of the BERT model goes through the forward LSTM layer $F = \{F_1, F_2, \dots, F_n\}$, and the backward LSTM layer; then, it is transformed to a matrix $P \in R^{n \times m}$, where P_{kz} denotes the probability of the output label z that corresponds to the input word feature x_{ijk} .

The LSTM model, illustrated in Figure 5, adds “forgetting gate f_t ,” “input gate i_t ,” “output gate o_t ,” and “cell state C_t ” to the structure of the RNN (recurrent neural networks) model, so it is known as improved RNN model. With the added units, the problems that long distance dependence and gradient disappearance of the recurrent neural network could be solved. Meanwhile, the units can adjust the memory function of the network help, which helps to maintain and update the state of the whole network.

The calculation formulas corresponding to the LSTM network units are shown as follows.

$$\begin{aligned}
 f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\
 i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
 \tilde{C}_t &= \tan h(W_c \cdot [h_{t-1}, x_t] + b_c), \\
 o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\
 C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t, \\
 h_t &= o_t * \tan h(C_t).
 \end{aligned} \tag{4}$$

In which, σ is sigmoid function, x_t is input vector, h_t is output vector, w is parameter matrix, and b is offset parameter.

4.3. *CRF Layer.* Dependency between tags is a common sense in the input sequence; here, we take the BIO labeling system to illustrate this sense. The starting label of each word in the input sequence is “B-” or “O,” and it is usually that “I-X” follows “B-X,” and “I-X” is used as the ending label of the word. However, “I-” cannot be used as the starting label. For example, a legal annotation sequence is “B-L I-L I-L,” which together represents a location information. Illegal labels such as “B-G I-L” may appear, if the labeling process is not controlled. Unfortunately, the BiLSTM layer focuses on the context information and sentence features of the input sequence and cannot learn these annotation rules.

The CRF layer takes the output characteristic matrix of BiLSTM layer as input and outputs the global optimal label sequence, that is $y = (y_1, y_2, \dots, y_n)$, the most possible sequence annotation. The CRF layer transforms the dependency information between tags into constraints when predicting tags, so as to ensure the accuracy of prediction. Such as the “I-L” label cannot appear after the “B-W” label, the dependency constraint relationship between the labels will be automatically learned by the CRF layer in the data training stage. The label constraint relationship is represented by the transfer matrix A , where A_{ij} represents the dependence intensity between the i th label and the j th label. The higher the score, the greater the intensity and vice versa. In the actual prediction process, the start state and end state will be added into the input sequence. Therefore, the actual matrix is $R^{(k+2) \times (k+2)}$. In a tag sequence with a length equals to the length of the input sequence, the model scores the tag y of the input sequence x , and this is calculated as follows:

$$\text{score}(x, y) = \sum_{i=1}^N P_{i y_i} + \sum_{i=1}^N A_{y_{i-1} y_i}. \tag{5}$$

After the score is calculated, and then the normalized probability is calculated with softmax unit:

$$P(y | x) = \frac{\exp(\text{score}(x, y))}{\sum_{y'} \exp(\text{score}(x, y'))}. \tag{6}$$

Let Y denotes the set of all labels, where $\forall y' \in Y$, the denominator of equation (6) denotes that all possible transfer scores are obtained. Take logarithms on both sides of equation (6) to obtain that the log likelihood with respect to the input sequence x .

$$\log(P(y|x)) = \text{score}(x, y) - \log\left(\sum_{y'} \exp(\text{score}(x, y'))\right). \tag{7}$$

In the training process, the set of maximum probability labels in the sequence is selected by obtaining the value

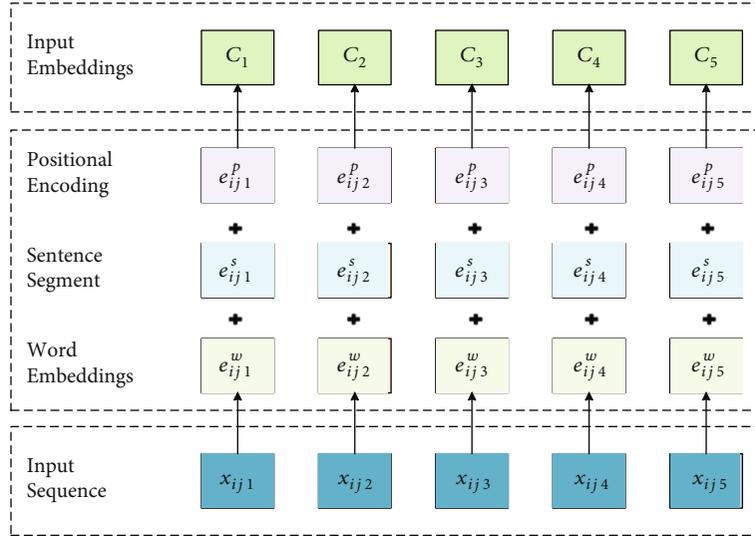


FIGURE 2: Input unit structure.

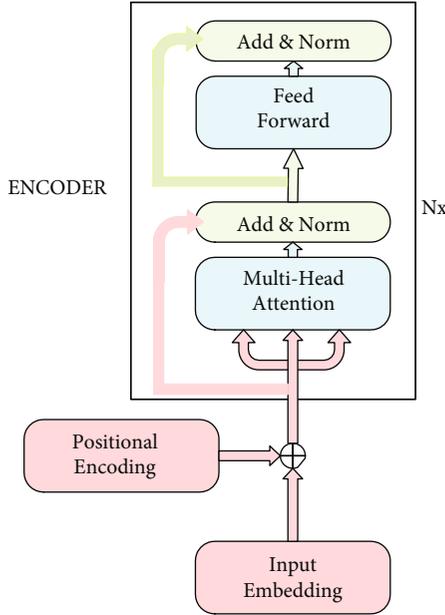


FIGURE 3: Transformer encoder unit.

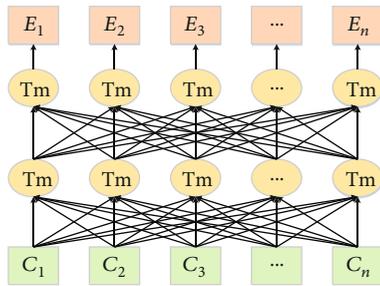


FIGURE 4: BERT pretraining language model.

of the maximum likelihood function, that is, the annotation sequence for the input sequence x predicted by CRF layer:

$$y^* = \operatorname{argmax}_{y \in Y} (\operatorname{score}(x, y)). \quad (8)$$

5. Experiments

We have conducted extensive literature research, and then, we find that few authors have considered the heterogeneous characteristics of military texts; thus, they have ignored the performance of these characteristics on the precision of deep learning model. Therefore, we have carried out relevant studies and conduct extensive experiments on related aspects. Experiments and discussions are shown as follows.

5.1. The Evaluation Metrics. In this work, precision (P), recall (R), and F -score (F) are selected as metrics for evaluating the performance of military entity extraction. The metrics are defined as follows:

$$\begin{aligned} P &= \frac{T_p}{T_p + F_p} \times 100\%, \\ R &= \frac{T_p}{T_p + F_n} \times 100\%, \\ F &= \frac{2 * P * R}{P + R} \times 100\%. \end{aligned} \quad (9)$$

In which, T_p is the number of positive entities identified by models, F_p is the number of negative entities identified by models, and F_n is the number of effective entities not detected by models.

5.2. Experimental Parameters. We carry out experimental with the BERT-Base model provided by an open-source project of Google. The information of hyperparameters in the

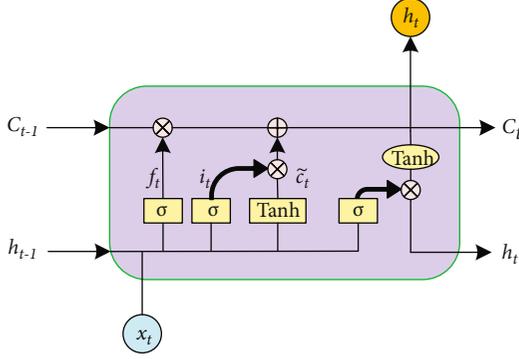


FIGURE 5: LSTM network structure.

TABLE 4: Model hyperparameters.

Parameters	Values
Batch size	64
Clip	0.5
Dropout rate	0.5
Learning rate	10^{-5}
Maximum sequence length	512
BiLSTM hidden layer dimension	256
Optimizer algorithm	SGD

training process is listed in Table 4. The deep learning framework and basic runtime environment are Pytorch V1.6.0 and Python V3.6.2. Experimental hardware configuration: 256G memory, 4 Nvidia GeForce RTX 3070 GPU.

5.3. Experimental Design and Experimental Result Analysis

5.3.1. Experiments on Corpus of Different Sizes. The three types of corpuses are divided into three data subsets of similar size by stratified sampling method, and experiments are carried out on them.

Experiment one (EP-ONE), one subset is extracted from each kind of corpus set to form a new corpus set, resulting in a total of 27 groups of experimental corpus sets. For each new corpus set, it is divided into five groups of similar size by the stratified sampling method, and then, using the cross-validation mechanism to carry out experimental, that four groups are selected as the training group, and the remaining one is used as the test group. With the cross-validation mechanism, five experiments are carried out on each new corpus set, and a total of 135 experiments are conducted. Finally, the mean of all experiments is taken as the test result.

Experiment two (EP-TWO), two subsets are extracted each time to form a new corpus set, resulting in 27 groups of experimental corpus sets. The following process is consistent with the EP-ONE.

Experiment three (EP-THREE), all the data is selected, and the corpus is divided into five groups of similar size by stratified sampling method. Then, the cross-validation mechanism is used as it is used in EP-ONE. A total of five

experiments are carried out, and finally, the mean of five experiments is taken as the test result. The experimental results are listed in Table 5.

The experimental results show that the performance of military entity recognition with BERT-BiLSTM-CRF model is constantly improving with the increase of the training corpus, where the F -value of the EP-TWO is 6.6% higher than that of the EP-ONE, the F -value of the EP-THREE is 10.97% higher than that of the EP-ONE, and the F -value of the EP-THREE is 4.37% higher than that of the EP-TWO. To explain this situation, we analyze the data features of the corpus constructed in this work. The analysis result shows that the military entities distributed in the corpus have characteristics of multiple types and variety of representations. These characters make the distribution of the entities sparse, and the smaller the corpus set is the data sparsity is more obvious. Therefore, the evaluation metric values are low with a small training corpus set, such as the result of EP-ONE, of which the F -value for military entity recognition is only 73.56%. However, with the data increase of the corpus set, the sparse problem is gradually solved. Such as the EP-THREE, the F -value of military entity recognition reaches 83.53%. Therefore, in order to ensure the performance of military entity recognition, it is necessary to build a relatively large corpus for training the BERT-BiLSTM-CRF model.

5.3.2. Cross-Validation Experiments. Cross-validation experiments are conducted on the three types of corpus; the experiments are divided into two parts: Selecting any one of them as the training set and the other two as the testing sets and selecting any two of them as the training set and the other as the test set.

The experiments are conducted according to the above two strategies, and for the convenience of representing the results, the corpus types of the experiments are represented by numbers, where “1” represents the abbreviated corpus, “2” represents the scientific or English name corpus, and “3” represents the novel and casual corpus, and the experimental results are listed in Table 6.

The experimental results show that it is not robust that the generalization ability of the BERT-BiLSTM-CRF model proposed in this work, when it is used on different corpus sets, since the large difference in entity representation between different corpus sets. Such as when the model is trained on the abbreviated corpus and tested on the other two types of corpuses, the recall rate is only 49.78%, and the F -value is 56.85%. However, when it was trained on the novel casual type corpus and tested on both the abbreviated and scientific or English name type corpus, it gets better performance; the recall rate and the F -value are 74.33% and 76.98%, an increase of 24.55% and 20.13%, respectively, over the first group. We sample entities from different corpus and make a comparison on the distribution. Then, we find that the novel casual corpus contains a larger span of entity types and entity representation types, so when the model trained on it performs well. When cross-corpus training is conducted, the generalization ability of the model is significantly improved, with F -values reaching 74.36%, 82.79%, and 87.39%, respectively. The purpose of this experiment is to

TABLE 5: Experimental results of the BERT-BiLSTM-CRF model on corpus sets with different size.

	<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)
EP-ONE	75.03	72.14	73.56
EP-TWO	81.29	79.06	80.16
EP-THREE	85.73	83.37	84.53

TABLE 6: Results of the cross-validation experiment on three corpus sets.

	<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)
1-(2,3)	66.25	49.78	56.85
2-(1,3)	70.18	58.36	63.73
3-(1,2)	79.83	74.33	76.98
(1,2)-3	77.94	71.09	74.36
(1,3)-2	84.37	81.26	82.79
(2,3)-1	89.13	85.72	87.39

illustrate that when it is necessary to extract entities across scenes and across corpus sets, one needs to pay attention to the entity features, and the distribution of entities in different scenes corpus sets and choose the training corpus set reasonably. The experimental results also show that when hardware conditions are limited, users could choose a corpus with more entity types as the training set to train the model and would achieve better performance.

5.3.3. Comparison Experiments. The comparison models are divided into horizontal and vertical two groups, where vertical group contains models, such as CRF, LSTM, and BiLSTM-CRF, and horizontal group contains models Lattice-LSTM-CRF, BERT-BiGRU-CRF, and BERT-IDCNN-CRF. Experiments are carried out on the whole corpus, and the experimental process is the same as the EP-THREE mentioned above. The models, for comparison, have achieved state of the art in the development of named entity recognition; thus, the experimental results would be more convincing when compared to these models.

In experimental process, the CRF model is trained and tested with the open source CRF++ (v0.58) tool. The LSTM-CRF model and BiLSTM-CRF model use word vectors trained with word2vec as input, and the vector dimension is 300 [22]. The experimental super parameters are consistent with the literature [13]. The Lattice-LSTM-CRF model adopts the super parameters in literature [23]. The input of CNN-BiLSTM-CRF model adopts word2vec word vectors as the input, the vector dimension is 300, and other super parameters are consistent with literature [24]. The experimental results are listed in Tables 7 and 8.

The experimental results show that the BERT-BiLSTM-CRF model applied in this work outperforms the other models listed above; it gets better performance for the recognition of military entities on the corpus. In the longitudinal comparison experiments, we compare the metric values of the BERT-BiLSTM-CRF model with any other model in

TABLE 7: Experimental results for longitudinal comparison.

Model	<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)
CRF	73.45	59.61	65.81
LSTM-CRF	74.93	71.72	73.29
BiLSTM-CRF	76.16	74.43	75.29
BERT-CRF	80.64	78.32	79.46
BERT-BiLSTM-CRF	85.73	83.37	84.53

TABLE 8: Experimental results for transverse comparison.

Model	<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)
Lattice-LSTM-CRF	79.57	76.29	77.90
CNN-BiLSTM-CRF	77.63	75.56	76.58
BERT-BiGRU-CRF	81.16	80.47	80.81
BERT-IDCNN-CRF	83.86	81.62	82.72
BERT-BiLSTM-CRF	85.73	83.37	84.53

the longitudinal group. Where compared to the CRF model, the recall rate and *F*-value are increased by 23.76% and 18.72%, respectively; compared to the LSTM-CRF model, the recall rate and *F*-value are increased by 11.65% and 11.24%, respectively; compared to the BiLSTM-CRF model, the recall rate and *F*-value are increased by 8.94% and 9.24%, respectively; and compared to the BERT-CRF model, the recall rate and *F*-value are increased by 5.05% and 5.07%, respectively. Word2vec is a static word vector, and its application in dynamic word sense transformation scenarios is poor for entity recognition, whereas dynamic word vectors can perform well in this scene. Therefore, the LSTM-CRF and BiLSTM-CRF entity recognition models are not as effective as the BERT-BiLSTM-CRF model with dynamic pretrained word vectors for entity recognition on a corpus with “five-plus” types. Although the BERT-CRF model introduces a pretraining mechanism, it is not as effective as the BERT-BiLSTM-CRF model with a bidirectional long and short-term memory network in learning the contextual features of sentences. In the transverse comparison experiment, the experimental process is the same as longitudinal experiment. And we analyze their performance as follows: compared to the Lattice-LSTM-CRF model, the recall rate and *F*-value are increased by 7.08% and 6.63%, respectively; compared to the CNN-BiLSTM-CRF model, the recall rate and *F*-value are increased by 7.81% and 7.95%, respectively; compared to the CNN-BiLSTM-CRF model, the recall rate and *F*-value are increased by 2.9% and 3.72%, respectively; and compared to the CNN-BiLSTM-CRF model, the recall rate and *F*-value are increased by 1.75% and 1.81%, respectively. Where the Lattice-LSTM-CRF model uses the improved grid LSTM element to integrate word and word order information, which can avoid word segmentation errors not transmitted in the network. The CNN-BiLSTM-CRF model uses CNN to extract character-level features of words, which improves the vector representation ability of words and reduces the influence of segmentation errors. Although these two models have been improved in character and word feature extraction, they cannot fundamentally

overcome the weakness of static word vector, and the performance improvement and application scenarios will be limited to some extent. The BERT model, which skillfully combines multiattentional mechanism and unsupervised subtask training task, can integrate characters, words, sentences, and word order to learn contextual information and then complete the comprehension task. The experimental results show that there is little difference between the models with pretraining structure, but BiLSTM performs well than IDCNN and BiGRU in the modeling of serialized text features, so the recognition performance of BERT-BiLSTM-CRF model is better in the entity recognition task of military language data.

6. Conclusion

In this work, we construct a set of corpuses and propose five rules for identifying the fuzzy boundary of military entities, which are applied to solve problems for military entity recognition, such as the lack of corpus, the single type of corpus, and the disunity of entity boundary division. These corpus have been divided into three types, abbreviation type, scientific or English name type, and novel and casual type. With these corpuses, we have conducted three types of experiments: (1) military entity recognition experiments using BERT-BiLSTM-CRF model on different sizes of corpus sets, (2) military entity recognition experiments using BERT-BiLSTM-CRF model across corpus sets, and (3) comparison experiments of multimodel military entity recognition. The experimental results illustrate the effect of different size data sets on the accuracy of the entity recognition model and the effect of data distribution on the accuracy of the recognition model and also validate the effectiveness of the BERT-BiLSTM-CRF model for military entity recognition.

At present, due to the limitation of data access, the amount of nonopen data available is limited, and many military entities are not yet covered. In the future, more new data samples will be generated by learning from existing sample data and combining the descriptions of some data patterns by domain experts using adversarial neural networks, which will be used to enrich the existing corpus. In addition, methods such as migration learning, knowledge distillation, and unsupervised learning are considered to reduce the reliance on corpus size and accuracy and to build lightweight military entity recognition models. In terms of practical applications, military knowledge graph plays a critical role in promoting the development of military intelligence, and military intelligence applications based on military knowledge graph will become more popular. In the future, intelligent resource management and scheduling technology will be widely used in the field of unmanned combat and in the field of military chess deduction.

Data Availability

The simulation experiment data of military texts used to support the findings of this study are restricted by the School Security Office in order to protect Military Intelligence

Information. Data is available from Hui Li, mcclane@njtu.edu.cn, for researchers who meet the criteria for access to confidential data.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The work by Ming Lyu is funded by the Natural Science Foundation of Jiangsu Province (Grant No. BK20180467).

References

- [1] S.-H. Jo, H.-J. Kim, S.-Y. Jin, and W.-S. Lee, "A study on building knowledge base for intelligent battlefield awareness service," *Journal of the Korea Society of Computer and Information*, vol. 25, no. 4, pp. 11–17, 2020.
- [2] W. Shu, K. Cai, and N. Xiong, "Research on strong agile response task scheduling optimization enhancement with optimal resource usage in green cloud computing," *Future Generation Computer Systems*, vol. 124, pp. 12–20, 2021.
- [3] J. Li and P. Wang, "Military text named entity recognition based on deep learning," in *International Conference on Applications and Techniques in Cyber Security and Intelligence*, pp. 808–816, Shanghai, China, 2018.
- [4] Y. Lu, R. Yang, X. Jiang, C. Yin, and X. Song, "A military named entity recognition method based on pre-training language model and BiLSTM-CRF," *Journal of Physics: Conference Series*, vol. 1693, no. 1, pp. 12161–12165, 2020.
- [5] B. Zhou, H. Zhang, R. Zhang, Y. Feng, and Y. Xu, "Construction of military corpus for entity annotation," *Computer Science*, vol. 46, no. 6A, pp. 540–546, 2019.
- [6] W. Jiang, J. Gu, and L. Cong, "Research on CRF and rules based military named entity recognition," *Command Control & Simulation*, vol. 33, no. 4, pp. 13–15, 2011.
- [7] Y. Feng, H. Zhang, and W. Hao, "Named entity recognition for military texts," *Computer Science*, vol. 42, no. 7, 2015.
- [8] H. Shan, H. Zhang, and Z. Wu, "A military named entity recognition method based on CRFs with small granularity strategy," *Journal of Armored Force Engineering Institute*, vol. 31, no. 1, pp. 84–89, 2017.
- [9] F. You, J. Zhang, D. Qiu, and M. Yu, "Weapon named entity recognition based on deep neural network," *Computer System Application*, vol. 27, no. 1, pp. 239–243, 2018.
- [10] C. Liu, Y. Yu, X. Li, and P. Wang, "Named entity recognition in equipment support field using tri-training algorithm and text information extraction technology," *IEEE Access*, vol. 9, pp. 126728–126734, 2021.
- [11] X. Wang, R. Yang, and W. Zhu, "Military named entity recognition method based on deep learning," *Journal of Armored Force Engineering Institute*, vol. 32, no. 4, pp. 94–98, 2018.
- [12] X. Zhang, X. Cao, and M. Zhang, "Military named entity recognition based on self-attention mechanism," *Command Control & Simulation*, vol. 41, no. 6, pp. 29–33, 2019.
- [13] X. Yin, H. Zhao, J. Zhao, W. Yao, and Z. Huang, "Multi-neural network collaboration for Chinese military named entity recognition," *Journal of Tsinghua University (Science and Technology)*, vol. 60, no. 8, pp. 648–655, 2020.

- [14] W. Liu, B. Zhang, W. Chen, C. Zhang, Y. Chen, and R. Pan, "Military named entity recognition based on transfer representation learning," *Command Information System and Technology*, vol. 11, no. 2, pp. 64–69, 2020.
- [15] Z. Zhang, S. Wu, D. Jiang, and G. Chen, "BERT-JAM: maximizing the utilization of BERT for neural machine translation," *Neurocomputing*, vol. 460, pp. 84–94, 2021.
- [16] L. Li and Y. Jiang, "Integrating language model and reading control gate in BLSTM-CRF for biomedical named entity recognition," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 3, pp. 841–846, 2020.
- [17] I. Ashrafi, M. Mohammad, A. S. Mauree et al., "Banner: a cost-sensitive contextualized model for Bangla named entity recognition," *IEEE Access*, vol. 8, pp. 58206–58226, 2020.
- [18] J. Qiu, Y. Zhou, Q. Wang, T. Ruan, and J. Gao, "Chinese clinical named entity recognition using residual dilated convolutional neural network with conditional random field," *IEEE Transactions on Nanobioscience*, vol. 18, no. 3, pp. 306–315, 2019.
- [19] T. Zhou, W. Wu, L. Peng et al., "Evaluation of urban bus service reliability on variable time horizons using a hybrid deep learning method," *Reliability Engineering and System Safety*, vol. 217, article 108090, 2022.
- [20] L. Dong, M. N. Satpute, W. Wu, and D.-Z. Du, "Two-phase multidocument summarization through content attention-based subtopic detection," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 6, pp. 1379–1392, 2021.
- [21] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 2017, pp. 5999–6009, 2017.
- [22] Z. Yin and Y. Y. Shen, "On the dimensionality of word embedding," *Advances In Neural Information Processing Systems 31 (NIPS 2018)*, vol. 31, pp. 1–12, 2018.
- [23] Y. Zhang and J. Yang, "Chinese NER using lattice LSTM," in *Proceedings of the 56th annual meeting of the association for computational linguistics (ACL), VOL 1*, pp. 1554–1564, Melbourne, Australia, 2018.
- [24] D. Gao, L. Peng, and Y. Bai, "HAZOP text named entity recognition using CNN-BiLSTM-CRF model," in *2020 Chinese Automation Congress (CAC)*, pp. 6159–6164, Shanghai, Peoples R China, 2020.