

## Article

# Multi-Feature Fusion Event Argument Entity Recognition Method for Industrial Robot Fault Diagnosis

Senye Chen <sup>1,\*</sup>, Lianglun Cheng <sup>1</sup>, Jianfeng Deng <sup>2</sup> and Tao Wang <sup>2</sup>

<sup>1</sup> School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China

<sup>2</sup> School of Automation, Guangdong University of Technology, Guangzhou 510006, China

\* Correspondence: 2112005024@mail2.gdut.edu.cn

**Abstract:** The advance of knowledge graphs can bring tangible benefits to the fault detection of industrial robots. However, the construction of the KG for industrial robot fault detection is still in its infancy. In this paper, we propose a top-down approach to constructing a knowledge graph from robot fault logs. We define the event argument classes for fault phenomena and fault cause events as well as their relationship. Then, we develop the event logic ontology model. In order to construct the event logic knowledge extraction dataset, the ontology is used to label the entity and relationship of the fault detection event argument in the corpus. Additionally, due to the small size of the corpus, many professional terms, and sparse entities, a model for recognizing entities for robot fault detection is proposed. The accuracy of the entity boundary determination of the model is improved by combining multiple text features and using the relationship information. Compared with other methods, this method can significantly improve the performance of entity recognition of the dataset.

**Keywords:** event logic knowledge graph; fault detection; event argument entity recognition; multi-feature Fusion



**Citation:** Chen, S.; Cheng, L.; Deng, J.; Wang, T. Multi-Feature Fusion Event Argument Entity Recognition Method for Industrial Robot Fault Diagnosis. *Appl. Sci.* **2022**, *12*, 12359. <https://doi.org/10.3390/app122312359>

Academic Editor: Yu-Dong Zhang

Received: 18 October 2022

Accepted: 28 November 2022

Published: 2 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In today's industrial world, industrial robots are widely used and play an important role in the production of large and medium-sized companies. In view of the complex structure and serious failures of industrial robots, a sudden failure of industrial robots can cause delays in construction because of the loss of production [1]. To ensure the safety of robots, intelligent and efficient fault detection is vital [2]. Data-driven methods have been widely used, but knowledge-based fault detection methods have not yet been explored [3–5]. Using knowledge graphs to represent knowledge in the field of mechanical fault detection can improve the efficiency and accuracy of mechanical fault detection.

We collected mechanical fault detection logs as a fault relevant corpus. There are numerous logical relationships between events in the fault relevant corpus which can be helpful for the fault diagnosis. For example, the high humidity in the operating environment of the motor causes the motor to generate excessive heat after starting. There is an event logic relationship in this sentence. However, most knowledge graphs generally take nominal entities and their relationships as the research objects, lack the mining of logical knowledge of events, and have poor reasoning ability. Using an event logic knowledge graph as a knowledge representation technique, we can track how events evolve and further examine what knowledge is. The event logic knowledge graph is more powerful, more in line with the objective laws of human cognition, and more useful for practical purposes. To construct a graph, it is necessary to extract events from a large number of fault relevant corpora according to the ontology model. However, the event classes defined by most ontology construction methods are typically divided into fault phenomenon classes and fault cause classes, which results in coarse knowledge granularity, which is not conducive

to knowledge reasoning. Consequently, the best way to obtain valuable fault detection knowledge is to decompose event classes through finer-grained ontology construction and to recognize entities effectively. As a result of the small corpus size, technical terms, and sparse entities, existing deep learning methods perform poorly in event argument entity recognition on mechanical fault relevant corpora. First of all, many existing methods enhance the recognition effect by improving the input features through vocabulary enhancement, which is easily susceptible to information loss. Second, the blurred lexical boundaries of units in Chinese can lead to word segmentation errors and propagation problems. Lastly, there are numerous relations among the event argument entities in the mechanical fault relevant corpus. Incorporating the relationship information between event argument entities can help the model to better understand the semantic structure, but most methods do not combine the entity-relationship information for prediction, and are unable to infer semantic meanings through the relationship.

In light of the above motivation, we propose MF-GCN, which is a model for recognizing event argument entities on fault relevant corpora. By combining the multi-text feature and entity-relationship of event argument entities, MF-GCN is able to achieve ideal results in event argument entity recognition on fault relevant corpora. MF-GCN does not employ a vocabulary enhancement method, but instead uses character-level embedding as the basic unit. MF-GCN utilizes pretrained Word2vec embedding layers to obtain character-level vector representations. In terms of multi-feature fusion, MF-GCN mainly extracts local context features, context features, and context salience features. Local context features are important in determining entity semantics. Context features enable each entity in a sentence to capture long-distance dependency information, and context salience features can integrate local context information with long-distance dependency information better. The MF-GCN model also extracts global context features with the aid of stacked bidirectional long short-term memory (Bi-LSTM), which improves its representation capability. The first-stage argument entity prediction is performed by combining multiple text features. To assist in the identification of the semantic structure, it is necessary to introduce entity-relationship information. MF-GCN uses a bidirectional graph convolutional network (Bi-GCN) to modify entity boundaries with manually labeled relationship information, thereby improving the accuracy of entity boundary determination in the second stage of the model. In summary, this paper makes the following contributions:

- We use the labeling method described in 3.1 to label the data to develop a fine-grained event dataset.
- We propose MF-GCN, which is a model for recognizing event argument entities on fault relevant corpora.
- In the case study, it is the first time that a large corpus of mechanical fault detection logs was collected and a fine-grained ontology of fault detection is constructed.

## 2. Related Work

### 2.1. Ontology Construction of Fault Detection

Researchers have conducted in-depth studies on the ontology construction of professional domains in order to realize knowledge extraction from professional domain corpora over the past few years. The current research work on ontology construction methods mainly includes the skeleton method [6], the methodological method [7], and the seven-step method [8]. Geng et al. [9] constructed an ontology model of the phenomenon, cause, source, and maintenance plan. Qiang et al. [10] proposed a method for building a machine tool fault diagnosis ontology, summed up four key concepts in fault detection, and established a machine tool fault detection core body. To sum up, in recent years, the research focus of fault detection ontology construction is the event concept of fault phenomenon, structure, and cause. However, device causes and phenomena are mainly described by unstructured patterns, and considering they are event entity knowledge, the structure is not optimal.

## 2.2. Entity Recognition Method

In order to construct a comprehensive and detailed fault detection event graph, it is necessary to extract knowledge from a large amount of data. Knowledge extraction relies on entity recognition as a key technology. The current entity recognition technology mainly improves the recognition effect in two ways: by improving the input features and by improving the neural network structure. The following works are related to improving input features. Zhang et al. [11] proposed to use of Lattice LSTM for Chinese NER tasks. Gui et al. [12] proposed lexicon rethinking convolutional neural network (LR-CNN). Sui et al. [13] proposed a collaborative graph network (CGN) to solve the information loss problem of Lattice LSTM. Gui et al. [14] proposed a dictionary-based graph neural network (LGN) that aggregates local information through a graph structure and adds global nodes to integrate global information. Li et al. [15] proposed a flat-lattice transformer (FLAT), which has excellent parallelization ability. Ding et al. [16] proposed a graph neural network method with a multi-graph structure, which can effectively integrate rich place name information. Ma et al. [17] proposed Simplified Lattice LSTM, a method that incorporates all matching words into a Chinese character-based NER model. Liu et al. [18] proposed word-character LSTM (WC-LSTM) to obtain vocabulary boundary information while reducing word segmentation wrong influence. The following related works have been performed in order to improve the neural network structure. Zhu et al. [19] proposed CAN to capture information from adjacent characters and sentence context. Liu et al. [20] proposed a method CN3, which can both dynamically construct a task-specific structure of sentences and exploit rich local dependency information. Chen et al. [21] proposed GRN, which fused the local context features into the global context features. Xu et al. [22] proposed ME-CNER to derive rich semantic information from various granularities. Xu et al. [23] proposed an attention-based neural network architecture to exploit document-level global information from an electronic medical record corpus. In summary, the existing methods can improve recognition accuracy by enhancing the vocabulary, but at the cost of information loss, and due to the blurred boundaries of the Chinese unit vocabulary, word segmentation errors are likely to occur, leading to the problem of error propagation. Furthermore, most methods do not include entity relations for semantic reasoning.

## 3. Methodology

Next, we will introduce the labeling method of entities and relationships for mechanical fault detection, and describe the details of MF-GCN.

### 3.1. Event Argument Entity and Relationship Labeling Strategies

In order to construct event logic ontologies for fault diagnosis, we use Protégé prior to data labeling, as shown in Figure 1. Inspired by [24–26], the task of fault detection event argument entity recognition is converted to a sequence labeling task. Entities and relationships are predefined based on recommendations from experts in the field of mechanical fault detection. An entity follows the BIO pattern, in which B indicates the start of the entity, I indicates the middle of the entity, and O indicates none. The labeling process is as follows. Firstly, we build the ontology model based on the expert's advice. Based on the ontology model, 7 argument entities Attribute, Attribute value, Equipment, Sub equipment, Component, Part and Status value and 4 argument relations Lead To, Consist Of, Has Attributes, and Appear are defined. Appear, Lead To are event trigger words. Consist Of, Has Attributes are qualitative knowledge relations. In the second stage, we pre-label the data samples according to the labeling method. We discuss the solution when uncertain words appear, and update the labeling method. In the third stage, we divided into two groups for formal annotation and compared the annotation results to ensure the accuracy of the annotation. Figure 2a,b illustrate our labeling method with a sample sentence as an example. A sample sentence is as follows: The winding of the motor has become damp and the insulation has deteriorated causing the housing to be overcharged. Seven triples are contained in the sentence. Lead To, Consist Of, Has Attributes, and Appear are predefined

relation categories. The relationship category is labeled as N if the entity does not have any relationships.

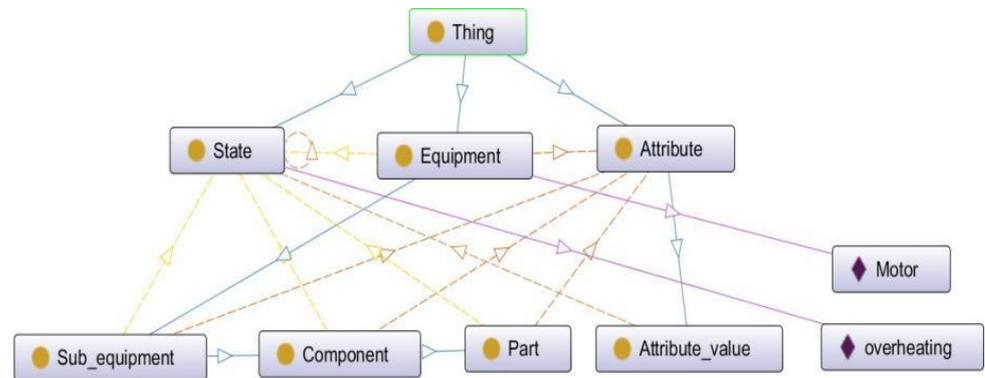


Figure 1. Mechanical fault detection ontology construction schematic.

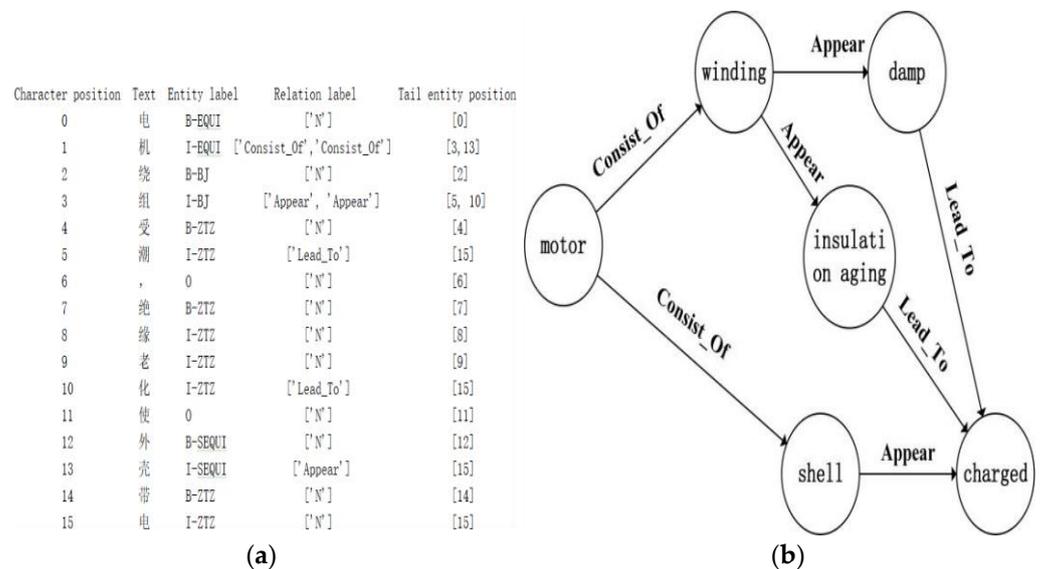
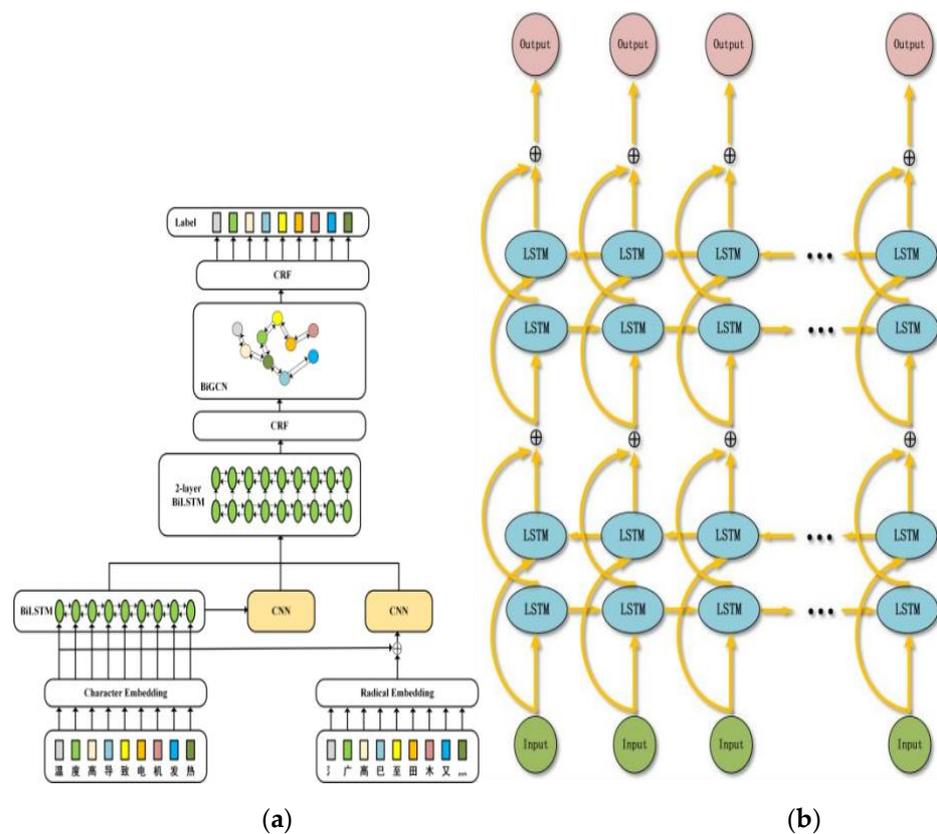


Figure 2. Example of argument entity and relation annotations and relation diagrams for sample sentences. The sentence in (a) means that the motor winding is affected with damp and the insulation aging makes the shell live. An entity follows the BIO pattern, in which B indicates the start of the entity, I indicates the middle of the entity, and O indicates none. A relationship is connected by an entity tail to another entity tail. (b) shows the entity relationship of the sentence in (a). It shows that the motor is composed of winding and shell. The shell is charged due to damp and insulation aging.

### 3.2. Model Structure

The MF-GCN consists of three layers: a feature fusion layer, a stacked BiLSTM layer, and a BiGCN entity relation inference layer. Figure 3a illustrates the model structure of this paper.



**Figure 3.** The structure of the model. (a) the text on the left means that high temperature causes the motor to heat up. (b) is the radical sequence of the left sentence.

*Feature Fusion Layer:*

*Context Features.* Context features enable entities to capture information about long-distance dependencies. For a text sentence  $T = t_1, t_2, \dots, t_n$ ,  $t_i$  is the character in the text sentence. Through the Word2vec embedding layer to obtain the character vector representation of the sentence  $C = (x_1, x_2, \dots, x_n)$ , after the vector  $C$  is Dropout processed, it is input to Bi-LSTM so that each character of the sentence obtains the context feature, and output this vector representation  $X = (o_1, o_2, \dots, o_n)$  with context features extracted.

*Context Salient Features.* Information on long-distance dependency and local context salient features can be combined better with context salient features. Using CNNs (Convolutional Neural Networks) for convolution operations, we obtain the vector representation  $Y$  of the salient context feature.

*Local Context Features.* Both of the above features only consider character-level embeddings and do not consider radical features. In order to maximize the benefit of radical features, this paper proposes a method for extracting local context features using radical information. Local context features are important in determining entity semantics. For each character  $t_i$  in the text sentence  $T = \{t_1, t_2, \dots, t_n\}$ , extract the radical  $r_i$  of each character, and obtain the radical sequence  $R = r_1, r_2, \dots, r_n$  of the sentence. Then, through the radical embedding layer, the radical sequence  $R$  is transformed into the radical-level vector representation  $D$ . The radical-level vector  $D$  and the character-level vector  $C$  are merged, and after Dropout processing is performed, they are input into the CNN to extract the local context feature vector  $Z$  that combines the radical information and character information. Local context features allow each character to obtain context information from adjacent characters, and these features are important for recognizing event argument entities in fault detection datasets. The formula for calculating the feature vector  $Z$  is as follows:

$$Z = Conv(D \oplus C) \tag{1}$$

*Stacked BiLSTM Layer.* The feature vectors  $X$ ,  $Y$ , and  $Z$  are merged and input into a stacked two-layer Bi-LSTM, and the output entity hidden layer represents the sequence  $H = (x_1, x_2, \dots, x_n)$ . Figure 3b illustrates the specific structure of Bi-LSTM stacked with two layers. By stacking stacked Bi-LSTMs, long-range dependencies can be captured, making up for the disadvantage that the Bi-GCN module, which will be used later, cannot encode long-range information. Finally, we input the output vector of the stacked Bi-LSTM into the conditional random field (CRF) to obtain the first-stage entity prediction result.

*BiGCN entity relation inference layer.* We utilize expert-defined and labeled event-argument-entity relations in the dataset to construct graph structures from text sequences. Given that there are head and tail entities involved in the entity-relationship, this paper uses BiGCN to construct a weighted connection graph for text characters, to obtain the relationship adjacency matrix between event argument entities, to extract regional node features, and to update global context features. By using the manually labeled relationship information, the entity boundary is modified in the second stage of the modeling process to improve the accuracy of the entity boundary determination. There are weighted-weighted connection graphs if there are  $k$  relations. Assuming that a text sequence of  $n$  characters is given, with characters as nodes, an  $n \times n$  adjacency matrix  $A$  can be used as the entity-relationship adjacency matrix of each node for a certain relationship. If a Bi-GCN has  $l$  layers,  $h^{l-1}$  is the hidden layer vector of node  $i$  at layer  $l-1$ , and  $h^l$  is the output vector of node  $i$  at layer  $l$ . Each node obtains the relationship information with other adjacent nodes through the relationship adjacency matrix and passes it as input to the next layer.

Finally, the output vector of Bi-GCN is input to CRF to obtain the final entity prediction result. The joint probability of label sequence and feature sequence is as Equation (2):

$$p(M, Y) = \sum_{i=1}^n A_{y_{i-1}, y_i} + p_{i, y_i} \quad (2)$$

$$\text{Loss}_{ent} = \underset{Y' \in f(M)}{\text{argmax}} \log(p(M, Y) - \log \left( \sum_{Y' \in f(M)} e^{p(M, Y')} \right)) \quad (3)$$

$$\text{loss}_{total} = \text{loss}_{1ent} + \text{loss}_{2ent} \quad (4)$$

Throughout the model training, the total loss is calculated as the sum of the two-stage event argument entity recognition losses, defined as equation.

## 4. Experiment

### 4.1. Experimental Settings

**Dataset.** This paper uses the labeled mechanical fault relevant text as the training data of the NER model, with a total of 900 sample sentences and a total of 5860 labeled entities. For entity labeling, the BIO labeling method is used. In addition, this paper also needs to combine relational information to predict entities. This paper uses four relation types to label the relation between event argument entities, as follows: Appear, Lead To, Consist Of and Has Attributes.

**Evaluation index.** In this paper, the accuracy rate  $P$  (Precision), the recall rate  $R$  (Recall), and the F1 score are used as indicators to evaluate the performance of the model.

**Model hyperparameter settings and environment configuration.** The model proposed in this paper is implemented based on the Tensorflow framework, and the experiment will perform 5-fold cross-validation. Adam optimizer is used during training. The dropout rate is set to 0.9 to prevent overfitting. The experiments performed in this paper all use the same configuration.

### 4.2. Model Performance Comparison

We compared the model MF-GCN with 5 models that perform well in public datasets. Five comparison models are presented: Lattice LSTM, WC- LSTM, Simplified Lattice LSTM,

LR-CNN, and Bert-CRF. In comparison with Lattice LSTM and WC-LSTM, the information loss problem of the latter two is not present in MF-GCN. In comparison with Simplified Lattice LSTMs and LR-CNN, although they solve the problem of information loss, they are not capable of exploiting the entity-relationship information that MF-GCN does. The scores of the above models are shown in Table 1. MF-GCN has achieved the best recognition effect in the mechanical fault detection dataset, with an F1 score of 93.58%.

**Table 1.** Comparison of the effect of different models on the event argument entity recognition of the mechanical fault detection dataset (unit:%).

Model	P	R	F1 Score
Lattice LSTM [11]	90.67	92.12	91.39
Simplified Lattice [17]	92.06	92.69	92.37
LR-CNN [12]	92.29	93.44	92.86
Bert-CRF [27]	92.31	93.72	93.01
<b>MF-GCN</b>	<b>93.78</b>	<b>93.38</b>	<b>93.58</b>

Additionally, we conduct comparative experiments to demonstrate the effectiveness of the MF-GCN feature fusion layer on the Weibo dataset. This dataset is a Chinese social media (Weibo) named entity recognition dataset (Weibo-NER-2015), which contains 1890 pieces of information collected from Weibo between November 2013 and December 2014, with a total of 1890 samples. The training set is 1350, the validation set is 270, and the test set is 270. The four labels are PER, LOC, GPE, and ORG. When the Weibo dataset is used for entity recognition tasks, the fusion of multiple text features can improve the recognition effect since the corpus and semantics are sparse. As the Weibo dataset lacks labeled event argument entity relationships, MF-GCN must remove the BiGCN entity relation inference layer for experiments. The final results are shown in Table 2. MF-GCN still achieves a good recognition effect after removing the BiGCN entity relation inference layer.

**Table 2.** Comparison of the recognition effects of different models on the Weibo dataset (unit:%). NE, NM, and Overall represent the F1 scores for named entities, nominal entities (excluding named entities), and both.

Model	NE	NM	Overall
Lattice LSTM [11]	53.04	62.25	59.79
WC-LSTM [18]	52.55	67.41	59.84
Simplified Lattice [17]	56.99	61.41	61.24
LR-CNN [12]	57.14	66.67	59.92
ME-CNER [22]	75.17	64.39	68.93
<b>MF-GCN(Remove Bi-GCN module)</b>	<b>74.93</b>	<b>64.84</b>	<b>69.52</b>

#### 4.3. Ablation Experiment

We demonstrate the effectiveness of each module in improving the performance of the model by removing each module from the model and comparing it with the original recognition effect. The results are shown in Table 3. When the context feature is removed, the F1 score is reduced to 92.95%, which shows that the con-text feature can enable each entity in the sentence to capture long-distance dependency information and improve the performance of the model. When the context salient feature is removed, the F1 score is reduced to 93.32%, which proves that combining local context information and long-distance dependency information can improve the effect of entity recognition. When the local context features are removed, although the Recall is increased to 94.35%, the Precision is only 92.53%, and the F1 score is reduced to 93.44%, which indicates that the use of local context features to infer entity semantic information has improved model performance. In the absence of the relationship information, that is, the Bi-GCN module, the F1 score is reduced to 90.30%, indicating that the addition of the event argument entity relationship

information greatly improves the effect of event argument entity recognition. In the case of Bi-LSTM with different stacked layers, when the stacked layer is 2, the F1 score is the highest. During the experiment, the model that removes the context feature, context highlight feature, local context feature, and Bi-GCN module all adopt the Bi-LSTM that is stacked with two layers.

**Table 3.** Ablation experiment of MF-GCN in mechanical fault detection dataset (unit:%).

Model	P	R	F1 Score
MF-GCN(Remove context features)	92.59	93.30	92.95
MF-GCN(Remove context salient features)	93.68	92.96	93.32
MF-GCN(Local context features)	92.53	94.35	93.44
MF-GCN(Remove Bi-GCN module)	90.26	90.33	90.30
MF-GCN (1 layer Bi-LSTM)	93.47	93.47	93.47
<b>MF-GCN(2 layers stacked Bi-LSTM)</b>	<b>93.78</b>	<b>93.38</b>	<b>93.58</b>
MF-GCN(3 layers stacked Bi-LSTM)	93.52	92.96	93.24

## 5. Conclusions

In the field of mechanical fault detection, the corpus is relatively small in size, with many technical terms and sparse semantics. The existing methods fail to integrate multiple text features and relational information well, and often perform poorly when performing event argument entity recognition on corpora in the field of mechanical fault detection. In this paper, we propose a model that fuses three textual features and incorporates entity-relationship information, achieving an F1 score of 93.58%. The method outperforms other typical models on the mechanical fault detection dataset. In our work, we demonstrate that by integrating more textual features and incorporating entity-relationship information, it is possible to improve the model's event argument entity recognition performance. Although our model has achieved good results in the recognition of event argument entities on the mechanical fault relevant corpus, there is still room for improvement. In our next step, we will expand the mechanical fault relevant corpus and use a portion of the corpus for pre-training character embedding layers in order to obtain a more accurate model.

**Author Contributions:** Writing—original draft, S.C.; Writing—review & editing, J.D.; Supervision, T.W.; Project administration, L.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was financially supported by the Supply chain collaborative service platform for large manufacturing enterprises (2019YFB1705500) and the Key Program of NSFC-Guangdong Joint Funds (U1801263).

**Data Availability Statement:** Data not available due to project restrictions.

**Acknowledgments:** The authors thank the editor and anonymous reviewers for their careful work and thoughtful suggestions that have helped improve this paper substantially.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Zhao, W.; Wang, K.; Xi, A.; Zeng, P.; Yang, S.; Sun, Y.; Guo, H. An industrial robot health assessment method for intelligent manufacturing. *Robot* **2020**, *42*, 460–468.
- Stuhlenmiller, F.; Weyand, S.; Jungblut, J.; Schebek, L.; Clever, D.; Rinderknecht, S. Impact of cycle time and payload of an industrial robot on resource efficiency. *Robotics* **2021**, *10*, 33. [[CrossRef](#)]
- Chen, L.; Cao, J.; Wang, X. Nonlinear spectrum based fault diagnosis method for industrial robot drive systems. *J. Xi'an Jiaotong Univ.* **2019**, *53*, 93–99.
- Huan-Kun, H.S.U.; Hsiang-Yuan, T.; Huang, M.; Huang, H. Intelligent fault detection, diagnosis and health evaluation for industrial robots. *Mechanics* **2021**, *27*, 70–79.
- Sun, Z.; Yuan, X.; Fu, X.; Zhou, F.; Zhang, C. Multi-scale capsule attention network and joint distributed optimal transport for bearing fault diagnosis under different working loads. *Sensors* **2021**, *21*, 6696. [[CrossRef](#)] [[PubMed](#)]

6. Zhou, L.; Zhang, C.; Karimi, I.A.; Kraft, M. An ontology framework towards decentralized information management for eco-industrial parks. *Comput. Chem. Eng.* **2018**, *118*, 49–63. [[CrossRef](#)]
7. Martinez-Garcia, J.R.; Castillo-Barrera, F.; Palacio, R.R.; Borrego, G.; Cuevas-Tello, J.C. Ontology for knowledge condensation to support expertise location in the code phase during software development process. *IET Softw.* **2020**, *14*, 234–241. [[CrossRef](#)]
8. Guan, H.; Xie, D. Research on building of acupuncture domain ontology. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017; pp. 1–4.
9. Geng, D.; Fu, X. Research on fault diagnosis mechanism of production line equipment based on semantic. In Proceedings of the 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 17–19 July 2020; pp. 220–223.
10. Zhou, Q.; Yan, P.; Xin, Y. Research on a knowledge modelling methodology for fault diagnosis of machine tools based on formal semantics. *Adv. Eng. Inform.* **2017**, *32*, 92–112. [[CrossRef](#)]
11. Zhang, Y.; Yang, J. Chinese ner using lattice lstm. *arXiv* **2018**, arXiv:1805.02023.
12. Gui, T.; Ma, R.; Zhang, Q.; Zhao, L.; Jiang, Y.; Huang, X. Cnn-based chinese ner with lexicon rethinking. *IJCAI* **2019**, 4982–4988.
13. Sui, D.; Chen, Y.; Liu, K.; Zhao, J.; Liu, S. Leverage lexical knowledge for chinese named entity recognition via collaborative graph network. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 3830–3840.
14. Gui, T.; Zou, Y.; Zhang, Q.; Peng, M.; Fu, J.; Wei, Z.; Xu, H. A lexicon-based graph neural network for chinese ner. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 1040–1050.
15. Li, X.; Yan, H.; Qiu, X.; Huang, X. Flat: Chinese ner using flat-lattice transformer. *arXiv* **2020**, arXiv:2004.11795.
16. Ding, R.; Xie, P.; Zhang, X.; Lu, W.; Li, L.; Si, L. A neural multi-digraph model for chinese ner with gazetteers. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019; pp. 1462–1467.
17. Ma, R.; Peng, M.; Zhang, Q.; Huang, X. Simplify the usage of lexicon in chinese ner. *arXiv* **2019**, arXiv:1908.05969.
18. Liu, W.; Xu, T.; Xu, Q.; Song, J.; Zu, Y. An encoding strategy based word-character lstm for chinese ner. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, MN, USA, 2–7 June 2019; pp. 2379–2389.
19. Zhu, Y.; Wang, G.; Karlsson, B.F. Can-ner: Convolutional attention network for chinese named entity recognition. *arXiv* **2019**, arXiv:1904.02141.
20. Liu, P.; Chang, S.; Huang, X.; Tang, J.; Cheung, J.C.K. Contextualized non-local neural networks for sequence learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 29–31 January 2019; Volume 33, pp. 6762–6769.
21. Chen, H.; Lin, Z.; Ding, G.; Lou, J.; Zhang, Y.; Karlsson, B. Grn: Gated relation network to enhance convolutional neural network for named entity recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 29–31 January 2019; Volume 33, pp. 6236–6243.
22. Xu, C.; Wang, F.; Han, J.; Li, C. Exploiting multiple embeddings for chinese named entity recognition. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 2269–2272.
23. Xu, G.; Wang, C.; He, X. Improving clinical named entity recognition with global neural attention. In *Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data, Macau, China, 23–25 July 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 264–279.
24. Zeng, X.; Zeng, D.; He, S.; Liu, K.; Zhao, J. Extracting relational facts by an end-to-end neural model with copy mechanism. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia, 15–20 July 2018; pp. 506–514.
25. Tian, J.; Lu, X.; You, X.; Xiao, G.; Han, J. Joint extraction of entities and relations based on hierarchical sequence labeling. *Beijing Da Xue Xue Bao* **2021**, *57*, 53–60.
26. Wu, S.; Liang, X.; Xie, N.; Zhou, A.; Hao, X. Annotation method for joint extraction of domain-oriented entities and relations. *J. Comput. Appl.* **2021**, *41*, 2858.
27. Chang, Y.; Kong, L.; Jia, K.; Meng, Q. Chinese named entity recognition method based on bert. In Proceedings of the 2021 IEEE International Conference on Data Science and Computer Application (ICDSCA), Dalian, China, 29–31 October 2021; pp. 294–299.