Multi-modal big data knowledge aggregation for advanced automobile intelligent manufacturing operation and maintenance

Yang Rui1*, Xu Zhijie1, Xie Tianjie2, Zhang Yu1

1Guangdong Institute of Intelligent Manufacturing, Guangdong Key Laboratory of Modern Control Technology, Guangzhou 510070, Guangdong, China
2Cloud Computing Center Chinese Academy of Sciences, Dongguan 523808, Dongguan China

*yangrui@casc.ac.cn *msyxzj@sina.com

Abstract-In this study, automatic knowledge extraction technology is analyzed for advanced automobile intelligent manufacturing operation and maintenance multi-modal information with knowledge graph construction method and complex knowledge reasoning method. An association representation model is proposed based on deep feature learning, a unified multi-modal semantic space is constructed, and a semantic matching technology for multi-modal equipment operation and maintenance is designed based on the reasoning method of deep recurrent neural network.

1. Introduction
The core of the "Made in China 2025" strategy refers to intelligent manufacturing. Equipment operation and maintenance is considered to be an important means to ensure the safe, stable and economic operation of equipment; it is critical to achieve advanced automotive intelligent manufacturing. As fueled by the continuous improvement of complexity and intelligence, equipment failures show multiple causes and effects, and conventional fault handling methods based on experience or mechanism are no longer comprehensively applicable. With more demands for environmental protection, energy consumption optimization and quality improvement, higher requirements are raised with respect to efficiency and safety experience. [1] Accordingly, equipment operation and maintenance is facing severe challenges, and massive, multi-source, heterogeneous equipment operation and maintenance big data should be integrated to achieve intelligent equipment operation and maintenance. [2] There are considerable multi-modal and valuable equipment operation and maintenance information in equipment historical operation and maintenance records and literature (e.g., various troubleshooting methods, equipment maintenance cases, and equipment operation data collected by various sensors). [3] The equipment operation and maintenance knowledge map is capable of extracting, organizing and managing knowledge from considerable equipment operation and maintenance multi-modal information, and offering users intelligent operation and maintenance services. However, it is subject to the problems below: first, the multi-modal information of voice, video, picture and other equipment operation and maintenance exhibits different levels of knowledge expression, making conventional text-oriented knowledge extraction methods no longer applicable;
second, different modal content is characterized. and huge differences exist in representation and
distribution, causing modal gaps, and different modal knowledge is difficult to effectively integrate;
third, the direct application of conventional knowledge reasoning methods to multi-modal equipment
operation and maintenance knowledge graphs is poor and costly, which is hard to migrate.

2. Challenge

2.1. Automatic knowledge extraction technology of multi-modal information
There is considerable equipment operation and maintenance information in the Internet (e.g., Industrial
Maintenance Network, China Maintenance Network), historical operation and maintenance records of
the equipment itself, and physical documentation (e.g., equipment management and maintenance work
experience summary exchanges, various troubleshooting methods, [4] equipment maintenance cases,
equipment operating procedures, and equipment operating data collected by various sensors). [5] The
mentioned information is usually multi-modal and multi-source heterogeneous in the form of text, voice,
video, images and others, containing valuable operation and maintenance knowledge; however, for text,
voice, video, images and other equipment operation and maintenance, multi-modal information has
different levels of knowledge expression to varying degrees. [6] Conventional knowledge extraction is
primarily text-oriented. There are few researches on the extraction technology of multi-modal
information (e.g., voice, video and images). The fusion and representation ability of knowledge of
different modal characteristics (e.g., poor, low accuracy and recall rate). To generate a knowledge graph
for industrial equipment operation and maintenance multi-modal information, the automatic extraction
technology of multi-modal information should be explored.

2.2. Method for constructing knowledge graph of multi-modal information
The so-called multi-modal knowledge map refers to the analysis and mining of multi-modal knowledge
via visual, audio, language and other perceptions for supplementing and expanding the conventional
text-based knowledge system to form intelligent processing capabilities for multi-modal data. The
purpose of establishing a multi-modal knowledge graph is to present a basic and computable knowledge
expression structure and conduct semantic relationship analysis and cognitive-level reasoning in a
multi-modal environment. The knowledge extracted from the multi-modal data of industrial equipment
operation and maintenance may have considerable noises and redundancy. [7] Only by organically
fusing the mentioned knowledge can a larger-scale multi-modal knowledge graph be established. By
knowledge fusion, the ambiguity of concepts, redundancy and wrong concepts can be eliminated, and
the quality of knowledge can be ensured.

2.3. Complex knowledge reasoning method based on multi-modal information
The industrial equipment operation and maintenance knowledge map extracts, organizes and manages
knowledge from considerable equipment operation and maintenance multi-modal information, hoping
to offer users intelligent services that help understand user needs (e.g., understanding search semantics
and providing more accurate search answers). [4] This involves knowledge reasoning oriented to
knowledge graphs, and conducts the in-depth analysis and reasoning of equipment operation and
maintenance multi-modal information. Most of the existing knowledge graphs adopt conventional
knowledge reasoning methods, whereas both rules and abstract-level ontology constraints should be
instantiated, which is relatively poor in calculability. Industrial equipment operation and maintenance
knowledge has considerable examples and a wide range of content. In terms of graphs, the cost is high;
effective and wide-covered rules and ontology constraints are also difficult to obtain, causing a low
recall rate of inference results. [8] On the other hand, statistical features rely too much on existing data
and are hard to migrate, which is difficult to deal with sparse samples; when the data is noisy, the
extracted features may even mislead the reasoning.

3. Intelligent operation and maintenance platform
To address the mentioned challenges, this study primarily analyzes automatic knowledge extraction technology, knowledge map construction method, complex knowledge reasoning method for advanced automobile intelligent manufacturing operation and maintenance multi-modal information. Besides, a set of industrial equipment intelligent operation is developed based on multi-modal knowledge map maintenance system for intelligent operation and maintenance of equipment (Fig. 1).

Figure 1: Realize advanced automobile intelligent manufacturing

Lastly, the following technologies are adopted to carry out the knowledge aggregation and application of multi-modal big data in the operation and maintenance of advanced automobile intelligent manufacturing:

3.1. Automatic knowledge extraction technology for multi-modal information of industrial equipment operation and maintenance

To obtain the massive equipment operation and maintenance multi-modal information in the network exhibiting high efficiency and high quality, and to effectively construct the knowledge entity corpus for the equipment operation and maintenance field, this study aims to adopt distributed parallel acquisition technology and incremental acquisition mechanism, as well as designing a conceptual semantic analysis-based Crawling strategies and algorithms. Protege-owl (an ontology construction tool) is adopted to construct the equipment operation and maintenance service ontology, web content and useful information are extracted from the web ontology language (owl) tags to be converted into metadata, and extended case-based reasoning ECBR (Extended Case-Based Reasoning) algorithm is employed to calculate the correlation between metadata and the concepts in the ontology. If the correlation is greater than the threshold, the correlation is stored in the metadata attribute. Besides, the multi-modal information of equipment operation and maintenance is stored in metadata attributes. Before the automatic knowledge extraction of equipment operation and maintenance multi-
modal information, semantic features should be extracted from equipment operation and maintenance multi-modal information. This study intends to use convolutional neural networks to extract visual features of images and videos, and use LSTM deep neural network to extract the features of text and speech, and employ the multi-modal deep belief network as the feature fusion model for the effective fusion of different modal features. By feature fusion, the codec framework of deep learning is adopted to achieve the associated modeling of multimodal information and language, the multimodal features are introduced into the bidirectional sequence LSTM encoder, and the LSTM is used to encode the information stream. The backward and backward bidirectional sequence encoding captures the context vector of future information, as well as using the attention mechanism to acquire significant information. The language generation module employs deep reinforcement learning and two sampling methods, i.e., random sampling and maximum sampling, to generate predicted sentences in the LSTM decoder. The language description of multi-modal information can be obtained with this model, as an attempt to extract knowledge elements (e.g., entities, relationships, classifications and attributes). For the particularity of the field of equipment operation and maintenance entities, the existing entity extraction model no longer applies to extracting equipment operation and maintenance entities. The existing entity extraction algorithm should be optimized. This study focuses on the lack of corpus in the professional field. A BiLSTM-CNN-CRFs network model is proposed based on deep learning of the attention mechanism, and a domain named entity recognition method is proposed based on feedback semantic transfer learning. First, part of the equipment operation and maintenance entity training data set are constructed, the equipment operation and maintenance domain corpus and the general domain corpus are trained to obtain the corpus document vector, the semantic similarity between the domain corpus and the general corpus is calculated, semantic transfer learning is performed on the domain samples, and multiple Transfer corpus is constructed. Subsequently, the BiLSTM-CNN-CRFs network model is adopted to perform domain named entity recognition on the migrated corpus. For relationship extraction, it is planned to adopt the method of unsupervised automatic construction of entity relationship trigger words, model the entity relationship data sets, filter by word probability weights, filter the trigger word sets again with syntactic analysis tools, and finally construct equipment operation dimensional entity relationship trigger word dictionary. To extract relational entity pairs in the form of entity-attribute-value pairs, an unsupervised extraction method based on pattern clustering is adopted to abstract the entity relationship into a relationship model, and then the relationship model is further generalized to automatically extract entity relationships and attributes the goal of.

3.2. Knowledge graph construction method for multi-modal information of industrial equipment operation and maintenance

The primary problem of constructing knowledge graphs for equipment operation and maintenance multimodal information refers to the study on the representation learning technology for multimodal knowledge graphs. With the success of representation learning technologies in fields (e.g., image, video, language and natural language processing), some researchers have begun to study representation learning technologies for multi-modal knowledge graphs, transforming entities and relationships into a low-dimensional space, the real-valued vector (i.e., the distributed semantic representation). This study intends to use multi-modal public representation learning to project the features of text, images, voice, video and other multi-modal data into the same representation space. Such structure can integrate the contextual information into the multi-modal semantic space, as well as ensuring the coherence in time sequence and the diversity of semantic embedding. By designing a reasonable representation learning model, and projecting the information extracted from different modal data into the identical semantic space, a unified representation space can be established, and multi-modal knowledge fusion can be achieved. The core of entity linking is to calculate the similarity between the mention and the entity in the knowledge base, and to select the target entity mentioned by a specific entity based on the similarity. The core of the mentioned process refers to mining the evidence information that can be used to identify and referencing the target entities, expressing these evidences in a form for computer processing, and developing high-performance algorithms to synthesize different evidences to make link
decisions. Compared with conventional statistical methods, deep learning is capable of learning task-specific representations, establishing information associations between different modal data, and achieving better entity analysis performance. The application of deep learning unifies the mathematical tools applied in natural language processing and speech, image and video processing, thereby breaking the barriers between the mentioned different modal information, which enables the processing and fusion of multi-modal information. By exploiting the feature extraction capabilities of deep learning, the effective representation of different modalities can be extracted at the bottom level, and the semantic association of different modalities can be developed at the high level, as an attempt to more effectively capture the nonlinear correlation between different modal content.

3.3. Complex knowledge reasoning method for industrial equipment operation and maintenance multi-modal information

Knowledge reasoning method is based on deep recurrent neural network. The powerful learning ability of deep neural network is exploited to simulate the reasoning of the human brain, design a deep recurrent neural network (DRNN) to model and reason, and construct an auxiliary storage matrix to store intermediate results or necessary information required for reasoning. DRNN is adopted to learn and memorize the known triples (i.e., head entity, relationship and tail entity) in the knowledge graph, model the fact triples of the knowledge graph, and infer new triples. In the training process, it aims to imitate the human brain to use the existing experience knowledge for learning and reasoning novel knowledge and continuously updating the existing knowledge. The knowledge graph triplet acts as input for training, and the DRNN Read and write the storage matrix and constantly update the storage matrix. During the prediction, input the incomplete triples to be predicted (leave blank elements to be estimated) into the trained knowledge inference model, continuously interact with the storage matrix through DRNN, perform multi-step inference, and output the completed triple group. The optimization strategy of knowledge reasoning model complies with LSTM-RNN. Facing the increasingly large-scale equipment operation and maintenance knowledge graph, when constructing knowledge reasoning models based on deep cyclic neural networks, To weaken the loss of gradients and deal with the adverse effects of long-distance dependencies on neural networks, long and short-term memory network (LSTM) chain units are used to replace recurrent neural network (RNN) chain units to enhance the long-term memory ability of deep recurrent neural networks. Long and short-term memory recurrent neural network (LSTM-RNN) constructs a more optimized knowledge reasoning model. Moreover, when using LSTM-RNN to train large-scale knowledge inference models, over-fitting problems can be encountered, which seriously affects the generalization ability of the trained deep recurrent neural network. It is planned to adopt Dropout technology to prevent overfitting and further increase the robustness and generalization ability of the deep recurrent neural network, as an attempt to more effectively adapt to large-scale knowledge graphs Reasoning in knowledge. Parallel acceleration technology for knowledge inference based on Spark-GPU distributed memory computing platform. To increase the knowledge reasoning speed of the large-scale equipment operation and maintenance knowledge graph and ensure the timeliness of reasoning, the Spark-GPU distributed memory computing platform is used to parallelize and accelerate the reasoning process. Combining the existing most advanced distributed memory computing framework Spark and mainstream general-purpose computing hardware acceleration equipment GPU, a Spark-GPU distributed memory computing platform is built, a parallel research is conducted on knowledge inference methods based on deep recurrent neural networks, and JNI and storage Management and other relevant technologies that can be applied in a single node of the Spark distributed memory computing platform are employed. Call single or multiple GPUs to further accelerate the inference process in parallel.

4. conclusion

Based on multi-modal big data analysis computing cluster system architecture, develop semantic matching. Functional modules (e.g., configuration, intelligent question and answer); the scale of the deployment computing cluster is greater than 100 physical nodes, The data scale of storage and
management is more than 1PB, supports offline analysis and mining of PB-level data, online analysis of TB-level data, and provides explainable instructions for intelligent operation and maintenance solutions; high-performance data interaction middleware for industrial equipment is achieved, which can be integrated no less than 5 types of data sources; no less than 3 types of knowledge graph extraction models; knowledge graphs are established in industrial equipment, domain knowledge graphs or semantic databases are established in Chinese or other languages, and integrated semantic data of no less than 100,000 entities (Entities), triples (N-triples). The number of entries is no less than 1M, and the overall amount of semantic database data is no less than 200M. A unified multi-modal semantic space is constructed, the joint representation of multi-modal information of industrial equipment operation and maintenance is achieved at the knowledge level, the multi-modal knowledge graph of industrial equipment is constructed based on the reasoning method of deep recurrent neural network, and the increasing scale reasoning problem of multi-modal equipment operation and maintenance knowledge graph is solved. The semantic matching technology is developed for multi-modal equipment operation and maintenance problems, and a precise operation and maintenance knowledge question and answer system is built for user intent, with the semantic matching accuracy ≥90% and the recall rate ≥60%. An intelligent operation and maintenance platform is developed for industrial equipment, equipment operation and maintenance multi-modal data are fully exploited for intelligent equipment operation and maintenance that can be interpreted and calculated. The number of industrial equipment operation and maintenance data records reaches 1 million, and it supports concurrent access by 10,000 users, and each user gives the response. The time is less than 5 sec. The joint representation of multi-modal information of advanced automobile intelligent manufacturing operation and maintenance is presented at the knowledge level, the problem of excessive cost caused by multi-modal information feature selection is solved, and multi-modal knowledge is realized through the knowledge learning of multi-modal knowledge graph integration to build a multi-modal industrial equipment operation and maintenance knowledge map. The parallel acceleration of knowledge inference on the distributed memory computing platform is achieved, so the operation and maintenance knowledge graph of industrial equipment can be effectively completed, and the reasoning speed can be significantly up-regulated. Then, the intelligent question answering algorithm is developed based on the knowledge graph, and the industrial equipment intelligence is built based on the multi-modal knowledge graph Operation and maintenance.

The implementation architecture diagram is illustrated in Fig. 2.
ACKNOWLEDGMENT
This study is supported in part by GDAS’ Project of Science and Technology Development (2017GDASCX-0115), GDAS’ Project of Science and Technology Development (2018GDASCX-0115), and GDAS’ Special action for building a first-class research institution in China (2020GDASYL-20200402007).

REFERENCES
[1] G. Zhang, C.H. Chen, P. Zheng, R.Y. Zhong. An integrated framework for active discovery and optimal allocation of smart manufacturing services. Journal of Cleaner Production, pp. 273, 2020.
[2] Y.H. He, Y.X Zhao, X. Han, D. Zhou, W.Z Wang. Functional risk-oriented health prognosis approach for intelligent manufacturing systems. Reliability Engineering and System Safety, pp. 203, 2020.
[3] X. Li, S.H Siahpour, J. Lee, Y.C Wang, J. Shi. Deep Learning-Based Intelligent Process Monitoring of Directed Energy Deposition in Additive Manufacturing with Thermal Images. Procedia Manufacturing, pp. 48, 2020.
[4] X.Y Li, C.J Zhang. Some new trends of intelligent simulation optimization and scheduling in intelligent manufacturing. Service Oriented Computing and Applications, 3rd ed., vol. 3, 2020.

[5] A. Simeone, A. Caggiano, Y.F Zeng. Smart cloud manufacturing platform for resource efficiency improvement of additive manufacturing services. Procedia CIRP, pp. 88, 2020.

[6] N. Choudhury, F. Faisal, M. Khushi. Mining Temporal Evolution of Knowledge Graphs and Genealogical Features for Literature-based Discovery Prediction. Journal of Informetrics, 3rd ed., vol.14, 2020.

[7] B.Z Gao, J.L Hong, Q.U Ting, Y.U Shuyou, H. Chen. An Output Regulator with Rejection of Time-Varying Disturbance: Experimental Validation on Clutch Slip Control. IEEE Transactions on Control Systems Technology, DOI: 10.1109/TCST.2019.2900211

[8] H.Q Chu, L.L Guo, Y.J Yan, B.Z Gao, H. Chen, N. Blan. Energy-efficient longitudinal driving strategy for intelligent vehicles on urban roads. Science China (Information Sciences), 6th ed., vol.62, pp. 160-162, 2019