IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 29 February 2024, accepted 13 March 2024, date of publication 22 April 2024, date of current version 3 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3392388

RESEARCH ARTICLE

A Weakly Supervised Chinese Named Entity Recognition Method Combining First-Order Logic

XI TANG AND DONGCHEN JIANG
School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China
Corresponding author: Dongchen Jiang (jiangdongchen@bjfu.edu.cn)

ABSTRACT Named entity recognition is a key prerequisite for many tasks. However, the high cost of entity annotation limits feature learning and generalization capabilities of models. To address this problem, this paper integrates the weakly supervised method with first-order logic for Chinese named entity recognition. Firstly, a knowledge base is established by using first-order logic, tailored to the characteristics of the Chinese named entity recognition dataset. Secondly, self-training approach is introduced to address the issue of suboptimal feature learning in the model, stemming from a limited number of entity types. Lastly, the first-order logic knowledge base is incorporated into self-training approach to rectify mislabeling in the training process, which improves the generalization ability. The F1-score on the public datasets ACE05 and MSRA are improved by 2.56% and 0.35% respectively.

INDEX TERMS Weakly supervised learning, self-training, named entity recognition, first-order logic.

I. INTRODUCTION

Named entity recognition (NER) is an important information extraction task in natural language processing. It serves as a prerequisite for many tasks, such as relation extraction [1] and knowledge graph construction [2]. Named entity recognition has received increasing attention because its performance directly impacts the results of subsequent tasks. In recent years, deep learning has become a widely used method for named entity recognition. The effectiveness of deep learning methods relies on high-quality labeled datasets. Unfortunately, the reality persists that acquiring high-quality labeled datasets is a costly endeavor, resulting in models that struggle to effectively learn the requisite features. Consequently, mainstream deep learning approaches for named entity recognition may exhibit suboptimal performance, particularly when applied to texts within specific domains. Thus, how to solve the problem of costly labelling is an important factor for the practice of named entity recognition.

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed Elhoseny.

In datasets of Chinese named entity recognition, an imbalance in the distribution of entities across various types frequently arises, which is primarily due to the costs associated with dataset annotation. This imbalance leads to incomplete supervision. Concurrently, disparities in the cognitive scope and awareness levels among annotators, along with challenges associated with entity nesting in entity annotations, may give rise to mislabeling, ultimately culminating in the provision of inaccurate supervision. In the field of machine learning, weakly supervised learning refers to the process of building models under limited supervision, which includes both incomplete and inaccurate supervision [3]. Hence, it is necessary to ascertain effective methodologies for addressing weakly supervised Chinese named entity recognition.

In weakly supervised learning, self-training (ST) [4] has been proposed to mitigate the challenge of incomplete supervision in scenarios with limited entity classes. Self-training utilizes a limited labeled dataset to initially train a teacher model; subsequently, this teacher model generates pseudo-labels for unlabeled data, facilitating the retraining of a student model. Self-training effectively leverages available data resources, resulting in improved model performance. However, the errors introduced in the self-training iterations frequently lead to issues of inaccurate supervision.

In this paper, we adopt the teacher-student model to address incomplete supervision, and utilize first-order logic knowledge base to rectify the pseudo-labeling errors generated. Based on the above, we propose a first-order logic based method within a weakly supervised framework for Chinese named entity recognition. The efficacy of the method is demonstrated by experiments. The contribution is as follows:

- An improved self-training approach combined with a first-order logic knowledge base is proposed for pseudo-labeling self-correction.
- An enhanced cross-entropy loss is introduced to address the issue of sample imbalance.
- Different experiments are conducted to show that our method performs better on both inaccurate and incomplete supervision.

II. RELATED WORK

A. WEAKLY SUPERVISED NAMED ENTITY RECOGNITION ALGORITHMS

Weakly supervised named entity recognition involves using unconventional or incomplete supervisory labels to guide model training and improve the model's generalization ability. Commonly used methods include heuristic rules and remote supervision.

Heuristic rule-based named entity recognition is an empirical and expertise-based method for identifying and categorizing named entities by analyzing local features and structure of textual data. McCallum and Li [5] proposed an approach that utilizes conditional random fields, feature induction, and web-enhanced dictionaries for named entity recognition. This approach reduces the need for manual feature engineering, and maximizes the benefits of statistical learning. Torisawa [6] utilized Wikipedia as an external source of knowledge to improve the accuracy of named entity recognition. This is accomplished by extracting category information and contextual features from the Wikipedia data. Ratinov and Roth [7] analyzed the impact of different features and used heuristic rules to address partial localization in named entity recognition. Nothman et al. [8] used Wikipedia as an external knowledge source to learn multilingual named entity recognition and successfully achieved the recognition and classification of entity classes by utilizing Wikipedia annotations in multiple languages.

Remote supervised named entity recognition is a set of methods that improve the performance of a model by utilizing a large amount of unlabelled textual data when there is only a limited amount of manually labelled data available. This is achieved through the automatic generation of training data using an existing knowledge base or other external resources. Peng et al. [9] proposed a remotely supervised named entity recognition method based on learning from positive and unlabelled samples. This approach utilizes unlabelled data to mitigate the impact of noise. Lison et al. [10] utilized entity information and rules from a knowledge base to provide weak supervision for unlabelled data, resulting in the generation of labelled data. Ying et al. [11] proposed a remote label refinement model that can provide modification suggestions for remote data without the need of supervised labels. This approach reduces the quality requirements for the knowledge base.

In summary, weakly supervised named entity recognition algorithms can effectively utilize heuristic rules and remotely supervised information to enhance the performance of algorithm. However, these algorithms cannot effectively utilize unlabelled data and still have limitations in the incompletely supervised case.

B. SELF-TRAINING SEMI-SUPERVISED ALGORITHMS

Self-training semi-supervised algorithms make better use of unlabeled data and achieve data enhancement through iterative training methods. Niu et al. [12] employed a confidence-based approach to generate soft labels for unlabeled data, in addition to self-training. They then used these soft labels for self-training in addition to the labeled data. Xie et al. [13] proposed a self-training with noisy student approach to improve the performance of image classification tasks on ImageNet. Meng et al. [14] proposed a method for remotely supervised named entity recognition that improves the model's performance by incorporating self-training and noise-robust learning, augmented with a language model. Du et al. [15] proposed a highly scalable approach using self-training approach to improve the performance of pre-trained models on natural language understanding tasks.

However, the pseudo-labeled data obtained through the self-training approach may still introduce noise into the model, resulting in incorrect learning and, consequently, performance degradation.

C. FIRST-ORDER LOGIC CORRECTION ALGORITHMS

First-order logic correction algorithms represent the logical structures in natural language sentences as first-order logic formulas. These algorithms use inference rules to detect and correct logical errors within the formulas. Huang et al. [16] proposed an inference method based on semi-supervised inverse deduction learning to obtain more accurate conclusions by leveraging known facts and rules to infer unknown causal relationships. Li et al. [17] proposed a weakly-supervised method for named entity annotation based on learnable logic rules to automatically annotate named entities in a corpus. This approach reduces the dependence on annotated data. Zhou et al. [18] proposed integrating logical reasoning and machine learning to rectify the obtained perceptual results.

The above studies have made significant progress. However, self-training mainly focuses on the enhancement of data samples, and first-order logic focuses on error correction, both targeting specific challenges. Given this, in order to effectively address multiple challenges in weakly supervised environments, we propose a comprehensive method that merges self-training approach with first-order logic knowledge base correction strategies.

III. METHOD

In this paper, we propose a weakly supervised method for Chinese named entity recognition. The method address the problems of incomplete and inaccurate supervision by incorporating the self-training approach to alleviate the weak feature learning ability that arises from a limited number of entity categories. First-order logic rules are also defined to address the problem of inaccurate supervision. The overall structure of the method is depicted in Figure 1. The method is based on the pre-trained BERT model, utilizes a self-training approach to augment the data, and integrates a first-order logic knowledge base to correct the generated pseudo-labels. Since the method incorporates BERT, self-training, and a first-order logic knowledge base, the three modules involved in the method will be introduced separately in what follows.

A. BERT PRE-TRAINING MODEL

By leveraging the BERT model as a foundation, the teacher model and the student model are trained iteratively in a self-training process. A brief introduction to its principles is provided in the following:

BERT is a bidirectional pre-trained language model. It is based on the Transformer [19] architecture, which captures dependencies on arbitrary distances in the input sequence through the attention mechanism. The Transformer consists of multiple layers stacked on top of each other. Each layer consists of a multi-head self-attention sublayer and a feed-forward neural network sublayer. Specifically, the calculation of the multi-head self-attention component within these layers is as follows:

$$\begin{aligned} \boldsymbol{Q} &= X\boldsymbol{W}^{\boldsymbol{Q}} \\ \boldsymbol{K} &= X\boldsymbol{W}^{\boldsymbol{K}} \\ \boldsymbol{V} &= X\boldsymbol{W}^{\boldsymbol{V}} \end{aligned} \tag{1}$$

In this formula, X denotes the word embedding of the input sequence. $\boldsymbol{W}^{\boldsymbol{Q}}$, $\boldsymbol{W}^{\boldsymbol{K}}$, $\boldsymbol{W}^{\boldsymbol{V}}$ denote the learned weight matrices respectively. Then calculate the attention score:

$$z = \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{2}$$

In this formula, d_k denotes the dimension of the Q and K vectors. Dividing $\sqrt{d_k}$ in the equation makes the model's gradient descent more stable during training. Ultimately, a softmax function was used to normalize the scores for each word, enabling the encoder to distinguish the level of attention given to words with different scores.

During the pre-training phase, BERT utilizes a masked language model and next sentence prediction. Firstly, certain words in the input sequence are randomly masked, and then the model is tasked with predicting the masked words based on contextual information. Secondly, when given two sentences A and B, the model needs to determine whether B is the next sentence of A or not. The input embedding of BERT consists of three separate embeddings that incorporate character, sentence, and positional information. The spliced embedding is then input into the BERT encoder to obtain the output hidden layer features.

In summary, we adopt BERT as the foundational backbone in self-training approach. By harnessing the power of BERT, our method benefits from its pre-trained knowledge, enabling the model to effectively learn from labeled data.

B. FIRST-ORDER LOGIC KNOWLEDGE BASE

In the self-training process, the quality of generated pseudo-labels can impact the performance of the model. We summarize common errors in pseudo-labels and develop a first-order logic knowledge base to correct these biases. Table 1 lists five common entity labeling errors.

TABLE 1. Common entity labeling errors.

Starting label	Ending label
B-X	B-X
B-X	B-Y
B-X	E-Y
E-X	E-X
E-X	E-Y

In the label, B denotes the starting boundary of an entity, E denotes the ending boundary of an entity, and X and Y denote different entity categories. To address the five types of errors mentioned above, we formulate corresponding first-order logic statements for correction. These statements are ultimately integrated to form a first-order logic knowledge base. An example in the first-order logic knowledge base is depicted in Figure 2.

The logical formula in figure 2 describes that for any positions i and j, if i is labeled B-X and j is also labeled B-X, with a distance d between them and no intervening E-X tags, one of the following two actions will be taken:

- If the string beginning at position i corresponds to an entity in the database, and there exists a position k (where k > i) labeled E-X following i, then the label at position k will be updated to E-X.
- If the string starting at position i does not match any entity in the database and its length exceeds 10, then the label at position i will be changed to O.

Based on the first-order logic expressions mentioned above, we can effectively correct noisy pseudo-labels that start with B-X and end with B-X. In this manner, the first-order logic knowledge base is capable of identifying and correcting most errors within pseudo-label data, while also allowing

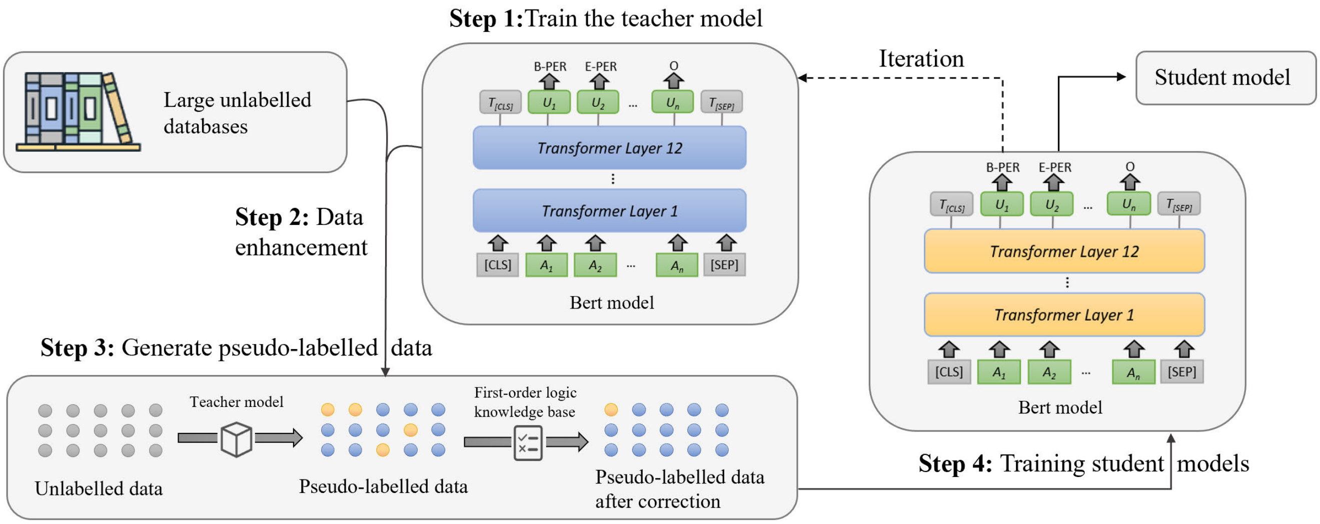


FIGURE 1. A weakly supervised Chinese named entity recognition method combining self-training.

```
∀ i, j ( L( i , "B-X" ) ∧ L( j , "B-X" ) ∧ Distance( i , j , d ) ∧ NoTagBetween( i , j , "E-X" ) →
  ( InDatabase( Substring( S , i , len ) ) →
    ∃ k ( k > i ∧ L( k , "E-X" ) )
  )
  ∧
  (¬InDatabase( Substring( S , i , len ) ) ∧ len > 10 →
    L( i , "O" )
  )
)
```

FIGURE 2. Example of first-order logic knowledge base.

for customization and modifications to meet specific requirements, thereby providing significant flexibility.

C. SELF-TRAINING CORRECTION METHOD WITH FIRST-ORDER LOGIC KNOWLEDGE BASE

As shown in Figure 1, the self-training correction method with first-order logic knowledge base is divided into four steps. Firstly, an initial teacher model is trained using a well-annotated dataset. This training is based on the Bert model and involves fine-tuning with an enhanced loss function to optimize performance. Secondly, unlabeled text data similar to the training set are added, and pseudo-labels are generated for this data using the trained teacher model. Thirdly, the original labeled data is combined with the generated pseudo-label data, and a student model is retrained. Finally, the student model is employed as a new teacher model, and the process is repeated. After multiple iterations, the final student model is produced.

Since the pseudo-labels generated by the self-training approach are not authentic labels, noise may be introduced to the subsequent steps by them, and the model may be misled into learning incorrect features. In order to minimize the noise effects caused by pseudo-labels on the model, we integrate logical inference with deep learning to correct the noise introduced by pseudo-labeling. The process, inspired by the application of inverse deductive learning to sentence prediction [20], is described as follows:

$$\begin{aligned} & X = \{x_1,\ x_2, \cdots, x_n\}, f \rhd \mathcal{O} \\ \text{s.t.} \quad & KB \vDash \mathcal{O}, \ \text{or} \\ & KB \vDash \Delta(\mathcal{O}), f \leftarrow \psi(f, \Delta(\mathcal{O})) \end{aligned} \tag{3}$$

In this formula, f represents a learned mapping function. It generates logical facts $\mathcal{O}$ from an unlabeled dataset X. Subsequently the consistency of $\mathcal{O}$ with knowledge base KB is checked. If $\mathcal{O}$ is consistent with KB, no changes are required. If inconsistencies arise, $\mathcal{O}$ is altered in accordance with KB to $\Delta(\mathcal{O})$, and f is updated to accommodate this change, ensuring that future outputs align with the knowledge base. This process of adjustment is facilitated by a specific function ψ, which recalibrates f to produce more accurate logical facts.

In our method, the correction process for pseudo-labels are represented by the following formula:

$$A_{LD} \leftarrow \{P_{LD}, K\ B, \mathcal{O}\} \tag{4}$$

In this formula, P_{LD} represents the pseudo-label data generated by the teacher model, $\mathcal{O}$ consists of logical facts extracted from P_{LD}, KB is a first-order logic knowledge base that validates these logical facts according to the rules of first-order logic, and A_{LD} denotes the corrected pseudo-labeled data.

During the process, the consistency between the facts in $\mathcal{O}$ and the knowledge in KB is checked. If inconsistencies are detected, P_{LD} is modified according to the logical discrepancies identified by KB, resulting in the generation of A_{LD}, which ensures alignment with the established knowledge base. Correction using first-order logic knowledge base enhance the annotation quality of pseudo-labeled data, effectively reducing the accumulation of noise during the self-training phase, which provides more robust data for the subsequent stages of model training.

In Chinese Named Entity Recognition, class imbalance is a common issue where some categories are significantly underrepresented in the training dataset. The imbalance can degrade the performance of model, as models tend to minimize overall errors during training, which typically results in good performance on majority classes but inadequate recognition of minority classes. To address this challenge, we integrate smoothing factors [21] and fading factors into the cross-entropy function:

$$p_t = \begin{cases} (1-p)^{\theta}, & y=1 \\ p^{\theta}, & \text{Other} \end{cases} \tag{5}$$

$$W_i = \frac{B}{T_i} \tag{6}$$

In this formula, p represents the probability of the predicted sample, B is the total number of entities, T_i is the number of entities of category i in all entities, and W_i is the static weight of a category in the current batch.

Equation (7) is defined by combining equations (5) and (6) with the cross-entropy method:

$$loss = W_i p_i CE_{loss} \tag{7}$$

By introducing smoothing factors, the model increases its focus on challenging samples. If the sample probability p approaches 1, the smoothing factor decreases, helping to prevent excessive loss accumulation on correctly classified samples with high confidence. Conversely, if p approaches 0, indicating a sample is challenging to classify, the smoothing factor increases, ensuring the model pays more attention to these samples during gradient updates. The introducing of fading factors makes the model adequately focuses on globally infrequent entity categories. Thus, fading factors are adjusted based on the frequency of occurrence of each category, balancing the weights across different categories.

The integration of smoothing and fading factors employs an intelligent strategy that considers aspects such as sample distribution and prediction difficulty. This approach allows the model to more effectively handle challenging samples, thus enhancing its performance in complex scenarios.

IV. EXPERIMENTS AND ANALYSES

A. TRAINING DETAILS

Our experimental data is from the public datasets ACE05 and MSRA, and the specific data details are presented in Table 2.

1) BERT MODEL TRAINING CONFIGURATION

We use BERT as our baseline model. The Adam optimizer is utilized with a learning rate schedule of 0.0001, batch sizes are set to 128, a dropout rate of 0.5 is applied, and 10 epochs of iteration are completed. A minimum running memory of 25GB is required by our model.

TABLE 2. Detailed statistics of ACE05 and MSRA.

Datasets	Label	Training set	Testing set
ACE05	PER	21771	2470
	ORG	10643	1261
	GPE	13327	2092
	LOC	2350	263
	FAC	2568	246
	VEH	1027	155
	WEA	628	42
MSRA	LOC	86849	7291
	ORG	103261	6977
	PER	51738	5824

2) EXPERIMENTAL EVALUATION METRICS

The metrics used in the experiment were precision (P), recall (R), and F1-score. Precision measures the proportion of correctly identified entities among recognized entities. Recall assesses the proportion of correctly identified entities out of all actual entities. The F1-score, calculated as the harmonic mean of Precision and Recall, serves as a comprehensive metric of model performance.

$$R = \frac{TP}{TP+FN} \tag{8}$$

$$P = \frac{TP}{TP+FP} \tag{9}$$

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

TP represents correctly predicted positive instances, FP represents negatives misclassified as positives, and FN represents positives misclassified as negatives.

B. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the weakly supervised Chinese named entity recognition method, comparative experiments between the baseline model BERT and our method are conducted, with the results presented in Table 3.

Experiment results demonstrate that, by incorporating the self-training approach and a first-order logic knowledge base into the baseline model, our method achieved an increase of 2.56% in F1-score on the ACE05 dataset and an increase of 0.35% on the MSRA dataset. The reasons are analysed as follows: (i) When there are insufficient samples of a specific label in the data, the model cannot adequately learn the features of that label, resulting in decreased recognition capabilities. Self-training expands the range of samples the model encounters, alleviating the problem of inadequate supervision and significantly improving the model's capability to learn the features of entities across various categories. (ii) The pseudo-labels generated by the self-training approach may be mislabeled, introducing noise to the training data of model. By integrating a first-order logic knowledge base into the self-training process, the problem of inaccurate supervision in the pseudo-labels is rectified. This noise reduction preventing the decline in model's recognition accuracy. (iii) By incorporating smoothing factors and fading factors into the loss function, the model alleviates the

TABLE 3. Comparative experiment results on ACE05 and MSRA.

Dataset	Model	Label	P/%	R/%	F1/%
ACE05	BERT	PER	81.70	86.79	84.17
		ORG	76.90	79.58	78.22
		GPE	76.55	84.71	80.42
		LOC	53.92	51.89	52.88
		FAC	58.33	63.64	60.87
		VEH	70.59	63.16	66.67
		WEA	52.38	66.67	58.67
		Total	76.75	81.59	79.10
	Our	PER	87.69	86.26	86.97
		ORG	76.69	84.99	80.63
		GPE	80.87	82.49	81.67
		LOC	58.91	52.77	55.67
		FAC	61.98	69.12	65.35
		VEH	73.45	64.71	68.65
		WEA	63.87	63.98	63.92
		Total	**81.02**	**82.31**	**81.66**
MSRA	BERT	LOC	92.66	93.19	92.92
		ORG	85.44	94.34	89.67
		PER	96.50	96.75	96.63
		Total	91.51	94.56	93.01
	Our	LOC	93.07	93.91	93.49
		ORG	85.81	94.56	89.97
		PER	97.08	96.41	96.74
		Total	**91.95**	**94.82**	**93.36**

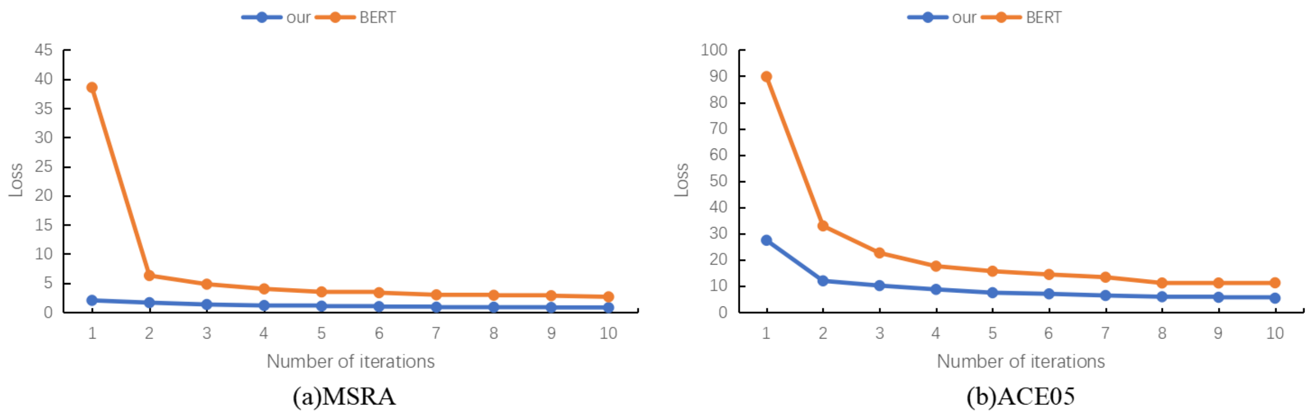


FIGURE 3. Speed variations of loss function convergence.

issue of over fitting, thereby enabling more efficient feature learning.

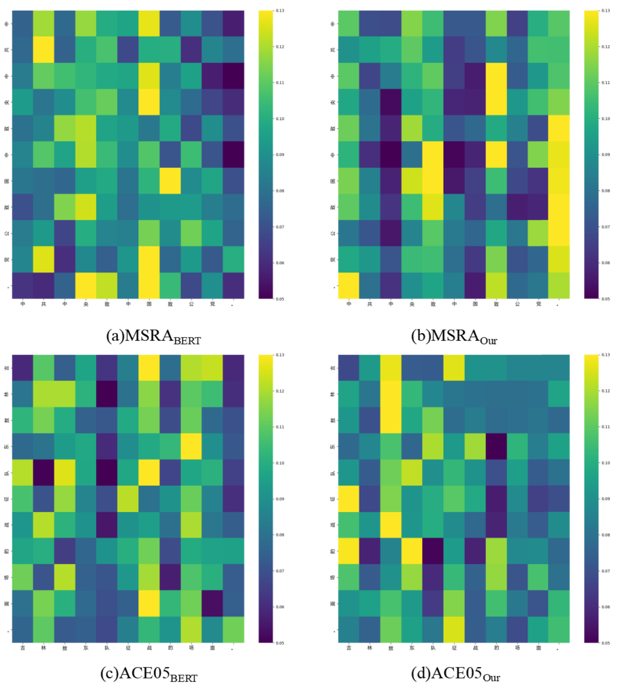
During the experiments, it was found that the method also accelerates the model's convergence speed on the loss function, as shown in Figure 3. This is due to our method significantly enhancing the model's ability to recognize various categories. The improvement in recognition capability directly boosts the overall learning efficiency of the model, thereby accelerating its convergence speed on the loss function.

Since our method aims to address the problem of Chinese named entity recognition in a weakly supervised environment, we also examine how its performance changes in an incompletely supervised setting. Four different levels of weakly supervised scenarios are designed: 10%, 20%, 40%, and 80% of the total sentences from the original ACE05 training set are selected as the training data, while the test set is kept completely unchanged. BERT is employed as the baseline model.

TABLE 4. Performance variations under incomplete supervision scenarios.

Rate	Model	P/%	R/%	F1/%
10%	BERT	65.00	49.49	56.20
	Our	**69.55**	**57.70**	**63.07**
20%	BERT	62.43	68.50	65.32
	Our	**72.05**	**67.45**	**69.67**
40%	BERT	68.43	73.70	70.97
	Our	**73.23**	**76.62**	**74.89**
80%	BERT	74.91	76.56	75.73
	Our	**75.18**	**79.64**	**77.34**

The experiment results are shown in Table 4. The combination of self-training and first-order logic knowledge base correction in the incompletely supervised environment leads to a significant improvement in performance. Among them, the most significant boost occurs when only 10% of



(a)MSRA$_{BERT}$ (b)MSRA$_{Our}$

(c)ACE05$_{BERT}$ (d)ACE05$_{Our}$

FIGURE 4. Visualization of attention weights.

the training samples are used. This may due to the following reasons: (i) With very few training samples, the BERT pre-trained model cannot effectively learn the features of the label categories. By employing the self-training approach, a larger number of labels from different categories are introduced, which enhances the model's ability to learn the features of these label categories. (ii) The error rate of pseudo-labeling generated by the very small training samples is also very high, resulting in a significant amount of noisy data that interferes with the model. Correcting the pseudo-labels using the knowledge base of first-order logic significantly reduces the impact of noisy data on the model.

After visualizing the attention weights of the various models, as shown in Figure 4, the more attention points converge towards yellow, the higher the attention weight. A more dispersed weight distribution is shown by the baseline supervised model on the ACE05 and MSRA datasets when self-training and a first-order logic knowledge base are not combined. By combining self-training with a first-order logic knowledge base, attention is then focused on entity categories. This approach is more aligned with human judgment. For example, in the visualization of ACE05 weights, the baseline model shows less attention to the nested entities in "Jilin Aodong team" whereas the method proposed in this paper assigns more attention weights to this entity.

To further validate the effectiveness of our method, we conducted comparative experiments with the Back Translation Method (BTM) [22], Synonym Replacement (SR), and Self-Training (ST) [4] methods using the ACE05 dataset, as shown in Table 5. From the experimental results, our method

TABLE 5. Performances of different data enhancement techniques on ACE05.

Model	P/%	R/%	F1/%
SR	78.12	80.97	79.51
BTM	78.66	81.01	79.81
ST	79.53	82.16	80.82
Our	**81.02**	**82.31**	**81.66**

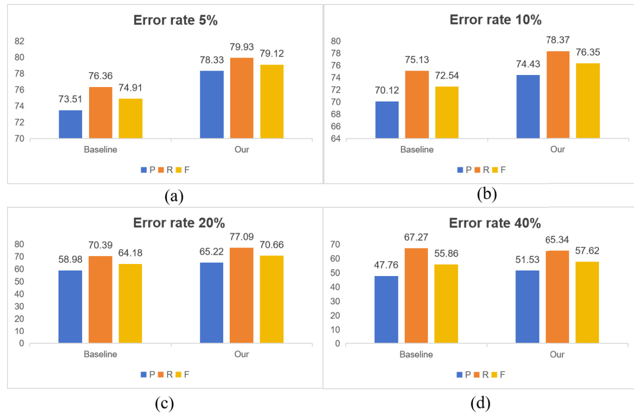


FIGURE 5. Performance variations under inaccurate supervision scenarios.

TABLE 6. Ablation experiment.

BERT	ST	first-order logic	smoothing factor	fading factor	F1/%
√					79.10
√	√				80.82
√	√	√			81.34
√	√	√	√		81.45
√	√	√	√	√	**81.66**

outperforms other data enhancement techniques, proving its effectiveness.

To verify the robustness of our method, comparative experiments are conducted on the ACE05 dataset. From all the labels in training set of ACE05 dataset, 5%, 10%, 20%, and 40% of them are randomly selected and then replaced with incorrect labels, which introduces errors in the label's location, type, or both. As experiment results shown in Figure 5, when faced with four different degrees of noisy labeled data, the F1 scores of our method consistently outperform the baseline model, demonstrating enhanced robustness against inaccuracies in labeling of our method.

Finally, to validate the proposed components in this paper, the model's ablation experiments on the ACE05 dataset are analyzed. The ablation experiments primarily use F1-score as the evaluation metric, and the experimental results are presented in Table 6. It can be seen that each component of the model is designed to enhance performance, which further validates the effectiveness of our proposed method.

V. CONCLUSION

This paper proposes a method for Chinese named entity recognition under the condition of weak supervision.

Specifically, this paper utilizes a self-training method to enhance the learning capability of model under insufficient supervision. A first-order logic knowledge base is defined to address the noisy pseudo-labels generated by the automatic labeling process, effectively reducing the negative impact of noise accumulation on model performance. Besides, by improving the loss function, the model can effectively learn features from sparse category samples.

Our method is evaluated based on ACE05 and MSRA datasets. Experiment results show that our method outperforms the baseline model, demonstrating an improvement in recognition performance. Our method also shows effectiveness in experiments of incomplete supervision, proving ability to leverage limited data; it exhibits significant robustness in experiments of inaccurate supervision, confirming suitability for handling noisy environments in weakly supervised scenarios.

VI. LIMITATION

We explore a weakly supervised named entity recognition method that utilizes pre-trained language models combined with self-training and first-order logic to address inaccuracies and incomplete supervision. As for the limitations, the first is that the potential bias introduced by the first-order logic knowledge base and its influence on model generalization may need to be optimized. Secondly, this paper does not thoroughly explore potential failure modes or edge cases in which the proposed method may not perform well. Finally, this paper does not address the interpretability of model outputs and their impact on downstream applications or decision-making processes.

REFERENCES

[1] C. Xiong, Z. Liu, J. Callan, and T.-Y. Liu, "Towards better text understanding and retrieval through kernel entity salience modeling," in *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jun. 2018, pp. 575–584.

[2] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 494–514, Feb. 2022.

[3] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *Nat. Sci. Rev.*, vol. 5, no. 1, pp. 44–53, Jan. 2018.

[4] B. Zoph, G. Ghiasi, T.-Y. Lin, Y. Cui, H. Liu, E. D. Cubuk, and Q. Le, "Rethinking pre-training and self-training," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 3833–3845.

[5] A. McCallum and W. Li, "Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons," in *Proc. 7th Conf. Natural Lang. Learn.*, 2003, pp. 188–191.

[6] J. Li, C. Zhu, S. Li, T. Zhao, and D. Zheng, "Exploiting Wikipedia priori knowledge for Chinese named entity recognition," in *Proc. 12th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD)*, Aug. 2016, pp. 698–707.

[7] L. Ratinov and D. Roth, "Design challenges and misconceptions in named entity recognition," in *Proc. 13th Conf. Comput. Natural Lang. Learn.*, 2009, pp. 147–155.

[8] J. Nothman, N. Ringland, W. Radford, T. Murphy, and J. R. Curran, "Learning multilingual named entity recognition from Wikipedia," *Artif. Intell.*, vol. 194, pp. 151–175, Jan. 2013.

[9] M. Peng, X. Xing, Q. Zhang, J. Fu, and X. Huang, "Distantly supervised named entity recognition using positive-unlabeled learning," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics*, 2019, pp. 2409–2419.

[10] P. Lison, J. Barnes, A. Hubin, and S. Touileb, "Named entity recognition without labelled data: A weak supervision approach," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, 2020, pp. 1518–1533.

[11] H. Ying, S. Luo, T. Dang, and S. Yu, "Label refinement via contrastive learning for distantly-supervised named entity recognition," in *Proc. Findings Assoc. Comput. Linguistics*, 2022, pp. 2656–2666.

[12] Y. Niu, F. Jiao, M. Zhou, T. Yao, J. Xu, and M. Huang, "A self-training method for machine reading comprehension with soft evidence extraction," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, 2020, pp. 3916–3927.

[13] Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le, "Self-training with noisy student improves ImageNet classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10684–10695.

[14] Y. Meng, Y. Zhang, J. Huang, X. Wang, Y. Zhang, H. Ji, and J. Han, "Distantly-supervised named entity recognition with noise-robust learning and language model augmented self-training," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2021, pp. 10367–10378.

[15] J. Du, E. Grave, B. Gunel, V. Chaudhary, O. Celebi, M. Auli, V. Stoyanov, and A. Conneau, "Self-training improves pre-training for natural language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol.*, 2021, pp. 5408–5418.

[16] Y.-X. Huang, W.-Z. Dai, J. Yang, L.-W. Cai, S. Cheng, R. Huang, Y.-F. Li, and Z.-H. Zhou, "Semi-supervised abductive learning and its application to theft judicial sentencing," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2020, pp. 1070–1075.

[17] J. Li, H. Ding, J. Shang, J. McAuley, and Z. Feng, "Weakly supervised named entity tagging with learnable logical rules," in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics 11th Int. Joint Conf. Natural Lang. Process.*, 2021, pp. 4568–4581.

[18] Z.-H. Zhou, "Abductive learning: Towards bridging machine learning and logical reasoning," *Sci. China Inf. Sci.*, vol. 62, no. 7, pp. 1–3, Jul. 2019.

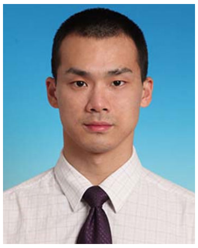
[19] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–9.

[20] L. Ouyang, R. Huang, Y. Chen, and Y. Qin, "A sentence prediction approach incorporating trial logic based on abductive learning," *Appl. Sci.*, vol. 12, no. 16, p. 7982, Aug. 2022.

[21] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[22] A. Wei Yu, D. Dohan, M.-T. Luong, R. Zhao, K. Chen, M. Norouzi, and Q. V. Le, "QANet: Combining local convolution with global self-attention for reading comprehension," 2018, *arXiv:1804.09541*.

XI TANG received the B.S. degree from Hebei Normal University, China, in 2020. She is currently pursuing the M.S. degree with Beijing Forestry University. Her main research interests include named entity recognition and terminology extraction.



DONGCHEN JIANG received the Ph.D. degree in computer software and theory from the State Key Laboratory of Software Development Environment, Beihang University, China, in 2013.

He is currently an Associate Professor with the School of Information Science and Technology, Beijing Forestry University. His main research interests include natural language processing and formal methods.