IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 9, 2021, accepted August 17, 2021, date of publication September 3, 2021, date of current version September 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3109911

Named Entity Recognition in Equipment Support Field Using Tri-Training Algorithm and Text Information Extraction Technology

CHENGUANG LIU [1], YONGLI YU [1], XINGXIN LI [1], AND PENG WANG [2]
[1] Army Engineering University of the People's Liberation Army, Shijiazhang 050051, China
[2] Test and Training Research Centre, Army Test and Training Base, Jinzhou 121000, China
Corresponding author: Xingxin Li (lxx_1226@sina.com)

ABSTRACT Weaponry equipment names belong to an important military naming entity that is difficult to identify because of features, such as complex components, miscellaneous, and scarce annotation corpus. Here, the automatic recognition of weaponry equipment names is specifically explored, a NER (Named Entity Recognition) algorithm is proposed based on BI-LSTM-CRF (Bi-directional Long Short Term Memory Conditional Random Field), thereby demonstrating the effectiveness of domain features in domain-specific entity recognition. Firstly, Chinese characters are represented by word embedding and input into the model. Then, the input feature vector sequence is processed by BI-LSTM (Bi-directional Long Short Term Memory) NN (Neural Network) to extract context semantic learning features. Finally, the learned features are connected to the linear CRF (Conditional Random Field), the NEs (Named Entities) in the equipment support field are labeled, and the NER results are obtained and output. The experimental results show that the accuracy of the NER algorithm based on the BI-LSTM-CRF model is 92.02%, the recall rate is 93.21%, and the F1 value reaches 93.88%. The effect of this model is better than the BI-LSTM NN model and LSTM-CRF (Long Short Term Memory Conditional Random Field) NN model. The proposed model provides some references for entity recognition in the field of equipment support.

INDEX TERMS Text information extraction, named entity recognition, equipment support, automatic word segmentation, information identification.

I. INTRODUCTION

With the rapid development of AI (Artificial Intelligence) and big data technology, the intellectualization of the general domain and vertical domain is developing steadily. Particularly, customized solutions are needed for the vertical domains because of their peculiar industrial features. NER (Named Entity Recognition), as an important semantic analysis tool, is a key branch of IE (Information Extraction), emotional analysis, and SU (Semantic Understanding). Meanwhile, NER is the foundation of such AI applications as knowledge mapping. Scholars have researched and summarized general domain NER methods, while the vertical domain NER should consider the specific features of the professional field. Here, the NER in the field of equipment support is studied. Technically, NER in the field of equipment support can effectively extract the field information and provide the basis for the subsequent construction of the domain knowledge map and the further development of big data applications. Besides, NER is regarded as a subtask of information extraction, and it can locate and classify the atomic elements in a group of words and then output the directory in a fixed format [1], [2].

Both domestic and international scholars have studied entity recognition and military entity recognition. Liao *et al.* [3] pointed out that military NER is one of the key technologies in military information extraction. Traditional MNER task methods relied on cumbersome feature engineering and professional domain knowledge. Then, an automatic military entity recognition method was proposed based on BI-LSTM (Bi-directional Long Short-Term Memory) NN (Neural Network) and self-attention mechanism. The experimental results showed that the self-attention mechanism could effectively improve task performance. The F-values of military literature and network military text recognition were 90.15% and 89.34%, respectively, which

The associate editor coordinating the review of this manuscript and approving it for publication was Ahmed A. Zaki Diab.

was better than other models. You *et al.* [4] proposed a DNN (Deep Neural Network) recognition method based on word vector features and state features and used for weaponry in electronic text, such as aircraft, tank vehicles, artillery, and missiles. The experiment showed that the F1 value on the test corpus is 0.9102. Englmeier [5] suggested that context classified specific environments through key attributes that served as a semantic markup role. When summarizing information at a more coarse-grained level, the context could be a recursive construct. Consequently, the recognition and standardization of context at different granularity levels were put forward to support faster and more accurate information retrieval. The system used the semi-automatic method of supervised learning to extract and standardize data from text. This method was based on NER and simple ontology to identify and eliminate context ambiguity. Shibuya and Hovy [6] argued that when entity names contained other names, identifying all name combinations might become difficult and expensive. Therefore, a new method was proposed, which could recognize the outermost NEs (Named Entities), but also the nested NEs. Meanwhile, a decoding method was established for reasoning, which iteratively extracted from the outermost entity to the internal entity in an external to internal way. Experiments suggested that the F1 score of this method on the GENIA data set was 85.82%.

In recent years, researchers have made many achievements in Chinese NER. But in some niche areas, such as military equipment support, the research on NER is rare due to factors like confidentiality. Therefore, a NER method is proposed here. The proposed method adds the word position information of a single Chinese character in the segmentation results to the word vector pre-trained through the large-scale corpus, then inputs the combined feature vector, and trains the input with the bidirectional threshold cycle unit CRF (Conditional Random Field) network. Finally, the proposed method completes the NER in the military field.

II. NER IN EQUIPMENT SUPPORT FIELD BASED ON BI-LSTM-CRF (BI-DIRECTIONAL LONG SHORT TERM MEMORY CONDITIONAL RANDOM FIELD) MODEL

A. OVERVIEW OF KEY TECHNOLOGIES OF NER IN THE EQUIPMENT SUPPORT FIELD

There are various NEs in the field of military equipment support, they have a unique grammatical structure and a complex composition, distinct from everyday texts. Specifically, the differences can be manifested from six aspects. 1. The format and narrative methods are fixed. 2. Sentence structure is simple [7]. 3. The terminology is accurate. 4. The usage frequency of military language is high. 5. There are usually parallel or progressive relations in segmentation or clause. 6. Contents are very normative [8]–[10]. Generally, in military equipment support, NEs can be classified into countries, time, weaponry equipment models, companies, and institutions. The analysis of the corpus implies that the naming of countries, time, weaponry equipment models, companies, and institutions follow specific rules, except for the naming of some complex companies and institutions. The following are the results of the feature analysis of NEs. Location entities are generally bounded to the country, province, city, county. Weaponry entities are generally composed of Chinese characters, letters, numbers, and symbols. The difficulty of NER in the field of military equipment support is that the boundary of Chinese text words is not clear enough, there are more nested combinations and more problems of reference elimination [11], [12].

The automatic word segmentation technology can separate the text into single words for subsequent processing, such as labeling and recognition [7], [13], [14]. Here, the texts in equipment support are collected as the source data, and the forward maximum matching method is used for word segmentation according to their features. Classified labeling can classify word segmentation and label them with the same identifier. Here, the key information is identified and extracted according to the text features in the equipment support field, and the targeted labeling method is adopted. The category labels are customized, and the words with distinct information are merged and classified [15], [16]. Compared with the part-of-speech tagging of semantic annotation in NLP (Natural Language Processing), the proposed tagging method also provides hints on the semantics and categories of word segmentation units.

B. BI-LSTM-CRF MODEL

LSTM network is an effective deformation of RNN (Recurrent Neural Network). The hidden layer of each moment in the LSTM network contains multiple memory modules, each module contains one or more memory cells, and each memory cell contains three memory gates [17], [18]. The LSTM network can effectively solve gradient explosion or gradient disappearance in the simple RNN, optimizes non-RNN networks, and can deal with context information within a large range. Therefore, the LSTM model is advantageous for the sequence labeling problem. The chronologically expanded diagram is shown in figure 1 [19], [20].

In Figure 1, $X_1, X_2, X_3 \ldots X_t$ represents the input of the network at each moment, $C_1, C_2, C_3 \ldots C_t$ denotes the internal state of the network at each moment, and $H_1, H_2, H_3 \ldots H_t$ stands for the external state of the network at each moment. Compared with the simple RNN, the LSTM network is improved from two aspects. The first improvement is that the internal state variable is introduced for linear cyclic information transfer, as shown in equation (1), and information is output to the external state, as shown in equation (2) [13], [21], [22].

$$c_t = f_t \odot c_{t-1} + i_t \odot c_t' \tag{1}$$

$$h_t = o_t \odot \tanh(\mathrm{c}_t) \tag{2}$$

In Eq. (1) and Eq. (2), the meaning of i_t and o_t represent the different gate structures in the LSTM network, $\odot$ denotes the product of vector elements, c_{t-1} stands for the memory

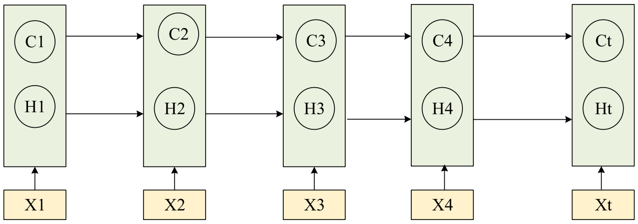


FIGURE 1. LSTM network structure.

unit at the last rise time, and c_t' stands for the candidate state obtained through the nonlinear structure.

The second improvement is that three control gates are introduced, including the forgetting gate, the input gate, and the output gate, controlling the path of information transmission. The calculation of the forgetting gate is shown in equation (3), the calculation of the input gate is shown in equation (4), and the calculation of the output gate is shown in equation (5).

$$f_t = \delta(W_f x_t + U_f h_{h-1} + b_f) \quad (3)$$
$$i_t = \delta(W_i x_t + U_i h_{t-1} + b_i) \quad (4)$$
$$o_t = \delta(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

In (3), (4), and (5), $\delta(.)$ represents a logarithmic function, x_t denotes the input at the current time, h_{t-1} indicates the external state at the previous time, W stands for the weight matrix, b means the offset vector, and U is the coefficient.

The forgetting gate, input gate, and output gate can control the information storage location, storage time, and reading mode. The operations of the forgetting gate, input gate, and output gate are slightly different. The forgetting gate multiplies the memory unit vector of the previous state to obtain the historical information that needs to be retained [23], [24].

The forward LSTM and backward LSTM are combined into BI-LSTM. The time window BI-LSTM can predict the step size of the next moment using the historical value of the time window length. The parameter value of the time window step represents how many pieces of historical data are used to predict future values. During BI-LSTM model training, the trained data are input, the predicted value is calculated and output in the previous direction, and then the predicted value is calculated and output in the backward direction. Afterward, the model gradient parameters and weight coefficients calculated in the forward direction and backward direction are updated based on the comparison with the actual value. The BI-LSTM model prediction process can be divided into four steps: firstly, the time window step of a sample data is set to three, its input is x_{t-1}, x_t, x_{t+1}, and there are two separate LSTM units. For the backward calculated LSTM unit, the samples are input into the LSTM unit in order of x_{t-1}, x_t, x_{t+1} to obtain the first set of state outputs. Then, for the backward calculated LSTM unit, the samples are input into the LSTM unit in order of x_{t-1}, x_t, x_{t+1} to obtain the first set of state outputs. Consequently, the two groups of state outputs are obtained, and the feature dimensions of each element are the same, and the two groups of state variables are spliced. Finally, for each input, an output vector with a length of $2 * x_t$ feature dimension is obtained [25], [26].

C. CRF ALGORITHM BASED ON TRI-TRAINING

CRF (Conditional Random Field) model is an undirected graph model based on the hidden Markov model, which is commonly used in sequence labeling tasks. The hidden Markov model with CRF layer can be expressed as in Eq. (6).

$$S(\mathrm{X}, \mathrm{Y}) = \sum_{i=0}^{n} A_{I,J} + \sum_{i=1}^{n} p_{i,y_i} \quad (6)$$

In Eq. (6), A denotes the transfer matrix, and p stands for a probability matrix of n * k.

The basic idea of the Tri-training algorithm is similar to a multi-classifier, which sets three classifiers C1, C2, and C3, and then randomly samples through Bootstrap algorithm from sample set LL to generate different training datasets and three classifiers. The classification ability of the three classifiers is different, and the data X in the unlabelled data set U is labeled [1], [27], [28]. If two datasets have the same classification effect for X, the X sample is classified as C2(X) and added to the C1 training set, and similarly, the other two classifiers repeat the above process. After multiple iterations, the training stops when the classification performance of the three classifiers does not change. Obviously, the advantages of Tri-training semi-supervised learning are that it can classify with a small amount of labeled corpus and a large number of the unlabelled corpus, the military equipment support information is mostly classified as confidential, and labeled corpus is difficult to obtain. Therefore, the semi-supervised learning Tri-training algorithm is applied to train a better classification model on fewer labeled corpus and a large number of unlabelled corpus [29]–[31].

The workflow of the CRF algorithm based on Tri-Training can be divided into four steps.

1. A small amount of annotated corpus can train the CRF model, and the trained CRF model C1 is generated.
2. Then, Model C2 can identify the NE in the field of equipment support for the unlabelled corpus U, and the confidence of the results of U is estimated.
3. The confidence threshold is set to 0.85, word segmentation is selected from the test corpus and set to subset u, then u is added to the training corpus L, and subset u is deleted.
4. The CRF model will be generated through the repetition of the above steps. Specifically, when the confidence difference between CK and CK-1 is no more than 0.5%, the optimal model CK will be generated.

D. NER PROCESS IN EQUIPMENT SUPPORT FIELD

Here the NER is based on the material particularity of the equipment support field, and the specific process is shown in figure 2. Figure 2 shows that the process is divided into four steps. The first step is the material text pre-treatment.

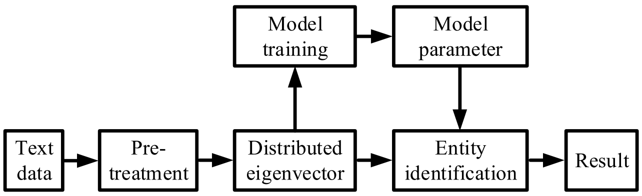


FIGURE 2. Model training steps.

The second step is to train the distributed eigenvector. The third step is to train the BI-LSTM network. The fourth step is to identify and extract entities.

E. BI-LSTM-CRF SEQUENCE LABELING MODEL

Figure 3 illustrates the sequence labeling model architecture for NER in the field of equipment support.

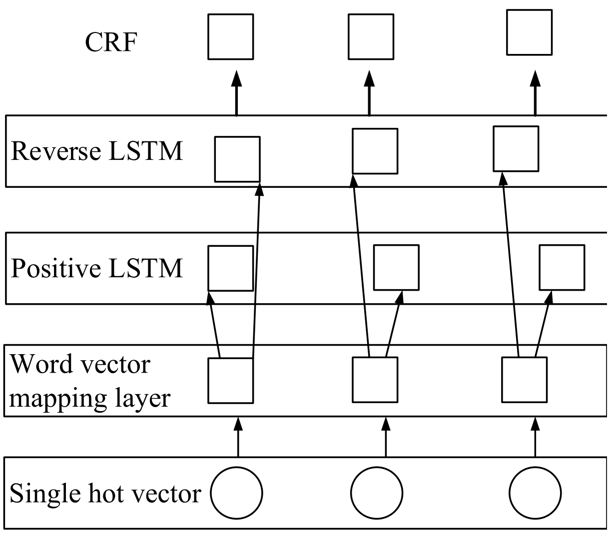


FIGURE 3. The architecture of the sequence labeling model.

The bottom layer of the model is the word vector mapping layer. The second layer of the model is the BI-LSTM layer, and it can automatically extract sentence features without manual interference. The third layer of the model is the CRF layer, and it can label the sentence in the sequence. If a label sequence with a length of the sentence is denoted as $y = (\mathrm{y}_1, \mathrm{y}_2, \mathrm{y}_3, \ldots \mathrm{y}_n)$, the score of the sentence x labeled with y can be expressed as in Eq. (7).

$$SCORE(\mathrm{x}, \mathrm{y}) = \sum_{i=1}^{n} P_{i,y_i} + \sum_{i=1}^{n+1} A_{y_{i-1},y_i} \tag{7}$$

The total score of each location adds up to the score of the whole sequence, and the score of each location is composed of the transition matrix A of CRF and the output of LSTM P_{i,y_i}. Model training is achieved through the maximization of the log-likelihood function.

The construction process of BI-LSTM-CRF follows. 1. The complete sentence sequence first enters the word vector layer, and the word vector corresponding to the sentence sequence is obtained by querying the word vector table as the input sequence. 2. Multiple parameter matrices of the BI-LSTM layer are randomly initialized. The input sequence obtained from the word vector layer enters the BI-LSTM layer, and the output of LSTM in two directions is spliced as the output of the hidden layer. 3. CRF layer is responsible for tagging the sequence of sentences, and each position is labeled using the parameter matrix.

F. EXPERIMENT SETTING

To better realize NER in the corpus of equipment support field, the source of military text corpus used in the experiment is the data set of equipment support field based on open-source data. The data source includes the public data set of Sina military news and CCKS 2020 technical evaluation task of Institute of Systems Engineering, Military Academy of Sciences. The final experimental data set is constructed by machine annotation and manual verification, which contains 200,000 words and 5 types of labeled entities. The labeling method is BIO. Then, 80% of them are randomly selected as the training corpus, and 20% of them are chosen as the test corpus. In the designed experiment, the hidden layer of the BI-LSTM network contains four layers, the number of neurons in the hidden layer is 120, 200, and 250, respectively. The word vector dimension is 100, and the length of the single input sequence of the model is 15. The NN model is trained through a small batch random gradient descent algorithm. The number of batch samples is 20, and the sample iteration is 100 times. The Adam optimizer is used in the training process, the dropout is set to 0.5, and the learning rate is 0.0001.

G. EVALUATION INDICES

Here, the model is evaluated through precision, recall rate, and F1 score, which is calculated through Eq. (8), Eq. (9), and Eq. (10), respectively.

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

$$F1 = \frac{2RP}{R + P} \tag{10}$$

III. EXPERIMENTAL RESULTS AND DISCUSSION OF NER

A. RESULTS OF NER IN EQUIPMENT SUPPORT FIELD BASED ON BILSTM-CRF

Figure 4 shows the results of the NER experiment.

In Figure 4, P denotes the precision, R represents recall rate, and F1 stands for F1 score.

Figure 4 demonstrates that when the NEs are place names, the precision rate of the network is 88.2%, the recall rate is 91.89%, and the F1 score is 90.12%. When the NEs are time, the precision rate of the network is 79.81%, the recall rate is 74.47%, and the F1 score of the network is 76.15%.

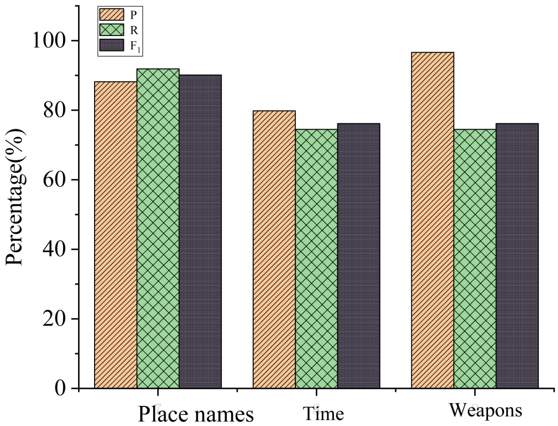


FIGURE 4. Experimental results of NER.

When the NEs are weaponry, the precision rate is 96.93%, the recall rate is 74.47%, and the F1 score is 76.15%.

B. ONTOLOGICAL FEATURE VECTOR EXPERIMENT

Figure 5 shows the statistical results of the precision, recall rate, and F1 score of the word vector, sentence vector, and sentence vector + ontology feature.

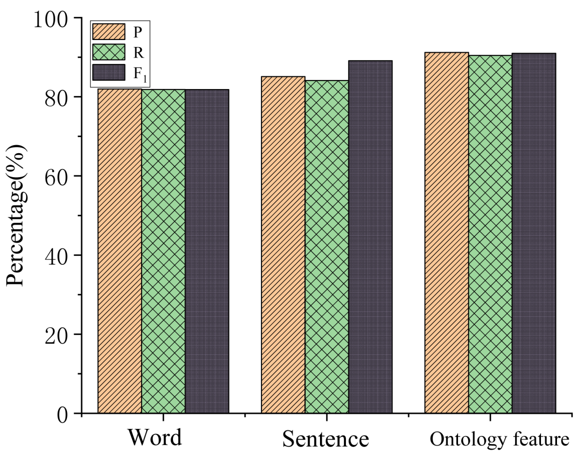


FIGURE 5. The experimental results of the ontology feature vector.

In Figure 5, P denotes the precision, R represents the recall rate, F1stands for the F1 score.

In the first group of experiments, only word vectors are input into the constructed model. The precision of the model is 81.95%, the recall rate is 81.83%, and the F1 score is 81.81%. The effect is poor, not as good as the general effect of NER, indicating that when the model input is only word vectors, the NER cannot be effectively realized in the field of equipment support.

In the second group of experiments, word vectors and sentence vectors are input into the constructed model. The precision of the model is 85.11%, the recall rate is 84.11%, and the F1 score is 81.81%. Compared with experiment 1, the precision, recall rate, and F1 score of experiment 2 have been improved. The input of the word vector is conducive to morphological feature acquisition within word segmentation and improves the recognition performance of the model.

In the third group of experiments, sentence vectors + ontology features are input into the constructed model. The precision of the model is 91.20%, the recall rate is 90.44%, and the F1 score is 90.98%. Compared with experiment 1 and experiment 2, the precision, recall rate, and F1 score of Experiment 3 have been improved, indicating that the addition of ontology features can greatly improve the recognition ability of NEs.

C. MODEL COMPARISON EXPERIMENT

The performance of the constructed model is compared with that of the BI-LSTM model and LSTM-CRF model to evaluate its effectiveness, as shown in Figure 6.

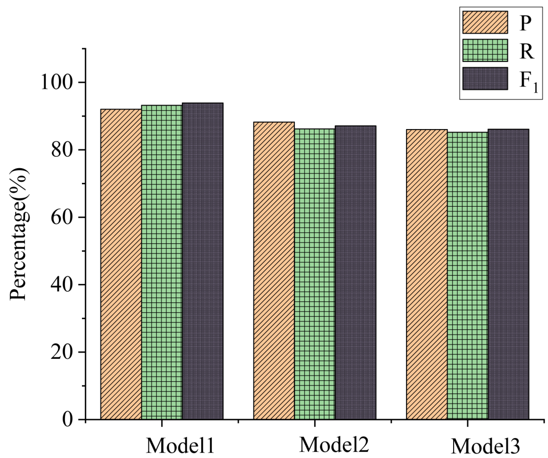


FIGURE 6. Experimental results.

In figure 6, Model 1 represents the proposed model, Model 2 denotes the BI-LSTM model, Model 3 is the LSTM-CRF model, P denotes the precision, R represents the recall rate, and F1 stands for the F1 score.

The results indicate that the precision of the NER method based on the BI-LSTM-CRF model is 92.02%, the recall rate is 93.21%, and the F1 score is 93.88. The precision of the NER method based on the BI-LSTM model is 88.21%, the recall rate is 86.21 %, the F1 score is 87.11%. The precision, recall rate, and F1 score of the NER method based on the LSTM-CRF model are 86.01%, 85.21%, and 86.11%, respectively. The precision and the recall rate of the NER method based on the proposed model are more than 80%, so the recognition performance is better than that of the BILSTM model and LSTM-CRF model. Hence, the proposed algorithm is effective.

To evaluate the iterative learning effect of the Tri-Training algorithm on the CRF model, the performance of the NER method based on the proposed BI-LSTM-CRF model is compared with that of the NER method based on BI-LSTM-CRF model without Tri-Training algorithm pre-learning. The experimental results are shown in Figure 7.

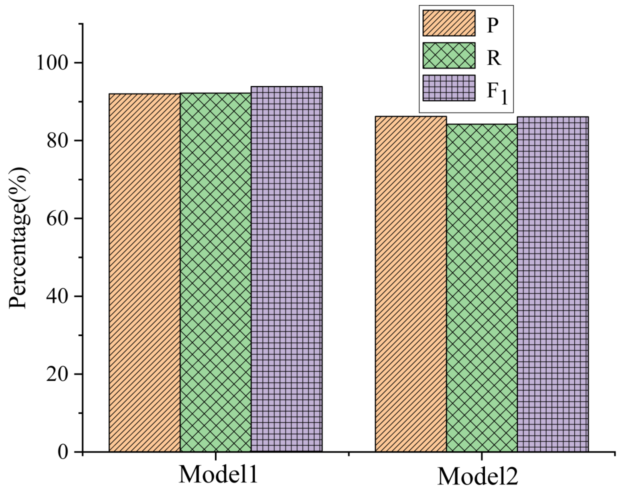


FIGURE 7. Performance comparison results.

In Figure 7, Model 1 represents the proposed model, Model 2 represents the BI-LSTM-CRF model without Tri-Training algorithm pre-learning, P represents the accuracy, R represents the recall rate, and F1 represents the F1 score. Figure 7 shows that the accuracy index of the proposed model is 6% higher than that of the BI-LSTM-CRF model without the Tri-Training algorithm. The recall rate and F1 score of the proposed model are also higher than those of the BI-LSTM-CRF model without the Tri-Training algorithm. The performance of the proposed model is better than that of the BI-LSTM-CRF model without the Tri-Training algorithm.

IV. CONCLUSION

Here, several key technologies of NER in the field of equipment support are studied, including word segmentation, classification, annotation, and information extraction. A tri-training algorithm is proposed to iteratively learn the CRF model, and then complete keyword recognition and extraction, which effectively improves the training effect of the algorithm. Finally, the model is trained and evaluated by combining the open-source data set and the self-built data set, which provides an effective solution for NER in the field of equipment support. Due to the particularity of the equipment support field, the effect of the model is very dependent on domain knowledge. Therefore, the proposed model still has some shortcomings. For example, in the process of tri-training, the probability generated by the CRF annotation sequence is selected as the condition for selecting high confidence annotation data. This method is relatively simple, and more feasible and effective methods need to be explored in the future. In future work, this paper will use the variational Bayesian algorithm (TARA) to cluster the data.

REFERENCES

[1] I. E. Bazi and N. Laachfoubi, "Arabic named entity recognition using word representations," *Int. J. Comput. Sci. Inf. Secur.*, vol. 14, no. 8, pp. 956–965, 2018.

[2] R. Catelli, F. Gargiulo, V. Casola, G. De Pietro, H. Fujita, and M. Esposito, "Crosslingual named entity recognition for clinical de-identification applied to a COVID-19 Italian data set," *Appl. Soft Comput.*, vol. 97, Dec. 2020, Art. no. 106779.

[3] F. Liao, L. Ma, J. Pei, and L. Tan, "Combined self-attention mechanism for Chinese named entity recognition in military," *Future Internet*, vol. 11, no. 8, p. 180, Aug. 2019.

[4] F. You, "Weapon named entity recognition based on deep neural network," *Comput. Syst. Appl.*, vol. 29, no. 8, pp. 21–43, 2018.

[5] K. Englmeier, "Named entities and their role in creating context information," *Procedia Comput. Sci.*, vol. 176, pp. 2069–2076, Jan. 2020.

[6] T. Shibuya and E. Hovy, "Nested named entity recognition via second-best sequence learning and decoding," *Trans. Assoc. Comput. Linguistics*, vol. 8, pp. 605–620, Dec. 2020.

[7] M. Cho, J. Ha, C. Park, and S. Park, "Combinatorial feature embedding based on CNN and LSTM for biomedical named entity recognition," *J. Biomed. Informat.*, vol. 103, Mar. 2020, Art. no. 103381.

[8] S. Trivedi, R. Gildersleeve, S. Franco, A. S. Kanter, and A. Chaudhry, "Evaluation of a concept mapping task using named entity recognition and normalization in unstructured clinical text," *J. Healthcare Inform. Res.*, vol. 4, no. 4, pp. 395–410, 2020.

[9] X. Liu, N. Yang, Y. Jiang, L. Gu, and X. Shi, "A parallel computing-based deep attention model for named entity recognition," *J. Supercomput.*, vol. 76, no. 2, pp. 814–830, Feb. 2020.

[10] C. Hu and J. Cheng, "Named entity recognition based on character-level language models and attention mechanism," *Int. Core J. Eng.*, vol. 6, no. 1, pp. 196–201, 2020.

[11] E. K. Akkaya and B. Can, "Transfer learning for Turkish named entity recognition on noisy text," *Natural Lang. Eng.*, vol. 27, no. 1, pp. 35–64, Jan. 2021.

[12] L. Gligic, A. Kormilitzin, P. Goldberg, and A. Nevado-Holgado, "Named entity recognition in electronic health records using transfer learning bootstrapped neural networks," *Neural Netw.*, vol. 121, pp. 132–139, Jan. 2020.

[13] G. Wen, H. Chen, H. Li, Y. Hu, Y. Li, and C. Wang, "Cross domains adversarial learning for Chinese named entity recognition for online medical consultation," *J. Biomed. Informat.*, vol. 112, Dec. 2020, Art. no. 103608.

[14] M. Marrero and J. Urbano, "A semi-automatic and low-cost method to learn patterns for named entity recognition," *Natural Lang. Eng.*, vol. 24, no. 1, pp. 39–75, 2018.

[15] P. Huang, H. Wang, and Y. Jin, "Offline data-driven evolutionary optimization based on tri-training," *Swarm Evol. Comput.*, vol. 60, Feb. 2021, Art. no. 100800.

[16] X. Guan, Y. Li, Q. Zeng, and C. Zhou, "An automatic text summary extraction method based on improved textrank and TF-IDF," in *Proc. IOP Conf., Mater. Sci. Eng.*, vol. 563, 2019, Art. no. 042015.

[17] J. Sun, K. Lei, L. Cao, B. Zhong, Y. Wei, J. Li, and Z. Yang, "Text visualization for construction document information management," *Autom. Construct.*, vol. 111, Mar. 2020, Art. no. 103048.

[18] B. Yu and K. Wang, "Research on express information extraction based on multiple sequence labeling models," in *Proc. IOP Conf.: Earth Environ. Sci.*, 2021, vol. 632, no. 2, Art. no. 022060.

[19] Y. Liu, "Optimization of conditional random field model based on circular neural network," *World Sci. Res. J.*, vol. 5, no. 7, pp. 32–38, 2019.

[20] Q. Jia, T. N. Wang, and Y. C. Zhang, "A new mode of army equipment support," in *Proc. J. Phys., Conf.*, 2020, vol. 1649, no. 1, Art. no. 012042.

[21] J. Chen, Y. Wang, and X. Jiang, "Research on key technologies of equipment support intelligent system," in *Proc. J. Phys., Conf.*, 2020, vol. 1575, no. 1, Art. no. 012127.

[22] D. Nozza, P. Manchanda, E. Fersini, M. Palmonari, and E. Messina, "LearningToAdapt with word embeddings: Domain adaptation of named entity recognition systems," *Inf. Process. Manage.*, vol. 58, no. 3, May 2021, Art. no. 102537.

[23] R. Zhang, Y. Gao, R. Yu, R. Wang, and W. Lu, "Medical named entity recognition based on overlapping neural networks," *Procedia Comput. Sci.*, vol. 174, pp. 27–31, Jan. 2020.

[24] Y. Chen, Y. Hu, Y. Li, R. Huang, Y. Qin, Y. Wu, Q. Zheng, and P. Chen, "A boundary assembling method for nested biomedical named entity recognition," *IEEE Access*, vol. 8, pp. 214141–214152, 2020.

[25] K. Zhang and G. Qian, "Research on named entity recognition technology for Chinese titles," in *Proc. J. Phys., Conf.*, 2021, vol. 1732, Art. no. 012016.

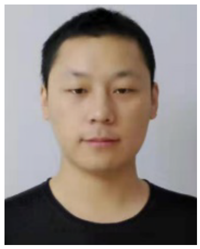
[26] R. Jain, "Named entity recognition in English text," in *Proc. J. Phys., Conf.*, 2020, vol. 1712, Art. no. 012013.

[27] J. Muré, "Optimal compromise between incompatible conditional probability distributions, with application to objective Bayesian Kriging," *Nephron Clin. Pract.*, vol. 23, pp. 271–309, Oct. 2019.
[28] P. S. Muhuri, P. Chatterjee, X. Yuan, K. Roy, and A. Esterline, "Using a long short-term memory recurrent neural network (LSTM-RNN) to classify network attacks," *Information*, vol. 11, no. 5, p. 243, May 2020.
[29] L. Hu, J. Zhang, X. Pan, H. Yan, and Z.-H. You, "HiSCF: Leveraging higher-order structures for clustering analysis in biological networks," *Bioinformatics*, vol. 37, no. 4, pp. 542–550, May 2021.
[30] L. Hu, K. C. C. Chan, X. Yuan, and S. Xiong, "A variational Bayesian framework for cluster analysis in a complex network," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 11, pp. 2115–2128, Nov. 2020.
[31] L. Hu and K. C. C. Chan, "Fuzzy clustering in a complex network based on content relevance and link structures," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 2, pp. 456–470, Apr. 2016.

CHENGUANG LIU was born in 1996. He received the degree from Shijiazhuang Campus of Army Engineering University, where he is currently pursuing the master's degree. His research interests include knowledge graph and equipment supportability verification.

YONGLI YU was born in 1962. He received the M.S. and Ph.D. degrees from Nanjing University of Science and Technology. He is currently a Professor with Army Engineering University and an Expert in equipment reliability and supportability. His research interests include equipment maintenance engineering, equipment comprehensive support, and virtual maintenance.

XINGXIN LI was born in 1980. He received the M.S. and Ph.D. degrees from Ordnance Engineering College, in 2005 and 2010, respectively. He is currently a Lecturer with Shijiazhuang Campus of Army Engineering University. His research interests include system fault detection and virtual maintenance.



PENG WANG was born in 1991. He received the B.S. degree from Beihang University, in 2013, the M.S. degree from Ordnance Engineering College, in 2015, and the Ph.D. degree from Shijiazhuang Campus of Army Engineering University, in 2021. He is currently an Assistant Engineer at Test and Training Research Centre, Army Test and Training Base. His research interests include system fault detection and control theory.