

# A knowledge Extraction Method of Power Equipment Operation and Inspection for Multimodal Data

Xinghui Zhang<sup>a</sup>, Tongyang Liu<sup>b</sup>, Heqi Zheng<sup>b</sup>, Fengda Zhang<sup>c</sup>, Qiang Zhang<sup>b\*</sup>

<sup>a</sup>State grid Corporation of China, Beijing, China

<sup>b</sup>State Grid Smart Grid Research Institute Co., Ltd., Beijing, China

<sup>c</sup>State Grid Shandong Electric Power Company, Jinan, China

\* Corresponding author's email: zhangqiang1@geiri.sgcc.com.cn

Supported by the Science and Technology Project of State Grid Corporation of China: Research on Key Technologies of Multimodal Knowledge Computing Engine for Power Grid Main Equipment Fault (5700-202258449A-2-0-ZN).

## Introduction

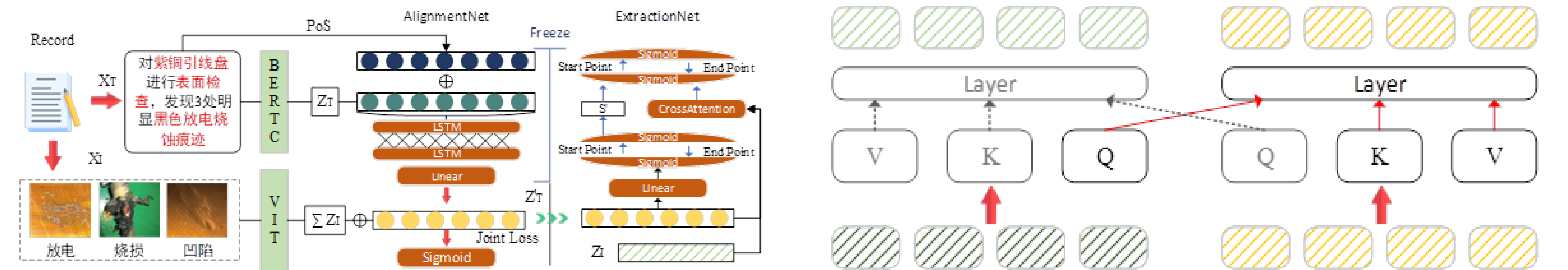
Power equipment operation and inspection is one of the critical links to maintaining the regular operation of electric power enterprises, and extracting the names of electric power equipment and the corresponding operation and inspection relationships to build a domain knowledge graph is a crucial technology for realizing knowledge precipitation in the electric power domain. However, considering MNER and MRE as pipeline architecture leads to model error propagation while ignoring the semantic relationship between power equipment entities and relations. In particular, in power equipment operation and inspection, the object in which a fault or defect occurs has a strong causal relationship with the cause of the fault. For example, the cause of the failure of “the main transformer No. 2” may be "high-temperature overheating" or "localized discharge," and the on-site inspection image shows "multiple ablation traces." Multiple traces of ablation". From this, it can be initially determined that the cause of the fault is "partial discharge."

## Contribution

- Analyse the multimodal data generated from the fault description, field inspection, disassembly inspection, and cause analysis of the power operation and inspection process.
- Propose a multimodal data entity relationship extraction model in the electric power field to improve the accuracy of professional knowledge extraction in the electric power field

## Methods

The overall model architecture is shown below. A multimodal device alignment network (Alignment Net) is constructed to find the name of the device body in the current text record and the image that agrees with the name of the device body from a collection of candidate O&M images, and the network is trained using the designed joint loss. Then, the obtained text vectors and image vectors are inputted into the knowledge extraction network (Extraction Net), which firstly extracts the device location information by using the two-pointer network classifier and then fuses the textual information and image information by the constructed multimodal cross-attention mechanism to enhance the semantic information about the reasons for the transportation inspection.



## Results

In this study, we designed a comparison experiment on the public dataset and a model structure ablation experiment on the power equipment record dataset for the proposed multimodal data-oriented power equipment operation and inspection entity relationship extraction method.

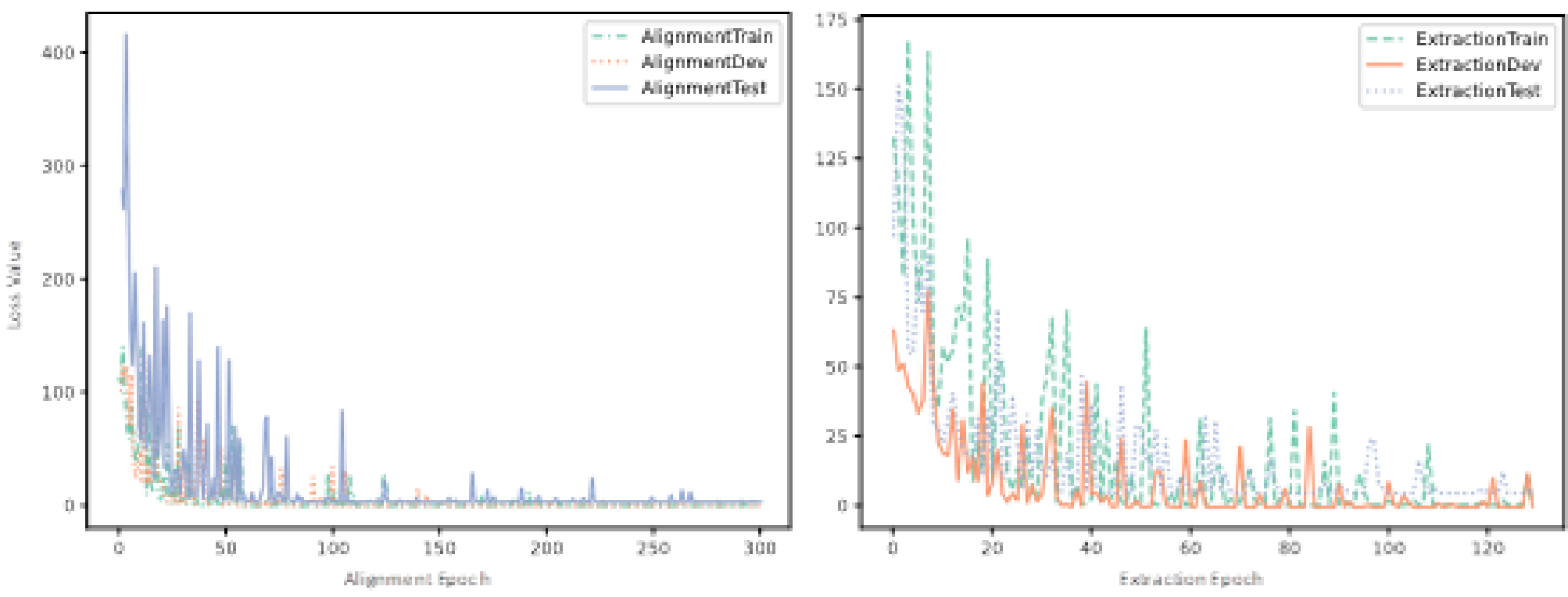


Figure 4. Loss function value.

Table 1. Electric power public opinion relationship.			
ID	Relation Name	Abbreviations	Volume
1	Failure Cause	FC	531
2	Parameters of the equipment	PA	203
3	Non-compliance	NC	54
4	Checking operations	CO	339
5	Factory	FA	152
6	Failure Consequences	CF	464
8	Is	IS	80
9	Other	O	20

Table 2. The result of public datasets.			
Model	Text Match Image		
	R@1	R@5	Rs
SCAN	48.6	77.7	126.3
DPRNN	55.5	81.3	136.8
SGRAF	58.5	83.0	141.5
Ours	64.3	88.1	152.4

Table 3. The results of the ablation tests.			
Model	P	R	F1
+Image	72.1	73.2	72.6
+Subject	63.1	66.8	64.8
ours	77.3	74.1	75.6

## Conclusions

- Compared with other entity-relationship extraction methods, we consider the multimodal data characteristics in the electric power operation and inspection domain and propose a network structure that specifically addresses image-text alignment and knowledge extraction, which improves the extraction accuracy of models in the domain.
- The proposed method is subjected to comparison and ablation experiments on independently constructed power equipment-related and public datasets, which validate the method's effectiveness in both public and domain-specific scenarios.

## References

[1] Zhou B, Zhang Y, Song K, et al. A Span-based Multimodal Variational Autoencoder for Semi-supervised Multimodal Named Entity Recognition, in Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, 2022, pp.6293-6302.

[2] Zhao Q, Gao T, Guo N. TSVFN: Two-Stage Visual Fusion Network for multimodal relation extraction. Information Processing & Management, vol.60 no.3, Art.no.103264, 2023.

[3] Yuan L, Cai Y, Wang J, et al. Joint multimodal entity-relation extraction based on edge-enhanced graph alignment network and word-pair relation tagging, in Proceedings of the AAAI conference on artificial intelligence. Vancouver, Canada, 2023, vol.37, no.9, pp.11051-11059.