Chinese Relation Extraction of Apple Diseases and Pests Based on BERT and Entity Information

Mei Guo[1], Jiayu Zhang[1], Nan Geng[1,2,3], Yaojun Geng[1], Yongliang Zhang[1], and Mei Li[1](✉)

[1] College of Information Engineering, Northwest A&F University, Yangling, China
{nangeng,gengyaojun,limei}@nwsuaf.edu.cn
[2] Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Northwest A&F University, Yangling, China
[3] Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Northwest A&F University, Yangling, China

Abstract. Most existing methods for Chinese relation extraction suffer from fine-grained relation categories and unbalanced category distribution in the field of apple diseases and pests. To solve above problems, we construct the AppleRE dataset, which contains 28 relation categories and 20060 relation instances with the characteristic of richer categories than existing agricultural datasets. Then, we propose a relation extraction model BE-ARE introducing BERT and entity information, in which dynamic character representations and entities that reflects the unique meaning of Chinese words are utilized to enhance the data features. The performance of BE-ARE on AppleRE achieved a precision of 98.44%, a recall of 96.75% and an F1-score of 97.59%. The F1- score is increased by 1.77%–12.38%, which outperforms the comparison models. Experimental results demonstrate the competitiveness of BE-ARE when considering fine-grained classification and unbalanced category distribution. In addition, the proposed model shows the generalization on the public datasets in distinct domains.

Keywords: Chinese relation extraction · Apple diseases and pests · BERT · Entity information

1 Introduction

Apple plays an important role in agricultural production. Diseases and pests hinder the apple industry and cause production losses. Therefore, their control must receive great attention. Most authoritative information in this domain is stored in unstructured text. Chinese relation extraction (CRE) of apple diseases and pests aims to extract semantic relations between the two entities marked in a given sentence and to convert unstructured text into structured data, which contributes to knowledge graph and intelligent Q&A. For example, the sentence

G. Memmi et al. (Eds.): KSEM 2022, LNAI 13370, pp. 579–592, 2022.
https://doi.org/10.1007/978-3-031-10989-8_46

"苹果白粉病危害苹果幼芽 (Apple powdery mildew harms the young buds of apples)" has two entities, "苹果白粉病 (apple powdery mildew)" and "幼芽 (young buds)", and the relation between them is "危害 (harm)".

Deep learning has been successfully applied in many fields [5,10], especially in relation extraction. CNN [15] and RNN are used to learn underlying features automatically. [16] proposed the BLSTM model based on RNN to deal with long-distance patterns. [18] developed the Att-BLSTM model to capture the most important semantic information in a sentence. With the emergence of the BERT model [2,12] designed the R-BERT model and utilized the entities to tackle relation extraction. [11] used entity context information to enhance BERT. Entities are important information in relation extraction. [3,4] introduced entity descriptions and entity types to achieve relation extraction. They have limitations of no entity descriptions and limited entity types. In CRE, entity information can overcome these limitations due to the unique meaning of Chinese words.

Relation extraction in agriculture has received extensive attention [7,14]. However, there are currently no publicly available datasets and models in the field of apple diseases and pests. The problems of fine-grained relation categories and unbalanced category distribution exist in this domain that need to be solved urgently. Therefore, the main contributions of this paper include:

- We construct a CRE dataset of apple diseases and pests named AppleRE. The dataset contains 28 categories and has the characteristics of fine-grained classification.
- We propose a CRE model named BE-ARE. The proposed model improves performance of relation extraction by introducing BERT and entity information.
- The BE-ARE model achieves the optimal results on the AppleRE dataset and also outperforms other models on the public datasets in distinct domains.

The remainder of this paper is organized as follows. Section 2 describes the details of AppleRE dataset. Section 3 presents the details of BE-ARE model. Section 4 and Sect. 5 introduce the experimental results and discussion. Section 6 concludes the paper.

2 AppleRE Dataset

2.1 Data Collection

Books and Internet texts, as carriers of apple disease and pest information, have recorded lots of data. Compared with Internet texts, books are more normative, professional and authoritative. Therefore, to ensure the quality of the dataset, we selected seven authoritative books as data sources under the guidance of experts from Northwest A&F University, as shown in Table 1. These books introduce more than fifty common apple diseases and sixty apple pests, and comprehensively cover the relations between entities of apple diseases and pests, which provide a knowledge base for the construction of the dataset.

Table 1. Data sources of dataset.

Book title	Author	Publisher	Pubdate
Integrated prevention and control of apple diseases and pests	Shishan Guo	Henan Science and Technology Press	1992
Prevention and control of apple diseases and pests	Jinyou Wang	Golden Shield Press	1992
Prevention and control of apple diseases and pests	Yizhi Sun	Shaanxi Science and Technology Press	1997
Prevention and control of apple diseases and pests with pictures	Zigang Cao	China Agriculture Press	1999
New diagnosis and control of apple pests and diseases	Youguang Chen	Shaanxi Science and Technology Press	2003
Integrated prevention and control technology of apple diseases and pests	Jinying Wu	Northwest Agriculture and Forestry University Press	2009
Prevention and control of diseases, pests and weeds in modern apple production	Dongping Li	Chemical Industry Press	2018

2.2 Relation Categories

The relations between entities of apple diseases and pests can be divided into three coarse-grained categories, namely “苹果 (apple)”, “病虫害 (disease and pest)” and “防治 (control)” after analysis of existing agricultural research [7,14] and guidance from experts. Based on this, we further divided the three categories into fine-grained categories. “苹果 (apple)” is divided into “别称 (alias)”, “分布 (distribute)” and “品种 (variety)”. “病虫害 (disease and pest)” is divided into “属于 (belong to)”, “包含 (include)”, “取食 (feed)”, “危害 (harm)”, “侵染 (infect)”, “寄生 (parasitism)” and “越冬 (overwinter)”. “防治 (control)” is divided into “农业防治 agricultural control (增施 increase、清除 clear、种植 plant、翻耕 plow、灌溉 irrigate、撒施 spread)”, “生物防治 biological control (保护 protect)”, “物理防治 physical control (涂抹 smear、捕杀 catch、诱杀 trap and kill、套袋 bagging)” and “化学防治 chemical control (喷洒 spray、消毒 disinfect、浸泡 soak、杀灭 kill、禁用 ban、防治 prevention and control)”. We also defined the category “unknown”, which describes entities that have no relation. Finally, 28 fine-grained relation categories are obtained in the dataset.

2.3 Relation Annotation

We annotated the relation instances based on relation categories. The specific annotation process is shown in Fig. 1. Given the raw sentence with the marked entities [17], the relation instances are composed of entity pairs, their corresponding relationship and the raw sentence. The format of instance is like “entity1 entity2 relation sentence”. When all the annotations are completed, we proofread the entire dataset to ensure the correctness of all instances. Finally, the AppleRE dataset contains 20060 relation instances.

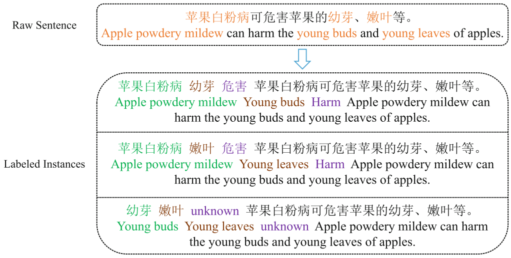

Fig. 1. An example from raw sentence to labeled instances.

2.4 Dataset Features

AppleRE is an agricultural dataset compared with general corpora, such as SanWen [13], RECDI [9], PCNN-A [14] and AgFeature [7]. Among them, SanWen belongs to the literature field. RECDI belongs to the medicine domain. PCNN-A and AgFeature belong to the agriculture field. Their comparison information can be found in Table 2.

Table 2. Comparison between AppleRE and other corpora.

Dataset	Domain	Category	Quantity	Max length
SanWen	Literature	10	21240	88
RECDI	Medicine	6	3429	-
PCNN-A	Agriculture	5	7085	-
AgFeature	Agriculture	7	24764	128
AppleRE	Agriculture	28	20060	159

"-" denotes unknown value.

Compared with these general corpora, the features of the AppleRE dataset are summarized as follows:

(1) AppleRE contains 28 relation categories. As shown in Table 2, AppleRE provides rich categories compared with corpora in different fields. The categories of other agricultural datasets belong to coarse-grained classification. AppleRE provides more fine-grained categories. Especially for "防治 (control)", fine division is committed to achieving precise prevention and control.

(2) The relation category distribution is unbalanced in AppleRE. For example, the instances of relations "撒施 (spread)", "灌溉 (irrigate)"and "消毒 (disinfect)"are 69, 41 and 38, all of which are less than 100. The proportions in the dataset are 0.34%, 0.2% and 0.19%, which account for less than 0.5%. They are divided into training set, test set, and validation set with few instances, which can easily cause extraction errors.

(3) The sentences of relation instances are long text sequences in AppleRE. Table 2 counts the lengths of sentences contained in relation instances of different dataset. The maximum sentence length in AppleRE is 159, which is longer than others. The sentence length provides a reference for the parameter value in subsequent experiments.

3 BE-ARE Model

In this section, we present our BE-ARE model in detail. As shown in Fig. 2, the BE-ARE model takes the character-based BiGRU structure as the basic framework and contains four layers. The BERT layer is used to obtain the dynamic character representations and entity representations of the text. The BiGRU layer can mine the context information and obtain the semantic feature vectors. The attention layer assigns different weights to the feature vectors to enhance the model performance. The classification layer splices the output of the attention layer and the entity representations obtained by the BERT layer to realize relation extraction. This section will focus on the BERT layer and the classification layer. The details of BiGRU and attention can be found in [1] and [18].

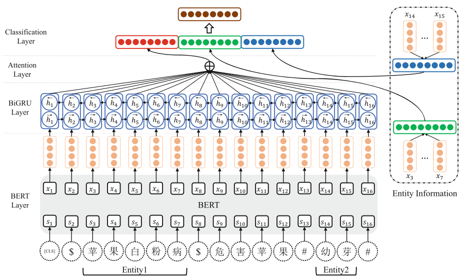

Fig. 2. BE-ARE model architecture.

3.1 BERT Layer

The BERT layer aims to obtain the character vector representations and entity vector representations of the apple disease and pest text. To identify the entity pairs in a sentence, we adopt the tokens "$" and "#" to mark the entity1 and entity2, and also add "[CLS]" to the beginning of each sentence mentioned in [12]. As shown in Fig. 3, after insertion of the tokens, the sentence with two entities will become to: "[CLS]$ 苹果白粉病 $ 危害苹果 #幼芽 #(Apple powdery mildew harms the young buds of apples)".

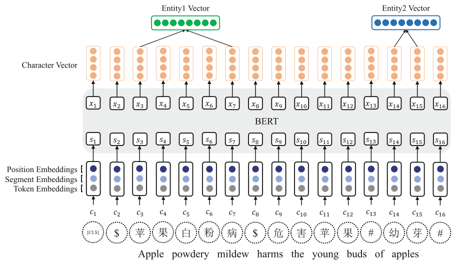

Fig. 3. Character vectors and entity vectors generated by BERT.

Given a sentence S consisting of n characters $S = \{c_1, c_2, \ldots, c_n\}$ and the two entities are $e_1 = \{c_a, c_{a+1}, \ldots, c_b\}$ and $e_2 = \{c_p, c_{p+1}, \ldots, c_q\}$, the input of BERT can be expressed as $\{[\text{CLS}], c_1, c_2, \ldots, \$, c_a, \ldots, c_b, \$, \ldots, \#, c_p, \ldots, c_q, \#, \ldots, c_n\}$, where c_i is i-th character of S; a and b are the start and end indices of e_1 in S; p and q are the start and end indices of e_2 in S. c_i is represented by token embedding, segment embedding and position embedding. The summation of these three embeddings denoted as s_i is fed into the BERT layer and the final output from BERT is x_i. Vectors x_a to x_b are from BERT for entity e_1, and x_p to x_q are from BERT for entity e_2. We utilize average operations to obtain vector representations of the two entities. The character vector E_S is shown in Eq. (1) and the entity vectors E_{e_1} and E_{e_2} are shown in Eq. (2) and Eq. (3):

$$E_S = [x_1, x_2, \ldots, x_n] \tag{1}$$

$$E_{e_1} = \frac{1}{b-a+1} \sum_{m=a}^{b} x_m \tag{2}$$

$$E_{e_2} = \frac{1}{q-p+1} \sum_{m=p}^{q} x_m \tag{3}$$

where m represents the position of the entity character in S.

3.2 Classification Layer

The classification layer is used to extract the relations between the entity pairs of apple diseases and pests. For example, the entity e_1 "苹果白粉病 (apple powdery mildew)" and e_2 "幼芽 (young buds)" are Chinese words with specific meanings. The former implies an apple disease and the latter refers to the parts of apple, both of which provide additional information for relation extraction. Therefore, the entity information in a sentence can help improve the capability of the BE-ARE model.

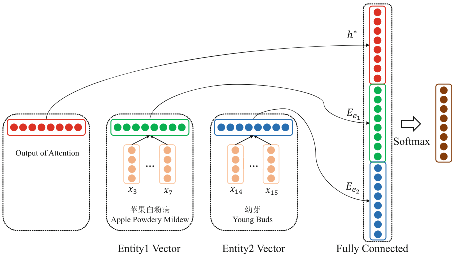

Fig. 4. Classification layer architecture.

As shown in Fig. 4, the feature fusion vector combines the output of the attention mechanism and the entity information, denoted as H_e, which is concatenated by h^*, E_{e_1} and E_{e_2} as shown in Eq. (4). Then the feature vector is fed into the softmax classifier to calculate the probabilities of each relation as shown in Eq. (5) and finally achieves relation extraction as shown in Eq. (6).

$$H_e = [h^* \oplus E_{e_1} \oplus E_{e_2}] \tag{4}$$

$$p = \mathrm{softmax}(WH_e + b) \tag{5}$$

$$\hat{y} = \arg\max p \tag{6}$$

where h^* is the output of the attention layer; p is the final probability expression; W is the weight parameter; b is the bias parameter and $\hat{y}$ is the predicted relation category. We adopt cross-entropy in [6] as the loss function.

4 Experiments

4.1 Experimental Settings

Dataset. We carry our experiments on different datasets, including AppleRE, SanWen and FinRE [6]. The AppleRE dataset contains 20060 relation instances with 12036, 4012 and 4012 instances for training, testing and validation respectively.

Experimental Environment. Ubuntu 16.04OS, GeForce RTX 2080Ti GPU, PyTorch1.2.0 and Python3.6.

Parameter Settings. The parameters of BERT layer refer to [2]. The initial learning rate is set to 2e-5 with AdamW optimization [8]. GRU size is set to 250. Dropout is set to 0.5 to avoid overfitting. Batch size is 16 and epoch is 10.

The character embedding of comparison model are pretrained by word2vec and the embedding size is 100 [6].

Evaluation Metrics. Precision, recall and F1-score are applied in the experiments.

4.2 Experiments on AppleRE

Several popular models, such as CNN [15], BLSTM [16], Att-BLSTM [18], BERT [2] and R-BERT [12] are compared to verify the effectiveness of BE-ARE on AppleRE. Table 3 shows that BE-ARE achieves optimal results among all the models. CNN performs the worst. The apple disease and pest corpus are long text sequences but CNN is inappropriate to exploit long-distance contextual information. The performance of the BERT-based models is better than the LSTM-based models. This is because the embeddings of the former produce richer and dynamic vector representations of characters than the character embeddings of the latter. Compared with BERT and R-BERT, BE-ARE utilizes BiGRU and attention mechanism to improve model performance.

Table 3. Results of different models on AppleRE.

Model	Precision	Recall	F1
CNN	82.94	87.61	85.21
BLSTM	91.99	89.87	90.92
Att-BLSTM	93.87	93.57	93.72
BERT	94.57	95.17	94.87
R-BERT	95.57	96.07	95.82
BE-ARE	**98.44**	**96.75**	**97.59**

Table 4. Performance of different models for coarse-grained category on AppleRE.

Relation	CNN	BLSTM	Att-BLSTM	BERT	R-BERT	BE-ARE
苹果 apple	97.49	98.65	98.99	99.83	99.74	**99.83**
病虫害 disease and pest	96.77	97.66	98.41	98.96	**99.31**	99.03
防治 control	92.75	96.58	98.17	99.12	**99.36**	99.28

Table 5. Performance of different models for fine-grained category on AppleRE.

Relation	CNN	BLSTM	Att-BLSTM	BERT	R-BERT	BE-ARE
别称 alias	97.80	99.50	99.26	100.00	**100.00**	99.75
分布 distribute	98.42	99.64	99.82	99.82	99.82	**99.82**
危害 harm	87.21	93.42	94.49	95.61	96.85	**97.96**
增施 increase	90.91	95.52	97.06	97.06	97.06	**97.06**
清除 clear	93.83	95.82	98.07	97.73	99.35	**99.35**
涂抹 smear	82.05	85.71	95.65	95.77	95.77	**100.00**
喷洒 spray	88.80	96.81	98.08	98.72	98.51	**98.73**
消毒 disinfect	72.73	61.54	72.73	66.66	66.66	**88.89**
越冬 overwinter	95.67	97.56	99.00	97.03	99.25	**99.50**
品种 variety	97.70	98.15	98.16	99.54	100.00	**100.00**
防治 prevention and control	81.08	93.33	90.00	97.37	97.37	**98.67**
侵染 infect	100.00	92.31	100.00	100.00	100.00	**100.00**
种植 plant	88.89	100.00	100.00	100.00	100.00	**100.00**
浸泡 soak	94.74	85.72	94.74	94.74	**100.00**	94.74
套袋 bagging	0.00	66.67	66.67	66.67	**90.91**	88.89
翻耕 plow	85.71	75.00	100.00	100.00	100.00	**100.00**
寄生 parasitism	91.80	93.10	94.74	**100.00**	90.56	98.24
灌溉 irrigate	80.00	80.00	90.91	90.91	90.91	**100.00**
禁用 ban	88.89	94.12	94.12	88.89	**94.12**	88.89
撒施 spread	66.67	76.92	75.00	88.89	70.59	**100.00**
保护 protect	75.00	93.33	93.33	94.12	**100.00**	93.33
包含 include	98.31	100.00	100.00	100.00	**100.00**	96.55
杀灭 kill	67.86	86.37	90.91	97.56	97.56	**97.67**
诱杀 trap and kill	88.09	94.44	98.63	97.30	100.00	**100.00**
取食 feed	86.92	93.33	94.78	95.45	97.34	**97.78**
属于 belong to	99.20	100.00	100.00	100.00	100.00	**100.00**
捕杀 catch	85.71	92.31	88.89	96.00	**100.00**	96.00

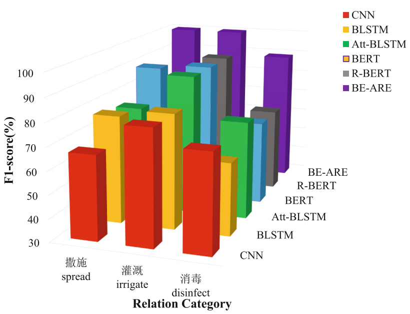

Fig. 5. F1-scores of unbalanced relation category.

Table 4 and Table 5 show the F1-scores of the models under different relation classification. The results show that under the coarse-grained classification, the performance of each model is similar because of few relation categories and a similar number of corresponding relation instances. However, under the fine-grained classification, the performance of models is different and the advantages of the BE-ARE model are more prominent. The F1-scores of different models for each category on AppleRE show similar differences. These models can fully learn the characteristics of relations during training when the relation category corresponds to many instances, such as "分布 (distribute)" and "危害 (harm)" . However, due to unbalanced distribution of relation categories, these models perform poorly at some relations, such as "撒施 (spread)", "灌溉 (irrigate)" and "消毒 (disinfect)", which have fewer instances resulting in models being difficult to fully learn the corresponding features. In comparison, BE-ARE is slightly impacted by the unbalanced distribution of relation category as shown in Fig. 5.

4.3 Experiments on the Public Datasets

To verify the generalization of BE-ARE, we selected two public datasets SanWen and FinRE. Table 6 lists the F1-scores of all models and the results of CNN, BLSTM and Att-BLSTM are from [6]. The BE-ARE model yields the highest F1-scores on the SanWen and FinRE datasets, which are 71.29% and 50.79%, respectively. There are positional overlaps between entity pairs in SanWen. The relation instances for each category in FinRE are insufficient. BE-ARE adopts BERT and the entity information to further enhance the capability to extract relations, thus achieving best performance. Experiments demonstrate the generalization of BE-ARE on the datasets in different fields.

Table 6. F1-scores of models on SanWen and FinRE.

Model	SanWen	FinRE
CNN	59.42	41.47
BLSTM	61.04	42.87
Att-BLSTM	59.48	41.48
BERT	70.26	48.07
R-BERT	67.89	50.01
BE-ARE	**71.29**	**50.79**

5 Discussion

5.1 Ablation Study

Effect of BERT Embedding. To analyze the effect of BERT, we conduct comparative experiments by considering randomly initialized embedding and pretrained character embedding. The randomly initialized embedding adopts the default PyTorch[1] method and its embedding size is the same as that of the pretrained character embedding.

The F1-score of the model with BERT is higher than those of models with randomly initialized and pretrained embedding, which increases by 4.4% and 2.5% respectively in Table 7. Randomly initialized and pretrained embedding can produce static and single vector representations of characters in sentences, which lack relevance analysis of semantic context. However, BERT provides dynamic vector representations of characters and contains richer semantic information, which has stronger text comprehension capabilities. Hence, the experimental results demonstrate that BERT contributes to improving the effective performance of BE-ARE.

Table 7. F1-scores of different embeddings on BE-ARE.

Model		F1	Improvement
BE-ARE	Random	93.19	4.40
BE-ARE	Word2vec	95.09	2.50
BE-ARE	BERT	97.59	-

Effect of Entity Information. To further explore the contribution of entity information used in the classification layer, we analyze apple disease and pest entities with different strategies, namely No entity, only Entity1, only Entity2 and joint Entity1 and Entity2.

Table 8 shows that the F1-score of strategy with two entities increases by 2.73%, 1.42% and 1.54% respectively compared with others. When there are no entities in the classification layer, the tokens "$" and "#" locate two entities and make the BERT output contain the location information of two entities. When one entity is introduced in this layer, the model effect is further improved. The model performs best when there are two entities in this layer. The results indicate that entity information in the classification layer can bring additional features to BE-ARE, which is conducive to the relation extraction.

Table 8. F1-scores of different entity information on BE-ARE.

Model		F1	Improvement
BE-ARE	No entity	94.86	2.73
BE-ARE	Entity1	96.17	1.42
BE-ARE	Entity2	96.05	1.54
BE-ARE	Entity1 + Entity2	97.59	-

[1] https://pytorch.org/docs/stable/generated/torch.nn.Embedding.html.

5.2 Effect of Sentence Lengths

To explore the effect of sentence lengths on BE-ARE, we selected 140, 150 and 160 as sentence lengths according to the sentence feature of AppleRE. As shown in Fig. 6, with the increase of the sentence length, the performance of BE-ARE first rises and then falls. The reason for this phenomenon is that when the sentence length is set too short, sentences will be intercepted and result in loss of entity and relationship information. When the sentence length is too long, too much invalid information will be filled into the sentences and cause a decrease in model performance. Although the performance of BE-ARE decreases as the sentence length increases, it still has great advantages and benefits more from long-distance text compared with other models.

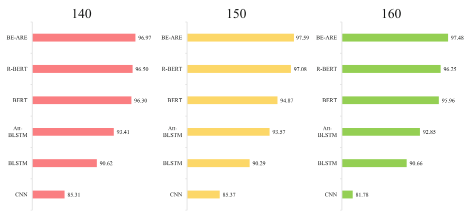

Fig. 6. F1-scores of models under different sentence lengths.

5.3 Case Study

To verify the superiority of BE-ARE, we analyzed an example of AppleRE. As shown in Table 9, the relation between “过氧乙酸 (peroxyacetic acid)”and “病菌 (pathogen)”is “杀灭 (kill)”. CNN cannot recognize the relation between the two entities. The comparison models are interfered by the relation “防治 (prevention and control)”between “过氧乙酸 (peroxyacetic acid)”and “腐烂病 (rot disease)” in the sentence and cause the relation to be incorrectly identified as “防治 (prevention and control)”. Thanks to BERT and entity information, BE-ARE obtains richer semantic representations and entity features, which strengthen the model capabilities to make correct predictions.

Table 9. Predictions for different models.

Sentence	过氧乙酸主要用于防治腐烂病和杀灭枝干上越冬的轮纹病、炭疽病病菌。 Peroxyacetic acid is mainly used to prevent and control rot disease and to kill the overwintering rot and anthracnose pathogens on the branches and trunks	
Entity	Entity1: 过氧乙酸 peroxyacetic acid	Entity2: 病菌 pathogen
Relation	杀灭 kill	
Prediction	CNN: unknown	BLSTM: 防治 prevention and control
	Att-BLSTM: 防治 prevention and control	BERT: 防治 prevention and control
	R-BERT: 防治 prevention and control	BE-ARE: 杀灭 kill

6 Conclusion

In this paper, the AppleRE dataset was constructed to solve the problem of limited data in the field of apple diseases and pests. Compared with other corpora, the relation categories in AppleRE are more fine-grained. Moreover, the BE-ARE model was proposed to explore the relations between the entity pairs and introduced BERT and entity information. BE-ARE outperforms previously proposed models on the datasets in distinct domains, such as Chinese apple diseases and pests (AppleRE), literature (SanWen) and financial (FinRE). Experiments show the effectiveness of BE-ARE when dealing with fine-grained relation categories and unbalanced category distribution. In future work, we will focus on introducing different features to solve the task of relation extraction.

Acknowledgements. This work is supported by the Key Research and Development Program of Shaanxi (No. 2019ZDLNY07-06-01).

References

1. Cho, K., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation (2014). arXiv preprint: arXiv:1406.1078
2. Devlin, J., Chang, M., Lee, K., Toutanova, K.: Bert: pre-training of deep bidirectional transformers for language understanding (2018). arXiv preprint arXiv:1810.04805
3. Ji, G., Liu, K., He, S., Zhao, J.: Distant supervision for relation extraction with sentence-level attention and entity descriptions. In: Proceedings of the AAAI Conference on Artificial Intelligence (2017)
4. Lee, J., Seo, S., Choi, Y.: Semantic relation classification via bidirectional LSTM networks with entity-aware attention using latent entity typing. Symmetry **11**(6), 785 (2019)
5. Li, Y., Song, Y., Jia, L., Gao, S., Li, Q., Qiu, M.: Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning. IEEE Trans. Industr. Inf. **17**(4), 2833–2841 (2020)
6. Li, Z., Ding, N., Liu, Z., Zheng, H., Shen, Y.: Chinese relation extraction with multi-grained information and external linguistic knowledge. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 4377–4386 (2019)
7. Liu, Z., Chen, Y., Dai, Y., Guo, C., Zhang, Z., Chen, X.: Syntactic and semantic features based relation extraction in agriculture domain. In: International Conference on Web Information Systems and Applications, pp. 252–258 (2018)
8. Loshchilov, I., Hutter, F.: Fixing weight decay regularization in adam (2017). arXiv preprint: arXiv:1711.05101
9. Qi, T., et al.: KeMRE: knowledge-enhanced medical relation extraction for Chinese medicine instructions. J. Biomed. Inform. **120**, 103834 (2021)
10. Qiu, H., Zheng, Q., Msahli, M., Memmi, G., Qiu, M., Lu, J.: Topological graph convolutional network-based urban traffic flow and density prediction. IEEE Trans. Intell. Transp. Syst. **22**(7), 4560–4569 (2020)
11. Wan, Y., Sun, L., Zhao, P., Wang, J., Tu, S.: Relation classification based on information enhanced BERT. J. Chin. Inf. Process. **35**(3), 69–77 (2021)

12. Wu, S., He, Y.: Enriching pre-trained language model with entity information for relation classification. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pp. 2361–2364 (2019)
13. Xu, J., Wen, J., Sun, X., Su, Q.: A discourse-level named entity recognition and relation extraction dataset for Chinese literature text (2017). arXiv preprint: arXiv:1711.07010
14. Yue, Y., et al.: Agricultural pest and disease relation extraction based on multi-attention mechanism and distant supervision. J. Anhui Agric. Univ. **47**(4), 5 (2020)
15. Zeng, D., Liu, K., Lai, S., Zhou, G., Zhao, J.: Relation classification via convolutional deep neural network. In: Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 2335–2344 (2014)
16. Zhang, D., Wang, D.: Relation classification via recurrent neural network (2015). arXiv preprint: arXiv:1508.01006
17. Zhang, J., Guo, M., Geng, Y., Li, M., Zhang, Y., Geng, N.: Chinese named entity recognition for apple diseases and pests based on character augmentation. Comput. Electron. Agric. **190**, 106464 (2021)
18. Zhou, P., et al.: Attention-based bidirectional long short-term memory networks for relation classification. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pp. 207–212 (2016)